

# Do External Threats Reduce Affective Polarization?

## An Experiment on Russia's Invasion of Ukraine

Jonas Pilgaard Kaiser\*

Markus Seier†

Aarhus University

Aarhus University

February 17, 2025

### Abstract

In many countries, partisans have become increasingly biased in how they evaluate others based on political affiliation. We suggest that this increase in affective polarization may in part be caused by changes in the global power distribution which caused many countries to experience a long period without external (military) threats. To study the importance of external threats, we conduct a priming experiment to examine how making Russia's full-scale invasion of Ukraine in 2022 salient causally influences affective polarization and collaboration in the U.S. We find that priming Americans with Russia's military aggression leads to a modest reduction in affective polarization and an increase in cooperativeness as measured by behavior in an incentivized coordination game. Surprisingly, the effect of making Russia's invasion salient does not depend on perceived cross-party disagreement about the conflict. These results suggest that researchers should also consider international relations to understand within-country polarization and willingness to collaborate.

**Keywords:** Affective polarization, social identity, common enemy effect, external threat, 2022 Russian invasion of Ukraine

---

\*Corresponding author. Current address: Technical University of Berlin, Faculty of Economics and Management; kaiser@finance.tu-berlin.de.

†Contact: Aarhus University, Department of Economics and Business Economics; mseier@econ.au.dk. Supplementary material for this article is available in the appendix in the online edition. The study is pre-registered on OSF (<https://osf.io/u82kq>). Replication files are available in the JOP Dataverse (<https://dataverse.harvard.edu/dataverse/jop>). The empirical analysis has been successfully replicated by the JOP replication analyst. The project was approved by the Institutional Review Board at Aarhus University (20.04.2022, BSS-2022-036). We gratefully acknowledge the financial support of William Demant Fonden and Knud Højgaards Fond.

Short title: Do External Threats Reduce Affective Polarization?

Many Western countries have seen an increase in affective polarization (Gidron, Adams, and Horne 2020; Iyengar et al. 2019), which is the tendency for partisans to view opposing partisans negatively and copartisans positively (Iyengar, Sood, and Yphtach, 2012, p. 406; Iyengar and Westwood, 2015, p. 691). This development is alarming, as it leads to a decrease in social cohesion and cross-party collaboration (Hetherington 2015; Layman, Carsey, and Horowitz 2006; MacKuen et al. 2010). Prior research has mainly focused on causes of increased polarization within countries, including social sorting (Mason 2015), partisan media (Lau et al. 2017), campaign messages (Hansen and Kosiara-Pedersen 2017), and moralization of politics (Garrett and Bankert 2020). Far less attention has been paid to external factors such as global changes in international relations. This is surprising as the increase in affective polarization – especially in the U.S. – coincided with the fall of the Soviet Union which fundamentally changed the global power distribution and the nature of conflicts. A common dictum is that external threats to a group strengthen group cohesion and collaboration (De Jaegher 2021; Sumner 1906) – at least when there is cross-party agreement regarding the threat (John and Dvir-Gvirsman 2015; Orian Harel, Maoz, and Halperin 2020).<sup>1</sup>

The aim of this paper is to study the influence of external threats on affective polarization and collaboration by using the context of Russia’s full-scale invasion of Ukraine in 2022. Specifically, we address the following research questions: How did Russia’s invasion of Ukraine influence affective polarization and willingness to collaborate in the U.S.? And how did these effects depend on perceived political agreement about how to handle the conflict?

To answer these questions, we conduct an online experiment with 1,403 U.S. citizens in May 2022 and make either the threat of Russia’s invasion or political disagreement about how to handle the threat salient. Specifically, we prime the participants by exposing them to one of three news articles. In the *Invasion* treatment, participants read about Russia’s invasion of Ukraine and how this poses a threat to the interests of the U.S.

---

<sup>1</sup>We use the term ‘threat’ in the broad sense that encompasses both existential and symbolic threats as both types have been found to increase group cohesion (Riek, Mania, and Gaertner 2006).

In the *Disagreement* treatment, participants also read about Russia’s invasion, but the article focuses on the cross-party disagreement about how to handle the crisis. Finally, participants in *Control* read an apolitical and emotionally neutral news article. After the news prime, we measure affective polarization by using feeling thermometers, and we measure participants’ willingness to compromise in an incentivized coordination game.

Testing pre-registered hypotheses, we find that priming participants with Russia’s invasion of Ukraine leads to a modest reduction in affective polarization as measured by feeling thermometers. The difference between own- and opposite-party feeling thermometer rating (henceforth ‘FT difference’) is 2.9 degrees lower in the *Invasion* treatment than in *Control* (from a baseline of an FT difference of 51.4 in *Control*). Looking at political disagreement, we expected that it would mitigate or even reverse the effect of the external threat. Yet, we find that no statistically significant differences between affective polarization in *Invasion* and *Disagreement*.

We then move beyond attitudinal measures and examine how making the external threat salient influences participants’ behavior in a coordination game with real monetary stakes. We first demonstrate that greater affective polarization as measured by feeling thermometers predicts a greater tendency to discriminate based on the party affiliation of the opposing player. Then, we show that priming participants with Russia’s invasion increases their willingness to compromise. Specifically, the *Invasion* treatment increases participants’ probability of cooperating by approximately 6.5 percentage points regardless of the partisan affiliation of the other player (from a baseline probability of 40.3 percent in *Control*). Again, we find no significant differences between *Invasion* and *Disagreement*.

We make two substantial contributions to the literature on political psychology. We are (to the best of our knowledge) the first to study how the salience of an external threat causally influences affective polarization and how this depends on perceived (dis)agreement about how to handle the threat. In doing so, we build on previous research that relates external threats to political identities and political polarization. Related to our study, Gehring (2021) uses observational data to show that the Russian invasion of Ukraine in 2014 made citizens in Eastern Europe identify more strongly with the EU

and support common EU policies to a greater extent. Bafumi and Parent (2012) examine how political polarization between Republican and Democratic House Members (so-called “elite polarization”) was greater when the U.S. had relatively greater military capabilities than the Soviet Union during the Cold War.

Our second contribution is that we provide evidence of how affective polarization matters for incentivized behavior in a coordination game that measures people’s ability to cooperate with supporters of both parties. Previous studies have shown that individuals are more selfish and less trusting towards supporters of the opposite party (Fowler and Kam 2007; Iyengar and Westwood 2015; Whitt et al. 2021). The novel feature of our study is that we use a coordination game in which two players must agree on an action in order to earn money. Our game measures willingness to compromise for achieving the most efficient outcome (the common good), and it therefore captures essential aspects of the political process.

Finally, our finding that the salience of external threats matter for affective polarization and collaboration has implications for politicians in times of crises. Affective polarization greatly influences how well governments function and respond to critical challenges like the COVID-19 outbreak (Flores et al. 2022; Hetherington 2015). That external threats reduce affective polarization can therefore be vital for the immediate policy response – even if the effect is only transient (Chong and Druckman 2007; Myrick 2021). A key finding in this paper is that this effect is not significantly different when people perceive greater political disagreement regarding the crisis. This suggests that as long as people agree that there is an external threat, it can be possible to have open discussions about how to best handle the threat without harming the unifying effect of the threat.

## Context

For most of the 20th century, the U.S. faced critical external threats, and many argue that such threats have helped unite the American people and increase social cohesion (e.g., Desch 1996). The most prominent threats were World Wars I and II and the Cold War. The latter was marked by a bipolar global power distribution (Waltz 1979) which

lasted for more than 40 years. In this period, the threat of the communist Soviet Union provided a common enemy that could unite Republicans and Democrats (Bafumi and Parent 2012; Huntington 1997). The Cold War defined the key priorities of U.S. politics – both foreign and domestic – and the focus on the Soviet Union often distracted from internal issues in the U.S. (Bafumi and Parent 2012; Myrick 2021).

Following the collapse of the Soviet Union, which brought an end to the Cold War, the U.S. lost its counterweight, and the global power balance shifted towards unipolarity. There was no longer an empire fighting for principles that opposed the cornerstones to the American society, including liberty, democracy, and individualism (Desch 1996; Huntington 1997). And some even argued that humanity had progressed to the endpoint of its ideological evolution: Western liberal democracy (Fukuyama 1992). Since then, a major task for the U.S. has been to adjust to its role as the world’s sole military superpower, and the U.S. has had to redefine its interests in international relations. And while the 9/11 attacks made salient the threat of terrorism, the increased bipartisan support of President George W. Bush was short-lived and faltered with the 2003 invasion of Iraq (Khazatsky 2021). Thus, the U.S. has been without a critical military threat for many years, and it is likely that this has increased the scope for internal divisions (Bafumi and Parent 2012).

We argue that to understand the rise of affective polarization in the U.S., one must consider the development in external threats to the U.S. As we cannot manipulate the presence of an external threat, we use a priming approach to obtain the causal impact of making an external threat salient on affective polarization in the U.S. To this end, we exploit Russia’s full-scale invasion of Ukraine on February 24, 2022.<sup>2</sup> Though the U.S. did not involve its military in the conflict, it took a definite stand in support of Ukraine and imposed economic sanctions on Russia. In addition, it increased the number of troops stationed in NATO countries near Ukraine. The American public found Russia’s invasion alarming: In a survey of U.S. adults conducted March 21 – March 27, 2022,

---

<sup>2</sup>In Appendix A.1, we describe the historical background for the Russian invasion, and we describe its development up until the time of our experiment.

by Pew Research Center, 70 percent responded that they now consider Russia an enemy of the U.S., whereas in January 2022, Americans were equally likely to describe Russia as a competitor or as an enemy. In another survey by Pew Research Center (2022a), conducted April 25 – May 1, 50 percent responded that they were “extremely” or “very” concerned that the support for Ukraine might lead to a U.S. war with Russia. Democrats and Republicans both approved of placing strict economic sanctions on Russia (80 and 73 percent, respectively). Yet, there was also some cross-party disagreement regarding the conflict: 55 percent of Republicans “somewhat” or “strongly” disapproved of the Biden administration’s response to Russia’s invasion of Ukraine, whereas 63 percent of Democrats “somewhat” or “strongly” approved of the Biden administration’s response.

In the next section, we explain theoretical frameworks from the social identity approach that are useful for understanding how an external threat such as the one posed by Russia’s military aggression may influence polarization in the U.S.

## Theory

Affective polarization is rooted in people’s social identities (Iyengar et al. 2019), and we therefore draw on theories from the social identity approach to inform our study (Tajfel et al. 1971; Tajfel and Turner 1979). With social identity, we refer to “that part of an individual’s self-concept which derives from his knowledge of his membership in a social group (or groups) together with the value or emotional significance attached to that membership” (Tajfel 1978, p. 63). In our context, examples of social groups include ‘Democrats’, ‘Republicans’, and ‘Americans’. When individuals identify with social groups, they distinguish between those who belong to the same group as themselves (ingroup) and those who do not (outgroup, cf. self-categorization theory, Turner et al., 1987). Such group identification enables individuals to enhance their self-esteem and sense of meaning (Crocker and Luhtanen 1990), obtain a feeling of distinctiveness compared to outgroups (Turner et al. 1987), and reduce uncertainty about the social world (Abrams and Hogg 1988). Yet, it also leads to social comparisons between the groups. Because the group identification is emotionally significant, individuals are motivated to think highly of their own group (“ingroup love”) and focus on the negative characteristics of

the other group (“outgroup hate”, Weisel and Böhm, 2015). In our political context, such motivated reasoning gives rise to affective polarization, whereby voters are more likely to distrust and dislike those affiliated with the other party (Druckman and Levendusky 2019; Iyengar, Sood, and Yphtach 2012). These sentiments are further strengthened by people’s tendency to like those who share their attitudes (cf. the similarity-liking effect, Byrne 1961). When individuals distrust and dislike each other, collaboration for the greater good becomes more difficult to achieve (Hetherington 2015), suggesting that greater affective polarization should, *ceteris paribus*, predict less efficient outcomes when individuals interact with members of opposing parties. Yet, the presence of an external threat may reduce affective polarization in the two ways we describe in the following.

**The Feeling of a Common Ingroup.** First, an external threat may influence what individuals perceive to be their ingroup and outgroup. Individuals most often belong to more than one social group, and how they perceive themselves is influenced by the context they are in. As noted by Turner et al. (1994), the context dependence of identity does not imply that the “true” identity is being distorted; rather, it is a feature of a person’s identity that it is adaptive, as it makes identities more accurate and useful. The adaptivity of a person’s identity may be particularly important in the current setting: Whereas voters may differ in their partisan identities (Democrats vs. Republicans), they share the common identity of being Americans (cf. the common ingroup identity model, Gaertner et al. 1989; Gaertner and Dovidio 2000). When individuals think of themselves not as partisans but rather as belonging to a nation, members of the opposing party are seen as fellow Americans and become part of the ingroup rather than the outgroup. Importantly, the motivated reasoning that leads people to think highly of their ingroup now also encompasses the political opposition, whereby attitudes towards the political opposition should improve. Consequently, as shown by Levendusky (2018), increasing the salience of national identity reduces affective polarization. In the current context, such an increase in the salience of national identity may stem from the presence of international conflicts, such as the war between Russia and Ukraine, because they highlight relations between countries.

**Intergroup Conflict Theory.** Second, an external threat may also function as a rallying cry for individuals to overcome their differences and unite against the threat, a phenomenon known as the common enemy effect (Sumner 1906; Tajfel and Turner 1979; De Jaegher 2021). Previous research has demonstrated that perceived threats may reduce gaps between groups (Dovidio et al. 2004), and this occurs through both attitudes (Bonanno and Jost 2006) and emotions (Porat et al. 2019). The responses to such threats depend on a wide range of factors, including the nature of the intergroup relation, situational factors, and the cultural dimensions on which the groups differ (see Stephan, Ybarra, and Morrison, 2009, for a review). Most important for the present study, a group is more prone to perceive a threat from another group if the groups have a history of conflict (Stephan et al. 2002). In the current context, this may enhance the perception of threat that Americans feel when hearing about Russian military aggression. And with the perceived threat from Russia, Republicans and Democrats may feel a need to overcome their differences and instead focus on the conflict between Americans and Russians (cf. intergroup threat theory, Stephan, Ybarra, and Morrison 2009).

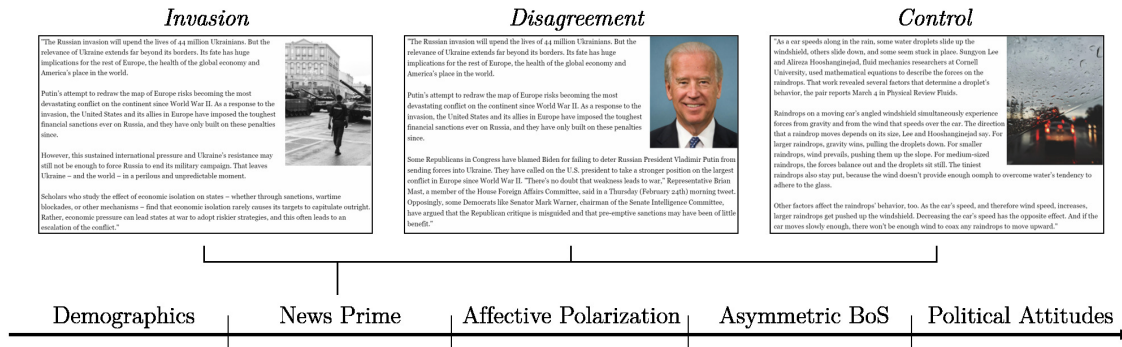
Importantly, the effect of international conflict may be *reversed* if individuals perceive an intense cross-party polarization related to the conflict. First, cross-party polarization may increase the salience of people’s partisan identity, which in turn may increase affective polarization (West and Iyengar 2020). Second, cross-party polarization could lead individuals to perceive their political opposition to be a hindrance for dealing with the threat, thereby reinforcing the polarization (John and Dvir-Gvirsman 2015; Orian Harel, Maoz, and Halperin 2020). As formulated by Brewer (1999, p. 436): “When negative evaluations of the outgroup such as contempt or fear are also already present, common threat in particular may promote scapegoating and blame rather than mutual cooperation”. An interesting question is therefore whether the effect of Russia’s military aggression changes when the perceived cross-party disagreement is high. We examine this in our experiment, which we describe in the following section.



## Experimental Design

Our experiment consists of five parts (see Figure 1), and these are completed in a single session using the online research platform Prolific. First, participants answer questions about their demographics which we use as controls in the analysis. Second, participants are randomly assigned to one of three treatments that vary in the content of a news prime. Participants either read about the threat of Russia’s invasion, the political disagreement regarding Biden’s response to the invasion, or an apolitical topic. Third, participants rate the Democratic and Republican parties on feeling thermometers and answer questions that reveal possible stereotypes regarding members of the two parties. Fourth, participants play an asymmetric Battle of the Sexes game; a coordination game in which one player can increase the total payoff by foregoing some personal earnings. Finally, participants answer questions about their political attitudes. The full set of instructions are available online and can be accessed here: <https://doi.org/10.17605/OSF.IO/6W829>

Figure 1: Timeline of the experiment



## Demographic Survey

Participants first answer demographic questions about their age, gender, ethnicity, education, and employment. This information provides us with details about the background of our sample. Moreover, we use demographics as control variables in testing our hypotheses because individuals’ demographics to some extent predict affective polarization (Iyengar et al. 2019), thereby increasing the efficiency of our coefficient estimates and our statistical power.

## Treatments: News Primes

In the second part of the experiment, participants read a brief news article, and they are asked to answer in 1-2 sentences a question about the content of the news article (the full articles are included in the online instructions).<sup>3</sup> We vary the topic and framing of the news article in a between-subjects design; in the discussion, we provide suggestive evidence that the effect of the news articles stems from subtle cues (priming) rather than from providing new information that changes participants' attitudes. We stratify participants based on their stated party affiliation on Prolific such that there is an equal number of Republicans and Democrats in each treatment.

In the *Invasion* treatment, participants first read an introduction to Russia's invasion of Ukraine. This describes how the fate of Ukraine "has huge implications for the rest of Europe, the health of the global economy and America's place in the world". The introduction also describes "Putin's attempt to redraw the map of Europe" and how "the United States and its allies in Europe have imposed the toughest financial sanctions ever on Russia". After this introduction, the article takes a threat perspective and describes that "economic pressure can lead states at war to adopt riskier strategies, and this often leads to an escalation of the conflict". Supporting the threat narrative, the article is accompanied by a picture of Russian tanks.

In the *Disagreement* treatment, participants read the same introduction to the Russian invasion as in *Invasion*. But the article continues to describe the disagreement between Republicans and Democrats about how well President Biden is handling the crisis. Specifically, it mentions how "some Republicans in Congress have blamed Biden for failing to deter Russian President Vladimir Putin from sending forces into Ukraine", quoting Representative Brian Mast saying that "there's no doubt that weakness leads to war." Participants then read how "some Democrats [...] have argued that the Republican critique is misguided". To emphasize the role of politics, the article is accompanied by a picture of President Joe Biden.

---

<sup>3</sup>The news articles draw heavily on the formulations by [Kingsley \(2022\)](#), [Morgan \(2022\)](#), and [Morales \(2022\)](#).

In *Control*, participants read an apolitical and emotionally neutral news article about how raindrops move on car windshields, illustrated with a picture. By making the content emotionally neutral, we avoid potential confounds with, e.g., individuals in a good mood perceiving the world in a more inclusive and integrative way (e.g., Bless and Fiedler 2006).

Upon reading the news article and writing 1-2 sentences about its content, participants in *Invasion* and *Disagreement* answer three questions about Russia’s invasion of Ukraine: How much they have followed the development of Russia’s invasion of Ukraine, how big a threat they consider Russia’s invasion of Ukraine to be for the U.S. and its interest, and to what extent they consider Democrats and Republicans to disagree/agree on how to handle Russia’s invasion of Ukraine. These questions serve as a manipulation check, as the *Invasion* treatment should increase perceived threat, and the *Disagreement* treatment should increase perceived disagreement compared to *Control*. In Appendix A.3.3, we show that this is indeed the case.

### **Affective Polarization**

In the third part of the experiment, participants answer standard questions used to elicit affective polarization. Participants start by rating the Democratic and Republican parties on feeling thermometers (American National Election Studies 1968), and we randomize the order of the parties to preclude order effects. A feeling thermometer is a scale from 0 to 100 where ratings between 0 and 49 degrees mean that one feels cold and unfavorable towards the party (with 0 being the coldest), and ratings between 51 and 100 degrees mean that one feels warm and favorable (with 100 being the warmest). A rating of 50 means that one neither feels warmly nor coldly towards the party.

We additionally elicit the participants’ stereotypes by asking them to rate the Democratic and Republican parties on five traits: Patriotism, selfishness, intelligence, open-mindedness, and honesty (Druckman and Levendusky 2019; Garrett et al. 2014; Iyengar, Sood, and Yphtach 2012). To avoid order effects, we randomize both the order of the party and the order of the traits.

## Asymmetric Battle of the Sexes Game

To obtain an incentivized, behavioral measure of participants’ ability and willingness to collaborate across party lines, the fourth part of the experiment consists of an asymmetric Battle of the Sexes game (Attanasi et al. 2016). Similar to the traditional Battle of the Sexes game, the asymmetric version is a one-shot coordination game in which two players (Player 1 and 2) must choose the same action of two (A and B) to earn money. For the asymmetric game, however, the total payoffs are different depending on what action the two players coordinate on (summarized in Table 1). Specifically, if both players choose A, Player 1 earns 35 cents, and Player 2 earns 5 cents. If both players choose B, Player 1 earns 15 cents, and Player 2 earns 35 cents. If the players choose differently, both Player 1 and Player 2 earn zero cents.

Table 1: Payoff matrix for the asymmetric Battle of the Sexes game

1\2	A	B
A	35,5	0,0
B	0,0	15,35

There are two pure-strategy Nash equilibria to the asymmetric Battle of the Sexes game: Both play A and both play B.<sup>4</sup> Yet, in this game, both players choosing B yields a greater total payoff. Importantly, this outcome does not Pareto dominate the other equilibrium; rather, Player 1 must forego some earnings to achieve the “greater good”, and she must agree to earn less than Player 2 to do so. Thus, analyzing participants’ decisions in the role of Player 1 informs about their willingness and ability to coordinate and compromise for the “greater good”.

Using the strategy elicitation method, participants make four potentially payoff-relevant decisions in this game. They decide both as Player 1 and as Player 2, and they play the game together with a Republican and a Democrat in random order. Participants are informed that we randomly match them with a supporter of the Republican or Democratic party after the survey is completed, and that one of their four decisions

---

<sup>4</sup>There is also a Nash equilibrium in mixed strategies, where Player 1 plays A with 87.5 percent probability and Player 2 plays B with 70 percent probability.

will be payoff relevant (randomly drawn and paid as a bonus via Prolific).<sup>5</sup>

Participants answer three control questions before making their decisions. Participants are allowed to continue only when they answer all three control questions correctly. If participants answer incorrectly, they are informed about this and are asked to try again.

An advantage of examining behavior in a coordination game is that the game creates a sense of dependency between the two players, as participants realize that they cannot unilaterally increase the social surplus.<sup>6</sup> In line with the idea that players are interdependent in the asymmetric Battle of the Sexes game, Attanasi et al. (2016) find that stronger social ties between the two players make it more likely that Player 1 will choose B.

We view the incentivized coordination game as complementary to the survey measures of affective polarization, as there are conceptual differences between the two approaches. Importantly, the feeling thermometers and elicitation of stereotypes concern general attitudes towards the Republican and Democratic parties. Previous research suggests that attitudes towards political parties are more comparable to attitudes towards party elites than towards ordinary voters (Druckman and Levendusky 2019). In contrast, the coordination game has the participants cooperate with individuals who support one of the two parties. A relation between, e.g., affective polarization in the feeling thermometers and decisions in the asymmetric Battle of the Sexes game would thus indicate that there is a relation between general attitudes towards political parties and small-scale interactions between individuals. As such, the incentivized coordination game also speaks to the broader debate about the relation between general attitudes and behavior (e.g., Fishbein and Ajzen 2010).

---

<sup>5</sup>A potential concern with using a within-subject design is that asking participants to play the game with persons from both parties could induce experimenter demand effects if participants, e.g., feel that they ought (not) to change their decisions. We nevertheless opted for a within-subject design to achieve sufficient power for our budget. Due to the random order of decisions, we are able to test for order effects (between-subjects). As described in Appendix A.3.5, we find no order effects, suggesting that the within-subject nature of our design does not introduce a bias into our analysis.

<sup>6</sup>This is in contrast to other games such as the public goods game, in which a single player can increase the social surplus regardless of what the other players do.

## Political Attitudes

In the final part of the survey, the participants answer general questions about their political preferences. Specifically, they indicate on a 5-point Likert scale how interested they are in politics, and they answer what party they support and how strongly they do so. For people indicating that they are Independent, we ask whether they consider themselves as closer to the Democratic Party or the Republican Party.<sup>7</sup> In our main analysis, we use self-reported political affiliation (and strength of this affiliation) as well as interest in politics as control variables because we expect these to predict some variation in affective polarization, thereby increasing the efficiency and statistical power of our analysis. Specifically, previous studies show that affective polarization is larger among Democrats than Republicans (Renström, Bäck, and Carroll 2021; West and Iyengar 2020), and strength of party support should correlate positively with affective polarization because the partisan identity is more emotionally significant for individuals who more strongly identify with either of the parties (Tajfel et al. 1971; Tajfel and Turner 1979).

After answering questions about their political attitudes, participants in *Control* answer the three questions about Russia’s invasion of Ukraine, which we use for our manipulation check in the analysis. We placed these questions at the very end of the survey for the participants in *Control* to avoid any priming effects of having the participants consider the invasion.

---

<sup>7</sup>Note that we ask about political preferences after eliciting affective polarization because making participants think about their political identity could influence our measurement of affective polarization and thereby distort our treatment effects. One concern is that the treatments could influence participants’ responses to the questions about political attitudes (cf. post-treatment bias, Montgomery, Nyhan, and Torres 2018). Yet, this is unlikely to be a concern in our case, as we find no differences across treatments in party affiliation, strength of party support, or interest in politics (all  $p$ ’s  $> 0.2$ , Kruskal-Wallis test). Moreover, our results are qualitatively robust to using the participants’ party affiliation as recorded on Prolific rather than their stated affiliation at the end of our experiment. In our main analysis, we follow our pre-registration and use party affiliation as reported in the experiment because this is the most recent information provided by the participants.

## Procedure

We recruited 1425 participants on Prolific between May 7 and 24, 2022, and the experiment was implemented in Qualtrics. As online experiments attract most respondents around the time when they are published, we started collecting data on a Saturday to ensure that we did not bias our sample against employed individuals. Online samples tend to be politically left-leaning and women are often over-represented (Paolacci, Chandler, and Ipeirotis 2010; Paolacci and Chandler 2014); therefore, we stratified the sample on Prolific to recruit an equal number of Democrats and Republicans as well as an equal number of men and women.<sup>8</sup>

We limited the sample to Americans who had completed 10 previous studies on Prolific with an approval rating of at least 98 percent. We applied several pre-registered screeners to ensure high-quality data, and we provide details for this in Appendix A.2.1. In total, we screened out 13 responses (0.9 percent). Furthermore, as our study concerns polarization between Democrats and Republicans, we excluded nine participants who identified as “true Independents”, leading to a main sample of 1403.<sup>9</sup> With this sample size, we expected based on power simulations to have a power of 0.8 to detect a treatment effect on feeling thermometer differences of approximately 5.4, equivalent to a standardized

---

<sup>8</sup>As we are interested in affective polarization, our main sample comprises only people who affiliate themselves with the Republican or the Democratic Party. To maximize the relevant sample given our budget, we therefore used the pre-recorded questions on Prolific in recruiting the most relevant participants. One might be concerned that we by doing so bias our sample towards the most politically interested sample. But the vast majority of participants on Prolific provide information about their political affiliation because doing so maximizes their chances of getting invited for studies. As of May 3, 2022, there were 37,916 active American participants on Prolific, of which 27,293 (72 percent) had reported their political affiliation. Of these, 17,821 (65 percent) declared support for either the Democratic or the Republican Party (Prolific 2022).

<sup>9</sup>Our results are robust to including all participants in the analysis. Furthermore, our results are robust to using RelevantID as a screener (Imperium 2022). RelevantID is an online fraud and duplicates detection service offered by Imperium, and it is embedded in Qualtrics. Note that we only use this for robustness as Zhang et al. (2022) demonstrate that combining individual screeners provides a similar fraud detection performance while making it more transparent why responses are flagged.

effect size of Hedge’s  $g_p = 0.15$  (see Appendix A.3.1 for details).

In our final sample, the mean age was 40 years, 49.9 percent were men, 82.3 percent were white or Caucasian, 62.2 percent were employed (part or full time), 10.2 percent were self-employed, 39.8 percent had obtained a Bachelor’s degree, and 13.6 percent had obtained a Master’s degree. The full set of sample characteristics is provided in Appendix A.3.2.

For completing the study, respondents earned on average USD 1.2 (min: 1.07, max: 1.41), and the median completion time was approximately 8 minutes (which is an upper bound as it relies on Qualtrics timing data that also counts time spent off task with the survey running in the background).

## Hypotheses

As described above, we rely on the social identity approach to form our (pre-registered) hypotheses. First, consider the *Invasion* treatment, which makes it salient that there is a threat and that the conflict is international. We posit that this makes participants think of their American identity as well as think of Russia as their common enemy, both factors contributing to a reduction in affective polarization. Thereby, we reach our first hypothesis:

**Hypothesis 1** *Affective polarization is lower in Invasion than in Control.*

Second, consider the *Disagreement* treatment. We expect that an emphasis on cross-party disagreement about how well President Biden is handling the crisis will counteract the two effects mentioned above. Specifically, the treatment increases the salience of political rather than national identities, and if individuals perceive an intense cross-party polarization related to the conflict, it is likely that they do not view their political opposition as a potential ally in dealing with the threat. We thereby reach our second hypothesis:<sup>10</sup>

**Hypothesis 2** *Affective polarization is greater in Disagreement than in Invasion.*

---

<sup>10</sup>We pre-registered to compare (i) *Invasion* with *Control*, (ii) *Disagreement* with *Control*, and (iii) *Invasion* with *Disagreement* if (i) and (ii) did not both show significant differences. In Appendix A.3.4, we report the results from (ii), which show that we fail to detect any difference between *Disagreement*



## How Do Threats Influence Affective Polarization?

In this section, we present the results on how the treatments influence affective polarization as measured by the feeling thermometers. In Appendix A.3.3, we show that the experimental manipulations work as intended: participants in *Invasion* perceive Russia to be a greater threat to the U.S. and its interests compared with participants in *Control*. Furthermore, participants in *Disagreement* perceive the same level of threat but more political disagreement about how to handle Russia’s invasion of Ukraine relative to participants in *Invasion*.

As pre-registered, our primary outcome measure is the difference in feeling thermometer ratings between own party and opposing party (henceforth FT difference). We planned to use stereotypes as secondary measures to provide more details on affective polarization. But in contrast to previous research (e.g., Garrett et al. 2014), we do not find affective polarization in stereotypes and refer instead the results of our pre-registered analyses to Appendix A.3.8.

### Descriptive Statistics

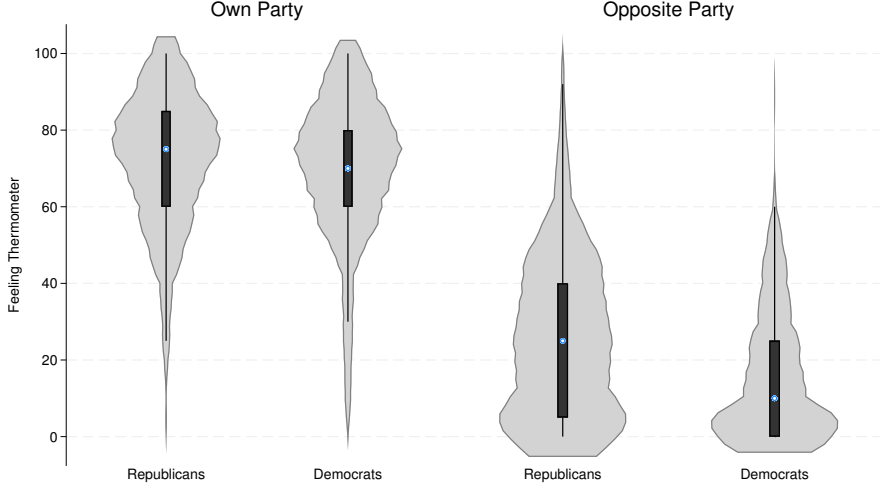
In line with previous research, we find affective polarization in the feeling thermometer questions. Across all participants, the average FT difference is 49.6 degrees. Participants on average rate their own party at 70.4 degrees and the opposite party at 20.8 degrees.<sup>11</sup> Figure 2 shows the distribution of answers to the feeling thermometer questions for Democrats and Republicans separately. As seen in the figure, we find a larger FT difference among Democrats than Republicans (in line with West and Iyengar, 2020, and Renström, Bäck, and Carroll, 2021): Both Democrats and Republicans rate their own party at around 70 degrees, but Democrats rate the Republican Party at 15.9 degrees while Republicans rate the Democratic Party at 25.8 degrees on average.

---

and *Control* for the full sample. This led us to test (iii), and we report (i) and (iii) in the main text for expositional purposes.

<sup>11</sup>Compared to the existing literature, we find a similar own-party feeling thermometer rating and around 5 degrees lower opposite-party rating (e.g., Iyengar, Sood, and Yphtach 2012; Levendusky 2018).

Figure 2: Affective polarization in Feeling Thermometer ratings



*Note:* This figure shows violin plots for how participants rate each party on the feeling thermometer. The plots show (epanechnikov) kernel density estimates, which depict the probability density of the data at different values. Within these estimates are standard box plots, which show the quartiles of the distributions (with the red dot as the median). The x-axis labels refer to partisan affiliation of the participants. For example, the left-most plot is the distribution of Republican participants rating their own party.

### Invasion Prime Reduces Affective Polarization

We now report the results related to our first hypothesis which states that affective polarization should be lower in *Invasion* than in *Control*. When conducting tests on FT difference, we follow the pre-analysis plan exactly and estimate OLS regressions “with (i) no controls, (ii) demographic controls (age, gender, ethnicity, education, and employment), and (iii) controls also for attitudes (party affiliation and interest in politics).” Our preferred specification is (iii) as we expect this to be the most efficient (as explained above). We also use non-parametric tests of treatment differences for robustness. Throughout and in line with our pre-registration, we rely on one-sided tests for the main tests as we have directional hypotheses, and we use two-sided tests otherwise.

In our preferred specification that includes all control variables, we find that affective polarization is lower in *Invasion* than *Control* as measured by FT difference, and this difference is statistically significant ( $t = -1.86, p = .034$ , one-sided test, cf. Table 2).<sup>12</sup>

<sup>12</sup>The treatment effect remains marginally significant when controlling for both the family-wise error

The size of the coefficient suggests that participants in *Invasion* on average rate a 2.86 degrees smaller FT difference relative to participants in *Control* (Hedge’s  $g_p = .088$ ). Arguably, this is a modest effect compared to the average FT difference of 51.4 in *Control*. It is somewhat smaller than the 5.6 degrees reduction that Levendusky (2018) finds from priming Americans with their national identity, but it is similar to the effect of 2.5 degrees that Boxell et al. (2020) find from priming individuals with the COVID-19 pandemic. It is, however, worth noting that we find this effect during the first months of the war when people were already influenced by the conflict (also in *Control*). This might also explain why our manipulation check only changes threat perceptions by .19 on a 5-point Likert scale (Hedge’s  $g_p = .180$ ).

In sum, our analysis of H1 leads us to conclude the following:

**Result 1** *Priming participants with Russia’s invasion of Ukraine leads to a modest reduction in affective polarization as measured by differences in feeling thermometer ratings.*

### **The Effect of Invasion Primes Is Unaffected by Disagreement**

Next, we report the results related to our second hypothesis, which states that affective polarization should be greater in *Disagreement* than in *Invasion*.

We find no difference in affective polarization between *Invasion* and *Disagreement* as measured by the FT difference (OLS:  $p = .779$ , cf. Table 2; MWU:  $p = .843$ ), and this result holds regardless of the level of controls. We return to these results in the discussion; for now, we simply conclude the following:

**Result 2** *When participants are primed with the threat of Russia’s military aggression, also priming participants with political disagreement about the conflict does not influence affective polarization.*

---

rate and the false discovery rate (both  $p$ ’s  $< .094$ ), cf. Appendix A.3.7). Yet, the nonparametric Mann-Whitney U-test just fails to reach marginal significance due to the increased noise from not including control variables ( $p = .104$ , one-sided test).

Table 2: Treatment effects on affective polarization

	Invasion vs. Control			Invasion vs. Disagreement		
Invasion	-2.504*	-2.476*	-2.859**	0.456	0.839	-0.444
	(1.860)	(1.855)	(1.563)	(1.850)	(1.852)	(1.581)
Constant	51.434***	46.806***	21.598***	48.474***	35.787***	14.960***
	(1.300)	(3.828)	(3.825)	(1.285)	(3.748)	(3.727)
N	926	926	926	933	933	933
Adj. R2	0.00	0.02	0.31	-0.00	0.03	0.29
Demographics	No	Yes	Yes	No	Yes	Yes
Attitudes	No	No	Yes	No	No	Yes

Notes: OLS regressions with FT difference as the dependent variable. Demographics include age, gender, dummies for ethnicity, and dummies for level of schooling. Attitudes include political interest (5-point Likert scale), partisan affiliation (dummy with value 1 if participant identifies as a Democrat), and strength of partisan affiliation (dummy with value 1 if participant is a strong supporter). Robust standard errors in parentheses.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (one-sided tests when in accordance with pre-registered hypotheses, two-sided otherwise)

## How Do Threats Influence Cooperation?

In this section, we examine participants’ behavior in the asymmetric Battle of the Sexes game to uncover how the salience of external threats influences people’s willingness to cooperate. We are particularly interested in the two decisions participants make in the role of Player 1 when facing either a Player 2 affiliated with their own or the opposite party. We first provide evidence that affective polarization as measured by the FT difference predicts participants’ behavior in the asymmetric Battle of the Sexes game. Then, we examine how the *Invasion* and *Disagreement* treatments influence cooperative behavior.

We use McNemar’s test to investigate within-subject changes in the proportions of participants who choose selfish/cooperative strategies. To investigate between-subjects treatment effects on behavior in the game, we use logit and multinomial logit as well as Fisher’s exact tests for robustness. We do not observe any order effects on Player 1 behavior (see Appendix A.3.5), and we therefore pool the data.

### Affective Polarization and Cooperation

Our main interest in the asymmetric Battle of the Sexes game is participants’ behavior in the role of Player 1. Participants display affective polarization in the sense that they choose differently depending on the partisan affiliation of the other participant: Pooling

all treatments, participants are more likely to choose the cooperative action B if Player 2 is an own-party member (47.1 percent) rather than an opposite-party member (39.3 percent). This difference is statistically significant ( $p < .001$ , McNemar’s test), and it is consistent with previous studies that show how partisans tend to behave more prosocially towards people from their own party in the Dictator Game and the Trust Game (Fowler and Kam 2007; Iyengar and Westwood 2015; Whitt et al. 2021).<sup>13</sup> The effects of ingroup-outgroup discrimination in our setting are large: If participants were as cooperative in the opposite-party condition as in the own-party condition, the total surplus would increase by 16.4 percent.<sup>14</sup>

As specified in our pre-analysis plan, we proceed to generate four “Player 1 types” that characterize how participants choose when facing an own- and opposite party members, respectively. We find that 49.2 (35.5) percent choose A (B) irrespective of the partisan identity of Player 2 (henceforth types *AA* and *BB*). Further, 11.6 percent of participants display affective polarization by playing the cooperative action B when Player 2 is from their own party, but the selfish action A when Player 2 is from the opposite party (*BA*). Only 3.7 percent display the opposite pattern of playing A when Player 2 is from their own party and B otherwise (*AB*).<sup>15</sup>

We expected that greater affective polarization would imply a greater tendency for participants to be of type *BA* rather than of type *BB*; accordingly, we find that affective polarization as measured by feeling thermometers predicts the participants’ incentivized

---

<sup>13</sup>In the decisions in the role of Player 2, there is not much indication of behavioral difference in own/opposite-party condition. Only around 8 percent of participants change their decision from own- to opposite-party condition in the role of Player 2 and a change in either direction is equally likely (McNemar’s  $\chi^2 = .08, p = .850$ ).

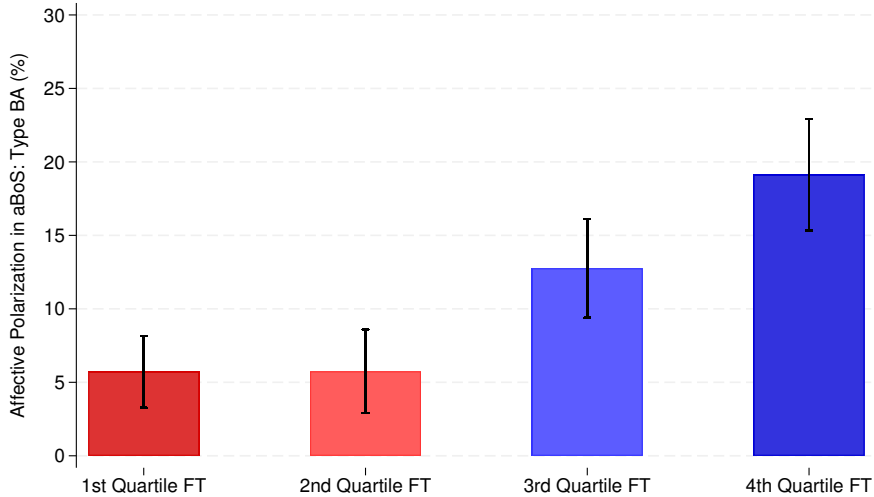
<sup>14</sup>To obtain this result, we use the frequencies from Player 2 behavior when facing a Player 1 affiliated with the opposite party. We then compute the expected total surplus given the average Player 1 behavior in the own-party and other-party conditions.

<sup>15</sup>Player 1 cooperation with own-party members is similar in our experiment to the level that Attanasi et al. (2016) find in their *University* treatment (47.1 vs. 49 percent) in which participants also share a generalized sense of belonging to an extended ingroup (students at the same university). Compared to Attanasi et al. (2016), we generally find a higher willingness to choose B in the role of Player 2.

behavior (see Figure 3). Specifically, a greater FT difference predicts a greater probability of being type *BA* and thus displaying affective polarization in the game (multinomial logistic regression,  $p < .001$ ). The marginal effect implies that a 10 degrees greater FT difference predicts a 2 percentage points greater probability of being of type *BA* in the asymmetric Battle of the Sexes game. This is a considerable effect considering that the baseline frequency of type *BA* is only 11.6 percent. In addition, a greater FT difference correlates significantly with a lower probability of being of type *BB* and a greater probability of being of type *AA*. As this relation holds across all treatments, we view this behavioral validation of feeling thermometers as an important, general result, and we summarize it as follows:

**Result 3** *Affective polarization as measured by feeling thermometers is behaviorally relevant: A greater difference in thermometer ratings predicts a greater tendency to discriminate based on party affiliation in an incentivized asymmetric Battle of the Sexes game.*

Figure 3: Share of types BA split by FT difference



*Note:* This figure divides participants into quartiles based on the size of the FT differences such that participants in the first quartile exhibit the smallest FT difference and participants in the fourth quartile exhibit the greatest FT difference. For each quartile, the bars indicate the share of participants who as Player 1 choose B when facing a person affiliated with own party and choose A when facing a person affiliated with the opposing party. Whiskers denote 95 percent confidence intervals.

## The Invasion Prime Increases General Willingness to Compromise

We now compare behavior in *Invasion* and *Control*. Cf. our pre-registration, we expected that a reduction in affective polarization would manifest in participants' willingness to compromise in the decision of Player 1. Accordingly, we find that participants in *Invasion* are more likely to choose the cooperative option B in the role of Player 1 compared to participants in *Control* (see Figure 4). Logistic regressions show that participants in *Invasion* are more likely to choose B both when facing a Player 2 who is affiliated with their own party (6.4 percentage points,  $p = .029$ , one-sided test) and the opposite party (6.7 percentage points,  $p = .021$ , one-sided test) compared to participants in *Control*.<sup>16</sup>

In addition, we expected that a reduction in affective polarization would cause participants to be less likely to discriminate based on the party affiliation of Player 2 (i.e., choose B in the own-party condition and A in the opposite-party condition). To test this, we perform a multinomial logistic regression with “Player 1 type” as the dependent variable (see Table 1). Contrary to our expectations, we find no difference between *Invasion* and *Control* in the likelihood of being of type BA ( $p = .612$ ). Rather, participants in *Invasion* are 7.7 percentage points less likely to be of type AA ( $p = .026$ ) and 6.1 percentage points more likely to be of type BB ( $p = .063$ ) relative to participants in *Control*. This supports the notion that salient threats increase general willingness to compromise. We return to this result in the discussion and here conclude as follows:

**Result 4** *The invasion prime makes participants more willing to compromise and choose B in the asymmetric Battle of the Sexes game irrespective of the partisan identity of the other player.*

---

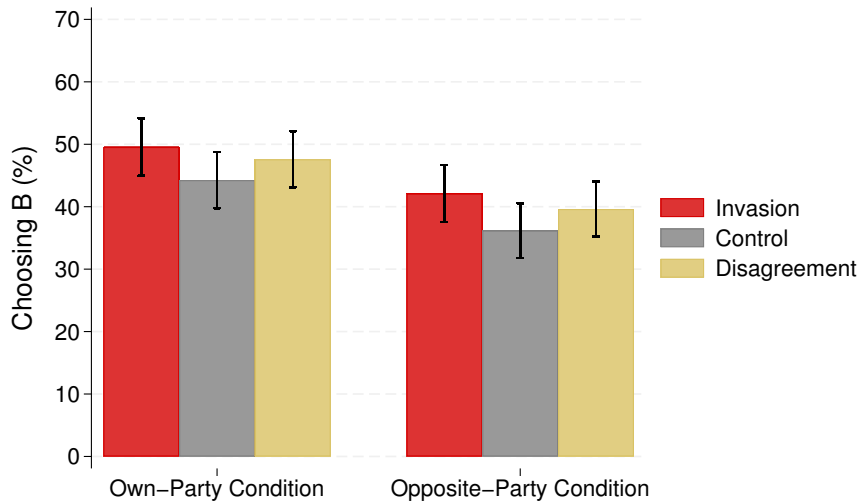
<sup>16</sup>Fisher's exact tests yield the same conclusion: Participants in *Invasion* are more likely to choose B when Player 2 is affiliated with their own party (49.6 vs. 44.3 percent,  $p = .060$ , Fisher's exact one-sided test) and with the opposite party (42.1 vs. 36.2 percent,  $p = .037$ , Fisher's exact one-sided test) compared to participants in *Control*.

## The Effect of Invasion Primes Is Unaffected by Priming Disagreement

We now compare Player 1 behavior in *Disagreement* and *Invasion*. Using logistic regressions, we find no statistically significant differences in the probability that participants play the cooperative option B – regardless of whether Player 2 is affiliated with the same party ( $p = .544$ , cf. Table A.3) or the opposite party ( $p = .393$ , cf. Table A.3). We also find no statistically significant effects on Player 1 types in multinomial regressions. Thus, we find that even though the *Disagreement* treatment successfully increases perceived political disagreement, this does not influence the effect of priming subjects with the Russian invasion (reflecting the results from before).

**Result 5** *When participants are primed with the threat of Russia’s military aggression, also priming participants with political disagreement about the conflict does not influence cooperative behavior.*

Figure 4: Decision of Player 1 across treatments



*Note:* This figure shows for each treatment the share of participants who choose B as Player 1 in the asymmetric Battle of the Sexes game when facing either a player affiliated with one’s own party or the opposite party. Whiskers denote 95 percent confidence intervals.

## Discussion

In the preceding analyses, we showed that priming Americans with Russia’s military aggression reduced affective polarization and increased cooperativeness. Against our



expectations, however, the effect of making the external threat salient was not affected by making internal disagreements about the threat salient as well. In the following section, we further discuss the effect of political disagreement and the effect of threats on cooperation, and we broaden the discussion to what can be learned from the priming method that we apply in this study.

We discuss further results in Appendix A.4. In particular, we show that there are treatment differences between Republicans and Democrats: Republicans rather than Democrats respond to the treatments with a reduction in affective polarization and an increase in willingness to cooperate in the coordination game. Further, we provide evidence that neither strength of partisan support, extent of political interest, nor the participants' gender nor age moderate our treatment effects.

### **The Effect of Political Disagreement**

As explained above, we find that when participants are primed with the threat of Russia's invasion of Ukraine, there is no additional effect of priming participants with political disagreement about how to handle the conflict. These results seem to be at odds with previous research that shows how making politics salient tends to increase affective polarization (e.g., Skytte [2021](#); West and Iyengar [2020](#)). In addition, they go against what one would expect from the importance of party cues for voter attitudes and behavior (Bullock [2019](#)): When voters are reminded that their own party has criticized the opposing party, the cue received from one's party is to feel animosity towards the opposing party.

One explanation for the absence of treatment differences between *Invasion* and *Disagreement* could be that the effect of Russia's military aggression dominates the effect of disagreement about how well President Biden is handling the crisis. As such, our study provides evidence that points in a different direction than the recent studies that explain how conflicts can make disagreements become even more pronounced and increase national divisions (John and Dvir-Gvirsman [2015](#); Orian Harel, Maoz, and Halperin [2020](#)). One possible reason for the mixed findings could be the type of threat that is being studied. From the perspective of the U.S., Russia's invasion of Ukraine primarily posed a

threat to American interests in Europe and to the geopolitical stability, with a widespread fear of the war escalating beyond the borders of Ukraine. In contrast, John and Dvir-Gvirsman (2015) study Israeli responses to the 2014 Gaza War which led to thousands of casualties and thus posed a much more existential threat. We view it as an interesting point for future research to explore what conditions determine if conflicts work to unite or divide a nation.

## **The Effect of Threats on Cooperation**

We have shown that the *Invasion* treatment makes participants more willing to compromise, and this result was not significantly different in *Disagreement*. In contrast to our expectations, willingness to compromise increased regardless of the partisan identity of Player 2. Thus, one cannot interpret the increase in cooperation as a result of reducing affective polarization. One can, however, understand this effect through the lens of the Perceived Target of Threat principle (Weisel and Zultan 2021). According to this principle, individuals tend to help the group if they perceive the group to be under threat, but they help themselves if they perceive themselves to be under threat. At the time of the present study, Americans arguably perceived the target of threat to be “the U.S.” rather than themselves as “individuals”. Hence, the *Invasion* treatment might make participants more cooperative in general, increasing their willingness to cooperate as Player 1 even when they are facing a Player 2 affiliated with their own party (whom group biases support also in *Control*).

## **Validity of Priming in Experiments**

As the priming method has become increasingly popular (Cohn and Maréchal 2016), researchers have become more aware of potential caveats with the method. One critique of priming is that effects tend to be short-lived. Indeed, the fact that we find an effect of priming participants with Russia’s invasion of Ukraine during the armed conflict suggests that the effect of the conflict is only present for individuals for whom the conflict is highly salient. It has long been known that while incumbent leaders become more popular following conflicts, this effect decays over time (Mueller 1973). Although we find causal

evidence of the effect of priming individuals with the conflict, our study provides no information about the long-run effects of the conflict or how this effect may change as the nature of the conflict changes.<sup>17</sup>

A related concern is whether our primes work as subtle situational cues as intended or whether the primes include new information that may change participants' attitudes (Cohn and Maréchal 2016). We address this in Appendix A.4.3.5 by investigating whether the observed treatment effects are moderated by the extent to which participants have been following the development of Russia's invasion of Ukraine. If the effect of the news articles was to provide new information, those who do not follow the invasion should respond more strongly as they know less about the conflict. Yet, we find no statistically significant differences in treatment effects based on how much participants follow the war; if anything, participants who have been following Russia's invasion of Ukraine actually reduce FT difference *more* in both the *Invasion* and *Disagreement* treatments. This suggests that the treatment effects in our study indeed follow from priming and not from new information.

## Conclusion

In this study, we have shown that making Russia's full-scale invasion of Ukraine in 2022 salient led to a modest reduction in affective polarization in the U.S. and that this effect did not depend on cross-party political disagreement. In addition, we have demonstrated that making Russia's military aggression salient increases general cooperativeness among Americans in an incentivized coordination game. Thus, we have shown that the presence of an external threat can reduce polarization and increase cooperation within a country. This suggests that global changes in international relations matter for within-country developments in polarization. As such, our study extends the literature that discusses why polarization has been on the rise in many countries (Iyengar et al. 2019)

---

<sup>17</sup>As for short-run effects, we can exploit that some participants took longer than others to complete the experiment. As we describe in Appendix A.4.2, we find that the treatment effects do not depend on the time between the news prime and answering the feeling thermometer questions (also when controlling for participants' speed).

and how polarization may be detrimental to collaboration (Hetherington 2015).

Yet, some factors challenge the external validity of our results. First, our sample differs from the American voting population on a number of observable characteristics. Specifically, our sample is younger, more educated, more likely to be unemployed, and Whites/Caucasians are over-represented. Yet, we control for all these background characteristics throughout the analyses, and we find no evidence that any of the demographics moderate our treatment effects (Appendix A.4.3). This corroborates the results from Snowberg and Yariv (2021) who find in a comparison between a representative sample and an MTurk sample that comparative statics are the same across samples even if averages change according to demographics (see also (Mullinix et al. 2015)).

A second limitation of our study is that we measure the causal effect of only one incident of military aggression, and we only measure this causal effect at one point in time. Our study provides no evidence for the generalizability of our effects across conflicts or time periods, and it does not speak to the long-run effects of the conflict.

An interesting avenue for future research is to elaborate on our surprising finding that priming participants with cross-party disagreement about how well President Biden is handling Russia’s invasion does not change the effect of the invasion prime. To us, this suggests that as long as people agree that there is an external threat, it can be possible to have open discussions about how to best handle the threat without harming the unifying effect of the threat. Future studies should explore how general this result is and shed light on what characteristics of a group determine if crises lead to less or more polarization.

## Acknowledgements

We especially thank Alexander Koch, Daniele Nosenzo, and Rasmus Skytte for valuable feedback and discussions. Also, we thank Hallgeir Sjøstad, Julia Nafziger, Nicola Maaser, Alexander Sebald, and seminar participants at the departments for Economics and Political Science at Aarhus University, the CNEE, the DGPE, and the Fairness and the Moral Mind Workshop at FAIR, NHH, for helpful comments.

## References

- Abrams, Dominic, and Michael A. Hogg. 1988. "Comments on the Motivational Status of Self-Esteem in Social Identity and Intergroup Discrimination." *European Journal of Social Psychology* 18 (4): 317–334.
- American National Election Studies. 1968. *ANES 1968 Time Series Study*. <https://electionstudies.org/data-center/1968-time-series-study/>.
- Attanasi, Giuseppe, Astrid Hopfensitz, Emiliano Lorini, and Frédéric Moisan. 2016. "Social Connectedness Improves Co-Ordination on Individually Costly, Efficient Outcomes." *European Economic Review, Social Identity and Discrimination*, 90:86–106.
- Bafumi, Joseph, and Joseph M Parent. 2012. "International Polarity and America's Polarization." *International Politics* 49 (1): 1–35.
- Bless, Herbert, and Klaus Fiedler. 2006. "Mood and the Regulation of Information Processing and Behavior." In *Affect in Social Thinking and Behavior*, edited by J. P. Forgas, 65–84. New York, NY, US: Psychology Press.
- Bonanno, George A., and John T. Jost. 2006. "Conservative Shift Among High-Exposure Survivors of the September 11th Terrorist Attacks." *Basic and Applied Social Psychology* 28 (4): 311–323.
- Boxell, Levi, Jacob Conway, James N. Druckman, and Matthew Gentzkow. 2020. *Affective Polarization Did Not Increase During the Coronavirus Pandemic*. Working Paper, Working Paper Series 28036. National Bureau of Economic Research.
- Brewer, Marilynn B. 1999. "The Psychology of Prejudice: Ingroup Love and Outgroup Hate?" *Journal of Social Issues* 55 (3): 429–444.
- Bullock, John G. 2019. "Party Cues." In *The Oxford Handbook of Electoral Persuasion*, edited by Elizabeth Suhay, Bernard Grofman, and Alexander H. Trechsel, 129–150. Oxford Handbooks, online edition.
- Byrne, Donn. 1961. "Interpersonal Attraction as a Function of Affiliation Need and Attitude Similarity." *Human Relations (US)* 14 (3): 283–289.
- Chong, Dennis, and James N. Druckman. 2007. "Framing Theory." *Annual Review of Political Science* 10 (1): 103–126.

- Cohn, Alain, and Michel André Maréchal. 2016. "Priming in Economics." *Current Opinion in Psychology* 12:17–21.
- Crocker, Jennifer, and Riia Luhtanen. 1990. "Collective Self-Esteem and Ingroup Bias." *Journal of Personality and Social Psychology* (US) 58 (1): 60–67.
- De Jaegher, Kris. 2021. "Common-Enemy Effects: Multidisciplinary Antecedents and Economic Perspectives." *Journal of Economic Surveys* 35 (1): 3–33.
- Desch, Michael C. 1996. "War and Strong States, Peace and Weak States?" *International Organization* 50 (2): 237–268.
- Dovidio, John F., Marleen ten Vergert, Tracie L. Stewart, Samuel L. Gaertner, James D. Johnson, Victoria M. Esses, Blake M. Riek, and Adam R. Pearson. 2004. "Perspective and Prejudice: Antecedents and Mediating Mechanisms." *Personality & Social Psychology Bulletin* 30 (12): 1537–1549.
- Druckman, James N, and Matthew S Levendusky. 2019. "What Do We Measure When We Measure Affective Polarization?" *Public Opinion Quarterly* 83 (1): 114–122.
- Fishbein, Martin, and Icek Ajzen. 2010. *Predicting and Changing Behavior: A Reasoned Action Approach*. New York: Psychology Press, Taylor & Francis Group.
- Flores, Alexandra, Jennifer C. Cole, Stephan Dickert, Kimin Eom, Gabriela M. Jiga-Boy, Tehila Kogut, Riley Loria, et al. 2022. "Politicians Polarize and Experts Depolarize Public Support for COVID-19 Management Policies across Countries." *Proceedings of the National Academy of Sciences* 119 (3): e2117543119.
- Fowler, James H., and Cindy D. Kam. 2007. "Beyond the Self: Social Identity, Altruism, and Political Participation." *The Journal of Politics* 69 (3): 813–827.
- Fukuyama, Francis. 1992. *The End of History and the Last Man*. Free Press.
- Gaertner, Samuel, and John Dovidio. 2000. *Reducing Intergroup Bias: The Common Ingroup Identity Model*. Philadelphia: Psychology Press.
- Gaertner, Samuel L., Jeffrey Mann, Audrey Murrell, and John F. Dovidio. 1989. "Reducing Intergroup Bias: The Benefits of Recategorization." *Journal of Personality and Social Psychology* (US) 57 (2): 239–249.

- Garrett, Kristin N., and Alexa Bankert. 2020. "The Moral Roots of Partisan Division: How Moral Conviction Heightens Affective Polarization." *British Journal of Political Science* 50 (2): 621–640.
- Garrett, R. Kelly, Shira Dvir Gvirsman, Benjamin K. Johnson, Yariv Tsfati, Rachel Neo, and Aysenur Dal. 2014. "Implications of Pro- and Counterattitudinal Information Exposure for Affective Polarization: Partisan Media Exposure and Affective Polarization." *Human Communication Research* 40 (3): 309–332.
- Gehring, Kai. 2021. "Can External Threats Foster a European Union Identity? Evidence from Russia's Invasion of Ukraine." *The Economic Journal*, ueab088.
- Gidron, Noam, James Adams, and Will Horne. 2020. "American Affective Polarization in Comparative Perspective." *Elements in American Politics*.
- Hansen, Kasper M, and Karina Kosiara-Pedersen. 2017. "How Campaigns Polarize the Electorate: Political Polarization as an Effect of the Minimal Effect Theory within a Multi-Party System." *Party Politics* 23 (3): 181–192.
- Hetherington, Marc. 2015. "'Why Polarized Trust Matters'." *The Forum* 13 (3): 445–458.
- Huntington, Samuel P. 1997. "The Erosion of American National Interests." *PALM COAST DATA* 76:28–49.
- Imperium. 2022. *RelevantID®* – Imperium. <https://www.imperium.com/relevantid/>. Accessed April 20, 2022.
- Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhotra, and Sean J. Westwood. 2019. "The Origins and Consequences of Affective Polarization in the United States." *Annual Review of Political Science* 22 (1): 129–146.
- Iyengar, Shanto, Guarav Sood, and Lelkes Yphtach. 2012. "Affect, Not Ideology: A Social Identity Perspective on Polarization." *The Public Opinion Quarterly* 76 (3): 405–431.
- Iyengar, Shanto, and Sean J. Westwood. 2015. "Fear and Loathing across Party Lines: New Evidence on Group Polarization." *American Journal of Political Science* 59 (3): 690–707.

- John, Nicholas A., and Shira Dvir-Gvirsman. 2015. ““I Don’t like You Any More”: Facebook Unfriending by Israelis during the Israel–Gaza Conflict of 2014.” *Journal of Communication* 65 (6): 953–974.
- Khazatsky, Andrei. 2021. “The Rise of Political Polarization in the U.S.” *The Universe*, June 4.
- Kingsley, Patrick. 2022. *The Roots of the Ukraine War: How the Crisis Developed*. <https://www.nytimes.com/article/ukraine-russia-putin.html>. Accessed April 21, 2022.
- Lau, Richard R., David J. Andersen, Tessa M. Ditonto, Mona S. Kleinberg, and David P. Redlawsk. 2017. “Effect of Media Environment Diversity and Advertising Tone on Information Search, Selective Exposure, and Affective Polarization.” *Political Behavior* 39 (1): 231–255.
- Layman, Geoffrey C., Thomas M. Carsey, and Juliana Menasce Horowitz. 2006. “Party Polarization in American Politics: Characteristics, Causes, and Consequences.” *Annual Review of Political Science* 9 (1): 83–110.
- Levendusky, Matthew S. 2018. “Americans, Not Partisans: Can Priming American National Identity Reduce Affective Polarization?” *The Journal of Politics* 80 (1): 59–70.
- MacKuen, Michael, Jennifer Wolak, Luke Keele, and George E. Marcus. 2010. “Civic Engagements: Resolute Partisanship or Reflective Deliberation.” *American Journal of Political Science* 54 (2): 440–458.
- Mason, Lilliana. 2015. ““I Disrespectfully Agree”: The Differential Effects of Partisan Sorting on Social and Issue Polarization.” *American Journal of Political Science* 59 (1): 128–145.
- Montgomery, Jacob M., Brendan Nyhan, and Michelle Torres. 2018. “How Conditioning on Posttreatment Variables Can Ruin Your Experiment and What to Do about It.” *American Journal of Political Science* 62 (3): 760–775.
- Morales, Marie. 2022. *Movement of Raindrops on Car Windshield Has Scientific Explanation As Gravity Forces Cause Them to Move*. <https://www.sciencetimes.com/art>



- [icles/36772/20220324/movement-raindrops-car-windshield-scientific-explanation-gravity-forces-cause-move.htm](https://www.reuters.com/world/us/republicans-target-biden-blame-over-putins-ukraine-invasion-2022-02-24/). Accessed March 24, 2022.
- Morgan, David. 2022. *Republicans Target Biden for Blame over Putin’s Ukraine Invasion*. <https://www.reuters.com/world/us/republicans-target-biden-blame-over-putins-ukraine-invasion-2022-02-24/>. Accessed February 25, 2022.
- Mueller, J. 1973. *War, Presidents and Public Opinion*. New York: John Wiley.
- Mullinix, Kevin J., Thomas J. Leeper, James N. Druckman, and Jeremy Freese. 2015. “The Generalizability of Survey Experiments.” *Journal of Experimental Political Science* 2 (2): 109–138.
- Myrick, Rachel. 2021. “Do External Threats Unite or Divide? Security Crises, Rivalries, and Polarization in American Foreign Policy.” *International Organization* 75 (4): 921–958.
- Orian Harel, Tal, Ifat Maoz, and Eran Halperin. 2020. “A Conflict within a Conflict: Intragroup Ideological Polarization and Intergroup Intractable Conflict.” *Current Opinion in Behavioral Sciences, Political Ideologies*, 34:52–57.
- Paolacci, Gabriele, and Jesse Chandler. 2014. “Inside the Turk: Understanding Mechanical Turk as a Participant Pool.” *Current Directions in Psychological Science* 23 (3): 184–188.
- Paolacci, Gabriele, Jesse Chandler, and Panagiotis G. Ipeirotis. 2010. “Running Experiments on Amazon Mechanical Turk.” *Judgment and Decision Making* 5 (5): 411–419.
- Pew Research Center. 2022a. “Americans’ Concerns About War in Ukraine: Wider Conflict, Possible U.S.-Russia Clash.” *Pew Research Center*, May 10.
- . 2022b. “Seven-in-Ten Americans Now See Russia as an Enemy.” *Pew Research Center*, April 6.
- Porat, Roni, Maya Tamir, Michael J. A. Wohl, Tamar Gur, and Eran Halperin. 2019. “Motivated Emotion and the Rally around the Flag Effect: Liberals Are Motivated to Feel Collective Angst (like Conservatives) When Faced with Existential Threat.” *Cognition and Emotion* 33 (3): 480–491.

- Prolific. 2022. *Prolific · Quickly Find Research Participants You Can Trust*. <https://www.prolific.co/#audience>. Accessed May 3, 2022.
- Renström, Emma A., Hanna Bäck, and Royce Carroll. 2021. “Intergroup Threat and Affective Polarization in a Multi-Party System.” *The Journal of Social and Political Psychology* 9 (2): 553–576.
- Riek, Blake M., Eric W. Mania, and Samuel L. Gaertner. 2006. “Intergroup Threat and Outgroup Attitudes: A Meta-Analytic Review.” *Personality and Social Psychology Review* 10 (4): 336–353.
- Skytte, Rasmus. 2021. “Dimensions of Elite Partisan Polarization: Disentangling the Effects of Incivility and Issue Polarization.” *British Journal of Political Science* 51 (4): 1457–1475.
- Snowberg, Erik, and Leeat Yariv. 2021. “Testing the Waters: Behavior across Participant Pools.” *American Economic Review* 111 (2): 687–719.
- Stephan, W. G., O. Ybarra, and K. R. Morrison. 2009. “Intergroup Threat Theory.” In *Handbook of Prejudice, Stereotyping, and Discrimination*, edited by T. D. Nelson, 43–59. New York: Psychology Press.
- Stephan, Walter G., Kurt A. Boniecki, Oscar Ybarra, Ann Bettencourt, Kelly S. Ervin, Linda A. Jackson, Penny S. McNatt, and C. Lausanne Renfro. 2002. “The Role of Threats in the Racial Attitudes of Blacks and Whites.” *Personality and Social Psychology Bulletin* 28 (9): 1242–1254.
- Sumner, W. G. 1906. *Folkways: A Study of the Sociological Importance of Usages, Manners, Customs, Mores, and Morals*. Boston, MA: Ginn.
- Tajfel, H. 1978. “Social Categorization, Social Identity and Social Comparisons.” In *Differentiation between Social Groups*, edited by H. Tajfel, 61–76. London, UK: Academic Press.
- Tajfel, H., and J. C. Turner. 1979. “An Integrative Theory of Intergroup Conflict.” In *The Social Psychology of Intergroup Relations*, edited by W. G. Austin and S. Worchel, 33–37. Monterey, CA, USA: Brooks/Cole.

- Tajfel, Henri, M. G. Billig, R. P. Bundy, and Claude Flament. 1971. “Social Categorization and Intergroup Behaviour.” *European Journal of Social Psychology* 1 (2): 149–178.
- Turner, John C., Michael A. Hogg, Penelope J. Oakes, Stephen D. Reicher, and Margaret S. Wetherell. 1987. *Rediscovering the Social Group: A Self-Categorization Theory*. x, 239. Cambridge, MA, US: Basil Blackwell.
- Turner, John C., Penelope J. Oakes, S. Alexander Haslam, and Craig McGarty. 1994. “Self and Collective: Cognition and Social Context.” *Personality and Social Psychology Bulletin* 20 (5): 454–463.
- Waltz, K.N. 1979. *Theory of International Politics*. New York: McGraw-Hill.
- Weisel, Ori, and Robert Böhm. 2015. ““Ingroup Love” and “Outgroup Hate” in Intergroup Conflict between Natural Groups.” *Journal of Experimental Social Psychology* 60:110–120.
- Weisel, Ori, and Ro’i Zultan. 2021. “Perceived Level of Threat and Cooperation.” *Frontiers in Psychology* 12.
- West, Emily A., and Shanto Iyengar. 2020. “Partisanship as a Social Identity: Implications for Polarization.” *Political Behavior* 44:807–838.
- Whitt, Sam, Alixandra B. Yanus, Brian McDonald, John Graeber, Mark Setzler, Gordon Ballingrud, and Martin Kifer. 2021. “Tribalism in America: Behavioral Experiments on Affective Polarization in the Trump Era.” *Journal of Experimental Political Science* 8 (3): 247–259.
- Zhang, Ziyi, Shuofei Zhu, Jaron Mink, Aiping Xiong, Linhai Song, and Gang Wang. 2022. “Beyond Bot Detection: Combating Fraudulent Online Survey Takers.” *Proceedings of the ACM Web Conference 2022*, 1–11.

## Biographical Statement

Jonas Pilgaard Kaiser is a postdoc at the Technical University of Berlin, 10623 Berlin, Germany. Markus Seier is an affiliated researcher at Aarhus University, 8200 Aarhus, Denmark.

# Online Appendix

## Did Russia’s Invasion of Ukraine Reduce Affective Polarization in the U.S.? Experimental Evidence

Jonas Pilgaard Kaiser

Markus Seier

Aarhus University

Aarhus University

February 17, 2025

### A.1 Background: Russia’s Invasion of Ukraine

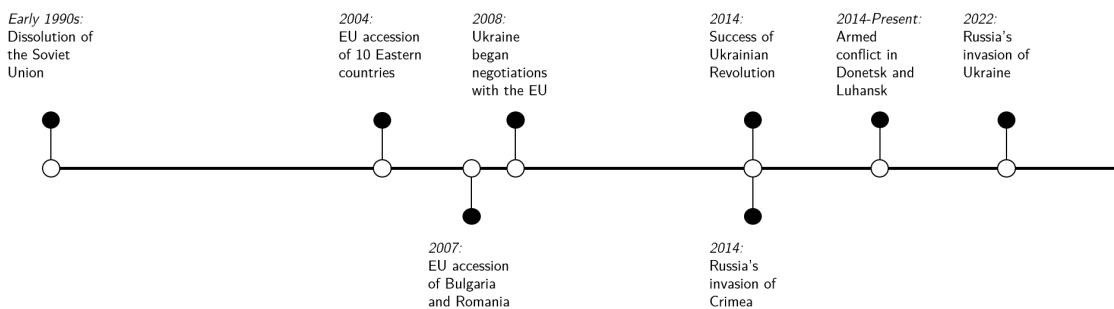
In the following section, we give a brief account of important developments for Ukraine’s recent history, and we describe how Americans viewed the Russian invasion at the time of our experiment. Our aim is not to provide full details of Ukraine’s complex history but rather to provide background information necessary for understanding the scope of the current conflict.

After Ukraine left the Soviet Union in 1991, Ukraine’s main political goals have been to ensure independence and sovereignty while simultaneously balancing cooperation with the European Union and Russia (Shyrokykh [2018](#)). This balance has been increasingly difficult to maintain as the European Union and the North Atlantic Treaty Organization (NATO) have expanded eastward. Specifically, the European Union enrolled 12 new states between 2004 and 2007 and initiated negotiations about an Association Agreement with Ukraine in 2008 (Gehring [2021](#)). Further, NATO announced at the Bucharest summit in 2008 plans of some day enrolling Ukraine and Georgia in the organization (Bebler [2015](#)). This eastward expansion has been viewed critically by Russia and President Vladimir Putin, who on several occasions has expressed his views of Ukraine being a part of Russia (Mackintosh [2022](#)).

The tensions between Russia and Ukraine increased drastically after the 2014 Ukrainian Revolution. The revolution followed the Russian-friendly President Viktor Yanukovich’s

refusal to sign an Association Agreement with the European Union, and it ended with Yanukovych being forced to flee the country (Gehring 2021). Afterwards, a separatist rebellion broke out in the east of Ukraine, and this gained support from Russia. In March 2014, Russia deployed military to Crimea and took over government buildings. Russian-backed authorities held a referendum shortly after this invasion, and Crimean voters overwhelmingly chose to join Russia. While Ukraine and Western countries called this referendum illegitimate, President Vladimir Putin finalized the absorption of the peninsula into Russia. The conflict then shifted to the Donetsk and Luhansk regions in eastern Ukraine. The pro-Russian separatists held a self-rule referendum claiming independence of the regions, but the Ukrainian government responded with a so-called “anti-terrorist operation” against the separatists. Throughout this operation, President Vladimir Putin denied Russian military involvement (Roman, Wanta, and Buniak 2017). The Ukrainian government and the separatists agreed on cease fire with the Minsk peace agreement in 2015, but there has not been stable peace in the regions since. By 2021, more than 13,000 people had been killed in this conflict and Western countries have responded with sanctions on Russia (HistoryExtra 2022).

Figure A.1: Timeline of the Russian/Ukrainian conflict



A dramatic escalation of the conflict occurred on February 24, 2022, when Russia launched a full-scale invasion of the Ukrainian mainland. The immediate response from the U.S. and European countries was to impose economic sanctions targeting Russian banks and Russia’s oil and gas industry. The U.S. did not employ any military in Ukraine, but they increased the number of troops in NATO countries near Ukraine. As of May 17 (around the time of our experiment), more than 3,380 civilians were confirmed to have

been killed (UN 2022), and 12 million refugees were believed to have fled Ukraine (BBC 2022).

## A.2 Experimental Design

### A.2.1 Screeners

In the following, we describe the pre-registered screeners that we applied to ensure high-quality data in our sample. As mentioned in the main text, we limited the sample to Americans who had completed 10 previous studies on Prolific with an approval rating of at least 98 percent (Matherly 2019). The following screeners led to the exclusion of in total 13 of 1425 responses (0.9 percent).

First, we used two ‘honeypots’ (coded in JavaScript) to detect bots. Honeypots are traps set up to engage and detect bad actors in a computer system. We followed Moss and Litman (2018a) and used survey items that were hidden from humans but would be read by a bot. Thus, only bots would be able to answer these questions. If any response was provided to one of these questions, the respondent was confirmed to be a bot, and we dropped it from the experiment.<sup>1</sup> We detected no bots in our sample.

Second, we followed Kennedy et al. (2020) and included a consistency check in the demographic questions. The first item asked participants about their age, and the last item asked about their year of birth (see also Zhang et al. 2022). This resulted in the exclusion of 9 participants who answered these questions inconsistently.

Third, we followed Chmielewski and Kucker (2020) and used the text prime as an additional screener for farmers (see also Dennis, Goodson, and Pearson 2020; Zhang et al. 2022). Farmers are respondents who manage to access the study despite not being in the U.S. (e.g., via server farms) and not being proficient in English (Moss and Litman 2018b). We did not identify any responses that misused the English language, used

---

<sup>1</sup>We prefer this honeypot to (re)CAPTCHAs because some bots are able to pass CAPTCHAs (Al-Fannah 2017; Sivakorn, Polakis, and Keromytis 2016), and the honeypot is unobtrusive, saving time and making it easier for people with vision impairment to complete the study (Bursztein et al. 2010). We applied the honeypots to both the consent form and the demographic survey as bots and humans may work in hybrid, using bots to complete only some parts of a survey.

nonsense phrases, or answered in single words unrelated to the question (e.g., “nice” and “good”). Neither did we find any responses that copied entire paragraphs from the news prime.

Finally, two respondents somehow managed to answer the survey twice. We exclude these four responses from our sample.

Note that we do not test for Virtual Private Servers (VPS) or Virtual Private Networks (VPN) although recommended by Kennedy et al. (2020) as Prolific regularly test for this to ban respondents (Prolific 2021). Also, our study did not use attention checks such as Instructional Manipulation Checks (Oppenheimer, Meyvis, and Davidenko 2009). These have been found to have limited diagnostic value as many participants in online samples are familiar with this type of question (Hauser and Schwarz 2016; Thomas and Clifford 2017), and they may change people’s behavior (Hauser and Schwarz 2015; Hauser, Ellsworth, and Gonzalez 2018).

## **A.3 Analysis**

### **A.3.1 Power Analysis**

The following section presents the a priori power analysis that informed our study and which we used in the pre-registration. With the available funding, we planned to recruit a sample of 1425 participants, which we planned to randomize equally to each of the three treatments. Our primary comparison is the difference in ratings on the feeling thermometer between participants’ own party and the opposing party. Based on Boxell et al. (2020), we expected that the participants rate their own party at a mean of 62 with a standard deviation of 24, and we expected that the participants rate the opposing party with a mean of 40 and a standard deviation of 26. Regarding the within-subject correlation between the two parties, one could a priori imagine two effects that work in opposite directions: On the one hand, when individuals strongly identify with one party, this may create a larger discrepancy between the two ratings, resulting in a negative within-subject correlation. On the other hand, if participants differ in their interpretation of the scale and the notions of “warm” and “cold”, this will create a positive

within-subject correlation. As we do not know which of these effect will be stronger, we assumed for the power analysis that there would be no correlation between ratings.

Based on the above assumptions, we ran power simulations in Stata (version 17) that showed us to expect a power of 0.8 to detect a treatment effect on FT difference of approximately 5.4, equivalent to a standardized effect size of Hedge’s  $g_p = 0.15$ . For comparison, Boxell et al. (2020) find that the party difference is 22 points. Also, Levendusky (2018) uses a national news prime to find an effect of 5.6 on an out-party feeling thermometer.

As explained in the paper, our final sample size was slightly below our target sample size (1,403 vs. 1,425). But such a small difference does not change the conclusions from our power simulations.<sup>2</sup>

### A.3.2 Sample Characteristics Table

The full set of sample characteristics are provided in Table A.1.

### A.3.3 Manipulation Check

Before turning to our main hypotheses, we first show that the experimental manipulations work as intended. That is, we first test whether participants in the *Invasion* treatment perceive Russia to be a greater threat to the U.S. and its interests relative to participants in *Control*. Second, we test whether participants in the *Disagreement* treatment perceive the same level of threat but more political disagreement about how to handle Russia’s invasion of Ukraine relative to participants in *Invasion*. Figure A.2 shows the average threat perception and disagreement perception in the three treatments.

**Invasion.** Participants in the *Invasion* treatment perceive Russia’s invasion of Ukraine as a greater threat to the U.S. and its interests relative to participants in *Control* (3.52 vs. 3.33,  $p = .006$ ,  $t$ -test). Instead of relying on the numerical rating, we can also see that participants in *Invasion* are more likely to rate the threat of Russia’s invasion of Ukraine as “somewhat severe” or “severe” relative to participants in *Control* (53.1 vs. 44.3 percent,  $p = .009$ , Fisher’s exact test). We find no difference in the perception of

---

<sup>2</sup>We refrain from changing any other inputs for the power analysis to avoid the problems of ex-post power calculations, (Hoenig and Heisey 2001).



Table A.1: Sample Characteristics

	Variable	Mean	Std. Dev.	Min.	Max.
	Age	40.00	14.38	18	80
	Male	0.50	.	0	1
<b>Ethnicity</b>					
	White	0.82	.	0	1
	Asian American	0.07	.	0	1
	Hispanic	0.05	.	0	1
	Black	0.04	.	0	1
	Other Ethnicity	0.01	.	0	1
<b>Employment Status</b>					
	Employed	0.72	.	0	1
	Self Employed	0.10	.	0	1
	Student	0.08	.	0	1
	Retired	0.07	.	0	1
	Other Employment	0.03	.	0	1
<b>Education Level</b>					
	Less Than High School	0.01	.	0	1
	High School	0.12	.	0	1
	Some College	0.20	.	0	1
	Associate Degree	0.10	.	0	1
	Bachelors Degree	0.40	.	0	1
	Masters Degree	0.14	.	0	1
	Doctorate Degree	0.04	.	0	1
<b>Political Attitudes</b>					
	Democrat	0.51	.	0	1
	Strong Supporter	0.48	.	0	1
	Political Interest	3.34	0.98	1	5
<b>Experimental Characteristics</b>					
	Duration in Seconds	569.15	353.65	134	4,050
	Fraud Score	0.86	5.08	0	75
	Prolific Score	99.83	0.39	98	100

disagreement between *Invasion* and *Control* (2.91 vs. 2.89,  $p = .703$ ,  $t$ -test).

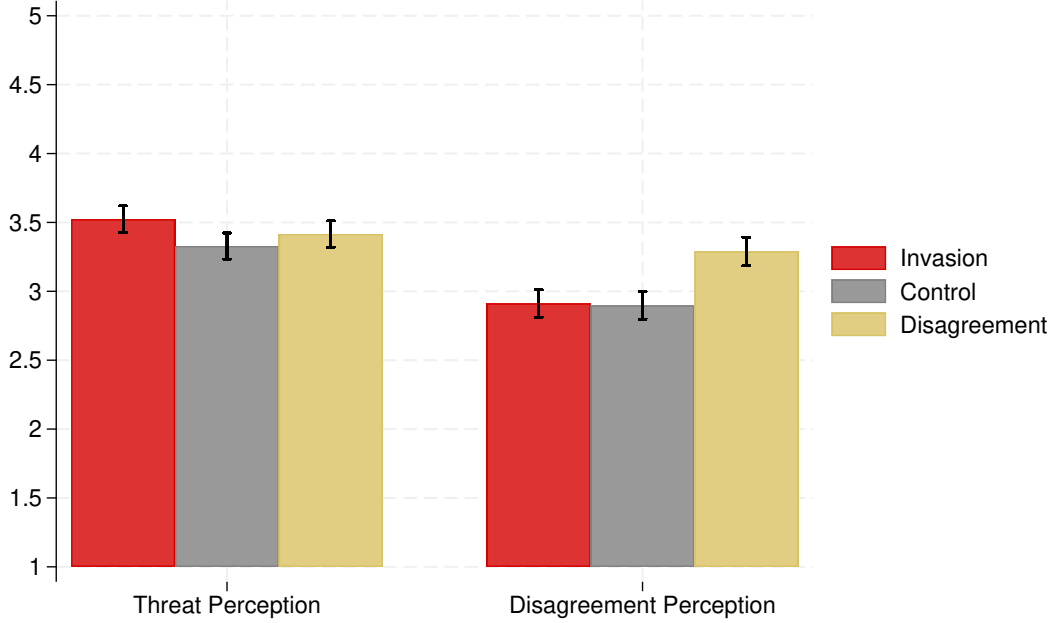
**Disagreement.** Participants in *Disagreement* perceive the same level of threat as in *Invasion* (3.42 vs. 3.52,  $p = .117$ ,  $t$ -test). Yet, participants in *Disagreement* perceive more political disagreement about how to handle Russia’s invasion of Ukraine relative to participants in *Invasion* (3.29 vs. 2.91,  $p < .001$ ,  $t$ -test). Similarly, participants in *Disagreement* are more likely to answer that Democrats and Republicans “strongly disagree” or “somewhat disagree” on how to handle Russia’s invasion of Ukraine relative to participants in *Invasion* (52.6 vs. 35.5 percent,  $p < .001$ , Fisher’s exact test).<sup>3</sup>

---

<sup>3</sup>The *Disagreement* treatment also successfully increases perceived disagreement compared to participants in *Control* (3.29 vs. 2.89,  $p < .001$ ,  $t$ -test). Similarly, participants in *Disagreement* are more likely to answer that Democrats and Republicans “strongly disagree” or “somewhat disagree” on how to

We conclude that our news primes succeeded in generating (i) a greater average threat perception in *Invasion* and (ii) a greater average disagreement perception in *Disagreement*.

Figure A.2: Manipulation check



*Note:* This figure shows for each treatment the average perceived threat and the average perceived political disagreement between the two parties. Whiskers denote 95 percent confidence intervals.

#### A.3.4 Affective Polarization in *Disagreement* and *Control*

In this section, we report the results from the pre-registered comparison between *Disagreement* and *Control*. We first report the results from the feeling thermometer and then turn to behavior in the asymmetric Battle of the Sexes game.

**Feeling Thermometers** We expected that making political disagreement salient would increase affective polarization compared to *Control*. Yet, contrary to our expectations participants report a 1.8 degrees greater FT difference in *Control* than in *Disagreement*. This indicates that affective polarization is reduced in *Disagreement*, but the difference is only marginally statistically significant in the specification with demographic controls (handle Russia's invasion of Ukraine relative to participants in *Control* (52.6 vs. 33.4 percent,  $p < .001$ , Fisher's exact test)).

( $p = .093$ ). The nonparametric Mann-Whitney U-test fails to reach statistical significance ( $p = .152$ , which is expected as this test does not incorporate control variables). Thus, we cannot confirm our pre-registered hypothesis and the effect even seems to move in the opposite direction.

**Asymmetric Battle of the Sexes.** In the coordination game, we expected to find that more people would compromise and choose B in *Control* than in *Disagreement*. But reflecting the results of the feeling thermometer, we find that the effect moves in the opposite direction as participants in *Disagreement* are 3 (3.2) percentage points more likely to choose B when Player 2 is affiliated with their own (the opposite) party. None of these effects reach statistical significance ( $p = .384$  and  $p = .332$ ).<sup>4</sup> Looking at “Player 1 types”, the marginal effects suggest that participants in *Disagreement* are more likely to be type *BB* and less likely to be type *AA* relative to participants in *Control*. However, the effects are not statistically significant (all  $p$ 's  $> .217$ ).

### A.3.5 Test for Order Effects in Asymmetric Battle of the Sexes

In the asymmetric Battle of the Sexes game, we are interested in participants' decisions in the role of Player 1. One concern when asking participants to play the game with a persons from both parties is that participants might feel that they ought (not) to change their decisions. As we randomized the order of the party affiliation of the first person the participants played the game with, we are able to test for order effects.

We test for the existence of order effects in participants' decisions in the role of Player 1. Participants make their first decision in the role of Player 1 either facing a person from their own party or from the opposite party. For both the own- and opposite-party condition, we test whether there are significant differences in Player 1's decision depending on which condition the participants is randomized into.

First, we investigate the own-party condition. 694 participants are randomized into

---

<sup>4</sup>Fisher's exact tests yield the same conclusion: The difference is statistically insignificant both when Player 2 is affiliated with the participants' own party (47.6 vs. 44.3 percent,  $p = .328$ , Fisher's exact test) and the opposite party (39.6 vs. 36.2 percent,  $p = .284$ , Fisher's exact test).

the own-party condition first and 709 participants are randomized into the own-party condition last. The frequency of choosing B is very similar across these two game orders (47.4 percent vs. 46.8 percent) and a Fisher’s exact test cannot reject that the decision in the role of Player 1 when the other person is from the same party is independent of game order ( $p = .831$ , Fisher’s exact test).

Secondly, we investigate the opposite-party condition. Mirroring the above, 709 participants are randomized into the opposite-party condition first and 694 are randomized into the opposite-party condition last. Again, the frequency of choosing B is very similar across these two game orders (39.9 percent vs. 38.6 percent) and a Fisher’s exact test cannot reject that the decision in the role of Player 1 when the other person is from the opposite party is independent of game order ( $p = .623$ , Fisher’s exact test).

Thus, we conclude that the within-subject design does not introduce a bias into our analysis and we pool the data in the analysis.<sup>5</sup>

### A.3.6 Tables for Asymmetric Battle of the Sexes

Table A.2: Player 1 types in the Battle of the Sexes game by party affiliation

<b>Player 1 Type</b>	<b>Republicans</b>	<b>Democrats</b>	<b>Total</b>
(AA)	55.2%	43.3%	49.2%
(AB)	4.6%	2.8%	3.7%
(BA)	9.6%	13.5%	11.6%
(BB)	30.6%	40.4%	35.6%

Notes: First letter refers to participants’ Player 1 decision when Player 2 is affiliated with own party, second letter refers to their Player 1 decision when Player 2 is affiliated with the opposite party.

---

<sup>5</sup>We actually find some evidence of order effects in the role of Player 2. Specifically, in the role of Player 2 in the own-party condition, a greater share of subjects choose A when they play the own-party condition last. However, we participants’ decisions in the role of Player 2 is not the focus of our analysis.

Table A.3: Treatment effects on Player 1 probability of choosing B

<b>Invasion vs. Control</b>						
	<b>Own-Party Condition</b>			<b>Opposite-Party Condition</b>		
Invasion	0.053*	0.062**	0.064**	0.059**	0.065**	0.067**
	(0.033)	(0.034)	(0.034)	(0.032)	(0.033)	(0.033)
N	926	926	926	926	926	926
Demographics	No	Yes	Yes	No	Yes	Yes
Attitudes	No	No	Yes	No	No	Yes
<b>Invasion vs. Disagreement</b>						
	<b>Own-Party Condition</b>			<b>Opposite-Party Condition</b>		
Invasion	0.020	0.021	0.021	0.025	0.026	0.028
	(0.033)	(0.034)	(0.034)	(0.032)	(0.033)	(0.033)
N	933	933	933	933	933	933
Demographics	No	Yes	Yes	No	Yes	Yes
Attitudes	No	No	Yes	No	No	Yes

Notes: The table shows marginal effects (average partial effects) from logistic regressions with Player 1 probability of choosing B as the dependent variable. Demographics include age, gender, dummies for ethnicity, and dummies for level of schooling. Attitudes include political interest (5-point Likert scale), partisan affiliation (dummy with value 1 if participant identifies as a Democrat), and strength of partisan affiliation (dummy with value 1 if participant is a strong supporter). Robust standard errors calculated using Delta method in parentheses.

\*  $p < .10$ , \*\*  $p < .05$ , \*\*\*  $p < .01$  (one-sided tests when in accordance with pre-registered hypotheses, two-sided otherwise)

### A.3.7 Multiple Hypothesis Testing

In accordance with our pre-registration, we control for multiple hypothesis testing to investigate the robustness of our findings (Cramer et al. 2016; List, Shaikh, and Xu 2019). We report here the results when we control for the family-wise error rate (FWER). Note that controlling for FWER also implies controlling for the false discovery rate (FDR). For this exercise, we consider as our ‘family’ the three main hypotheses with the test for our primary measure (Farcomeni 2008), which is FT difference and we do not control for multiple hypothesis testing in exploratory analysis (Bender and Lange 2001). In controlling for the FWER, we adjust the p-values from the regression with all control variables as this is our preferred specification.

We use the Romano-Wolf resampling procedure to control for FWER (Romano and Wolf 2005a; Romano and Wolf 2005b, 2016), and we implement this in Stata using the package developed by Clarke, Romano, and Wolf (2020). The advantage of this approach

is that it is a resampling procedure and thus takes the dependency between tests into account. We conduct this adjustment using 1000 resamples, which gives a Romano-Wolf adjusted  $p$ -value of .094 (unadjusted  $p = .040$ ) in the *Invasion vs. Control* comparison. The adjustment does not change the conclusion in the *Invasion vs. Disagreement* comparison ( $p = .758$ ). Thus, when controlling for FWER and FDR, we conclude that the treatment effect in the *Invasion vs. Control* comparison remains marginally significant while the effect in the *Invasion vs. Disagreement* comparison remains insignificant.

### A.3.8 Stereotypes Analysis

We present in this section the analyses related to the stereotypes that we elicited in the experiment. We present descriptive statistics that document very little existence of affective polarization in trait ratings. Further, we present analyses in accordance to our pre-registration, which shows that the treatments had little impact on trait ratings.

This finding diverges from prior research (e.g., Iyengar, Sood, and Yphtach [2012](#); Druckman and Levendusky [2019](#); Garrett et al. [2014](#); Levendusky [2018](#)), which shows affective polarization in trait ratings. We posit two reasons for this discrepancy. First, studies vary in how they operationalize affective polarization in trait assessments, and our method is not comparable to all previous studies. For example, Garrett et al. ([2014](#)) examine stereotypes in trait ratings by having participants assign eight traits to out-party supporters, using a dichotomous scoring system. They find a bias in that participants are more likely to assign negative than positive traits to out-party supporters. Similar approaches of only looking at out-party ratings have been used by e.g. Levendusky ([2018](#)). Conversely, our methodology, more similar to Druckman and Levendusky ([2019](#)), asks participants to rate both parties across all traits on a 1-5 scale. The participants evaluate both parties on the same traits on the same screen, and this makes it very salient to the participants that deviations between their responses amounts to having stereotypes. This leaves less scope for implicit biases to have an effect as compared to an approach where only attitudes towards the out-party are measured.

Second, we speculate that the meanings attached to specific traits may have evolved over time, possibly explaining why our results are not identical to e.g. Druckman and

Levendusky (2019), who collected data in 2017. For example, in a time where the credibility of news media is debated, the trait “open-mindedness” could for some participants be ambiguous if they believe that too much open-mindedness amounts to being naive. Similarly, attitudes towards a very high degree of “patriotism” may not have a positive connotation for all participants if it is reminiscent of the isolationist ideas of President Donald Trump to put “America First”. This line of thought is, however, only speculative, and it opens for future research about how trait perceptions have shifted over time, and how this might impact measures of affective polarization.

#### A.3.8.1 Descriptives: Baseline Affective Polarization

Participants rate both parties on a 5-point Likert scale on the following traits: Patriotism, selfishness, intelligence, open-mindedness, and honesty. Across all traits, we find that participants on average rate their own and the opposite party similarly: 3.13 vs. 3.07 for patriotism, 3.25 vs. 3.26 for selfishness, 3.04 vs. 3.01 for intelligence, 2.59 vs. 2.49 for open-mindedness, and 2.49 vs. 2.47 for honesty. Further, there is little difference between Republicans and Democrats. The only significant difference we find is that Republicans rate the Republican Party as significantly more open-minded than the Democratic Party (2.63 vs. 2.47,  $t = -2.4$ ,  $p = .016$ ).

#### A.3.8.2 Treatment Effects

In the following, we present regression outputs related to our three pre-registered hypotheses with net trait rating of each of the traits as the dependent variable. We report only the results from the regressions with the maximum number of controls. The treatment effects remain insignificant in specifications with fewer controls and the conclusion is the same in Mann-Whitney U-tests.

**H1: Affective Polarization in *Invasion* and *Control*.** We first examine whether the net trait ratings differ between *Invasion* and *Control*. We find no statistically significant difference for any of the traits, and all coefficients are small. In particular,  $\beta = -.094$ ,  $p = .513$  for patriotism,  $\beta = -.020$ ,  $p = .887$  for selfishness,  $\beta = -.005$ ,  $p = .963$

for intelligence,  $\beta = .057, p = .696$  for open-mindedness, and  $\beta = -.020, p = .861$  for honesty.

**H2: Affective Polarization in *Disagreement* and *Control*.** Next, we compare net trait ratings between *Disagreement* and *Control*. We generally find no statistically significant treatment effects and small coefficients:  $\beta = -.170, p = .228$  for patriotism,  $\beta = .158, p = .224$  for selfishness,  $\beta = -.180, p = .104$  for intelligence, and  $\beta = -.122, p = .399$  for open-mindedness. The only exception is honesty, we do find a statistically significant effect of  $\beta = -.221$  ( $p = .049$ ). The coefficient estimate is negative and suggests that participants in the *Disagreement* treatment rate a .2 smaller difference in how well their own- and the opposite-party is described by honesty. This should, however, be interpreted in the context that there is no affective polarization in honesty in the first place as participants on average rate the other party as more honest than their own party.

**H3: Affective Polarization in *Invasion* and *Disagreement*.** Finally, we compare net trait ratings between *Invasion* and *Disagreement*. We again find no statistically significant effects and small coefficient estimates:  $\beta = -.071, p = .620$  for patriotism,  $\beta = .184, p = .181$  for selfishness,  $\beta = -.178, p = .116$  for intelligence, and  $\beta = -.206, p = .160$  for open-mindedness. For honesty, however, we do find a marginally significant effect of  $\beta = -.208$  ( $p = .067$ ). Yet, this comes with same caveat as before with there not being affective polarization on this trait in the first place.

## A.4 Discussion

### A.4.1 Heterogeneous Treatment Effects by Partisan Affiliation

An exploratory analysis reveals substantial differences in how Democrats and Republicans respond to the invasion prime: As we detail below, both the *Invasion* and *Disagreement* treatments reduce affective polarization among Republicans, but Democrats display the same level of affective polarization across all three treatments. Notably, this



occurs even though both groups pass the manipulation checks.<sup>6</sup> In the following, we test for treatment effects for the two parties separately, and we then discuss possible reasons for why only Republicans are responsive to making the Russian invasion salient.

Looking first at feeling thermometers, we find a statistically and practically significant 6.2 degrees reduction in the FT difference in *Invasion* among Republicans (OLS:  $p = .008$ ). In contrast, we find no treatment effect among Democrats ( $p = .984$ ). For neither the Republicans nor the Democrats is there any difference between the *Invasion* and *Disagreement* treatments ( $p = .749$  and  $p = .722$ ). Consequently, Republicans exhibit a 5.0 degrees reduction in FT difference in *Disagreement* compared to *Control* ( $p = .023$ , OLS), but there is no effect among Democrats ( $p = .387$ ).

Likewise for the asymmetric Battle of the Sexes game, only Republicans are influenced by the *Invasion* treatment. Compared to Republicans in *Control*, Republicans in *Invasion* are more likely to choose B both when Player 2 is affiliated with the participant's own party (logit: 10.6 percentage points,  $p = .031$ ) and the opposite party (10.4 percentage points,  $p = .027$ ). In contrast, Democrats are not influenced by the *Invasion* treatment (all  $p$ 's  $> .322$ ). Again, there is no difference between *Invasion* and *Disagreement* for any of the parties. As for Player 1 types, the *Invasion* treatment makes Republicans 13.5 percentage points less likely to be of the selfish type *AA* (multinomial logistic regression:  $p = .004$ ) and 7.3 percentage points more likely to be of the cooperative type *BB* ( $p = .081$ ). In contrast, there is no effect among Democrats (all  $p$ 's  $> .291$ ). The effects of *Disagreement* are again not significantly different from those in *Invasion* for either party.

In sum, both treatments effectively reduce affective polarization among Republicans, and they increase Republicans' willingness to cooperate in the coordination game. None of the treatments have any effect on Democrats. First, one might speculate that this result is due to the characteristics of our sample. In our sample, Republican participants are on average older, less educated, and more likely to be White/Caucasian compared

---

<sup>6</sup>Specifically, we find that perceived threat is significantly greater in *Invasion* than *Control* for Democrats (3.73 vs. 3.53,  $p = .024$ ) and marginally so for Republicans (3.31 vs. 3.14,  $p = .099$ ). Moreover, perceived disagreement is significantly greater in *Disagreement* than *Invasion* for both Democrats (3.37 vs. 2.90,  $p < .001$ ) and Republicans (3.21 vs. 2.94,  $p = .010$ ).

to Democratic participants. But these sample differences reflect differences between Republicans and Democrats in the general population (Pew Research Center 2018), and we control for these characteristics throughout the analysis.

Second, one might speculate that our sampling procedure led to differences between Republicans and Democrats. Specifically, because there are more Democrats than Republicans on Prolific,<sup>7</sup> more Democrats were recruited early in the data collection. Thus, almost all Democratic participants were recruited on the first day of data collection, but it took longer to recruit the Republican participants (96 vs. 34.6 percent of responses collected on the first day). If something happened after we started the data collection on May 7, 2022, this would have disproportionately large effects on the Republican sample. Ex ante, we accounted for this by making sure that equal shares were randomized into each treatment on every day, implying that any developments in the conflict would influence participants across all treatments. Ex post, we find no evidence of any systematic change in the estimated treatment effect among Republicans over the course of the data collection.<sup>8</sup>

Third, it is possible that citizens respond differently to the conflict depending on what party the President represents. Studies of previous conflicts have found that the rally-around-the-flag effect (Baker and Oneal 2001; Mueller 1970), in which the President’s popularity increases as a response to a threat, is more pronounced among the opposition. This may occur because people affiliated with the President and his party are “at the flag” already before the threat. For instance, Fox (2009) finds that the popularity of President Bush increased more for Democrats than Republicans in response to both the 9/11 attacks and the Iraq War. Similarly, Callaghan and Virtanen (1993) show that in response to the Iranian hostage crisis in 1979, support for President Carter increased more among Republicans than Democrats. In this way, it is possible that those who are not in power and thus dislike the (party of the) President the most at the outset respond more

---

<sup>7</sup>As of August 16, 2022, 80 percent of the participant pool on Prolific who had stated their support for one of the two parties supported the Democratic Party.

<sup>8</sup>Specifically, we looked at the estimated treatment effect on each day of the data collection and fail to find any pattern in the changes.

strongly to a threat. This would make Republicans more responsive to our treatments.<sup>9</sup>

Fourth, it is possible that the threat of Russia has different effects for Republicans and Democrats if Democrats associate the Republican Party with Russia, e.g. due to Russia’s interference in the 2016 election in favor of Donald Trump. In this election, Russia i.a. waged social media campaigns to favor Donald Trump, fabricated articles and disinformation, and conducted cyberattacks on the Clinton campaign (Mueller 2019). As demonstrated by Darr et al. (2019), this scandal caused Democrats to evaluate the Trump administration less favorably, and it is possible that the controversial relation between the Republican Party and Russia lead Democrats to not think of Republicans as an ally for the present conflict. Indeed, as formulated by Jahani et al. (2020, p. 4): “People will only consider the enemy of their enemy to be a friend if they can see one of the enemies as a potential ally” (cf. social balance theory).

#### A.4.2 Fade-out of Treatment Primes

In this section, we investigate whether the treatment effects depend on the time between participants receive the news prime and answer the feeling thermometer questions. One critique of the priming method is that it only uncovers effects that do not last long. Since our study runs online, it is possible that some respondents take breaks and therefore spend more time than required to complete the study. Although these breaks do not occur at random, we test whether the treatment effects correlate with the time respondents take from receiving their prime to answering the questions related to affective polarization as any correlation will indicate a rapid fade-out of the treatment effect. In conducting this test, we are aware that some participants may be faster than others; so, we use participants’ response times for the demographic questions as a control for speed. We find no significant interaction effect between time spent and the treatment indicator (*Invasion* vs. *Control*:  $\beta = .023, p = .727$ ; *Invasion* vs. *Disagreement*:  $\beta = .019, p = .709$ ), suggesting no immediate fade-out in the treatment effect.

---

<sup>9</sup>A similar effect could occur if “being American” is more closely aligned with the Republican Party than the Democratic Party. In this case, the appeal to one’s national identity that could follow from a geopolitical threat would be more effective for Republicans than Democrats.

### A.4.3 Moderation Effects

We present in the following tests on whether some of our control variables moderates the effects of our treatments. We investigate the role of strength of partisan identity, political interest, age, gender, education level, fade-out of treatment primes, and the extent to which the participants have been following Russia’s invasion of Ukraine.

#### A.4.3.1 Strength of Partisan Identity and Political Interest

Drawing on the social identity approach (Tajfel et al. 1971; Tajfel and Turner 1979), we expect that the partisan identity is more emotionally significant for the individuals who more strongly identifies with either of the parties relative to people who merely lean towards a party. We therefore expect that affective polarization is more pronounced for the former. We test this and possible moderating effects of strength of party support in the following.

Almost half (48 percent) of our sample identifies as “strong supporters” of either of the parties. These participants display both more “ingroup love” and “outgroup hate”: They rate their own party significantly better (78.5 vs. 62.9,  $p < .001$ ) and the opposite party significantly worse (13.8 vs. 27.3,  $p < .001$ ). In investigating whether the strength of partisan affiliation moderates the treatment effects, one could imagine an effect in either direction: On the one hand, the lower opposite-party and higher own-party feeling thermometer rating leaves more room for reducing the FT difference. On the other hand, if individuals’ partisan identity is sufficiently strong, their attitudes might be rather immovable. We test this by including an interaction between our treatment indicator and our strong supporter indicator in an OLS regression. The coefficient estimates suggest that strong supporters reduce their FT difference by 2.65 degrees more in *Invasion* ( $p = .395$ ) relative to *Control*.<sup>10</sup> These effects are statistically insignificant, but we also note that

---

<sup>10</sup>We perform this test of moderation effects for Democrats and Republicans separately as well. The interactions are statistically insignificant (all  $p$ ’s  $> .460$ ), but the signs suggest that following the *Invasion* prime, strong supporters of both the Democratic and Republican Party reduces FT difference relative to partisan “leaners”.

our sample size is chosen to ensure power for the main comparisons and not moderation effects of these.

We next examine the moderating effect of being interested in politics. In our sample, 42.3 percent rate themselves “very” or “extremely” interested in politics. Similar to “strong supporters”, we find suggestive evidence that the politically interested participants reduce their FT difference relatively more following the primes (4.7 degrees more in *Invasion vs. Control*). These interactions are, however, also statistically insignificant (all  $p$ 's  $> .137$ ).<sup>11</sup>

#### A.4.3.2 Age

In the following, we test whether different age cutoffs has a moderating effect on the influence of our treatments. Surveys have demonstrated that age matters for people's attitudes towards the war (The Economist 2022). This is understandable from the intergroup conflict theory, which suggests that individuals are more likely to perceive a threat between two groups if these have a history of conflict, and the older generation has first-hand experience of living during the Cold War.

As expected, we find that threat perception is positively correlated with age (Spearman's  $\rho = .19$ ,  $p < .001$ ). To test for the moderating effect of age, we report two sets of regressions that use indicator variables for either (i) the participant being older than 30 (69.1 percent of the sample) or (ii) the participant being older than 40 (42.0 percent of the sample). We interact this with our treatment indicators to investigate whether age moderates the effect of the treatment primes. We find no evidence that participants' age moderates our treatment effect in either the *Invasion vs. Control* or *Invasion vs. Disagreement* comparison.<sup>12</sup>

---

<sup>11</sup>Looking at Democrats and Republicans separately, we do not find any statistically significant moderation effect.

<sup>12</sup>We tested this for Democrats and Republicans separately as well. We did not observe any significant moderating effect of age.

#### A.4.3.3 Gender

It has long been well-known that there are large gender differences in attitudes towards war and that men tend to be more favorable towards war than women (Covell 1996; Dupuis and Cohn 2011; Lester 1994). As our primary treatment involves priming participants with an armed conflict, we therefore test for a possible moderating role of gender in the following. We stratified our sample to include an equal number of men and women. Men and women rate the average level of threat perception similarly (3.38 vs. 3.46,  $p = .169$ ) and they show similar levels of affective polarization in FT difference (49.8 vs. 49.4,  $p = .811$ ). We test whether gender moderates the effect of the treatments by including an interaction between our treatment indicator and the male indicator in an OLS regression. The coefficient estimates suggest that men reduce their FT difference 1.9 degrees less than women in *Invasion* relative to *Control* ( $p = .549$ ). Further, we find a statistically insignificant difference between men and women in the *Invasion vs. Disagreement* comparison.<sup>13</sup>

#### A.4.3.4 Education Level

In the following, we test whether participants' educational attainment moderates the effects of our treatments. Compared to the American voting population, a large share of our sample has obtained at least a Bachelor's degree (54.7 %). We generate an High-education indicator taking the value one for participants who have obtained at least a Bachelor's degree. We test whether education level moderates the effect of the treatments by including an interaction between our treatment indicator and the Higher-education indicator in an OLS regression. The coefficient estimates suggest that highly educated participants reduce their FT difference 3.4 degrees less than lower educated in *Invasion* relative to *Control* ( $p = .273$ ). In the *Invasion vs. Disagreement* comparison, the coefficient estimate suggests that highly educated reduce their FT difference 7 degrees less

---

<sup>13</sup>We perform this test of gender moderation effects for Democrats and Republicans separately as well. The interactions are statistically insignificant ( $p$ 's  $> .663$ ).

than lower educated ( $p = .027$ ).<sup>14</sup>

#### A.4.3.5 Followers and Non-Followers

We now test whether the extent to which participants have been following the Russian invasion of Ukraine has moderating effect on the influence of our treatments. In designing the experiment, our intention was not to provide participants with any new information, but rather to only provide subtle situation cues by drawing participants' attention towards threat or political disagreement. As we provide basic information from mainstream media, this informational effect is especially plausible for the participants who have not been very attentive to the development during the first month of the invasion. We therefore test for informational effects by comparing treatment effects among followers and non-followers.

Fewer than half of the participants (39.8 percent) rate that they have been following the invasion “to a large extent” or “to the fullest extent”. We find no statistically significant difference in the treatment effects between these participants and the participants who have paid less attention to the conflict in neither the *Invasion vs. Control* or the *Invasion vs. Disagreement* comparison.<sup>15</sup> However, the signs on the coefficient estimates suggest that participants who have been following the invasion a lot respond to the *Invasion* treatment by decreasing their FT difference more relative to participants who have paid less attention to the conflict.<sup>16</sup>

---

<sup>14</sup>We perform this test of education-level moderation effects for Democrats and Republicans separately as well. The interactions in the *Invasion vs. Disagreement* comparison is statistically significant for Democrats and suggest that highly-educated Democrats reduce their FT difference around 7 degrees less.

<sup>15</sup>A similar result is found when we test this separately for Democrats and Republicans.

<sup>16</sup>Note that this effect is not causal as individuals decide by themselves how closely they wish to follow the invasion. For instance, it is likely that people who care more about the development in Ukraine pay more attention to it. Their responsiveness to the prime may thus be a result of the fact that they care more about the development rather than the fact that they have been following the war more closely.

## References

- Al-Fannah, Nasser Mohammed. 2017. “Making Defeating CAPTCHAs Harder for Bots.” In *2017 Computing Conference*, 775–782. London: IEEE.
- Baker, William D., and John R. Oneal. 2001. “Patriotism or Opinion Leadership?: The Nature and Origins of the “Rally ’Round the Flag” Effect.” *Journal of Conflict Resolution* 45 (5): 661–687.
- BBC. 2022. “How Many Ukrainian Refugees Are There and Where Have They Gone?” *July 4*.
- Bebler, Anton. 2015. “Crimea and the Russian-Ukrainian Conflict.” *Romanian Journal of European Affairs* 15 (1): 35–54.
- Bender, Ralf, and Stefan Lange. 2001. “Adjusting for Multiple Testing—When and How?” *Journal of Clinical Epidemiology* 54 (4): 343–349.
- Boxell, Levi, Jacob Conway, James N. Druckman, and Matthew Gentzkow. 2020. *Affective Polarization Did Not Increase During the Coronavirus Pandemic*. Working Paper, Working Paper Series 28036. National Bureau of Economic Research.
- Bursztein, Elie, Steven Bethard, Celine Fabry, John C. Mitchell, and Dan Jurafsky. 2010. “How Good Are Humans at Solving CAPTCHAs? A Large Scale Evaluation.” In *2010 IEEE Symposium on Security and Privacy*, 399–413. Oakland, CA, USA: IEEE.
- Callaghan, Karen J., and Simo Virtanen. 1993. “Revised Models of the “Rally Phenomenon”: The Case of the Carter Presidency.” *The Journal of Politics* 55 (3): 756–764.
- Chmielewski, Michael, and Sarah C. Kucker. 2020. “An MTurk Crisis? Shifts in Data Quality and the Impact on Study Results.” *Social Psychological and Personality Science* 11 (4): 464–473.
- Clarke, Damian, Joseph P. Romano, and Michael Wolf. 2020. “The Romano–Wolf Multiple-Hypothesis Correction in Stata.” *The Stata Journal* 20 (4): 812–843.
- Covell, Katherine. 1996. “National and Gender Differences in Adolescents’ War Attitudes.” *International Journal of Behavioral Development* 19 (4): 871–883.



- Cramer, Angélique O. J., Don van Ravenzwaaij, Dora Matzke, Helen Steingroever, Ruud Wetzels, Raoul P. P. P. Grasman, Lourens J. Waldorp, and Eric-Jan Wagenmakers. 2016. “Hidden Multiplicity in Exploratory Multiway ANOVA: Prevalence and Remedies.” *Psychonomic Bulletin & Review* 23:640–647.
- Darr, Joshua P., Nathan P. Kalmoe, Kathleen Searles, Mingxiao Sui, Raymond J. Pin-gree, Brian K. Watson, Kirill Bryanov, and Martina Santia. 2019. “Collision with Collusion: Partisan Reaction to the Trump-Russia Scandal.” *Perspectives on Politics* 17 (3): 772–787.
- Dennis, Sean A., Brian M. Goodson, and Christopher A. Pearson. 2020. “Online Worker Fraud and Evolving Threats to the Integrity of MTurk Data: A Discussion of Virtual Private Servers and the Limitations of IP-Based Screening Procedures.” *Behavioral Research in Accounting* 32 (1): 119–134.
- Druckman, James N, and Matthew S Levendusky. 2019. “What Do We Measure When We Measure Affective Polarization?” *Public Opinion Quarterly* 83 (1): 114–122.
- Dupuis, Erin C., and Ellen S. Cohn. 2011. “A New Scale to Measure War Attitudes: Construction and Predictors.” *Journal of Psychological Arts and Sciences* 14.
- Farcomeni, Alessio. 2008. “A Review of Modern Multiple Hypothesis Testing, with Particular Attention to the False Discovery Proportion.” *Statistical Methods in Medical Research* 17 (4): 347–388.
- Fox, Gerald T. 2009. “Partisan Divide on War and the Economy: Presidential Approval of G. W. Bush.” *Journal of Conflict Resolution* 53 (6): 905–933.
- Garrett, R. Kelly, Shira Dvir Gvirsman, Benjamin K. Johnson, Yariv Tsfati, Rachel Neo, and Aysenur Dal. 2014. “Implications of Pro- and Counterattitudinal Information Exposure for Affective Polarization: Partisan Media Exposure and Affective Polarization.” *Human Communication Research* 40 (3): 309–332.
- Gehring, Kai. 2021. “Can External Threats Foster a European Union Identity? Evidence from Russia’s Invasion of Ukraine.” *The Economic Journal*, ueab088.
- Hauser, David J., Phoebe C. Ellsworth, and Richard Gonzalez. 2018. “Are Manipulation Checks Necessary?” *Frontiers in Psychology* 9.

- Hauser, David J., and Norbert Schwarz. 2016. "Attentive Turkers: MTurk Participants Perform Better on Online Attention Checks than Do Subject Pool Participants." *Behavior Research Methods* 48:400–407.
- . 2015. "It's a Trap! Instructional Manipulation Checks Prompt Systematic Thinking on "Tricky" Tasks." *SAGE Open* 5 (2): 2158244015584617.
- HistoryExtra. 2022. *Russia-Ukraine Crisis: 9 Milestone Moments in History That Explain Today's Invasion*. <https://www.historyextra.com/period/general-history/russia-invade-ukraine-history-relationship-crimea-why-conflict-facts/>. Accessed January 10, 2024.
- Hoenig, John M, and Dennis M Heisey. 2001. "The Abuse of Power." *The American Statistician* 55 (1): 19–24.
- Iyengar, Shanto, Guarav Sood, and Lelkes Yphtach. 2012. "Affect, Not Ideology: A Social Identity Perspective on Polarization." *The Public Opinion Quarterly* 76 (3): 405–431.
- Jahani, Eaman, Natalie McDaniel Gallagher, Friedolin Merhout, Nicolo Cavalli, Douglas Guilbeault, Yan Leng, and Christopher A. Bail. 2020. *Exposure to Common Enemies Can Increase Political Polarization: Evidence from a Cooperation Experiment with Automated Partisans*. SocArXiv.
- Kennedy, Ryan, Scott Clifford, Tyler Burleigh, Philip D. Waggoner, Ryan Jewell, and Nicholas J. G. Winter. 2020. "The Shape of and Solutions to the MTurk Quality Crisis." *Political Science Research and Methods* 8 (4): 614–629.
- Lester, David. 1994. "Factors Affecting Student Attitudes toward War." *The Journal of Social Psychology* 134 (4): 541–543.
- Levendusky, Matthew S. 2018. "Americans, Not Partisans: Can Priming American National Identity Reduce Affective Polarization?" *The Journal of Politics* 80 (1): 59–70.
- List, John A., Azeem M. Shaikh, and Yang Xu. 2019. "Multiple Hypothesis Testing in Experimental Economics." *Experimental Economics* 22:773–793.
- Mackintosh, Eliza. 2022. "What Does Putin Want in Ukraine? The Conflict Explained." *CNN*, March 1.

- Matherly, Ted. 2019. "A Panel for Lemons? Positivity Bias, Reputation Systems and Data Quality on MTurk." *European Journal of Marketing* 53 (2): 195–223.
- Moss, Aaron, and Leib Litman. 2018a. "After the Bot Scare: Understanding What's Been Happening with Data Collection on MTurk and How to Stop It." *CloudResearch*.
- . 2018b. "Concerns about Bots on Mechanical Turk: Problems and Solutions." *CloudResearch*.
- Mueller, John E. 1970. "Presidential Popularity from Truman to Johnson." *The American Political Science Review* 64 (1): 18–34.
- Mueller, Robert S. 2019. *Report On The Investigation Into Russian Interference In The 2016 Presidential Election*. Technical report. U.S. Department of Justice.
- Oppenheimer, Daniel M., Tom Meyvis, and Nicolas Davidenko. 2009. "Instructional Manipulation Checks: Detecting Satisficing to Increase Statistical Power." *Journal of Experimental Social Psychology* 45 (4): 867–872.
- Pew Research Center. 2018. "Trends in Party Affiliation among Demographic Groups." *Pew Research Center, March 20*.
- Prolific. 2021. *Why Participants Get Banned*. <https://blog.prolific.co/why-participants-get-banned/>. Accessed April 20, 2022.
- Roman, Nataliya, Wayne Wanta, and Iuliia Buniak. 2017. "Information Wars: Eastern Ukraine Military Conflict Coverage in the Russian, Ukrainian and U.S. Newscasts." *International Communication Gazette* 79 (4): 357–378.
- Romano, Joseph P., and Michael Wolf. 2016. "Efficient Computation of Adjusted P-Values for Resampling-Based Stepdown Multiple Testing." *Statistics & Probability Letters* 113:38–40.
- Romano, Joseph P., and Michael Wolf. 2005a. "Exact and Approximate Stepdown Methods for Multiple Hypothesis Testing." *Journal of the American Statistical Association* 100 (469): 94–108.
- Romano, Joseph P., and Michael Wolf. 2005b. "Stepwise Multiple Testing as Formalized Data Snooping." *Econometrica : journal of the Econometric Society* 73 (4): 1237–1282.

- Shyrokykh, Karina. 2018. “The Evolution of the Foreign Policy of Ukraine: External Actors and Domestic Factors.” *Europe-Asia Studies* 70 (5): 832–850.
- Sivakorn, Suphannee, Iasonas Polakis, and Angelos D. Keromytis. 2016. “I Am Robot: (Deep) Learning to Break Semantic Image CAPTCHAs.” In *2016 IEEE European Symposium on Security and Privacy (EuroS&P)*, 388–403. Saarbrücken: IEEE.
- Tajfel, H., and J. C. Turner. 1979. “An Integrative Theory of Intergroup Conflict.” In *The Social Psychology of Intergroup Relations*, edited by W. G. Austin and S. Worchel, 33–37. Monterey, CA, USA: Brooks/Cole.
- Tajfel, Henri, M. G. Billig, R. P. Bundy, and Claude Flament. 1971. “Social Categorization and Intergroup Behaviour.” *European Journal of Social Psychology* 1 (2): 149–178.
- The Economist. 2022. “Americans’ Views on the War Reveal a Striking Generational Divide.” *April 2*.
- Thomas, Kyle A., and Scott Clifford. 2017. “Validity and Mechanical Turk: An Assessment of Exclusion Methods and Interactive Experiments.” *Computers in Human Behavior* 77:184–197.
- UN. 2022. “Ukraine: UN Rights Office Probe Spotlights Harrowing Plight of Civilians.” *UN, May 10*.
- Zhang, Ziyi, Shuofei Zhu, Jaron Mink, Aiping Xiong, Linhai Song, and Gang Wang. 2022. “Beyond Bot Detection: Combating Fraudulent Online Survey Takers.” *Proceedings of the ACM Web Conference 2022*, 1–11.