

When Opponents Silently Become Enemies: Explaining the Psychological Roots of Polarization in Democratic Societies

—
Matúš Grežo  |  matus.grezo@savba.sk

Institute of Experimental Psychology, Centre of Social and Psychological Sciences, Slovak Academy of Sciences, Bratislava, Slovakia

Abstract | Introduction: Affective polarization, when opposing ideological groups begin to view each other with hostility and moral contempt, represents a growing challenge in many democratic societies. Although democratic systems are based on pluralism of opinions, recent years have seen an escalation in value-laden conflicts that increasingly threaten social cohesion, particularly in culturally divided societies like Slovakia. Objectives and methods: This theoretical paper explores the psychological mechanisms behind the polarization, with a focus on how symbolic threats trigger defensive responses that may deepen ideological divides. The study builds upon the Intergroup Value Protection Model, which explains affective polarization as a process wherein individuals seek to protect their social embeddedness when they perceive an out-group's actions as threatening core moral values. The IVPM identifies multiple affective and behavioral responses to such threats. However, while the IVPM offers a robust framework for explaining group-level responses to value threats, it does not encompass the full spectrum of threat-defense mechanisms that may play a role in polarization. To address this gap, this study integrates the General Process Model of Threat and Defense (GPMTD) into the IVPM framework. Results: Drawing on the GPMTD model, the present study proposes three novel defensive responses to symbolic threat that may have the potential to escalate between-group conflicts and increase affective polarization in societies. First, the study highlights how people first engage in avoidance-oriented reactions, governed by the Behavioral Inhibition System. These proximal reactions – such as anxiety, withdrawal, or hypervigilance – may seem unrelated to polarization at first. Yet, research shows they can contribute to “silent radicalization” whereby individuals disengage from public discourse and increasingly seek out ideologically congruent and extremist narratives online. Second, exposure to threats can prompt individuals to engage in distal defenses of the Behavioral Activation System, such as reinforcing existing beliefs, intensifying moral convictions, and relying on confirmation bias and motivated reasoning. These responses may result in increased rigidity of worldviews and reduced openness to alternative perspectives. Third, the paper discusses how symbolic or existential threats can drive individuals to conspiratorial thinking, which in turn intensifies intergroup prejudice, institutional distrust, and support for antisocial or violent behavior. Discussion and conclusion: Integrating these three new defensive reactions into the IVPM allows for a richer account of how symbolic threats produce cognitive, emotional, and behavioral shifts that collectively drive polarization. The study calls for future research to empirically test this broader set of defensive responses. Understanding these psychological processes is essential for developing more comprehensive strategies to mitigate polarization in divided societies.

Keywords | polarization, intergroup value protection, threat, defensive reactions

Introduction

A plurality of opinions is a fundamental feature of democratic societies, where disagreement over ideas is seen as a driving force behind open debate and societal progress. Thus, such opinion-based polarization seems to emerge as a natural – and even beneficial – aspect of healthy democracies (Häusermann & Bornschieer, 2023). However, opinion divergence sometimes intensifies and escalates into affective polarization, wherein opposing factions not only disagree on issues but also come to perceive each other as hostile out-groups, often assigning negative stereotypes and moral failings to those holding contrary views.

This escalation is particularly evident when the conflict extends beyond politics and touches upon broader cultural and symbolic domains. In these domains, groups often interpret ideological differences as existential threats to their core identities, triggering their value-protective responses. This dynamic is observable in contemporary Slovakia, where ideological groups are increasingly clashing over polarizing social issues such as migration, universal rights, gender equality, and the role of traditional values (e.g., Beláňová, 2020; Bozogáňová et al., 2022; Očenášová, 2021; Guasti & Bustikova, 2023). These conflicts appear to transcend policy disagreements, reflecting deeper moral and identity-based divisions within society. The consequences of this affective polarization are not merely theoretical but have manifested in Slovakia through acts of politically and ideologically motivated violence, such as the shooting and killing of two LGBTQ+ members in 2022 in Bratislava, or the attempted assassination of Slovakia's Prime Minister in 2024 in Handlová. These incidents demonstrate how affective polarization can erode democratic norms and threaten the safety of individuals, emphasizing the urgent need to address the underlying psychological factors contributing to this polarization.

To explain the psychological mechanisms of affective polarization, a recent Intergroup Value Protection Model (IVPM) was proposed by van Zomeren et al. (2024). The model posits that people are strongly motivated to protect their social embeddedness and, therefore, continuously monitor the out-group's behavior to determine whether it threatens this embeddedness. The advantage of this model is that it integrates the extensive knowledge from moral and group psychology to explain which specific factors trigger the so-called value protection process and what specific forms of protective responses people tend to use to protect their embeddedness. Despite its inevitable contribution, I believe the model does not fully capture the range of possible protective responses that could contribute to the escalation of conflicts between social groups and, thus, increase polarization in societies.

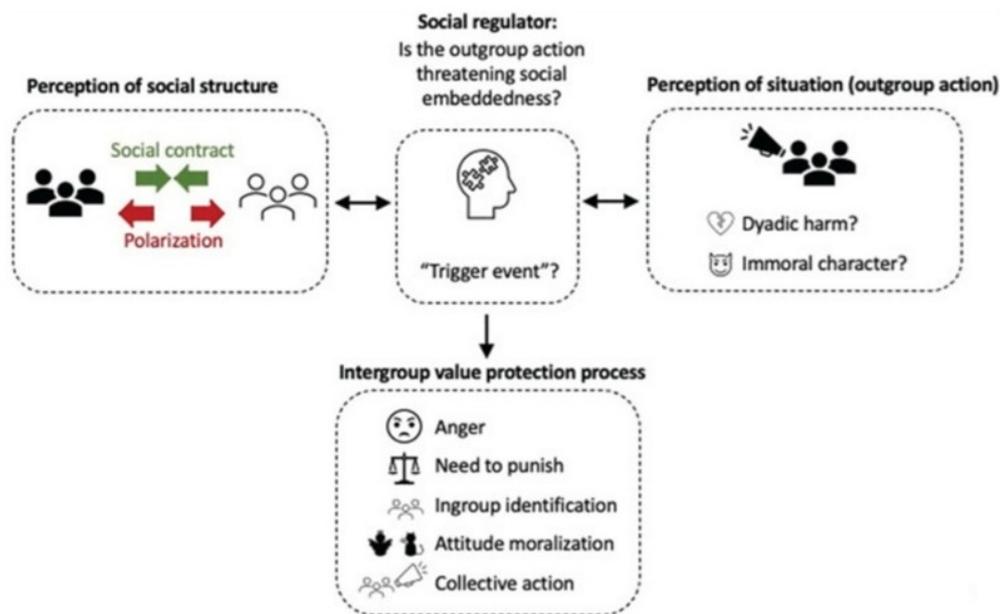
Objectives and Methods

In the present study, I propose the extension of the IVPM by employing the General Model of Threat and Defense (GPMTD, Jonas et al., 2014; Stollberg et al., 2024). The GPMTD complements the IVPM by detailing the proximal and distal threat-defense mechanisms, such as BIS- and BAS-driven responses, which can subtly escalate polarization over time. Using this theory, I propose additional mechanisms that extend the range of possible value protection responses people may utilize to preserve their social embeddedness. This integration provides a more complete explanatory framework than either model alone, clarifying not only how polarization emerges but also why these processes intensify even before intergroup conflict becomes visible. Thus, by combining the IVPM with the GPMTD, the present study helps to better understand the psychological mechanisms of affective polarization and offers a more nuanced framework for identifying the antecedents and consequences of intergroup value conflicts in culturally and ideologically divided societies.

The Intergroup Value Protection Model: A psychological framework for understanding affective polarization

The IVPM proposed by van Zomeren et al. (2024) offers a comprehensive framework for understanding how and why intergroup conflicts escalate in democratic societies. Rooted in the intersection of moral and group psychology, the IVPM posits that individuals are not merely passive observers of social life, but active social regulators motivated by a fundamental need to protect their social embeddedness – a sense of being securely and morally integrated within a valued social group. At the heart of the IVPM lies the idea that group members are constantly monitoring the social environment for signs that their core group values are threatened by out-groups. When a member of the in-group cognitively interprets an out-group’s behavior as a violation of shared moral or social values, this represents a “trigger event” that activates a value protection process. This process motivates individuals to engage in protective responses aimed at reaffirming the moral legitimacy of the in-group and preserving the cohesion and identity of their social group (see Figure 1).

Figure 1: Intergroup Value Protection Model proposed by van Zomeren et al. (2024)



Source: van Zomeren et al. (2024)

One of the central contributions of the IVPM is its ability to delineate the conditions under which this value protection process is triggered. According to van Zomeren et al. (2024), whether an individual perceives a specific action by the out-group as a value-threatening event depends on two broad evaluative domains: *perception of the social structure* and *perception of the situation (out-group action)*. The *perception of the social structure* involves two key components. First, individuals form beliefs about the degree of societal polarization – that is, whether society is divided into opposing ideological groups, as often observed in debates over contentious issues such as abortion rights or migration policy (Koudenburg & Kashima, 2022). Second, individuals evaluate the perceived strength of the social contract, which reflects the extent to which they feel a sense of mutual obligation and commitment to maintaining social cohesion across group boundaries, even in the presence of disagreement (Owuamalam et al., 2019).

The *perception of the situation* shifts the evaluative focus from society at large to the specific context of the out-group’s action. This domain also comprises two components. First, individuals

assess the perceived immorality of the out-group's behavior – namely, whether the out-group is seen as evil, lacking fundamental moral qualities (Brambilla & Leach, 2014). Second, individuals evaluate the perceived dyadic harm in the out-group's action, or the extent to which the out-group's actions are viewed as intentionally causing damage or suffering to a victim (Schein & Gray, 2018).

The combination of these perceptions – how immoral and harmful the out-group is seen to be, and how polarized and fragmented the society is perceived – directly influences whether the value protection mechanism is activated and how it manifests. Once triggered, the value protection process can lead to a range of affective, cognitive, and behavioral responses, all aimed at defending the moral integrity of the in-group. In particular, Van Zomeren et al. (2024) identify five core protective responses: (1) negative moral emotions such as anger, contempt, disgust, outrage or hatred toward the out-group, (2) a desire to punish the out-group or its members, also framed as retributive justice, (3) moralization of one's own attitudes, wherein personal attitudes are reframed as moral convictions, (4) stronger identification with the in-group, which can intensify the differences in perceptions of “who we are and what we stand for” against “who they are and what they stand for”, (5) collective action against the out-group, including protests, campaigns, or even hostile mobilization.

Importantly, these responses are mutually reinforcing. For instance, a protest initiated by one group can be interpreted by the opposing group as a further moral violation, thus triggering their own value protection process. This loop of mutual value protection can lead to escalating a vicious cycle of polarization, where compromise becomes less likely and social divisions deepen.

Results

Integrating the General Model of Threat and Defense into the model

The IVPM (van Zomeren et al., 2024) offers an elegant framework that captures the psychological underpinnings of many contemporary social movements and conflicts in Western societies, illustrating how a sense of threatened values can mobilize individuals toward defensive actions. Yet, while positing that value protection responses fundamentally stem from the perception of threat, the IVPM seems not to cover a complete array of possible threat-management strategies that have been documented in social-psychological research.

A closer look at the extensive literature on threat and defense reveals that people deploy a much broader repertoire of defensive processes that go beyond IVPM. The General Process Model of Threat and Defense (Jonas et al., 2014) was developed as a theoretical framework integrating the extensive and diversified threat and defense research. The model postulates that people experience feelings of threat every time they perceive a discrepancy between their expectations or needs and the actual reality.

Behavioral Inhibition System reactions

An important aspect of the GPMTD is that active defensive reactions are not the first direct and immediate responses to the threat. Instead, the activation of the behavioral inhibition system (BIS) is the first proximal reaction. The BIS is a neurobehavioral system that suppresses people's ongoing behavior and gives rise to a feeling of anxiety, heightened arousal, nervousness, and worry (Gray & McNaughton, 2003; Jonas et al., 2014). This BIS-driven alarm then precipitates proximal defensive responses – manifesting in avoidance motivation, hypervigilance, and inhibition of all ongoing behaviors – that are aimed at mitigating the immediate discrepancy and reducing arousal (Corr et

al., 2013). Jonas et al. (2014) consider this process to be mostly adaptive because it is targeted at resolving the discrepancy immediately by disengaging or switching one's goals. Simultaneously, these proximal reactions are associated with the increased vigilance for new information that helps to eliminate the negative feelings.

Since these initial proximal reactions are inhibitory and inwardly focused on reducing negative emotions, one might assume that they do not significantly contribute to polarization. Unsurprisingly, the IVPM overlooks these reactions and does not consider them as part of the defensive responses. However, I argue that such reactions may, in fact, contribute to the escalation of intergroup conflict over a longer and less visible trajectory. Specifically, research suggests that the withdrawal motivations triggered by BIS activation can underlie "silent" radicalization occurring in online spaces. Studies have found that individuals with heightened BIS sensitivity tend to experience less meaning and purpose in life, focus more on pessimistic aspects of situations, feel a greater sense of identity threat, and are particularly sensitive to societal rejection (Hirsh & Kang, 2016; McGregor et al., 2013; Slegers et al., 2015; Klackl et al., 2018). Combined with increased vigilance for new information, these tendencies may lead such individuals to search for ways to resolve the perceived threat by engaging with the groups that either point to the threat or claim to be able to resolve this threat for themselves (McGregor et al., 2013). This may lead to the willingness to engage with extremist narratives or radical groups that often present themselves in this way. Indeed, several theoretical and empirical studies support the idea that the activation of the BIS increases engagement with extremist content, radicalism, or activism in support of terrorist groups (McGarry & Shortland, 2023; Shortland & McGarry, 2022; Shortland et al., 2022; McGregor et al., 2013). McGarry and Shortland (2023) suggest that the BIS may be responsible for the phenomenon of so-called "bedroom radicals" (see Frissen, 2021) – individuals who seem normal throughout much of their lives but become radicalized in cyberspace and may even carry out extremist or terrorist attacks. This evidence has led me to propose an extension of the IVPM. Specifically, while the IVPM posits that one of the key value-protection defenses is increased ingroup identification, I hypothesize that *individuals may not only reinforce their identification with existing groups but may also begin to identify with new – and potentially more extremist – groups.*

Behavioral Activation System reactions

Naturally, people do not use only proximal avoidant reactions associated with the activation of BIS to reduce their feelings of threat. Often, they try to overcome the BIS activation by shifting to the Behavioral Activation System (BAS). The BAS helps individuals to transform their affective state from a threat-induced anxiety to a defense-induced approach-related positive affect (Stollberg et al., 2024). Importantly, the GPMTD distinguishes between direct resolution defenses – immediate cognitive, emotional, or behavioral reactions intended to regulate distressing arousal – and distal defenses, which serve to alleviate distress through indirect or palliative reactions (Jonas et al., 2014).

Reinforcing pre-existing beliefs and values. The GPMTD identifies a plethora of BAS-triggered defensive reactions, most of which do not appear to have the potential to escalate conflicts and polarize societies. However, there are some that could carry such potential. In particular, the exposure to threat – whether in the form of interpersonal conflict, uncertainty, or moral dilemmas – triggers specific defensive reactions at the cognitive level, primarily aimed at preserving a sense of internal consistency, certainty, and value alignment. Empirical findings demonstrate that exposure to threat promotes more abstract forms of thinking and increases the perceived connection between one's current actions and long-term, personally meaningful goals (Landau et al., 2011; Landau et al., 2009; McGregor et al., 2001). Consequently, individuals become more certain of their

opinions and more strongly committed to personal norms and values, which they describe as more important, identity-relevant, and value-congruent (McGregor et al., 2010; McGregor et al., 2013). This defensive cognitive reassurance can also involve increased reliance on personal opinions and moral convictions, especially when these ideals are threatened. For instance, McGregor and Jordan (2007) demonstrate that individuals confronted with self-relevant threats may respond with defensive zeal, i.e., an intensified, value-driven commitment to personal beliefs. This response is marked by two specific tendencies: first, the extremization of attitudes, meaning that individuals' opinions become more rigid, polarized, and less open to alternative viewpoints; and second, the illusion of consensus, whereby individuals overestimate the extent to which their views are shared by others – a phenomenon known as the false consensus effect. Thus, exposure to threat may lead individuals to adopt more extreme views and to falsely assume that their beliefs are widely shared, providing a sense of psychological security and coherence.

This dynamic closely mirrors the entrenching hypothesis described by Greenberg and Jonas (2003), which posits that existential threats increase individuals' commitment to their pre-existing worldviews, ideologies, and value systems. When faced with threat or uncertainty, people tend to reinforce their beliefs – not necessarily by changing opinions, but by holding onto them more rigidly and defensively. The hypothesis builds on the foundations of Terror Management Theory, arguing that the preservation of a coherent worldview serves as a psychological defense against existential anxiety. The entrenching hypothesis explains how threat motivates not openness or flexibility, but rather attitudinal rigidity and the rejection of counter-attitudinal information, particularly on value-laden issues.

This pattern was confirmed in a recent study by Leota et al. (2023), who investigated how existential threats affect beliefs in moral purity depending on political ideology. They found that after being exposed to a threat, politically conservative participants increased their endorsement of moral purity norms, while liberal participants showed a decrease in such endorsement. This demonstrates that threat exposure intensifies people's existing moral and ideological commitments – precisely as the entrenching hypothesis predicts.

Moreover, exposure to threat facilitates motivated reasoning and confirmation bias – people become more likely to assess or search for information that confirm their preexisting beliefs (Munro & Stansbury, 2009), as well as a stronger tendency toward verification and weaker tendency toward falsification (Muris et al., 2014; Remmerswaal et al., 2010). This process may reinforce ideological rigidity and the illusion of consensus with others, particularly on value-laden issues such as religion, national identity, or moral debates.

Overall, threat exposure triggers a psychological need for cognitive coherence and conviction, leading individuals to reinforce their existing beliefs and selectively attend to information that affirms their worldview. At the societal level, I believe this dynamic may contribute to the formation of polarized groups with increasingly extreme and rigid views. The more these groups confront their opposing perspectives, the more deeply entrenched their convictions become. Thus, I hypothesize that after exposure to threat, *individuals may cognitively reinforce their preexisting beliefs and values, contributing to the greater divergence of opinions on polarizing issues and potentially leading to greater affective polarization between social groups.*

Searching for explanations and illusory patterns. When individuals are confronted with perceived threats – such as existential insecurity, threatened identity, or social exclusion – they experience a heightened psychological need for meaning, order, and control. To restore this sense of stability, they actively search for patterns that help them perceive their situation as more

predictable and understandable (van Prooijen et al., 2018). Importantly, this search for patterns often extends beyond rational or accurate perceptions. Evidence shows that people exposed to threats are more likely to misperceive patterns in randomness, see images in noise, form illusory correlations, and – most relevant to polarization – incline toward conspiracy beliefs (Whitson & Galinsky, 2008; Van Harreveld et al., 2014; Spasovski et al., 2021).

Conspiracy theories are particularly effective at reducing threats because they offer simplified, agent-based explanations for complexity and ambiguity, often attributing events to intentional acts of malevolent out-groups. In particular, van Prooijen's (2020) Existential Threat Model posits that existential threats activate deep-seated psychological mechanisms related to sense-making and control. According to the model, these threats trigger a motivated cognitive response in which people seek explanations that reduce uncertainty and anxiety. When an antagonistic out-group is salient, individuals are more likely to interpret ambiguous events as intentional, fueling conspiratorial thinking.

Indeed, Heiss et al. (2021) found that in high-uncertainty environments such as the COVID-19 pandemic, threat perceptions were associated with reduced factual knowledge acquisition and increased belief in conspiracy claims. Rather than engaging in problem-focused coping strategies, individuals often turned to non-scientific narratives that offered emotionally satisfying but inaccurate explanations. This pattern aligns with the findings by Liekefett et al. (2022), who demonstrated that existential threat, anxiety, and aversion to uncertainty predict sustained endorsement of conspiracy beliefs over time, indicating their role as a compensatory mechanism for psychological distress. The inclination to conspiratorial thinking also increases in the context of social exclusion—Poon et al. (2020) showed that individuals who felt ostracized were significantly more likely to endorse conspiracy theories, likely as a means of reasserting meaning. Moreover, a threat can also arise from perceived intergroup conflict. In this regard, Krüppel et al. (2021) found that individuals experiencing symbolic or realistic intergroup threats were more inclined to adopt conspiratorial narratives, especially when these narratives served to defend the ingroup against perceived external hostility.

Importantly, a growing body of empirical research suggests that conspiracy beliefs intensify polarization, while this effect may operate via both direct and indirect pathways. In particular, some studies show that conspiracy beliefs directly heighten prejudice and intergroup hostility. For example, Jolley et al. (2020) experiments showed that exposure to an anti-immigrant conspiracy exacerbated prejudice toward immigrants, while exposure to an antisemitic conspiracy increased negative attitudes toward Jews and reduced willingness to vote for a Jewish candidate. Additionally, Jungkunz et al. (2025) found that conspiracy beliefs are associated with an increased willingness to justify political violence. In another study, Šrol et al. (2022) surveyed Slovak adults about COVID-19 conspiracies and found that individuals endorsing such beliefs (e.g., blaming a foreign country for the virus) reported more prejudice toward the blamed groups, lower trust in government regulations, and a greater willingness to justify violent and non-compliant actions.

The association between conspiracy beliefs and a lack of trust in government aligns with van Prooijen et al.'s (2022) model, which suggests that conspiracy beliefs are linked to distrust in traditional institutions. This distrust disrupts individuals' sense of security, leading them to feel marginalized and deprived of power. These feelings, in turn, result in changes in social relationships, reduced cooperation, and non-compliance with social norms, ultimately manifesting in antisocial behavior, increased prejudice, intergroup conflict, polarization, and extremism.

Pummerer's (2022) model further explains these links by hypothesizing that institutional distrust leads to non-normative behavior due to altered perceptions of social norms. Specifically, individuals who adhere to conspiratorial thinking and express suspicion toward institutions and traditional authorities are more likely to underestimate descriptive norms and question injunctive norms that prescribe appropriate behavior in specific situations. Similar to van Prooijen et al. (2022), Pummerer argues that these distorted perceptions of social reality directly contribute to a decline in interpersonal cooperation and prosocial behavior, alongside an increase in non-normative actions. These tendencies can have harmful consequences for both individuals and society, such as participation in illegal or violent political demonstrations and other forms of antisocial behavior (Lewis, 2013; O'Connor & White, 2009; Gulliver et al., 2023).

Taken together, the evidence indicates that conspiracy beliefs may exacerbate polarization via both direct and mediated pathways. Directly, exposure to conspiratorial claims triggers negative attitudes toward the targeted out-group. Indirectly, conspiracy beliefs corrode trust. Importantly, the spread of conspiracy content often occurs within segregated media environments, which further magnify polarization. Social media platforms and partisan outlets can create "echo chambers" in which conspiratorial messages reinforce like-minded views (Del Vicario et al., 2016; Jiang et al., 2021). In practice, this means that individuals with conspiratorial thinking often self-select into media that validate those beliefs, creating a vicious cycle that amplifies affective polarization.

Overall, the evidence leads me to the suggestion that *perceived threat may contribute to rising polarization by encouraging illusory or conspiratorial thinking, which, in turn, increases prejudice and hostility toward out-groups, fosters support for violent and non-normative actions, and undermines cooperation and social cohesion within societies.*

Discussion and Conclusion

Integrating the GPMTD (Jonas et al., 2014) into the IVPM framework advances our understanding of how threat-related processes can actively contribute to polarization. While the IVPM highlights the broad set of defensive reactions ranging from moral outrage to collective action, it overlooks some cognitive intra-individual mechanisms through which symbolic threats may escalate into polarized worldviews. The GPMTD fills this gap by identifying additional threat responses that may have significant polarizing potential. As I suggested in this study, BIS-related threat inhibition may initially lead to avoidance, but in the long term, it may foster *silent radicalization* in which individuals retreat from public discourse, seek ideologically aligned online content, and gradually adopt more extreme attitudes. These latent processes may remain invisible, yet they play a crucial role in deepening societal divides. Furthermore, the GPMTD elucidates how threat exposure distorts information processing, heightening confirmation bias, increasing belief rigidity, and *reinforcing personal opinions and convictions*. Such cognitive defenses may also be associated with individuals' tendencies to reject "mainstream viewpoints" and *adhere to conspiracy beliefs*. These processes escalate polarization not through direct confrontation, but by intensifying affective and epistemic distance between social groups. Integrating these insights into the IVPM allows for a richer account of how symbolic threats produce cognitive, emotional, and behavioral shifts that collectively drive polarization.

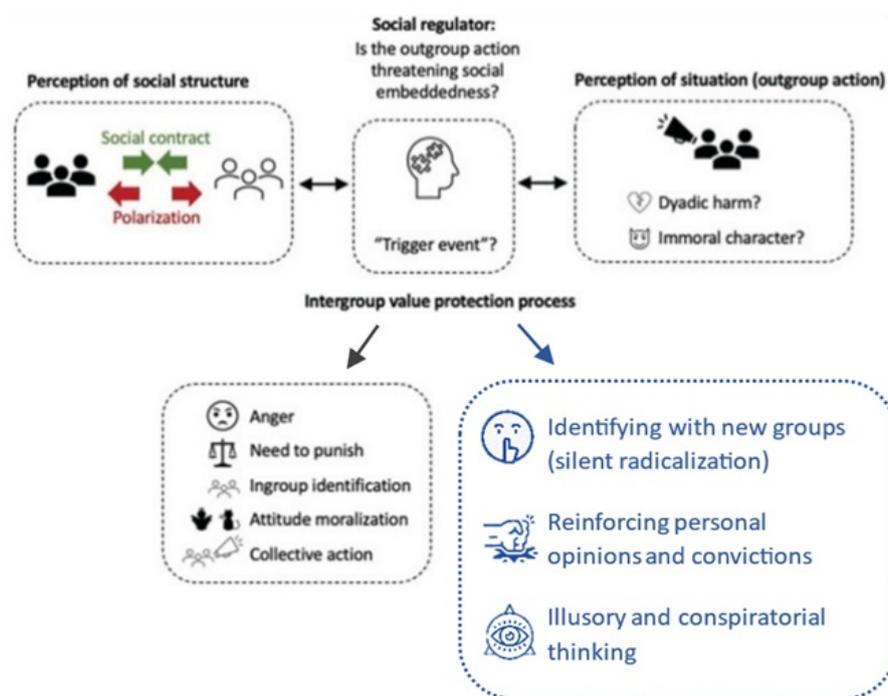
Proposed Integrated Threat-Value Protection Model (ITVPM)

To synthesize the insights from the two models, I propose an integrated model that combines the strengths of the IVPM and the GPMTD, hereafter referred to as the Integrated Threat-Value Protection Model (ITVPM). The ITVPM extends the IVPM by incorporating the proximal and distal

threat-defense mechanisms identified in the GPMTD. Specifically, the model visually and conceptually adds BIS-driven proximal avoidance and hypervigilance responses, as well as BAS-driven distal cognitive and behavioral defenses, to the original IVPM framework. This integration highlights how symbolic threats may trigger reactions ranging from immediate anxiety and withdrawal to reinforced moral convictions, selective information processing, and conspiratorial thinking that may escalate affective polarization over time.

By explicitly mapping these threat responses onto the IVPM structure, the ITVPM provides a comprehensive baseline for future empirical research. It allows for the identification of new areas where interventions may mitigate polarization before intergroup conflicts occur. Figure 2 illustrates the ITVPM, showing how the GPMTD responses (outlined in blue) interconnect with the IVPM value protection processes, thereby offering both a visual and conceptual synthesis.

Figure 2: *Integrated Threat-Defense Value Protection Model*



Implications and directions for future research and practice

Recognizing these new mechanisms calls for broadening our empirical lens. Future research may aim to capture a broader view of how individuals defensively process threat on cognitive, affective, and behavioral levels. A promising direction lies in experimental designs that systematically manipulate *real* or *symbolic* threats – such as moral norm violations, value conflict, or social exclusion – and track this broader spectrum of defensive responses. This includes not only those proposed by the IVPM, but also those stemming from the GPMTD, such as reinforcing opinions and convictions, selective information-seeking, confirmation bias, media-selection habits, engagement in value-affirming rituals, or online radicalization tendencies. By integrating this perspective, polarization research can more fully explain how people become psychologically entrenched long before the between-group escalation occurs.

Although this study is primarily theoretical, it also carries implications for preventive strategies in societies vulnerable to polarization. Programs that strengthen media and extremism-related civic literacy may help citizens recognize extremist and conspiratorial content before it becomes

internalized, thereby limiting the risk of “silent radicalization” in online environments (see Nienierza et al., 2021). Beyond this, educational initiatives aimed at cultivating perspective-taking and critical thinking could create psychological resilience against the cognitive rigidity that often follows from threat exposure (Scheibling-Sève et al., 2022). Lastly – and most importantly – addressing the erosion of institutional trust seems indispensable. Conspiracy-driven distrust does not merely distort information processing but corrodes the very foundations of social cooperation (Pummerer, 2022; van Prooijen, 2022). Policies that enhance institutional transparency, community-level trust, and cross-group solidarity, therefore, represent the most crucial buffer against the escalation of symbolic threats into affective polarization. Without rebuilding trust, even the most sophisticated literacy or educational programs may struggle to achieve lasting impact.

Acknowledgements

The author has no conflict of interest. All references are transparently indicated. The manuscript has not been submitted to another journal.

The study was supported by the Slovak Research and Development Agency as part of the research project APVV-23-0127: Polarized society: From political orientation to protecting moral values. The study was supported by the VEGA under Grant no. 2/0030/24: Distrust mindset and non-normative behavior: From perceptions of social reality to violations of social norms.

References

- Beláňová, A. (2020). Anti-abortion activism in the Czech Republic and Slovakia: 'Nationalizing' the strategies. *Journal of Contemporary Religion, 35*(3), 395–413. <https://doi.org/10.1080/13537903.2020.1836813>
- Bozogáňová, M., Kentoš, M., & Piterová, I. (2022). *Postoje k migrácii: Od predsudkov k populizmu*. Spoločenskovedný ústav CSPV SAV, v. v. i. <https://doi.org/10.31577/2022.9788089524761>
- Brambilla, M., & Leach, C. W. (2014). On the importance of being moral: The distinctive role of morality in social judgment. *Social Cognition, 32*(4), 397–408. <https://doi.org/10.1521/soco.2014.32.4.397>
- Corr, P. J., DeYoung, C. G., & McNaughton, N. (2013). Motivation and personality: A neuropsychological perspective. *Social and Personality Psychology Compass, 7*(3), 158–175. <http://dx.doi.org/10.1111/spc3.12016>
- Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G., et al. (2016). Echo chambers: Emotional contagion and group polarization on Facebook. *Scientific Reports, 6*, Article 37825. <https://doi.org/10.1038/srep37825>
- Frissen, T. (2021). Internet, the great radicalizer? Exploring relationships between seeking online extremist materials and cognitive radicalization in young adults. *Computers in Human Behavior, 114*, Article 106549. <https://doi.org/10.1016/j.chb.2020.106549>
- Gray, J. A., & McNaughton, N. (2003). *The neuropsychology of anxiety: An enquiry into the function of the septo-hippocampal system*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198522713.001.0001>
- Greenberg, J., & Jonas, E. (2003). Psychological motives and political orientation—The left, the right, and the rigid: Comment on Jost et al. (2003). *Psychological Bulletin, 129*(3), 376–382. <https://doi.org/10.1037/0033-2909.129.3.376>
- Guasti, P., & Bustikova, L. (2023). Varieties of illiberal backlash in Central Europe. *Problems of Post-Communism, 70*(2), 130–142. <https://doi.org/10.1080/10758216.2022.2156889>
- Gulliver, R., Chan, C. S., Tam, K. Y. Y., Lau, I. S. K., Hong, Y. Y., & Louis, W. R. (2023). Political distrust, perceived threat, and intentions to engage in normative and violent collective action: A mixed-methods study. *European Journal of Social Psychology, 53*(2), 401–417. <https://doi.org/10.1002/ejsp.2910>
- Häusermann, S., & Bornschier, S. (2023). *Democratic conflict and polarization: Healthy or harmful?* UBS Center for Economics in Society at the University of Zurich. <https://doi.org/10.5167/uzh-254747>
- Heiss, R., Gell, S., Röthlingshöfer, E., & Zoller, C. (2021). How threat perceptions relate to learning and conspiracy beliefs about COVID-19: Evidence from a panel study. *Personality and Individual Differences, 175*, Article 110672. <https://doi.org/10.1016/j.paid.2021.110672>
- Hirsh, J. B., & Kang, S. K. (2016). Mechanisms of identity conflict: Uncertainty, anxiety, and the behavioral inhibition system. *Personality and Social Psychology Review, 20*(3), 223–244. <https://doi.org/10.1177/1088868315589475>

- Jiang, J., Ren, X., & Ferrara, E. (2021). Social media polarization and echo chambers in the context of COVID-19: Case study. *JMIRx Med*, 2(3), e29570. <https://doi.org/10.2196/29570>
- Jolley, D., Meleady, R., & Douglas, K. M. (2020). Exposure to intergroup conspiracy theories promotes prejudice which spreads across groups. *British Journal of Psychology*, 111(1), 17–35. <https://doi.org/10.1111/bjop.12385>
- Jonas, E., McGregor, I., Klackl, J., Agroskin, D., Fritsche, I., Holbrook, C., ... & Quirin, M. (2014). Threat and defense: From anxiety to approach. In J. M. Olson, M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. 49, pp. 219-286). Academic Press. <https://doi.org/10.1016/B978-0-12-800052-6.00004-4>
- Jungkunz, S., Fahey, R. A., & Hino, A. (2025). Populist attitudes, conspiracy beliefs and the justification of political violence at the US 2020 elections. *Political Studies*, 73(2), 592–611. <https://doi.org/10.1177/00323217241259229>
- Klackl, J., Jonas, E., & Fritsche, I. (2018). Neural evidence that the behavioral inhibition system is involved in existential threat processing. *Social Neuroscience*, 13(3), 355–371. <https://doi.org/10.1080/17470919.2017.1308880>
- Krüppel, J., Yoon, D., Fieg, K., Sharma, P., & Mokros, A. (2021). On the relationship between right-wing attitudes, conspiracy beliefs, and intergroup threat: Introducing an indirect measure for intergroup threat. *Journal of Theoretical Social Psychology*, 5(4), 354–365. <https://doi.org/10.1002/jts5.103>
- Koudenburg, N., & Kashima, Y. (2022). A polarized discourse: Effects of opinion differentiation and structural differentiation on communication. *Personality and Social Psychology Bulletin*, 48(7), 1068–1086. <https://doi.org/10.1177/01461672211030816>
- Landau, M. J., Greenberg, J., Sullivan, D., Routledge, C., & Arndt, J. (2009). The protective identity: Evidence that mortality salience heightens the clarity and coherence of the self-concept. *Journal of Experimental Social Psychology*, 45(4), 796–807. <http://dx.doi.org/10.1016/j.jesp.2009.05.013>
- Landau, M. J., Kosloff, S., & Schmeichel, B. J. (2011). Imbuing everyday actions with meaning in response to existential threat. *Self and Identity*, 10(1), 64–76. <http://dx.doi.org/10.1080/15298860903557243>
- Leota, J., Simpson, D., Mazidi, D., & Nash, K. (2023). Purity, politics, and polarization: Political ideology moderates threat-induced shifts in moral purity beliefs. *British Journal of Social Psychology*, 62(2), 806–824. <https://doi.org/10.1111/bjso.12596>
- Lewis, N. (2013). The effects of perceived normative pressure and observability of behavior on intention to perform sun protection and nutrition behaviors on behalf of young children among parents. *Journal of Applied Social Psychology*, 43(1), E97–E108. <https://doi.org/10.1111/jasp.12038>
- Liekefett, L., Christ, O., & Becker, J. C. (2022). Can conspiracy beliefs be beneficial? Longitudinal linkages between conspiracy beliefs, anxiety, uncertainty aversion, and existential threat. *Personality and Social Psychology Bulletin*, 49(2), 167–179. <https://doi.org/10.1177/01461672211060965>
- McGarry, P., & Shortland, N. (2023). Anxious activism: The role of behavioral inhibition system in the radicalization process. *Computers in Human Behavior*, 139, Article 107550.

<https://doi.org/10.1016/j.chb.2022.107550>

McGregor, I., Jordan, C. H. (2007). The mask of zeal: Low implicit self-esteem, threat, and defensive extremism. *Self and Identity*, 6(2-3), 223-237. <http://dx.doi.org/10.1080/15298860601115351>

McGregor, I., Nash, K., Mann, N., & Phillips, C. E. (2010). Anxious uncertainty and reactive approach motivation (RAM). *Journal of Personality and Social Psychology*, 99(1), 133-147. <https://doi.org/10.1037/a0019701>

McGregor, I., Prentice, M., & Nash, K. (2013). Anxious uncertainty and reactive approach motivation (RAM) for religious, idealistic, and lifestyle extremes. *Journal of Social Issues*, 69(3), 537-563. <https://doi.org/10.1111/josi.12028>

McGregor, I., Zanna, M. P., Holmes, J. G., & Spencer, S. J. (2001). Compensatory conviction in the face of personal uncertainty: Going to extremes and being oneself. *Journal of Personality and Social Psychology*, 80(3), 472-488. <https://doi.org/10.1037/0022-3514.80.3.472>

Munro, G. D., & Stansbury, J. A. (2009). The dark side of self-affirmation: Confirmation bias and illusory correlation in response to threatening information. *Personality and Social Psychology Bulletin*, 35(9), 1143-1153. <https://doi.org/10.1177/0146167209337163>

Muris, P., Debipersad, S., & Mayer, B. (2014). Searching for danger: On the link between worry and threat-related confirmation bias in children. *Journal of Child and Family Studies*, 23, 604-609. <https://doi.org/10.1007/s10826-013-9727-0>

Nienierza, A., Reinemann, C., Fawzi, N., Riesmeyer, C., & Neumann, K. (2021). Too dark to see? Explaining adolescents' contact with online extremism and their ability to recognize it. *Information, Communication & Society*, 24(9), 1229-1246. <https://doi.org/10.1080/1369118X.2019.1697339>

O'Connor, E. L., & White, K. M. (2009). Intentions and willingness to use complementary and alternative medicines: What potential patients believe about CAMs. *Complementary Therapies in Clinical Practice*, 15(3), 136-140. <https://doi.org/10.1016/j.ctcp.2009.03.003>

Očenášová, Z. (2021). Poison in the juice: "Gender Ideology" and the Istanbul Convention in Slovakia. *Politické Vedy*, 24(4), 38-55. <https://doi.org/10.24040/politickevedy.2021.24.4.38-55>

Owuamalam, C. K., Rubin, M., & Spears, R. (2019). Revisiting 25 years of system motivation explanation for system justification from the perspective of social identity model of system attitudes. *British Journal of Social Psychology*, 58(2), 362-381. <https://doi.org/10.1111/bjso.12285>

Poon, K.-T., Chen, Z., & Wong, W.-Y. (2020). Beliefs in conspiracy theories following ostracism. *Personality and Social Psychology Bulletin*, 46(8), 1234-1246. <https://doi.org/10.1177/0146167219898944>

Pummerer, L. (2022). Belief in conspiracy theories and non-normative behavior. *Current Opinion in Psychology*, 47, Article 101394. <https://doi.org/10.1016/j.copsyc.2022.101394>

Remmerswaal, D., Muris, P., Mayer, B., & Smeets, G. (2010). "Will a Cuscus bite you, if he shows his teeth?" Inducing a fear-related confirmation bias in children by providing verbal threat information to their mothers. *Journal of Anxiety Disorders*, 24(5), 540-546. <https://doi.org/10.1016/j.janxdis.2010.03.012>

- Shortland, N., & McGarry, P. (2022). The personality and propaganda puzzle: Exploring the effect of personality on exposure to extremist content online. *Psychology of Violence, 12*(1), 1-10. <https://doi.org/10.1037/vio0000396>
- Shortland, N., Portnoy, J., McGarry, P., Perliger, A., Gordon, T., & Anastasio, N. (2022). A reinforcement sensitivity theory of violent extremist propaganda: The motivational pathways underlying movement toward and away from violent extremist action. *Frontiers in Psychology, 13*, Article 858392. <https://doi.org/10.3389/fpsyg.2022.858392>
- Scheibling-Sève, C., Pasquinelli, E., & Sander, E. (2022). Critical thinking and flexibility. In E. Clément (Ed.), *Cognitive Flexibility: The Cornerstone of Learning* (pp. 77-112). ISTE Ltd and John Wiley & Sons, Inc.
- Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review, 22*(1), 32-70. <https://doi.org/10.1177/1088868317698288>
- Sleegers, W. W., Proulx, T., & van Beest, I. (2015). Extremism reduces conflict arousal and increases values affirmation in response to meaning violations. *Biological Psychology, 108*, 126-131. <https://doi.org/10.1016/j.biopsycho.2015.03.012>
- Spasovski, O., Demuthova, S., & Kuzmanovic, V. (2021). Anxiety, illusory pattern perception and conspiracy beliefs during the COVID-19 pandemic. *Teorija in Praksa, 58*(4), 1133-1159. <https://doi.org/10.51936/tip.58.3.1333-1149>
- Šrol, J., Čavojová, V., & Ballová Mikušková, E. (2022). Finding someone to blame: The link between COVID-19 conspiracy beliefs, prejudice, support for violence, and other negative social outcomes. *Frontiers in Psychology, 12*, Article 726076. <https://doi.org/10.3389/fpsyg.2021.726076>
- Stollberg, J., Klackl, J., & Jonas, E. (2024). Empirical test of a general process model of threat and defense: A systematic examination of the affective-motivational processes underlying proximal and distal reactions to threat. *Journal of Experimental Social Psychology, 110*, 104526. <https://doi.org/10.1016/j.jesp.2023.104526>
- van Prooijen, J. W. (2020). An existential threat model of conspiracy theories. *European Psychologist, 25*(1), 16-25. <https://doi.org/10.1027/1016-9040/a000381>
- van Prooijen, J. W., Douglas, K. M., & De Inocencio, C. (2018). Connecting the dots: Illusory pattern perception predicts belief in conspiracies and the supernatural. *European Journal of Social Psychology, 48*(3), 320-335. <https://doi.org/10.1002/ejsp.2331>
- van Prooijen, J. W., Spadaro, G., & Wang, H. (2022). Suspicion of institutions: How distrust and conspiracy theories deteriorate social relationships. *Current Opinion in Psychology, 43*, 65-69. <https://doi.org/10.1016/j.copsyc.2021.06.013>
- van Zomeren, M., d'Amore, C., Pauls, I. L., Shuman, E., & Leal, A. (2024). The Intergroup Value Protection Model: A theoretically integrative and dynamic approach to intergroup conflict escalation in democratic societies. *Personality and Social Psychology Review, 28*(2), 225-248. <https://doi.org/10.1177/10888683231192120>
- van Harreveld, F., Rutjens, B. T., Schneider, I. K., Nohlen, H. U., & Keskinis, K. (2014). In doubt and disorderly: Ambivalence promotes compensatory perceptions of order. *Journal of Experimental*

Psychology: General, 143, 1666–1676. <https://doi.org/10.1037/a0036099>

Whitson, J. A., & Galinsky, A. D. (2008). Lacking control increases illusory pattern perception. *Science*, 322, 115–117. <https://doi.org/10.1126/science.1159845>