# COUNTER-PROJECTION IN A POLARIZED POLITICAL CLIMATE: EXAMINING THE ROLE OF THREAT AND WAYS TO REDUCE IT.

by

KATHRYN R. DENNING

A DISSERTATION

Presented to the Department of Psychology
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2022

DISSERTATION APPROVAL PAGE

Student: Kathryn R. Denning

Title: Counter-Projection in a Polarized Political Climate: Examining the Role of Threat and Ways to Reduce It.

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Psychology by:

| | |
|---|---|
| Sara Hodges | Chairperson |
| Sara Weston | Core Member |
| Sanjay Srivastava | Core Member |
| Ashley Angulo | Institutional Representative |

and

| | |
|---|---|
| Krista Chronister | Vice Provost for Graduate Studies |

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded June 2022

DISSERTATION ABSTRACT

Kathryn R. Denning

Doctor of Philosophy

Department of Psychology

June 2022

Title: Counter-Projection in a Polarized Political Climate: Examining the Role of Threat and Ways to Reduce It.

The term "polarization" is often associated with "opposition." Yet, how do people come to see others as their opposites? There is likely a social cognitive process at play. One such process is counter-projection, where people attribute traits, morals, and preferences that are opposite their own onto another person. To better understand when people counter-project, a series of studies were conducted to explore the possible mechanisms, especially threat, behind counter-projection with political out-group targets. These studies also examined if counter-projection can be reduced through analogous perspective-taking – an intervention that asked participants to remember a time when they experienced a similar circumstance to a target individual. The results suggest that counter-projection is triggered by perceptions of threat – either threat associated with an out-group the target belongs to (found in Study 1) or associated with the target themselves (found in Study 2B, which was run after piloting test measures in Study 2A). However, there appears to be an additional mechanism at play in counter-projection – stereotyping. Results related to when stereotyping explains counter-projection indicate that there may be multiple routes to counter-projection. Results regarding the analogous perspective-taking intervention were inconsistent across studies, suggesting the intervention is likely ineffective at reducing polarization.

CURRICULUM VITAE

NAME OF AUTHOR: Kathryn R. Denning

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon
Allegheny College, Meadville, Pennsylvania

DEGREES AWARDED:

Doctor of Philosophy, Psychology, 2022, University of Oregon
Master of Science, Psychology, 2018, University of Oregon
Bachelor of Science, Psychology, 2016, Allegheny College

AREAS OF SPECIAL INTEREST:

Perspective Taking
Political Intergroup Relations
Social Cognition
Data Science
Earned *Specialization in Educational Data Science* from the Educational Leadership Department, University of Oregon, Eugene, Oregon
Quantitative Methods

PROFESSIONAL EXPERIENCE:

Graduate Employee, Department of Psychology, University of Oregon, 2016 – 2022
Research & Data Science Intern, NWEA, Portland, Oregon, June 2021 – August 2021

GRANTS, AWARDS, AND HONORS:

Psychology Department Center for Inclusive Community Travel Grant, University of Oregon, 2020

Departmental Honors in Psychology, Allegheny College, 2016

Herbert Klion Faculty Prize, Allegheny College, 2016

Award for Outstanding Contributions to URSCA, Allegheny College (2016)

Distinguished Alden Scholar, Allegheny College (2015)

PUBLICATIONS:

Denning, K. R., & Hodges, S. D. (2022). When polarization triggers out-group "counter-projection" across the political divide. *Personality and Social Psychology Bulletin, 48*(4), 638-656. doi: 10.1177/01461672211021211

Hodges, S. D., Lieber, S., & Denning, K. R. (2021). Where perspective taking can and cannot take us (pp. 17-44). In O. Kramer & M. Gottschlag (Eds.), Recontextualized Knowledge: Rhetoric - Situation - Science Communication. Berlin: De Gruyter.

Hodges, S. D., Denning, K. R., & Lieber, S. (2018). Perspective taking: Motivation and impediment to shared reality. *Current Opinion in Psychology.* https://doi.org/10.1016/j.copsyc.2018.02.007

ACKNOWLEDGMENTS

Do you ever think about paths not chosen in life – if you had made a different choice, where would you be? Well, I can say with confidence that if I had not chosen this exact path, with these exact people, I would not have completed this dissertation. For that, to all of you, I am eternally grateful.

To my advisor, Sara Hodges, I have cherished our conversations and valued your mentorship. You have challenged me throughout my time in this program, and I would not be the researcher I am today without you. I am also thankful for the support of my other dissertation committee members, Sara Weston, Sanjay Srivastava, and Ashley Angulo. I would like to extend special thanks to Sara Weston, whose statistical mentorship has not only improved this dissertation but helped me grow into a more confident, well-versed statistician in my own right. I cannot thank you enough. To my labmates, I thank you for your feedback and encouragement.

I also extend my deepest gratitude to my family and friends. To my parents, thank you for being my biggest cheerleaders and instilling in me that I could achieve my wildest dreams. To one of my closest friends, Sarah Donaldson, I could not have gotten through the highs and lows of graduate school without you. To my other friends in graduate school – Sara Lieber, Bradley Hughes, and Sarah Horn – thank you for your support and friendship through the years.

Finally, I could not have completed this dissertation without my little family – Alex and Lily (whose puppy snuggles were an absolute requirement). Alex, without your endless support, this dissertation would not have been possible. Thank you for picking me up when I was down, for doing every chore so that I could continue working, for listening to my research more than anyone. When I think of paths taken in life, I feel so lucky this path led me to you. This dissertation is not just mine; it is ours. I look forward to what else we do together, my love.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

# CHAPTER I

## INTRODUCTION

Imagine you are surrounded by hundreds of people. They are waving hand-written posters above their heads and chanting their outrage, the echoes of which bounce off nearby buildings. Some may even be threatening violence and grabbing heavy objects to shatter windows. This description probably triggers a memory of recent political strife, such as a protest close to home or the attack at the U.S. Capitol building on January 6, 2021. Sadly, this political turmoil has become a staple of American life, with political polarization – seeing your political out-group as your direct opposite – increasing with each past U.S. Presidential Election (Deane & Gramlich, 2020; Gramlich, 2016; Pew, 2012). In fact, political polarization has become so commonplace that America is considered one of the most polarized countries in the world (Deane & Gramlich, 2020). This degree of polarization results in negative social consequences, such as increased anger (Mackie et al., 2000), disagreements (Chambers et al., 2006), and perceptions of the out-group as unintelligent and immoral (Pew, 2019). Moreover, political polarization has been increasingly regarded as a threat to American democracy, with the media publishing articles titled "A divided America is a national security threat" (Rice, 2020), "Even Canadians fear US democracy could end soon" (Obeidallah, 2022), and "America's coming age of instability: Why constitutional crises and political violence could soon be the norm" (Levitsky & Way, 2022).

With this degree of political polarization, it stands to reason that polarization's negative impact is not just visible in explicit social outcomes but also permeates those outcomes' underlying cognitive processes. One process of particular interest is projection, where people perceive others as like themselves and, thus, attribute their own traits to others (Robbins &

Krueger, 2005). Projection is typically regarded as a default social cognitive process that people employ with *everyone* (i.e., regardless of their in-group status) because it is a heuristic (e.g., it is cognitively less effortful to see people as like oneself; Krueger, 2007). However, with largely negative feelings toward political out-groups and increasing polarization causing people to see their political out-group as their *opposite*, people may no longer employ projection because they do not want to see their political out-group as *similar* to themselves. Thus, polarization may have created the ideal environment to undermine projection, which in turn may further political division. Yet, if we can identify why projection has changed under polarized circumstances, we may be able to develop methods to reduce political polarization.

**The perfect incubator for *counter*-projection**

When considering how projection may have changed to perpetuate political polarization, one must consider whether people simply stopped projecting or if they started attributing *opposite* traits instead, a process known as "counter-projection" (Denning & Hodges, 2022). One primary theory in support of counter-projection regards it as a cognitive heuristic – people automatically apply opposite characteristics to their out-group (Cadinu & Rothbart, 1996; Davis, 2017) to save themselves the cognitive resources necessary to determine *how* much they differ. Yet, consistent findings of counter-projection are new within the projection literature. Instead, more research has supported the former argument – that people stopped projecting. One theory in line with this argument contends that projection occurs through anchoring-and-adjustment, with people initially anchoring on themselves and projecting less as they adjust their perceptions of the other person further away from the self (but without counter-projecting; Krueger, 2007). This theory was supported by an experiment in which people had the opportunity to project onto members of three different groups as defined by a minimal group paradigm – an in-group, an

out-group, and a mixed group (which had both in- and out-group members; DiDonato et al. 2011). Participants projected the most to their in-group, somewhat to the mixed group, and did not project to the out-group, supporting the anchoring-and-adjustment theory. However, the minimal group paradigm may have made counter-projection unlikely, as the targets were not disparate enough from the self.

Though previous papers rarely found counter-projection (for a detailed review, see the introduction of Denning & Hodges, 2022), as polarization has increased, so have available examples of people's inclinations to counter-project. For instance, researchers have demonstrated (to their surprise) that people counter-project to idiographically disliked targets (Davis, 2017) and to targets whose descriptions have been manipulated to be negatively valenced (Machunsky et al., 2014). In a similar vein, people sometimes elicit *behaviors* opposite their interaction partner in social interactions – based on the personality traits the partner exhibits – resulting in assumed *dissimilarity* (Hughes et al., 2021). Counter-projection has even been found in politically polarized contexts, with participants counter-projecting onto out-groups that disagreed with them on a specific political issue – e.g., support for a new environmental policy (Gromet & Van Boven, 2020) or whether the U.S. should push the Soviet Union to release previously annexed countries (a polarizing issue at the time of the study; Mullen et al., 1992). Furthermore, people counter-project the *birth month* furthest from their own birth month onto out-group political candidates, while attributing their own birth month to in-group candidates (Castelli et al., 2009). The types of information people counter-project onto political out-group members also extend to everyday characteristics. For instance, people projected their morals, personality traits, and everyday preferences (e.g., how people felt about cats versus dogs or the color yellow) onto target individuals who voted differently than them in the 2016 U.S. Presidential Election

(Denning & Hodges, 2022). With the wide variety of information and circumstances under which counter-projection occurs, counter-projection clearly has a pervasive effect on many aspects of social interaction.

However, Denning and Hodges (2022) found that some variance associated with counter-projection can be explained by out-group stereotyping. Specifically, after including participants' idiographic stereotype of their political out-group in the same model with counter-projection, variance unique to counter-projection occurred only (1) on measures asking about people's everyday preferences or (2) when participants were highly identified with their in-group. Thus, a large portion of counter-projection in political examples is explained by use of out-group stereotypes that are opposite stereotypes of one's own group, but there is still a smaller portion of residual counter-projection that still lingers after stereotyping has been accounted for and seems to have a different underlying mechanism (Denning & Hodges, 2022).

Interestingly, both the circumstances under which most previous studies found counter-projection *and* when residual counter-projection occurs appear related to people's experiences of intergroup threat. For instance, those who highly identify with their political in-group likely do so due to perceptions that the out-group poses a threat to them and their in-group's values. Moreover, it seems that perceptions of threat would be a requirement for people to counter-project with personal characteristics that are orthogonal to political traits (e.g., everyday preferences, such as those used in Denning & Hodges, 2022). It is likely easier for people to see others as similar on everyday preferences, meaning they would project if they did not perceive threat (e.g., "that target is threatening, therefore they can't have *my* favorite color!" versus "I have no feelings toward that person, they are probably like me").

Past theories support the possibility that threat may contribute to counter-projection. Specifically, Clement and Krueger (2002) proposed – when they unexpectantly found counter-projection with out-group targets after participants received negative feedback – that people counter-project when they experience *threats to their sense of self*, choosing to project opposite characteristics because they feel threatened by a disliked out-group. This theory is consistent with Machunsky and colleagues' (2014) argument that people counter-project due to a *motivation* to create distance between themselves and a negative target, thus arguing that counter-projection may be an affective process instead of a cognitive one.

Interestingly, neither Clement and Krueger (2002) or Machunsky and colleagues (2014) studied counter-projection in the context of political polarization when they developed their theories. Yet, threat likely plays an even greater role in studies that have found counter-projection with political out-group targets (e.g., Denning & Hodges, 2022), especially as political outcomes have the potential to cause real harm (e.g., lack of military support could increase external threats to safety, lack of support for universal healthcare or social security programs could make it harder for people to survive). As such, polarized political contexts may be the perfect incubator for the threat level that is required to consistently trigger counter-projection (i.e., out-group targets are more threatening, and people feel the need to create more distance from them). Moreover, these contexts are likely to be the ones demonstrating the most negative impacts of people counter-projecting. For example, if we already are polarized from someone or dislike them, putting even more distance between ourselves and them is only going to increase division. This increased division may then lead us to see that person and their group more negatively, which may ultimately explain why increased self-segregation (Brown & Enos, 2021) and political violence (Jilani & Smith, 2018) has been occurring.

Consequently, it would seem fruitful to empirically test current theories that explain counter-projection as the result of threat, as this will improve our understanding of when people psychologically distance themselves from others. Moreover, specifically testing this relationship using an extremely contentious circumstance – like the current state of political intergroup relations in the U.S. – allows for a test of whether counter-projection takes place in a context in which threat *should* trigger it to occur.

**Measures of threat**

Finding an accurate measure of threat to test the role of political threat in counter-projection is a difficult task. For instance, though a significant portion of previous research has demonstrated that political conservatives experience more threat than liberals (see meta-analysis by Jost et al., 2017), many of these studies had significant confounds (for review, see Brandt et al., 2021; Crawford, 2017). Specifically, these studies used issues that were more threatening to conservatives (e.g., terrorism and national security), but the authors assumed that conservatives experienced more threat than liberals on *all* issues (a blind spot that conceivably occurred because the authors themselves *were* liberal; Brandt et al., 2014; Brandt et al., 2021; Crawford, 2017). When methods were altered to use issues threatening to both political sides, conservatives no longer experienced greater threat (Brandt et al., 2014). Further, it has also been demonstrated that liberals experience equal to and sometimes greater threat than conservatives about issues they see as more important (e.g., threats to the economy and poverty, healthcare, climate change; Brandt et al., 2021; Eadeh & Chang, 2020). As such, political threat is clearly pervasive, but with most previous measures of political threat focusing on issues that are more threatening to conservatives, knowledge of liberals' experience with threat is limited. When considering the extent that political polarization is increasing, this lack of research on liberal threat is detrimental

to our understanding of what is really going on and it limits the prior knowledge available to researchers who are trying to develop interventions to reduce polarization.

Further, methodologies for collecting more general measures of intergroup threat (i.e., not just those employed in studies of political relations) vary widely, resulting in inconsistent conclusions about the best approach to measure threat. For instance, results from measures explicitly asking about threat (e.g., "How much of a threat do you feel terrorists pose to the country?") are not typically related to results from indirect measures (e.g., questions measuring the distance people put between themselves and a threatening target, either physically or psychologically; Tally & Bettencourt, 2008). Thus, the underlying concept researchers are trying to measure may vary from study to study, introducing a potential confound. Complicating matters more, some measures fall in-between explicit and indirect (i.e., direct-but-not-explicit threat), measuring self-report statements by asking directly about the potentially threatening target's impact on the participant's life *but without* using language that is explicitly indicative of threat (e.g., using words like "harm" and "worry" instead of "threat"; Stephan et al., 2002, 2016). Studies that correlate these latter measures with both fully explicit and fully indirect measures are at best elusive and likely do not currently exist. The relationship between the differing measures of threat is also complicated by another issue – implicit measures of threat are more likely to measure threat towards individuals (e.g., Cesario & Navarrete, 2014) than both explicit and direct-but-not-explicit measures (e.g., Stephan et al., 2002, 2016). It is possible differences between the measures are explained by the level of threat (i.e., target versus group) explored, not how directly the measure of threat is operationalized (i.e., explicit versus implicit).

One of the most popular measures of threat stems from the direct-but-not-explicit category. The Intergroup Threat Theory by Stephan and colleagues (2002, 2016) purports that

there are two types of threat: symbolic – threats to culture, values, and way of life – and realistic – threats of physical harm or loss of resources (physical or economic). Realistic threats seem to coincide with the dimensions of issues typically tested in political threat studies: fear of physical violence for conservatives, and fear of loss of financial resources or access to healthcare for liberals (Brandt et al., 2021; Eadeh & Chang, 2020). Threats to safety have been found to lead to intolerance of opposing extremist groups for both liberals and conservatives (i.e., right-wing extremists for liberals, left-wing extremists for conservatives; Crawford et al., 2014).

However, there has been little work examining the effect of symbolic threats in the context of political intergroup relations. The most prevalent study that has explored symbolic political threat examined it at both the "macrolevel" – threats to the entire country – and the individual level (Brambilla & Butz, 2013). The authors found that threats of country-level cultural change led to prejudice toward individual targets that also posed symbolic threats to the participant. This same pattern was found for macrolevel realistic threats: threats to the economic security of one's own country led participants to be prejudiced towards targets that threatened their economic security. This study not only highlights people's sensitivity to symbolic political threats at the macro and individual levels, but that threats at one level (i.e., macrolevel) can impact perceptions of threat at another (i.e., individual level). The finding that individuals can pose symbolic and realistic threats to others is critical, as Intergroup Threat Theory is often used to assess group-level threat exclusively. However, people interact more with other *individuals* than *groups* generally, and thus further research examining the degree people identify other individuals as threatening is warranted, especially regarding political relations.

Fortunately, some measures of *indirect* target-related threat already exist outside of the Intergroup Threat Theory. The primary indirect measure of target-related threat measures the

distance people perceive between themselves and the threatening target, such as hypothetical estimates of the physical distance between the participant's location and that of the threatening target (Cesario & Navarrete, 2014), or the amount of space people place between themselves and a threatening target in the lab (Cook & Cottrell, 2021).

However, as socially intelligent beings, we also put distance between ourselves and potential threats psychologically. Studies that employ psychological distancing typically operationalize threat as the amount of distance between the participant's ratings for themselves and those of the targets on a series of traits or beliefs (e.g., Pyszczynski et al., 1995). Psychological distancing can also be measured by people's theoretical acceptance – or lack thereof – of out-group members in different parts of their life (e.g., being willing to work with them; Crawford, 2014). Though psychological distancing shares characteristics with counter-projection, it differs in a few critical ways. First, distancing does not necessitate opposition, as is the case with counter-projection. Furthermore, a common method of operationalizing psychological distancing is to provide the target's responses on the traits and beliefs, and then ask participants to judge how much the self's traits and beliefs differ. Thus, the participant is responding with prior knowledge of the threatening target's responses as reference, and then choosing to distance themselves when responding for their own traits. In this circumstance, participants may state that a trait does not describe them in the study – even when it actually is descriptive of them – to distance themselves from the responses they have been provided for the threatening target. In contrast, projection measures generally ask participants to rate themselves first before responding on behalf of the target individual, at which point participants are deciding if they want to attribute their own traits or the opposite of their traits onto the target. Yet, one important similarity between counter-projection and psychological distancing is that threat likely

plays a role in both. Specifically, if psychological distancing is an indirect measure of threat ("you have this trait, so I can't have it!"), it makes sense that distancing in the opposite direction should also be affected by threat ("I have this trait, so you can't have it!").

Studies have employed psychological distancing to measure political threat, finding that conservatives distance themselves from left-wing groups while liberals distance themselves from right-wing groups (e.g., people didn't want the other group to marry into their family, work with them, or be their friend; Crawford, 2014). Outside of politics, people psychologically distance themselves from targets that remind them of threats to their health (e.g., cancer), but not targets whose less-severe health issues do not trigger feelings of threat (e.g., a sprained ankle; Pyszczynski et al., 1995). However, psychological distancing with targets that reminds participants of threats to their own health can be reduced via certain interventions (Pyszczynski et al., 1993). Specifically, when participants were asked to discuss their own fear of death, as opposed to explicitly sympathizing with a cancer patient, they did not distance as much as those in the sympathy condition. Thus, it appears that prompting people to think of circumstances similar to their own can reduce feelings of threat, whereas more traditional socioemotional responses – like sympathy – do not. This result is promising for potentially reducing counter-projection, as the close relationship between counter-projection and psychological distancing suggests that similar interventions may work for both. As such, testing to see if threat directly triggers counter-projection is a valuable step towards considering possible interventions to reduce counter-projection.

**Interventions to reduce counter-projection**

When looking for an intervention that prompts thoughts of similarity between a target and the self, perspective taking naturally comes to mind. Traditional perspective-taking

instructions ask people to think about another person's thoughts or experiences as if they are that person, or "in their shoes" (Ku et al., 2015; Sassenrath et al., 2016), and have been found to increase similarity and self-other overlap (Laurent & Myers, 2011). However, traditional perspective taking can sometimes backfire, resulting in greater distance or more animosity between the perspective taker and the target of perspective taking or the target's group (for review, see Sassenrath et al., 2016). Interestingly, traditional perspective taking appears most likely to backfire when people experience some amount of threat to the self: e.g., conservatives taking the perspective of a gay couple (Mooijman & Stern, 2016), participants taking the perspective of a competitor (Pierce et al., 2014), or individuals highly identified with their in-group taking the perspective of an out-group member (Tarrant et al., 2012). It seems that taking the perspective of a threatening target – like sympathizing with a threatening target (Pyszczynski et al., 1995) – may simply highlight how much the target's perspective is different or harmful, subsequently prompting individuals to focus on how that different perspective could hurt them.

Yet, a different perspective-taking intervention, called "analogous perspective taking," may attenuate that sense of threat to the self-concept. Instead of putting oneself in another person's shoes, analogous perspective-takers try "to better understand the target by recalling a parallel situation from their own experience" (Gehlbach & Brinkworth, 2012, p. 16). There are two key differences between traditional and analogous perspective taking that may help the latter to better benefit prosocial outcomes. First, analogous perspective-taking instructions focus on one similar context, which does not encourage participants to think broadly about the facets of the target's perspective that might differ from their own (as is the case in traditional perspective taking). Second, analogous perspective-taking instructions are "self-focused," such that they ask participants about a *personal* experience that is related to the target's current experience to

promote understanding, allowing people to choose how they want to connect the self to the target. Interestingly, Pyszczynski and colleagues' (1993) intervention (i.e., prompting participants to think of their own health circumstances and fear of death) that was effective in reducing distance between the participant and target similarly prompted participants to be self-focused in how they connected to the target.

These key features of analogous perspective taking appear promising for prosocial outcomes. For instance, a study by Brockman and Kalla (2016) tested the analogous perspective-taking instructions in a field experiment, finding reduced prejudice towards transgender individuals and increased backing for policy that supported LGBTQ members up to three months after the intervention. However, the contribution of analogous perspective taking was confounded with other experimental elements that could have influenced these results. Specifically, Brockman and Kalla (2016) had canvassers ask participants about their views regarding transgender rights after the canvassers shared their personal experiences – known as *narrative persuasion*. In addition, they trained the canvassers to employ *non-judgmental listening* to the participants' responses (a confound the authors themselves criticized in their follow-up paper, Kalla & Brockman, 2020). Thus, though the conversation included an analogous perspective-taking manipulation – where the canvasser shared a time when they felt judged and asked the participant to share a time when they or someone they care about were "judged or treated differently" (Kalla & Brockman, 2016, pp. 1-2) – this piece of the overall manipulation may not be the only explanation for reduced threats to the self and positive prosocial outcomes.

Kalla and Brockman (2020) attempted to parse apart the influence of narrative persuasion and non-judgmental listening in three follow-up studies, but the exchange of narratives was again

confounded with analogous perspective taking. Specifically, the researchers had the canvasser and participant share personal experiences that were related to the intergroup issue being examined (i.e., immigration, transgender rights). Based on these findings, it is still unclear which part of their intervention really drove the effect and if analogous perspective taking has an independent effect at all. A quite different limitation of Kalla and Brockman's (2020) work is that their intervention would be difficult to scale up. Their attempts at generalizing their method – using videos showing the exchange of narratives while canvassing and exchanging narratives over the phone – requires one-on-one interactions of a canvasser with a participant. Unfortunately, in the current highly polarized political climate, one-on-one interactions may be too slow or resource-intensive to have the wide-spread impact needed to reduce biased intergroup relations.

Saveski and colleagues (2021) employed a scaled analogous perspective-taking instruction when examining engagement across politically opposite Twitter feeds. Subsequent results indicated that participants who received the analogous perspective-taking prompt beforehand – i.e., "if you see something you disagree with, think of a time you and a friend had a disagreement" (p. 4) – were less likely to see the opposing feed as different from their own, in comparison to those who did not receive the prompt. However, they did not find an effect of analogous perspective-taking on their other outcomes, such as interest in talking to the owner of the out-group feed.

Saveski and colleagues' (2021) experiment was a valiant first attempt to focus solely on the effect of analogous perspective-taking instructions, while also using a more scalable methodology. This scalability is critical, as testing whether the intervention can be condensed into a brief, written format will allow it to be employed beyond one-on-one in-person

interventions. However, Saveski and colleagues' (2021) study is the only one currently known to test analogous perspective taking as the sole intervention and it was limited in terms of conclusions that could be drawn about socioemotional outcomes – most of the outcomes in the study focused on engagement with the Twitter feed and the results of the socioemotional outcomes were mixed. As such, it is necessary to further test the replicability of a scalable analogous perspective-taking intervention across multiple experiments, to see if the intervention works on its own (without other methodological confounds) and whether it generalizes to improve a variety of intergroup outcomes, such as reducing counter-projection.

**Present Research**

Now that we have reviewed the literature on analogous perspective taking and its relevance to counter-projection, we can address a pressing research need – testing whether political polarization can be reduced (using counter-projection as a proxy). To address this need, my first research question asks, *(1) Does analogous perspective-taking reduce counter-projection?* Answering this question includes testing the analogous perspective-taking intervention within the political relations context *and* without previous confounds, like narrative persuasion and in-person interactions, in order to eliminate other explanations for previously found prosocial outcomes.

To reduce counter-projection, we also need to consider the underlying mechanisms that trigger it – namely *threat*. Consequently, my following research questions are two pronged: (2) *Is counter-projection driven by threat?* and (3) *Is the effect of analogous perspective-taking on counter-projection mediated by reducing threat?* By first exploring if counter-projection is driven by threat, an important potential underlying mechanism of counter-projection will be directly tested, one that has only been theorized post-hoc in past studies (Clement & Krueger,

2002; Manchusky et al., 2014). If counter-projection is driven by threat, testing whether threat mediates the effect of analogous perspective taking on counter-projection will be helpful to find an effective prosocial intervention to reduce counter-projection. Specifically, analogous perspective taking may work to reduce counter-projection, but if it fails in the same way that traditional perspective-taking interventions have (i.e., sometimes backfiring when threats to the self are high), it may not be the best route to follow for a scalable intervention to reduce intergroup bias. These research questions will be tested within the context of political intergroup relations, since the highly contentious nature of political polarization is likely to create threatening circumstances. If an intervention can be found that reduces threat and counter-projection in political intergroup relations, this intervention may likely also work in less contentious contexts.

The first two research questions were explored in Study 1, while all three research questions were addressed in Study 2B, which follows Study 2A, a pilot study testing specific measures. Study 1 used a pre-post design in which participants responded to questions about two targets, with one target introduced before the analogous perspective-taking prompt (i.e., a pre-instruction target for comparison) and one after participants were prompted to write about a time they felt misunderstood (i.e., the analogous perspective-taking prompt). To assess threats to the self, participants provided baseline self-report ratings of how threatening they regarded the out-group to be at the beginning of the study (i.e., before the intervention). Preliminary results from Study 1 informed the design of Studies 2A and B – particularly an unexpected finding in Study 1 that effects of the analogous perspective-taking instructions were moderated by target characteristics.

To continue to probe the effect of analogous perspective taking on counter-projection and the effect of threat, Study 2 had two parts. Study 2A tested different measures and manipulations to be employed in Study 2B. Study 2A compared different targets and the degree of threat they elicited. This step directly followed results from Study 1 showing that the targets may have unintentionally varied in threat when they were meant to be similar. Study 2A also focused on threat related to a specific *target*, with two distinctions from the group-level threat measured in Study 1: (1) the target was named directly when asking about threat and (2) the items used in the measure were asked after presenting participants with their target (Studies 2A and 2B used a between-subjects design, so participants were only provided one target), rather than providing baseline ratings of how threatening the overall target *group* was before the manipulation. The adapted measure tested indirect, direct, and direct-but-not-explicit measures of target-related threat (based on the reviewed literature above), and the relation among these measures was examined through factor analysis. The adapted measure expanded upon past literature by (1) being related to the target (something many previous measures, especially those stemming from Intergroup Threat Theory, have not done), and (2) on liberal participants who were often not examined in studies exploring the effects of and perceptions of political threat (i.e., researchers often previously studied conservatives, assuming they experienced more political threat; for review, see Brandt et al. 2021; Crawford, 2017). After exploration, a final measure of target-related threat was developed to test the effect of threat on counter-projection in Study 2B and to test whether threat tied to the target moderated counter-projection differently than group-focused threat. In addition, Study 2B examined whether target-related threat mediated the effect of analogous perspective taking on counter-projection.

Finally, to replicate Denning and Hodges' (2022) finding that stereotyping explains some of the variance in counter-projection, additional models were tested in Study 2B that included idiographic stereotyping as a predictor. Not only does this additional analysis allow for possible replication of Denning and Hodges' results, but it also explores whether threat is related to the circumstances in which Denning and Hodges previously found counter-projection variance above and beyond stereotyping. Thus, including stereotyping allows for the possibility to answer one final research question: (4) *When counter-projection persists after accounting for stereotyping, does threat predict the residual counter-projection variance? If so, can the residual counter-projection explained by threat be reduced by analogous perspective-taking?*

# CHAPTER II

## STUDY 1

Using a pre-post design, Study 1 explored the first two research interests: (1) examining whether analogous perspective-taking reduces counter-projection with political out-group targets and (2) exploring if counter-projection is triggered by perceptions of group-level threat. The hypothesis was that people would counter-project to a political out-group target less after the analogous perspective-taking instructions (i.e., post-instructions) than before (i.e., pre-instructions; the variable called "instruction timing"). This first prediction was founded on prior findings that analogous perspective-taking reduced both prejudice (e.g., against transgender individuals; Brockman & Kalla, 2016) and feelings of dissimilarity (e.g., Saveski et al., 2021) in contentious contexts.

It was also expected that there would be a main effect of perceived group-level threat on counter-projection, such that as people's perceptions that their out-group was threatening to their way of life increased, counter-projection would also increase. As previously mentioned, Study 1 operationalized threat at the *group-level* – a baseline measure of threat recorded *before* the instructions. A similar effect of in-group identification on counter-projection was also expected as threat, under the assumption that high in-group identifiers also experience higher levels of perceived threat (and replicating Denning and Hodges's 2022 finding that in-group identification predicts counter-projection). It was not expected that either group-level threat or in-group identification would interact with instruction timing (i.e., the analogous perspective-taking instruction pre/post-order), due to past findings that analogous perspective taking increased people's positive attitudes toward out-group targets equally, regardless of *previous* (i.e., baseline) feelings toward the group (Audette et al., 2020). As such, it was anticipated the

instructions would be effective regardless of previously held biases. Finally, since the two targets in the pre-post design varied in gender and other everyday descriptors (i.e., hobbies, job), exploratory analyses were conducted to examine whether the order in which the two targets were presented (i.e., the variable labeled "target order") moderated the results.

**Method**

Participants were recruited from the University of Oregon Psychology and Linguistics Human Subjects Pool. Sample size was determined by the pragmatic factor of how many participants signed up for the study in one academic term, and this resulted in an initial sample size of 362 participants. However, the final sample size was 300 participants after removing participants who did not complete the analogous perspective-taking narrative (described below; $n$ = 58); who incorrectly completed the manipulation check ($n$ = 4); or who had missing data ($n$ = 1). There were 203 female participants, 88 male participants, and 9 participants who reported something other than female or male. Most participants were white (64.3% white; 13.0% were Asian; 13.0% were Hispanic or Latino/a; 4.0% were Black or African American; 3.0% were Multiracial or other; 1.0% were American Indian/Alaska Native; < 1.0% were Middle Eastern or North African; 1.0% chose the response "prefer not to answer"). On average, participants were 19.4 years old ($SD$ = 2.03). Most participants did not support Trump in the 2020 U.S. Presidential Election (only 8.67% were Trump supporters; the rest, 91.33%, were not). Most of the sample identified as "progressive" or "liberal" (62.33%, with 24.67% identifying as "moderate"; 6.33% identifying as "conservative"; 6.67 % reporting they identified politically with some other party by selecting "other")[1].

[1] Study 1 did not include demographic data on participants' level of education, given that the subject pool that was used is made up of undergraduate students. However, this information was gathered in Studies 2A and 2B, which used mostly non-student samples.

**Procedure.** Participants began by responding for themselves to the Big Five Inventory 2 – Extra Small (Soto & John, 2017) plus four additional honesty items (Thalmayer et al., 2011). Both the Big Five Inventory and the honesty items are described in the "Measures" section below, available in Appendix A, and henceforth referred to as "BFI".

The study was conducted shortly after the U.S. 2020 Presidential Election (data were collected between January and March 2021). In-groups and out-groups were defined by who participants voted for or supported in the election. Participants were asked to report whether they voted for or supported Trump in the 2020 U.S. Presidential Election (if they did not vote but still preferred Trump to Biden or any other candidate, they were instructed to answer "Yes" to the question about supporting Trump). Participants also reported how highly they were identified with others who had the same voting preferences as their own in the 2020 U.S. Presidential Election (1 = *Not at all identified*; 5 = *Extremely identified*; henceforth, this variable is called "in-group identification"), and to what extent those who had different voting preferences from their own threatened their beliefs and way of life (1 = *Not at all*; 5 = *Extremely*; this variable is called "group-level threat").

Based on their voting preference, participants were next shown descriptions of two out-group targets (i.e., target people who reported political preferences that were the opposite of the participants' own preference). One target was presented before the analogous perspective-taking instructions and the other after. One of the targets was male and the other female, and the order in which the targets were presented was counterbalanced (i.e., "target order"). Both targets were white, appeared to be approximately in their mid-30's (images were selected from Minear & Park's 2004 database and are available in Appendix B), and were presented along with a description that included their job and one hobby. The female target was described as follows:

"Above is an image of a woman named Jen. Jen enjoys running, works as a business manager, and voted for [did NOT vote for] Trump in the 2020 U.S. Presidential Election." The male target was described as follows: "Above is an image of a man named Mike. Mike enjoys cooking, works as a high school teacher, and voted for [did NOT vote for] Trump in the 2020 U.S. Presidential Election."

After being introduced to the first target but before receiving the analogous perspective-taking instructions, participants were asked to provide ratings for the first target on the same BFI measure they had answered for themselves earlier. Following the BFI, participants were given three filler questions about their COVID-19 behaviors.[2] Next, participants were provided with the analogous perspective-taking prompt: "Please write about a time when you did not feel understood by another person because you held a different opinion." Following the prompt, participants were introduced to the second target and then asked to respond to the BFI items for that target. Finally, participants answered demographic questions.

**Measures.**

***Trait Measure.*** The Big Five Inventory 2 - Extra Small (BFI; Soto & John, 2017) includes 15 items, three each for extraversion, agreeableness, conscientiousness, negative emotionality, and open-mindedness. Also included are four items measuring a sixth personality trait of honesty from the 24-Question Big 6 scale (Thalmayer et al., 2011), which includes items such as, "Is not good at deceiving other people." All responses were on a 5-point Likert scale (1 = *Disagree Strongly* to 5 = *Agree Strongly*). All items are available in Appendix A.

[2] Specific COVID-19 questions and exploratory analyses are available on Github (Appendix C).

## Results

The analysis plan was pre-registered on OSF, with data and reproducible analyses available on Github (links in Appendix C)[3]. The first analyses tackled the pre-post analogous perspective-taking effect on counter-projection. The base model nested responses for the self (i.e., the participant) on the BFI at Level 1 (the item level) within participant (Level 2) and predicted responses for the targets on the BFI[4]. To differentiate between the targets, an additional predictor was included designating whether the response on each item was for the pre-instruction target or post-instruction target (i.e., "instruction timing"). Within this model, a significant interaction between instruction timing and responses on the BFI for the self would indicate there was an effect of the analogous perspective-taking instruction on projection/counter-projection.

When distinguishing between projection versus counter-projection, slopes are used as a diagnostic measure. Specifically, if a negative slope between the responses for the self and the target(s) occurs, this indicates counter-projection, with participants seeing the target(s) as having opposite responses to them on the BFI. If a positive slope is found, this indicates projection (i.e., regular positive projection) and that the participant sees a direct correspondence between their responses on the BFI and that of the target(s). Finally, if there is a flat slope (i.e., not significantly different from zero), this indicates that the participant is neither projecting nor counter-projecting to the target(s). When discussing self ratings predicting target ratings (as main effects or interactions), these effects will generally be referred to as "projection," as the directionality is not always clear when multiple conditions or variables are being examined at once (e.g., a negative slope may indicate projection is decreasing, but only simple slopes can

[3] Amendments to the original plan are noted both on OSF and throughout this paper.

[4] Initially, it was pre-registered that level 2 of the base model should be target order, not subject level. Thus, this model is a deviation from the original pre-registration.

show definitively if counter-projection occurred). As projection is the more common social cognitive process, it will be defaulted to. Once directionality and difference of slopes from zero have been determined (i.e., through simple slopes), the terms "projection" and "counter-projection" will be used appropriately to indicate directionality.

In addition to the base model, models also tested the predicted moderators of counter-projection. Specifically, these models tested whether there were main effects of in-group identification and group-level threat on counter-projection and if either moderator interacted with instruction timing (the latter was not expected). Both predictor variables were added at the participant level (Level 2). Exploratory analyses also were conducted to examine (1) whether there was an interaction effect of in-group identification and group-level threat, and (2) if target order (included at Level 2) altered the effect of the analogous perspective-taking instructions (neither effect was expected). In all models, instruction timing was dummy-coded, with the pre-instruction target as the reference condition. All analyses were conducted in R (version 4.0.3; R Core Team, 2020) and the multilevel models were calculated using "lme4" (Bates et al., 2015). Other packages employed in addition to base R included "psych" (Revelle, 2021), "rio" (Chan et al., 2021), "sjPlot" (Lüdecke, 2021), "Tidyverse" (Wickham et al., 2019), "emmeans" (Lenth, 2021), "parameters" (Lüdecke et al., 2020), "nlme" (Pinheiro et al., 2022), and "effects" (Fox, 2019).

**Effect of analogous perspective taking.** First, a base model was conducted testing only instruction timing on counter-projection. This model revealed a significant main effect of instruction timing on projection, $b = 0.08$, $p < 0.001$, 95% CI [0.06, 0.10]. However, since the "main effect" is technically an interaction between participants' responses on the BFI for the self and instruction timing (this also applies to all other main effects in Study 1, and those examining

projection in Study 2B), it must be decomposed by calculating simple slopes. Thus, simple slopes were calculated for the interaction effect of instruction timing with participants' responses on the BFI for the self on participants' responses on the BFI for the target (a corresponding visualization is available in Figure 1).

The simple slopes analysis indicated that participants counter-projected to the pre-instruction target, as illustrated by the negative slope between responses on the BFI for the self and responses for the target, $b$ = -0.07, $p < 0.001$, 95% CI [-0.10, -0.04]. However, the simple slope did not differ from zero for the post-instruction target, $b$ = 0.01, $p = 0.585$, 95% CI [-0.02, 0.04], indicating participants were neither counter-projecting nor projecting. Pairwise comparisons demonstrated that the simple slopes significantly differed from one another, $b$ = -0.08, $p < 0.001$. Subsequently, it appears that the analogous perspective-taking instructions *reduce counter-projection* (i.e., people counter-project less post-instruction relative to what they did with the pre-instruction target) but do not result in *positive* projection.

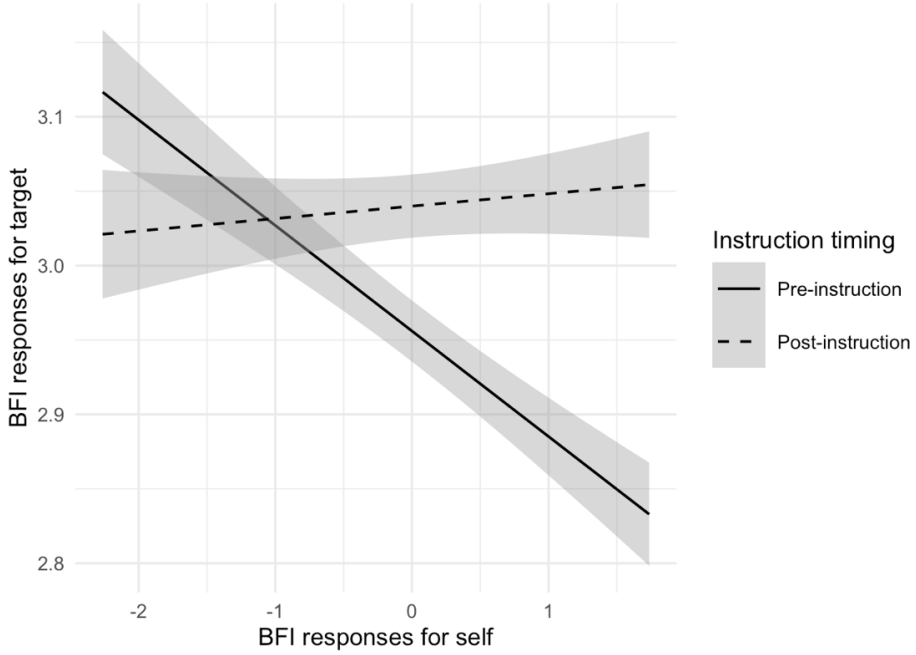



*Figure 1.* Main effect of instruction timing on projection in Study 1.

**Confirmatory moderators.** As predicted, adding in-group identification as a Level 2 moderator did not result in an interaction of instruction timing and in-group identification on projection, $b = 0.002$, $p = 0.832$, 95% CI [-0.02, 0.02]. However, significant main effects were found for both instruction timing, $b = 0.08$, $p < 0.001$, 95% CI [0.06, 0.10], and in-group identification, $b = -0.06$, $p = 0.002$, 95% CI [-0.09, -0.02], on projection. Decomposing the relationship of in-group identification and counter-projection, simple slopes were calculated using three groups: average in-group identifiers, high in-group identifiers (one standard deviation above average), and low in-group identifiers (one standard deviation below average). As seen in Figure 2, the simple slopes analysis indicated that high in-group identifiers counter-projected to a political out-group target, $b = -0.08$, $p < 0.001$, 95% CI [-0.13, 0.04]. However, the simple slopes for average in-group identifiers, $b = -0.03$, $p = 0.063$, 95% CI [-0.06, 0.002], and low in-group identifiers, $b = 0.02$, $p = 0.324$, 95% CI [-0.02, 0.07], did not differ from zero, indicating that these participants neither projected nor counter-projected. Pairwise comparisons indicated all simple slopes differed from one another, $bs = 0.05 – 0.11$, $ps = 0.003$. Thus, high in-group identifiers counter-projected more than average in-group identifiers. Moreover, though the simple slopes for average and low in-group identifiers did not differ from zero, they did differ from one another.

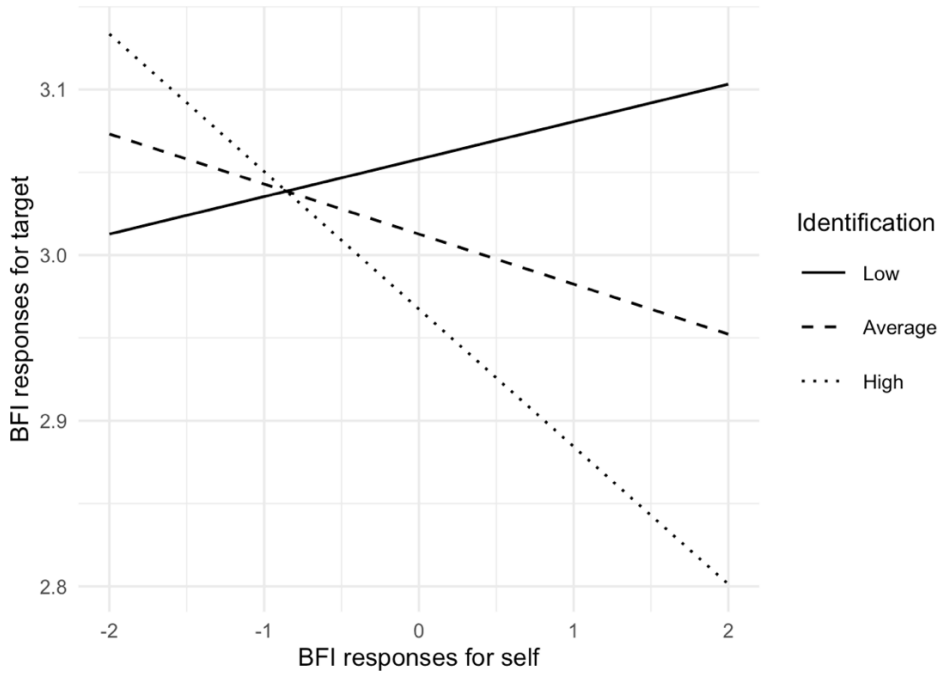



*Figure 2.* Main effect of in-group identification on projection in Study 1.

As predicted, when adding group-level threat as a Level 2 moderator, we again did not see a significant interaction, $b = 0.001$, $p = 0.862$, 95% CI [-0.01, 0.02]. However, there were significant main effects for both instruction timing, $b = 0.08$, $p < 0.001$, 95% CI [0.06, 0.10], and group-level threat on projection, $b = -0.07$, $p < 0.001$, 95% CI [-0.10, -0.05]. Once more, simple slopes were calculated using three groups: determined by average perceivers of threat, high perceivers of threat (one standard deviation above), and low perceivers of threat (one standard deviation below). As seen in Figure 3, high perceivers of threat counter-projected, $b = -0.12$, $p < 0.001$, 95% CI [-0.17, -0.08], while average perceivers of threat neither counter-projected nor projected, $b = -0.03$, $p = 0.058$, 95% CI [-0.06, 0.001]. Low perceivers of threat positively projected, $b = 0.06$, $p = 0.003$, 95% CI [0.02, 0.11]. Pairwise comparisons indicated that all slopes differed from one another, $bs = 0.09 – 0.19$, $ps < 0.001$. Thus, average perceivers of threat counter-projected less than high perceivers of threat.

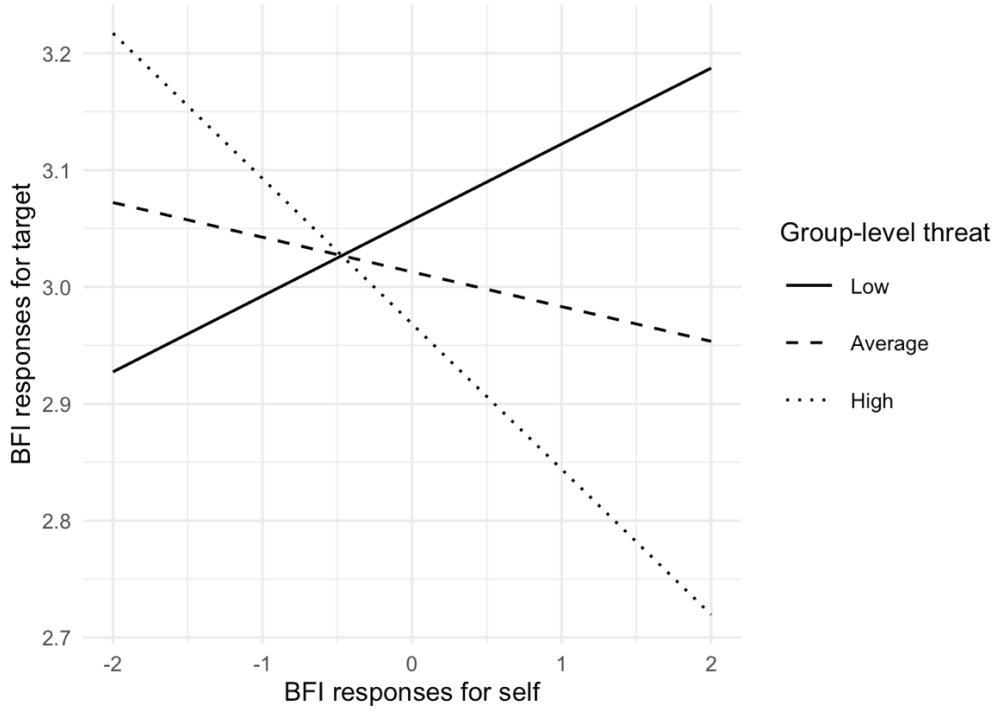



*Figure 3.* Main effect of group-level threat on projection in Study 1.

**Exploratory moderators.** When exploring interaction effects, neither in-group identification nor group-level threat interacted with instruction timing or one another, *ps* > 0.443. Again, the main effect of instruction timing on projection remained significant, b = 0.08, *p* < 0.001, 95% CI [0.06, 0.10]. Interestingly, although the main effect of group-level threat remained significant, *b* = -0.07, *p* < 0.001, 95% CI [-0.09, -0.04], the main effect of in-group identification was no longer significant, b = -0.03, *p* = 0.136, 95% CI [-0.06, 0.01]. Simple slope analyses were conducted decomposing the effect of group-level threat after accounting for in-group identification. Both high perceivers of threat, *b* = -0.12, *p* < 0.001, 95% CI [-0.16, -0.07], and average perceivers of threat, *b* = -0.03, *p* = 0.04, 95% CI [-0.06, -0.001], counter-projected to an out-group target, while low perceivers of threat positively projected, *b* = 0.05, *p* = 0.028, 95% CI [0.01, 0.10]. Pairwise comparisons indicated that the simple slopes all significantly differed from one another, *bs* = 0.08 – 0.17, *ps* < 0.001. Of particular interest, high perceivers of threat

counter-projected more than those who perceived average threat, $b = 0.085$, $p < 0.001$. Interestingly, after accounting for variance attributable to in-group identification, the slope for those who perceived average threat now significantly differed from zero, indicating the presence of counter-projection (as a reminder, when in-group identification was not accounted for in the model, the slope did not differ from zero). Overall, it appears that how threatening the out-group is to a person's way of life explains unique variance in counter-projection that is not explained by other variables.

Finally, a model was run that included target order as a moderator (i.e., whether participants saw Jen – the female target – first, prior to the analogous perspective-taking instructions, and then Mike – the male target – after the instructions; or whether they saw Mike before the instructions, and Jen after the instructions). The main effect of instruction timing remained significant, $b = 0.14$, $p < 0.001$, 95% CI [0.11, 0.17], while the main effect of target order was not significant, $b = 0.02$, $p = 0.627$, 95% CI [-0.05, 0.08]. However, there was an unexpected significant interaction between target order and instruction timing, $b = -0.13$, $p < 0.001$, 95% CI [-0.17, -0.08]. The interaction was decomposed with simple slopes, and a corresponding illustration is available in Figure 4.

Simple slopes indicated that participants counter-projected to the pre-instruction target regardless of whether it was the female target, $b = -0.08$, $p = 0.002$, 95% CI [-0.13, -0.03], or the male target, $b = -0.06$, $p = 0.014$, 95% CI [-0.11, -0.01]. However, for the post-instruction target, participants positively projected to the male target, $b = 0.06$, $p = 0.010$, 95% CI [0.01, 0.11], while they neither projected nor counter-projected to the female target, $b = -0.05$, $p = 0.063$, 95% CI [-0.09, 0.002]. Pairwise comparisons indicated that the simple slopes for the male and female target did not significantly differ pre-instruction, $b = -0.02$, $p = 0.962$, indicating participants

counter-projected similarly with both targets prior to the instructions. However, there was a significant difference between the slopes for the male and female targets post-instruction, $b$ = 0.11, $p$ = 0.009, demonstrating that people projected significantly more to the male target post-instruction than they did to the female target. Pairwise comparisons also show that there was a significant difference between the slopes for the female target pre-instruction and the male target post-instruction, $b$ = -0.14, $p < 0.001$, but no difference when the targets were seen in the reverse order (pre-instruction male target, then post-instruction female target), $b$ = -0.01, $p$ = 0.79.

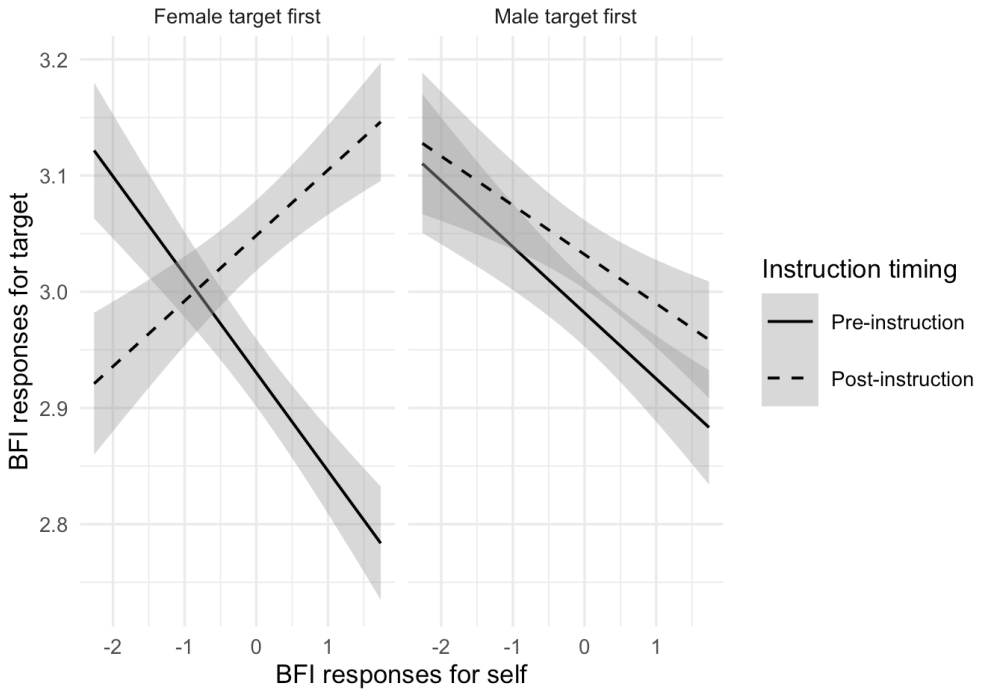



*Figure 4.* Exploratory interaction of target order and instruction timing on projection in Study 1.

Relevant demographics were explored as possible variables that could explain these results. Surprisingly, participant gender did not moderate this interaction, $b$ = 0.06, $p$ = 0.195, 95% CI [-0.03, 0.16][5]. To explore if this effect persisted when Trump supporters were removed

[5] No other main effects were significantly moderated by participant gender either. The variable of participant gender was filtered to include only participants who identified as male or female and did not include the 9 participants who identified as non-binary (n = 7), preferred to self-describe (n = 1), or did not want to answer (n = 1).

(i.e., using only liberal or liberal-leaning participants due to the liberal skew in the sample), the analysis was conducted again using only non-Trump supporting participants. However, the interaction of target order with instruction timing on counter-projection remained significant, $b$ = -0.14, $p < 0.001$, 95% CI [-0.18, -0.09]. Unfortunately, it was not possible to look at political preference as a moderator because of how few Trump supporters participated ($n$ = 26).

**Discussion**

The results of the pre-post design support the prediction that analogous perspective-taking reduces counter-projection. Moreover, the significant effect of analogous perspective-taking instructions was unaffected by the two moderators that were tested (in-group identification and group-level threat), even though both variables predicted counter-projection individually. Of note, the results of the two moderators replicate and build upon prior findings by Denning and Hodges (2022). Specifically, Denning and Hodges' (2022) found that in-group identification predicts counter-projection, which was replicated in the current study. The current results also expanded upon their findings by demonstrating that perceived group-level threat explained when in-group identification predicted counter-projection. Regardless, these results provide initially promising support for the analogous perspective-taking instructions, as they appear to reduce counter-projection regardless of baseline predispositions. For example, even if someone is highly identified with their own in-group, asking them to think about a time they felt misunderstood appears to help them see members of their political out-group as less disparate from themselves.

However, Study 1 produced an unexpected result that target order moderated the effects of the analogous perspective-taking instructions, demonstrating that the instructions' ability to reduce counter-projection appears to differ depending on target characteristics. Though it is

possible that participant demographics could explain this result, tests for likely demographic moderators – participant gender and voting preference – revealed no moderating effects.

Another possible explanation for these unexpected results could stem from the descriptions of the targets. Though the targets' jobs (i.e., business manager for the female target and teacher for the male target) and hobbies (i.e., running for the female target and cooking for the male target) in these descriptions were selected because they were relatively common, the descriptions and gender were not counterbalanced. As such, the selected descriptions may have systematically varied in terms of stereotype content, valence, or even the feelings of threat they provoked. The Stereotype Content Model (Fiske et al., 2002) may help explain these unexpected results, as it predicts that out-group target individuals who seem more competent – which perhaps the successful but not particularly "warm" female target might have seemed – are seen as threatening, while targets from out-groups associated with traditional female gender roles are often pitied and seen as non-threatening – like the male target who liked to cook and was a teacher. Thus, it is possible the analogous perspective-taking instructions were only effective with the male target because that target was less threatening, and so it follows that the threat-reducing power of the instructions may be more effective in reducing counter-projection only when there is less threat to reduce overall.

Threat as an explanation of Study 1's results seems increasingly likely due to the additional finding that group-level threat predicted counter-projection and did so more than in-group identification. Group-level threat was collected *before* participants saw the targets, making it a baseline measure of how threatening participants regarded the out-group to be in general. As such, this measure would not theoretically be expected to matter when testing whether one *target* was more threatening than another. The fact that group-level threat did not interact with target

order does not exclude threat, in general, as an underlying explanation for why there are differences between the two targets. Instead, these results demonstrate that the effect of the analogous perspective-taking instructions does not depend on baseline self-report *group* ratings.

The degree the targets triggered threat may have been amplified by the pre-post study design that directly *compared* the targets to one another. The resulting effect may have stemmed not only from the male target being less threatening than the female target, but because his presentation *followed* the more threatening female target, making him appear less threatening in comparison. As such, the successful "intervention" to reduce counter-projection may have been the product of both the analogous perspective-taking instructions *and* the presentation of a contrasting target.

Overall, the fact that counter-projection for political out-group targets varies due to target characteristics is both concerning and promising. On one hand, it is concerning that people may counter-project more to members of out-groups that present a greater threat to the self-concept. However, the suggestion that lower threat to the self-concept can be altered with some targets (i.e., the analogous perspective-taking instructions led to *positive* projection, at least with the male out-group target) is promising, as this may guide the development of new methods to ultimately reduce counter-projection and other harmful intergroup outcomes. Thus, it is necessary to further probe the threatening circumstances when counter-projection occurs, and under which circumstances analogous perspective-taking reduces counter-projection.

# CHAPTER III

## STUDY 2A

Due to Study 1's promising but inconclusive support for the idea that analogous perspective-taking instructions can reduce counter-projection (i.e., the instructions were effective, but for only one target individual), it was necessary to further probe the effect of analogous perspective taking on counter-projection in political contexts. One major area of focus was on the relationship of threat and counter-projection, as the results in Study 1 may have stemmed from one of the targets being described in a way that was seen as more threatening than the other, making the less threatening target more sensitive to the analogous perspective-taking instructions. However, it was unclear whether analogous perspective taking reduced *target*-related threat for either target since participants reported their feelings of threat *before* both seeing the target and the introduction of the analogous perspective-taking instructions.

To address the limitations posed by Study 1, two additional studies were run. Study 2A served as a pilot test for the measures employed in Study 2B, which further explored the relationship of threat and counter-projection and the question of whether counter-projection can be reduced via analogous perspective taking. As such, the first goal of Study 2A was to develop a measure of *target*-related threat that was measured *in response* to the presentation of a specific target (i.e., *after* the target was presented) and that was tied to the target (i.e., mentioned the target by name). This contrasts with the group-level measures used in Study 1 that measured threat towards the out-group in general. The second goal of Study 2A was to develop targets that differed in the degree of threat they prompted in participants, to provide more precise variation in perceived threat in Study 2B.

The aim to develop a measure of target-related threat was complicated by the fact that most threat scales measure threat at the group-level (e.g., they ask about threat posed by the out-group, not a specific individual; e.g., Stephan et al., 2002, 2016). Further, previous measures of threat that are target-centric frequently measure threat implicitly through physical distancing (Cesario & Navarrete, 2014; Cook & Cottrell, 2021), which is a very different methodology in comparison to group-level measures. Moreover, implicit threat measures (including these distancing measures) have been found to be uncorrelated with explicit threat measures – i.e., those that use the term "threat" when asking about people's perceptions (Tally & Bettencourt, 2008). This difference seems to stem from the measures prompting different degrees of cognitive awareness in the participant. However, this explanation is complicated by the fact that the implicit and explicit threat measures appear to systematically vary regarding the focus of the threat measure (target versus group, respectively). It seems likely that threat posed by an individual is on a different scale than that posed by a group, as a group has collective power to cause societal change, but an individual may personally affect someone's daily life (or in the case of a hypothetical target, represent an individual that participants know who poses a threat to their daily life). Thus, the different scales of threat may simply be unrelated to one another because they focused on different things. Matters are further complicated by the fact that some threat measures are neither clearly explicit nor implicit, and instead ask directly about threatening scenarios without using the word "threat" (i.e., direct-but-not-explicit as discussed in the introduction to this dissertation; an example includes measures introduced by Stephan et al.'s 2002 and 2016 work). The relationship of this latter category of direct-but-not-explicit measures with explicit and implicit measures has not been tested.

As such, Study 2A adapted previous measures that are explicit, implicit, and direct-but-not-explicit to develop a composite measure of threat that was target-related (heretofore referred to as "target-related threat"). Two of the items measured threat indirectly by examining physical distance (i.e., how close participants report being willing to sit to the target and how close they estimate that they live to the target). The items were inspired by the social distancing measures previously used by Cook and Cottrell (2021) and previous findings (and theory) that people physically avoid those who pose a threat to them (e.g., Wyner & Calvini, 2011). Two direct-but-not-explicit measures were also included, corresponding to items on the highly cited Intergroup Threat Theory scale by Stephan and colleagues (2002, 2016) measuring *realistic* threat (i.e., threats to physical safety, economic resources) and *symbolic* threat (i.e., threats to culture, way of life). By including both implicit and direct-but-not-explicit target-related threat measures, Study 2A measured threat more indirectly than in Study 1, which may guard against a socially desirable response bias that sometimes occurs with explicit threat measures (Tally & Bettencourt, 2008). In addition, participants were asked to explicitly rate how threatening they found both the target and the target's group (the term "explicit threat" is used in Studies 2A and 2B to indicate explicit threat and to distinguish between it and the other measures of threat).

In addition, Study 2A investigated whether the different threat items (physical distancing, symbolic and realistic threat, explicit threat that was target-related and group-related) were intercorrelated. However, analyses comparing the explicit threat items to the other threat items were considered exploratory, as it was expected the explicit threat items would not be related to the less direct items (Tally & Bettencourt, 2008). For that reason, the explicit threat measures were also not included in the development of the composite measure of target-related threat. As

Study 2A was primarily exploratory, there were no *a priori* predictions about which implicit and direct-but-not-explicit items would be combined into target-related threat.

Study 2A also tested different target descriptions for the degree of threat they elicited. Five different target descriptions were tested, with an aim to find sufficiently variable levels of threat in order to test for target-related threat effects. The goal was to identify at least three target descriptions that significantly differed from each other for use in Study 2B and described in Chapter 4. Moreover, developing differentially threatening targets expanded upon Study 1's result that the effectiveness of the analogous perspective-taking instructions differed depending on the target (the differences found in Study 1 may have been due to the different targets eliciting variable degrees of threat). The target descriptions used in Study 2A (found in Appendix D) were developed based on theory (Fiske et al., 2002; Stephan et al., 2016) and current polarizing events (e.g., the 2020 election and COVID-19 vaccination). The goal of pilot testing in Study 2A was to produce three targets that varied in terms of target-related threat: the "most threatening target," the "moderately threatening target," and the "least threatening target." Apart from deriving a composite measure of target-related threat, all threat items were compared across the different targets[6].

**Method**

Data were collected using Prolific Academic. A power analysis was conducted in R (package "pwr" by Champely et al., 2020) with the following criteria: a five-level categorical variable representing the target variable, a medium effect size ($f = 0.25$), an alpha level of 0.05, and a beta level of 0.80. Results indicated that an $n$ of 39 participants per condition was

[6] The measures not found to load into our one factor were analyzed in additional separate exploratory analyses (i.e., additional ANOVAs were conducted).

necessary to reach 80% power. Anticipating missing data, 50 participants were collected per condition (total $n = 250$). The link for the R code is available in Appendix C.

An initial sample of 251 participants was collected. However, the final sample resulted in 209 participants after removing participants who did not meet the pre-screen criteria ($n = 3$); who incorrectly answered the manipulation checks about the target ($n = 37$; described in the "Procedure"); or who may have participated multiple times ($n = 2$; identified using IP addresses, which is considered a possible indictor of bots on online crowdsourcing websites; Kennedy et al., 2020). Only one condition had less than the needed sample size of 40 after removing participants (cell sizes are provided with the target descriptions in the "Procedure"). The sample included 139 participants who identified as women, 64 who identified as men, and 6 participants who identified as something other than a woman or man. Participants were, on average, 35.10 years old ($SD = 12.84$). In terms of racial breakdown[7], 72.24% of participants were White or European American; 11.96% were Asian/Asian American; 7.18% were Latina/o, Hispanic, Chicano/a, or Puerto Rican; 5.26% were Black/African American; 2.87% were Multiracial; and less than 1.0% were American Indian/Alaska Native. In comparison to the U.S. demographics (United States Census Bureau, 2021), this sample was whiter (60.1% in the U.S. identify as white without identifying as Hispanic or Latino/a) and more Asian (5.9% in the U.S.), while less Black/African American (13.4% in the U.S.) and less Hispanic/Latino/a (18.5% in the U.S.).

Education was relatively normally distributed, with many people reporting their highest level of education as a bachelor's degree (36.36% bachelor's degree; less than 1.0% earned less than a high school diploma or its equivalent; 9.09% had completed high school; 20.57% had

[7] Study 1 used a collaborative data collection process within the University of Oregon Psychology department with pre-determined demographic questions. As such, the categories for race/ethnicity differ from those used in Studies 2A and 2B, where the categories were altered to reflect a wider variety of backgrounds.

completed some college; 9.09% had an associate's degree; 2.84% had completed some graduate school; 17.70% had a master's degree; 3.83% had a doctorate or professional degree). Our sample was disproportionately educated compared to the total U.S. demographics (United States Census Bureau, 2021), where 88.5% of people have completed high school or higher (in total, over 99% of the current sample had completed high school or higher) and only 32.9% have completed their bachelor's degree or higher in the U.S. (in the current sample, 60.73% had completed their bachelor's degree or higher). Participants' reports of the parents' education had a similar-shaped distribution to their own education, although for parents' education, slightly more people reported either that their parents did not earn a high school degree or that they had graduate-level degrees (4.78% earned less than a high school diploma or its equivalent; 13.40% had completed high school; 14.83% had completed some college; 9.09% had an associate's degree; 24.88% had a bachelor's degree; 1.91% had completed some graduate school; 22.01% had a master's degree; 9.09% had a doctorate or professional degree). Relevant to the political element of the study, most participants were born (93.78%) and raised in the U.S. (97.61%). Those primarily raised in other countries still reported having lived in the U.S. for between 12 and 35 years before participating in this study and were all U.S. citizens who were registered to vote. On a measure of political orientation (described in the "Measures" section), most participants identified as "extremely liberal" ($n$ = 75) or "liberal" ($n$ = 110). Some participants identified as neither liberal nor conservative ($n$ = 21) and very few identified as "conservative" ($n$ = 3)[8].

**Procedure.** Participants were first screened for voting preference by self-identifying that they voted for Joe Biden in the 2020 U.S. Presidential Election. Participants were screened for

[8] Participants were not required to identify as liberal, so long as they voted for Biden.

voting preference in order to gather a sample like that used in Study 1, which was a predominantly liberal sample who largely voted for Biden in the 2020 U.S. Presidential Election. After collecting 20 participants, the remaining participants were screened for location (i.e., living in the U.S.) out of concern that international locations may interfere with the validity of the implicit threat measure, operationalized as reports from participants about their distance from a certain location (described in the "Measures" section)[9].

Participants were shown the target's picture and description, and then asked to answer a series of three manipulation checks about the target (i.e., who the target preferred in the 2020 U.S. Presidential Election[10], what the target's opinion on COVID-19 was, and what the target's profession was). Depending on their condition, for some participants, the correct response to the question could be that the information was not provided. Immediately afterwards, participants responded to the measures of threat: (1) implicit threat operationalized through physical distancing from the target, (2) direct-but-not-explicit threat operationalized through questions about the symbolic and realistic threats posed by the target, and then (3) two explicit threat questions, one identical to the measure used to assess *group*-level threat in Study 1 (for replication and comparison) and the other measuring participants' explicit perceptions of *target*-related threat. Finally, participants answered questions about their political orientation and demographics.

[9] Removing the participants ($n$ = 6) who were living in a foreign country at the time of participation did not affect the results. Analyses are available in the Supplemental Materials document on Github for Study 2A (link in Appendix C).

[10] The political loss target condition did not explicitly say the target voted for Trump in 2020, only that she wanted to vote for him *again* in 2024. A larger number than expected of the participants in this condition ($n$ = 12) said the target did not vote for Trump in 2020 while answering the other manipulation check questions correctly. For power and because they answered the other questions correctly, these participants were retained, and the condition was updated in Study 2B to avoid unclear answers on the manipulation check.

**Measures.** Only new measures are described in this section (e.g., because the BFI was described in Study 1, it is not described here).

***Implicit threat.*** Participants responded to two items that operationalized implicit threat by asking participants to estimate the physical distance between themselves and the target individual (Cook & Cottrell, 2021; Wyner & Calvini, 2011). The first item showed participants a schematic of a coffee shop (adapted from Ravenscroft's 2019 schematic and available in Appendix E) and stated, "Imagine you and Jen are at the same coffee shop. Jen sits down at one table to work. Where would you sit in the coffee shop? While answering this question, please imagine what you would do if COVID-19 did not exist, or in pre-COVID times. Choose your seat by selecting the seat's corresponding letter listed in multiple-choice format below." The second item asked participants to estimate the distance between where they and the target individual live in town, "Based on what you know about Jen, where do you imagine Jen lives in relation to you? Select your answer from the multiple-choice options below." Answer options ranged from "Same neighborhood, but not next door" to "Different state." These items were used in order to measure threat without triggering a socially desirable response bias (Talley & Bettencourt, 2008).

***Direct-but-not-explicit threat.*** Direct-but-not-explicit threat was operationalized using two items that measured the amount of threat participants felt that people like the target pose to their physical safety and the culture of the country (based on Intergroup Threat Theory by Stephan et al., 2002, 2016). One of these asked about *realistic* threat (i.e., threat posed by people like the target to participants' physical safety or resources): "How worried are you that people like Jen support policies that could harm the country's safety or economic well-being now or in the future?" The other item asked about *symbolic* threat (i.e., threat to values, culture, and way of life): "How worried are you that people like Jen support policies that could harm American

culture and what America stands for?" (Both items used the same scale; 1 = *Not at all*, 4 = *Very*). These items were also intended to prevent a socially desirable response bias (Talley & Bettencourt, 2008). Specifically, they were altered from Study 1 with this goal in mind by (1) associating the target with like-minded individuals ("people like Jen") to reduce discomfort responding directly about one person, (2) employing terminology that prompted participants to consider specific situations that may be harmed (physical safety, American culture), and (3) not using the actual term "threat" in the items. As participants responded to threat from "people like Jen," this measure is technically not target-*level* but target-*related.* The measure does, however, differ from the group-level measures used in Study 1 in that it mentions the target by name and furthermore – in comparison to Study 1 – it was presented after the target is introduced. However, it asks about group members "like Jen," making the measure not solely about Jen (as a purely "target-level" measure would be). Thus, this measure is related to the target more so than at the group-level, but not only operationalized at the target-level.

***Explicit Threat.*** Participants were asked two explicit threat questions. The first – group-level threat – was identical to the question asked in Study 1 (i.e., the question about how threatening the political out-group is). The second question – target-related threat – asked, "To what extent do you find Jen threatening to your beliefs and way of life?" (1 = *Not at all*, 5 = *Extremely*). Though not all past studies have found a relationship between explicit and indirect measures of threat (Talley & Bettencourt, 2008), these measures were included to explore the relationship between explicit and indirect threat in our sample. Moreover, including these explicit measures provided continuity with Study 1, although the placement of the group-level threat questions differed in Study 2A, with its measurement occurring after the target was introduced.

***Political Orientation.*** This three-item questionnaire was based on items used in a political psychology study by Carney and colleagues (2008). The three items assessed different facets of political orientation: (1) overall political orientation ("Overall, where would you place yourself on the following scale of liberalism-conservatism?"), (2) political orientation based on social issues ("In terms of social and cultural issues (e.g., abortion, separation of church and state, affirmative action), where would you place yourself on the following scale?"), and (3) political orientation based on economic issues ("In terms of economic issues (e.g., taxation, welfare, privatization of social security), where would you place yourself on the following scale?"). Each item used the same scale; 1 = *Extremely liberal* to 5 = *Extremely conservative*.

**Results and Discussion**

Analyses employed the following strategy: (1) relationships among the measures of threat were explored and a composite factor of threat was developed (i.e., target-related threat), and (2) a one-factor ANOVA was conducted examining target-related threat by target condition, in order to identify which targets would be used in the Study 2B. Packages used in the Study 2A analysis not employed in Study 1 included "corrplot" (Wei & Simko, 2021) and "factoextra" (Kassambra & Mundt, 2020).

**Identifying target-related threat.** A raw correlation matrix is available in Figure 5, which demonstrates that the measures of the adapted Intergroup Threat Theory scale (i.e., realistic and symbolic threat) were highly correlated with one another. However, the two implicit threat items (both operationalized through physical distancing) showed little correlation with one another and only weak correlations with realistic and symbolic threat.

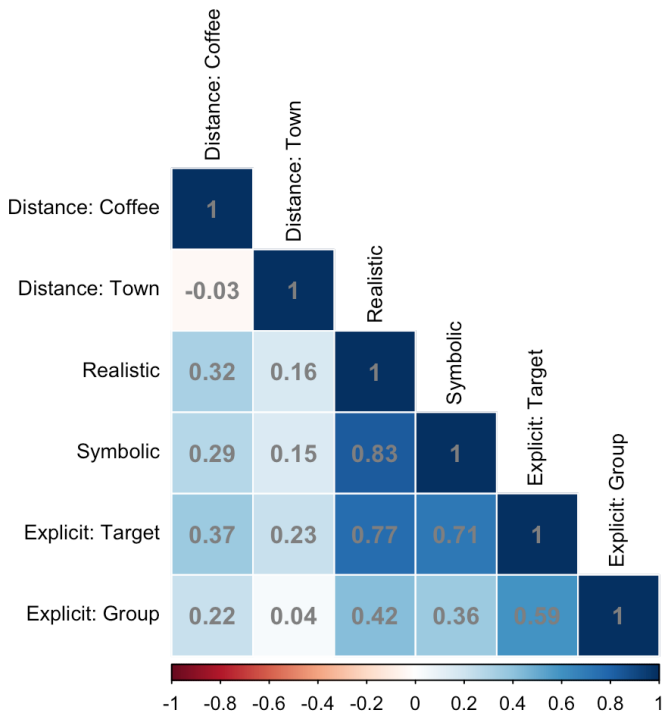



*Figure 5.* Raw correlation matrix for all measures of threat – including explicit – in Study 2A.

Because the different measures of threat were not highly intercorrelated, an exploratory factor analysis (EFA) was calculated to identify whether the two implicit measures and the two direct-but-not-explicit measures all loaded onto one factor. The EFA was restricted to one factor to develop a single composite measure of threat[11]. The factor loadings were rotated using oblimin rotation. As demonstrated in Table 1, the results supported one latent factor of target-related threat primarily consisting of realistic and symbolic threat, and this was what is used below in the analyses that follow as the measure of target-related threat. The measures of distance in the coffee shop and distance in town had much lower factor loadings and higher uniqueness. The reliability between the measures of symbolic and realistic threat were tested (α =

[11] The number of factors was initially explored using"nfactors" in R and looking at the number of recommended factors based on the root mean residual, with one factor being the most supported. As such, factor loadings for only one factor were calculated in the EFA. This analysis is available on Github (Appendix C).

0.91, 95% CI [0.88, 0.93]), before being averaged into a composite measure ($M$ = 2.86, $SD$ = 1.09) to be employed as the primary outcome measure in the ANOVA portion of the analysis.

**Table 1**

*Factor Loadings for Exploratory Factor Analysis Examining Threat in Study 2A*

| Items | Factor Loadings | Uniqueness |
|---|---|---|
| Realistic Threat | 0.98 | 0.05 |
| Symbolic Threat | 0.86 | 0.26 |
| Distance: Coffee | 0.32 | 0.89 |
| Distance: Town | 0.16 | 0.98 |

For theoretical validation purposes, the correlation of the measures of implicit threat (i.e., the distance measures) and direct-but-not-explicit threat with explicit threat were explored. As seen in Figure 5, the target-related explicit threat item was highly correlated with symbolic and realistic threat. The group-level explicit threat item was only moderately correlated with symbolic and realistic threat. The correlations of both explicit measures of threat with the implicit threat measures (operationalized through physical distancing) were weaker and sometimes non-existent. The lack of relationship between the implicit and explicit measures of threat is not surprising, as they have previously been found to be uncorrelated (Tally & Bettencourt, 2008). However, the relationship between the direct-but-not-explicit and explicit measures of threat was less certain *a priori*. In retrospect, the high correlation between target-related explicit threat and symbolic and realistic threat is not entirely surprising, especially since they were the only questions that asked participants to provide an opinion related to Jen (so some of the overlap may be due to method variance). Moreover, the structure of all three questions was similar (i.e., they all asked about Jen and people like her), with only slight variations in wording to operationalize different concepts (i.e., the symbolic and realistic threat vary in terms of

whether they are talking about cultural or physical threats; the target-related explicit threat measure asked about general threat). Thus, all three may be capturing the threat posed that is related to the target, while the other measures addressed different aspects of threat (group-level threat independent of Jen; desired distance from the target, without any reference to political opinions or affiliation). Overall, these high correlations suggest these measures are tapping into a concept of target-related threat.

**ANOVAs predicting threat by condition.**

***Target-related threat.*** Using target-related threat (i.e., the composite measure of realistic and symbolic threat) as the outcome, a one-way ANOVA with condition as the predictor was calculated. Target variable was dummy-coded with the non-political control target condition as the reference group because it was expected to be the condition with the least threat. The ANOVA was significant, $F(4, 204) = 42.07$, $p < 0.001$. Marginal means were calculated for target-related threat predicted by the target variable (available in Table 2).

**Table 2**

*Marginal Means for Target-related Threat Predicted by Target Variable in Study 2A*

| *Target Variable* | *Marginal Mean* | *SE* |
|---|---|---|
| Control Target | 1.47 | 0.12 |
| Warm Target | 2.93 | 0.12 |
| COVID Target | 3.28 | 0.12 |
| Competent Target | 3.21 | 0.14 |
| Political Loss Target | 3.49 | 0.13 |

*Note:* The target-related threat measure consisted of realistic and symbolic threat only. "Control target" refers to "non-political control target," but was shortened for space and consistency across tables.

Pairwise comparisons indicated that the non-political control target condition evoked significantly less threat in comparison to all other conditions (estimates and p-values available in

Table 3). The warm target condition also evoked significantly less threat than the political loss target condition, but did not differ significantly from the other two target conditions (i.e., the COVID target condition and the competent target condition). The political loss target condition did not lead to significantly more threat than the COVID target or the competent target conditions, nor did the COVID target condition and the competent target condition differ from one another.

As such, it appears that there are three targets consistent with the desired differing levels of threat to be used in Study 2B: Specifically, the political loss target was found to be the most threatening target, which is consistent with theory that loss of power is extremely threatening (Stephan et al., 2016). The warm target was found to be the moderately threatening target, which corresponds to previous findings that out-group members with traditional female gender roles are often pitied and regarded as non-threatening (Fiske et al., 2002). Finally, the non-political control target was the least threatening target. These three targets were thus chosen to be presented as the between-subjects target condition variable as a way of manipulating threat level in Study 2B.

In addition to identifying the target conditions, the results of the target-related threat measure also indicate that realistic and symbolic threat can be successfully adapted to be tied to target the target and within political contexts. This is important as the Intergroup Threat Theory items were typically employed at the group-level and did not test political threats, often being employed more to measure threat around immigration or racial differences (Stephan et al., 2002, 2016).

**Table 3**

*Estimates, Standard Error, t-values, and p-values for Pairwise Comparisons for Target-related Threat by Target Variable in Study 2A*

| *Comparison* | *b* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Control vs Warm target | -1.47 | 0.17 | -8.43 | <0.001 |
| Control vs COVID target | -1.81 | 0.17 | -10.41 | <0.001 |
| Control vs Competent target | -1.75 | 0.19 | -9.40 | <0.001 |
| Control vs Political Loss target | -2.02 | 0.18 | -11.34 | <0.001 |
| Warm vs COVID target | -0.34 | 0.17 | -2.00 | 0.270 |
| Warm vs Competent target | -0.28 | 0.18 | -1.53 | 0.547 |
| Warm vs Political Loss target | -0.55 | 0.18 | -3.14 | 0.016 |
| COVID target vs Competent target | 0.06 | 0.18 | 0.34 | 0.997 |
| COVID target vs Political Loss target | -0.21 | 0.18 | -1.19 | 0.757 |
| Competent target vs Political Loss target | -0.27 | 0.19 | -1.45 | 0.593 |

*Note:* "Control target" refers to "non-political control target," but was shortened for space and consistency across tables.

***Implicit threat.*** The ANOVA results (above) on the target-related threat measure were used to determine the three targets to be employed in Study 2B. However, further analyses were conducted exploring if the targets also produced different results on the implicit threat measures operationalized through physical distancing. Thus, two additional ANOVAs were calculated, one examining distance in the coffee shop as the dependent variable, and the other examining distance in town as the dependent variable. Both models were significant: distance in the coffee shop, $F(4, 204) = 6.70$, $p < 0.001$, distance in town, $F(4, 204) = 3.63$, $p = 0.007$. Marginal means are provided in Table 4.

**Table 4**

*Marginal Means for the Implicit Threat Measures by Target Variable in Study 2A*

| *Target Variable* | *Marginal Mean $_{coffee}$ (SE)* | *Marginal Mean $_{town}$ (SE)* |
|---|---|---|
| Control Target | 3.33 (0.14) | 3.00 (0.24) |
| Warm Target | 3.69 (0.14) | 4.24 (0.24) |
| COVID Target | 4.02 (0.14) | 3.93 (0.24) |
| Competent Target | 3.83 (0.15) | 3.86 (0.27) |
| Political Loss Target | 4.29 (0.14) | 3.68 (0.25) |

*Note:* "Control target" refers to "non-political control target," but was shortened for space and consistency across tables.

Pairwise comparisons indicated that the non-political control target condition resulted in significantly less reported seating distance in the coffee shop than the COVID target condition, $b$ = -0.70, $p$ = 0.004, and the political loss target condition, $b$ = -0.97, $p$ = <0.001. The warm target condition also resulted in significantly less seating distance than the political loss target condition, $b$ = -0.60, $p$ = 0.020. No other pairwise comparisons were significant, $ps$ > 0.11.

The results comparing targets in terms of participants' estimates of how far a target lives from them in town contained a surprising result: The warm target condition prompted the most distancing. Pairwise comparisons for the ANOVA predicting how far the target lived from the participant in town revealed that the non-political control target condition (who prompted the least distancing on this measure) resulted in significantly less reported distance from the warm target condition, $b$ = -1.24, $p$ = 0.003. None of the other comparisons were significant[12], $ps$ > 0.053.

Overall, some of the patterns from the physical distancing measures were consistent with the results found with target-related threat (i.e., the composite measure consisting of symbolic

[12] The comparison of the non-political control target condition versus the COVID target condition was marginally significant, $b$ = -0.93, $p$ = 0.053. As this pattern differs from the other measures of threat, it may reflect a tendency to social distance from someone who poses a threat to one's health from a lack of COVID precautions. However, as the result is marginal, this should be interpreted with caution.

and realistic threat). Specifically, both the non-political control target and the warm target conditions evoked less threat than the political loss target condition using the coffee shop distance measure, and the non-political control target condition differed from the warm target condition in both distancing measures. These results suggest that the non-political control target condition is the least threatening target across measures and the political loss target condition is the most threatening across *most* measures.

***Explicit Threat.*** For validation purposes, the patterns predicted by the target variable conditions for explicit threat were explored to see if they resembled those found with target-related threat (i.e., the composite of the realistic and symbolic threat items). After finding explicit threat at the target and group levels were reliable as one measure ($\alpha = 0.73$), the two were combined into one composite measure of explicit threat ($M = 2.75$, $SD = 0.91$). The composite measure of explicit threat was highly correlated with target-related threat ($r = 0.68$, $p < 0.001$, 95% CI [0.60, 0.74]). The ANOVA predicting the composite measure of explicit threat by target variable was significant, $F(4, 204) = 9.79$, $p < 0.001$. The resulting marginal means are available in Table 5.

**Table 5**

*Marginal Means for Explicit Threat Predicted by Target Variable in Study 2A*

| *Target Variable* | *Marginal Mean* | *SE* |
|---|---|---|
| Control Target | 2.13 | 0.13 |
| Warm Target | 2.69 | 0.12 |
| COVID Target | 3.04 | 0.12 |
| Competent Target | 2.74 | 0.14 |
| Political Loss Target | 3.16 | 0.13 |

*Note:* "Control target" refers to "non-political control target," but was shortened for space and consistency across tables. The marginal means reported above are for the composite measure of explicit threat.

Pairwise comparisons revealed that the non-political control target condition resulted in significantly less composite explicit threat than the warm target condition, $b = -0.56$, $p = 0.016$, the COVID target condition, $b = -0.92$, $p < 0.001$, the competent target condition, $b = -0.61$, $p = 0.013$, and the political loss target condition, $b = -1.03$, $p < 0.001$. The warm target condition only resulted in marginally less explicit threat than the political loss target condition, $b = -0.50$, $p = 0.075$. All other comparisons were not significantly different, $ps > 0.201$.

The relationship between the composite measure of explicit threat and target-related threat was mostly consistent with the idea that both measures of threat tap into overlapping constructs and supported the strategy of using the measure of target-related threat to select targets for the next study (Study 2B). Specifically, both the composite measure of explicit threat and target-related threat found that the non-political control target condition is the least threatening target. Moreover, both the composite measure of explicit threat and target-related threat found differences (though marginal with explicit threat) between the warm target and the political loss target conditions, indicating that the political loss target condition is the most threatening target condition, and the warm target is moderately threatening (i.e., more threatening than the non-political control target condition but less threatening than the political loss target condition).

# CHAPTER IV

## STUDY 2B

Study 2B employed the measures developed in Study 2A to expand upon the findings of Study 1, particularly the effect of analogous perspective-taking on counter-projection and the explanatory role of threat in counter-projection. As such, three goals of Study 2B included exploring (1) whether *target*-related threat directly related to when counter-projection occurred (e.g., Are perceptions of high threat necessary for counter-projection to occur?) (2) whether analogous perspective-taking reduces counter-projection, and (3) whether perceptions of target-related threat mediate the effect of analogous perspective taking on counter-projection. Additional goals of Study 2B included (4) replicating Denning and Hodges' (2022) previous finding that counter-projection persisted above and beyond stereotyping, (5) expanding upon their results by testing if any residual counter-projection is predicted by target-related threat, and (6) examining if this residual counter-projection can be reduced by analogous perspective taking. Finally, as the finding of residual counter-projection indicates a potentially unique social cognitive process separate from stereotyping used in perceiving out-group targets, Study 2B also (7) explored if residual counter-projection predicted behavioral intentions in response to political out-group targets.

In comparison to Study 1, there are two key differences in how threat is measured in Study 2B. First, threat is again tied to the *target* (as opposed to related to the group, as in Study 1), and it is measured by the composite measure of target-related threat developed in Study 2A. Second, variation in the target-related threat measure is manipulated by systematically varying how threatening the out-group target is (i.e., "target variable" in Study 2A). Three different targets were selected for use in Study 2B after examining between-target differences in levels of

the target-related threat measure in Study 2A. In the primary analyses in Study 2B, the manipulated target variable is not employed as a predictor, since the primary goal of including target-related threat was to ensure variability in responses to the measure of perceived target-related threat. However, the target variable will be used in a manipulation check to ensure this manipulation was successful in producing differences in target-related threat and in analyses to examine whether target-related threat and the manipulated target variable function similarly in relation to counter-projection.

The structure of the target-related threat measure identified in Study 2A was analyzed again in Study 2B and was expected to replicate, indicating reliability across samples. To this end, all measures of threat (i.e., implicit measures, direct-but-not-explicit measures, explicit measures) tested in Study 2A were also included in Study 2B, even if they were not expected to load onto a single factor of target-related threat. The three targets identified via pilot testing in Study 2A were employed in Study 2B, with the political loss target serving as the most threatening target, the warm target as the moderately threatening target, and the non-political control target serving as the least threatening target (together, these formed the three-level between-subjects target condition variable). In addition, Study 2B also manipulated whether participants received analogous perspective-taking instructions (i.e., the "instruction variable") or no instructions. This contrasted with Study 1 where *all* participants received analogous perspective-taking instructions due to the pre-post design. As such, the instruction variable had two conditions – the analogous perspective-taking instructions condition (i.e., the instructions used in Study 1) and the no instructions control condition.

## Predictions

Participants' target-related threat ratings were analyzed both as a predictor of counter-projection and as an outcome affected by the instruction variable; thus, there are predictions that reflect the variable as both a predictor and an outcome. Otherwise, the primary outcome measure is projection/counter-projection (i.e., whether self-ratings positively predict target ratings, as is seen with projection – or negatively predict target ratings, as an instance of counter-projection). Participants' willingness to have a conversation with the target was used as the measure of downstream behavioral intentions, and additional measure of interest explored in Study 2B. The predictions below are listed in the order the results are reported.

**Threat measure replication.** It was expected that the structure of target-related threat from Study 2A would replicate (i.e., that symbolic and realistic threat would form a single composite factor). In addition, it was predicated that the three different targets would significantly differ from one another in the amount of target-related threat they prompted, with the political loss target resulting in the most threat and the non-political control target resulting in the least threat.

**Effect of the instruction variable mediated by target-related threat.** Regarding counter-projection, it was expected people who reported higher target-related threat would counter-project the most, while those who reported little to no target-related threat would not counter-project (or even positively project), indicating that threat is necessary for counter-projection. A main effect of analogous perspective taking was also predicted, with those in the analogous perspective-taking condition counter-projecting less than those in the no instructions control condition. Furthermore, the effect of analogous perspective taking was anticipated to be mediated by perceptions of target-related threat (a pseudo-mediation was also conducted

confirming the relationship between the manipulated target variable and measured target-related threat, as well as the target variable's effect on counter-projection in a model including analogous perspective taking). Thus, the mechanism behind analogous perspective-taking's reduction in counter-projection was predicted to be via threat reduction[13].

**Residual counter-projection.** Drawing on an intriguing result from Denning and Hodges (2022), Study 2B tested whether it could replicate residual counter-projection (i.e., counter-projection above and beyond out-group stereotypes of a target) in the current sample, and if so, whether this residual counter-projection can be explained by the amount of threat associated with a target. Thus, this analysis helps address whether counter-projection is a distinct social cognitive process, or if it is explained by people using stereotypes of their out-group that also happen to be the opposite of how they rate themselves. To explore this, the measurement of stereotyping and counter-projection both employed two scales – the BFI (as in Study 1) and a measure composed of "Everyday Likes" items (ELI; a measure based on the items used in Denning & Hodges, 2022, and described in more detail below). Based on the results from Denning and Hodges (2022), it was expected that there would be less residual counter-projection for the BFI after accounting for stereotyping (because people hold stereotypes about how members of different political groups will respond on the BFI). In contrast, it was predicted that counter-projection would be unaffected after accounting for stereotyping for the ELI (the same results for counter-projection for occur for residual counter-projection). This was predicted because people were thought to have previously established stereotypes of political out-group members' personality traits (for

[13] It was initially predicted this would also depend on the manipulated target variable, through an interaction. However, upon reconsideration that the target variable was used to manipulate target-related threat, it was decided that only target-related threat would be employed, since both variables would be expected to show similar effects on counter-projection (explorations comparing multilevel models confirmed this) and could potentially be collinear. The preregistration for this study was adapted to reflect this change.

review, see Gerber et al., 2011) but not of their everyday likes (the ELI was specifically designed to be politically neutral; Denning & Hodges, 2022).

It was also predicted that the amount of residual variance remaining after accounting for stereotyping would depend on the amount of target-related threat people perceive. For the BFI, it was expected residual counter-projection would occur only when people perceived high threat (as high those who those who perceive high threat may need an additional process to create further distance) and that some counter-projection variance would be accounted for by stereotyping (based on findings by Denning and Hodges, 2022 using morals and personality traits with high in-group identifiers). However, Denning and Hodges (2022) previously found residual counter-projection regardless of high in-group identification using the ELI. Thus, it was expected that the same pattern of counter-projection prompted by perceived threat found prior to accounting for stereotyping would be found after. In other words, target-related threat would continue to predict residual counter-projection with the ELI in the same manner, regardless of stereotyping.

**Effects of residual counter-projection on behavioral intentions.** Finally, if residual counter-projection is found after accounting for stereotyping – indicating a distinct social cognitive process, maybe even possibly a heuristic – it is important to explore the effect of residual counter-projection on people's intended future behaviors. To address this in the current study, data were collected about participants' willingness to have a conversation with the out-group target. It is predicted that those who show residual counter-projection (after accounting for stereotypes of the out-group) will be the least willing to have a conversation with the target, across both the BFI and the ELI.

**Method**

Data were collected using both a sample of undergraduate students from the University of Oregon Psychology and Linguistics Human Subjects Pool and a sample from Prolific Academic. A power simulation to determine sample size was conducted using estimates of fixed effects, random effects, and residuals based on the results from Study 1 as well as the results of Denning and Hodges' (2022) Study 2. Denning and Hodges' results were used to estimate variance attributable to stereotyping and to account for any differences due to a between-subjects design. If there were more conservative estimates found in Study 2 of Denning and Hodges (2022; e.g., the residuals for subject-level variance were larger in the between-subjects design than the pre-post design, or the coefficients were smaller than the pre-post design), the more conservative estimates were utilized.

To further ensure the power simulation was conservative, all effect estimates were slightly underestimated (small effect sizes: 0.05 to 0.20) and larger residuals were approximated (i.e., 8.0 instead of 1.2 or 0.9). To be conservative, power was examined for the largest interaction of *possible* predictors of interest, even though we did not look to include the target variable as a predictor of counter-projection in the actual analyses. Specifically, the power simulation calculated one multilevel model that included target variable, instruction variable, and target-related threat predicting projection after accounting for stereotyping. We were not interested in interactions with stereotyping, just that its main effect was accounted for in the model. Stereotyping was also allowed to interact with all variables to make the simulation more conservative. Therefore, power was estimated on a smaller interaction in the model that did not include stereotyping as a predictor (i.e., the most complex interaction in the model did include stereotyping, but we used the second most complex interaction – the one without stereotyping).

A simulation was run 100 times at each of the following sample sizes: 400, 500, 550, 600, 650, and 700. Figure 5 shows 80% power was surpassed at 400 participants. To anticipate missing data and removal of data due to failed manipulation checks (there was a 15% removal rate of participants in Study 1), data were collected to surpass 90% power (500 participants, which allowed for removal of over 15% of participants). As many student participants as possible were collected via the Human Subjects Pool during the Winter Term of 2022 ($n$ = 44). In order to reach the target sample size of 500, the rest of the participants were collected via Prolific Academic ($n$ = 456).

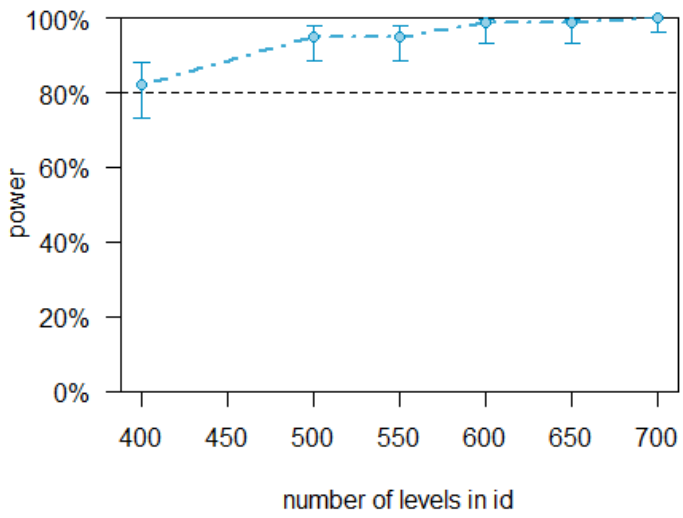



*Figure 6.* Power simulation for Study 2B.

A sample of 424 participants (405 from Prolific Academic, 19 from the Human Subjects Pool[14]) remained after removing participants who did not meet the pre-screen criterion of having

[14] Multilevel models testing the effects of the predictors – target-related threat, target variable, and instruction variable – on counter-projection and residual counter-projection were conducted with and without the participants from the Human Subjects Pool. However, the results did not change, so the full sample will be used in all analyses. The mediation was not conducted separately with and without student participants due to computer processing time needed to run this analysis.

supported President Biden in the U.S. 2020 Presidential Election ($n$ = 4); who incorrectly answered the manipulation checks about the targets (i.e., same questions as asked in Study 2A) or the analogous perspective-taking intervention ($n$ = 5; i.e., the question asked if they were asked in the study to write about a time they were misunderstood); or who incorrectly followed the analogous perspective-taking instructions ($n$ = 4); or who participated multiple times[15] ($n$ = 11). The sample included 270 participants who identified as women[16], 142 who identified as men, and 12 participants who identified as something other than a woman or man for their gender. Participants were, on average, 37.70 years old ($SD$ = 14.60). In terms of racial breakdown, 72.64% of participants were White or European American[17]; 10.85% were Asian/Asian American; 7.31% were Black/African American; 4.48% were Latina/o, Hispanic, Chicano/a, or Puerto Rican; 4.24% were multiracial; and less than 1.0% identified as Middle Eastern/North African or as falling in the "other" category. As in Study 2A, these demographics vary from the typical racial breakdown of the U.S., which is slightly less white (i.e., 60.1% white without identifying as Hispanic or Latino/a in the U.S.; United States Census Bureau, 2021) and Asian (5.9% in the U.S.), more Black/African American (13.4% in the U.S.), and more Hispanic/Latino/a (18.5% in the U.S.).

[15] This was determined though repeat IP Addresses, as in Study 2A. Testing for repeat IP Addresses is recommended when using online crowdsourcing websites like Prolific Academic to identify bots (Kennedy et al., 2020); it was also useful within the Human Subjects Pool. Specifically, repeat IP Addresses still occurred and were most likely a sign of participants having issues with the subject pool's method of granting credit to students for having participated, resulting in student participants taking part in the study multiple times. However, there is a smaller likelihood that participants were using shared computers to participate (e.g., a computer lab or a friend's computer). Regardless, within the Human Subjects Pool, only the first response per IP Address was retained to be safe.

[16] As noted earlier, Studies 2A and 2B asked about participants' gender instead of their sex.

[17] Data from student participants in Study 1 were collected using a subject pool mass-data collection protocol with pre-determined race/ethnic categories for data collection. As such, the categories differ from those used in Studies 2A and 2B, where the categories were able to be altered to reflect a wider variety of backgrounds.

Education was relatively normally distributed, with the largest category of people reporting their highest level of education as a bachelor's degree (37.50% Bachelor's degree; fewer than 1.0% had earned less than a high school diploma or its equivalent; 9.43% had completed high school; 23.35% had completed some college; 7.78% had an associate's degree; 4.01% had completed some graduate school; 14.62% had a master's degree; and 2.83% had a doctorate or professional degree). As in Sample 2A, our sample was more educated on average than the U.S. population, as seen by comparisons with total U.S. demographics (United States Census Bureau, 2021), where 88.5% of people have completed high school or higher (in comparison to over 99% in the current sample) and only 32.9% have completed their bachelor's degree or higher (in comparison to over 58.96% in the current sample). As in Study 2A, parents' education followed a similar pattern to participants' own education level, though with some slight variations (4.01% earned less than a high school diploma or its equivalent; 20.28% had completed high school; 11.08% had completed some college; 8.25% had an associate's degree; 28.30% had a bachelor's degree; 1.89% had completed some graduate school; 19.34% had a master's degree; 6.60% had a doctorate or professional degree).

Most participants were born in the U.S. (96.0%) and were raised in the U.S. (98.1%). For those who were not raised in the U.S., they had lived in the U.S. for between 10 and 35 years, making it likely that most participants were aware of American politics. Most participants identified as either "extremely liberal" ($n$ = 180) or "liberal" ($n$ = 188), which was not surprising given that a requirement for participation was that participants voted for Biden in the 2020 U.S. Presidential Election (see below). Nevertheless, some participants identified as "neither liberal nor conservative" ($n$ = 43) and a few identified as "conservative" ($n$ = 13). No one identified as "extremely conservative."

**Procedure.**

***Pre-manipulation procedure.*** Participants were screened for location (Prolific Academic participants had to live in the U.S.). To gather a sample that was politically similar to that used in Study 1 (i.e., a predominantly liberal sample), participants were prescreened for voting preference in the 2020 U.S. Presidential Election. Data were collected over a year after the election, in February and March 2022. On Prolific Academic, participants' previous self-identification that they voted for Biden was used as a pre-screen (a service provided by Prolific Academic), though participants were asked to confirm their voting history within the study. Within the Human Subjects Pool, participants identified whether they had voted for *or* supported Joe Biden in the 2020 U.S. Presidential Election, as some of the student participants may not have been old enough to vote when the election occurred. Student subjects who did not vote for/prefer Biden were allowed to complete the study, but their data were not included in the analyses.

Participants first responded to the BFI for themselves (as in Study 1). They were also asked to respond to ten Everyday Likes Items ("ELI"), a series of ten Likert-scale questions asking about everyday preferences, such as their inclination for the color yellow, their preference for tea versus coffee, or their preference for cats versus dogs (see "Measures" section for more information; five of the Everyday Likes Items had been previously used in Denning and Hodges, 2022). After responding to the BFI and ELI for themselves, participants answered both scales again, but this time for a stereotypical member of the out-group: those who supported Trump in the 2020 U.S. Presidential Election. This measure of stereotyping is considered idiographic – it is a participant's personal stereotype of out-group members on these items, not an average stereotype across everyone in the study.

***Instruction variable manipulation.*** Participants were randomly assigned to one of two instruction conditions, either the analogous perspective-taking instructions condition or the no instructions control condition. Those in the analogous perspective-taking instructions condition received the same instructions and writing task as in Study 1, where they were asked to write about a time they felt misunderstood. They were then randomly assigned to a target condition. Those in the no instructions control condition did not complete a writing task and were instead immediately randomly assigned to a target variable condition.

***Target variable manipulation.*** As previously discussed, target condition was manipulated by assigning participants one of three targets identified in Study 2A: The most threatening target was the "political loss target," the moderately threatening target was the "warm target," and the least threatening target was the "non-political control target." Initially, data collected from student participants in both Studies 2A and Study 2B were collected simultaneously, due to a limited amount of time available for data collection using the Human Subjects Pool. As such, once the three target conditions were identified using data collected from Study 2A, the data for Study 2B were filtered to include only participants assigned to those three conditions for analysis. Prolific Academic participants in Study 2B were assigned only to these previously determined three target variable conditions.

In running Study 2A, it was discovered that some participants in the most threatening target condition (i.e., the political loss target) in Study 2A were unsure if the target voted for Trump in 2020 (in the description of this target, it was implied, but not explicit in the wording: "she wanted to vote for Trump *again* in 2024"). Before collecting data from Prolific Academic participants, the description of this target was adapted to state clearly that Jen voted for Trump in 2020 ("Jen wants liberals to lose the Presidency and the U.S. House of Representatives in 2024.

She voted for President Trump in 2020 and hopes she can vote for him again in 2024.")[18]. However, the description was not changed for the student participants due to data collection already being in progress. Thus, the description of the most threatening target condition varied slightly for the participants in the Human Subjects Pool and Prolific Academic, making the manipulation check asking about who the target voted for in 2020 less ambiguous for the Prolific Academic participants in Study 2B. In addition to the manipulation check about voting preference, participants responded to the same manipulation checks from Study 2A asking about the target's opinion on COVID-19 and profession immediately after they were introduced to the target (participants had the option of answering that these pieces of information were not provided, which in some target variable conditions was actually correct).

***Post-manipulation procedure.*** Participants then responded to a series of measures about the target: (1) the two implicit measures of threat (the estimated physical distance participants would sit from the target in a coffee shop and the estimated physical distance participants thought they lived from the target); (2) the two direct-but-not-explicit measures of threat (symbolic and realistic threat); and (3) the two explicit threat questions (i.e., the questions that asked participants directly asked about "threat" in relation to the target – target-related explicit threat – and the participants' out-group – group-level explicit threat – were to their well-being). All six threat items were previously used in Study 2A. Participants then responded to the BFI and ELI for the target, and a question about how willing the participants would be to have a conversation with the target (described in the "Measures" section).

These measures were followed by three additional manipulation checks about the instruction variable. The first of these two questions asked how much participants took the

[18] Results with and without the participants from the Human Subjects Pool were calculated and were not different.

target's perspective and the second asked how often they thought the target was misunderstood. Participants in the analogous perspective-taking instructions condition were expected to have regarded the target as more misunderstood because they had written about a time when they personally felt misunderstood, which should trigger the realization that, if they – the participant – can be misunderstood, so too can the target (Brockman & Kalla, 2016). The third manipulation check asked participants whether they had been asked to write about a time they were misunderstood earlier in the study (those in the analogous perspective-taking instruction condition should respond "yes," while those in the control condition should respond "no"). At the end of the study, participants indicated their political orientation (on a scale of "extremely liberal" to "extremely conservative," using the same measures as in Study 2A) and provided demographics (same questions as in Study 2A).

**Measures.** Only new measures are described in this section (e.g., because the BFI was described in Study 1 and the threat items were described in Study 2A, they are not described again here).

***Everyday Likes Items.*** The ten ELI questions asked participants to report their preferences for everyday things and activities and were designed to be orthogonal to political stereotypes. Five of the items were previously used in Denning and Hodges' (2022) study exploring the role of stereotyping in counter-projection. The five prior items asked about participants' likes and preferences for stand-up comedy, cats versus dogs, condiments (e.g., ketchup, mayo, pickles), the color yellow, and winter versus summer. The five new items asked about pineapple as a pizza topping, coffee versus tea, playing video games, soccer versus basketball, and reading books for pleasure. All items were answered on 5-point Likert scales (e.g., for liking items, 1 = *Dislike a great deal*, 5 = *Like a great deal*; for preference items, 1 =

*Greatly prefer Winter to Summer*, 5 = *Greatly prefer Summer to Winter*) and can be found in Appendix F.

***Willingness to have conversation with target.*** To measure the effect of the analogous perspective-taking instructions on behavioral intentions, participants were asked how long they would be willing to have a conversation with the target individual, specifically: "If you met Jen while going about your day tomorrow, how long of a conversation would you be willing to have with her?" Response options included: 0 minutes, 15 minutes, 30 minutes, 45 minutes, and 60 minutes. For analysis purposes, the responses were converted to a 5-point scale (1 = *0 minutes* to 5 = *60 minutes*).

**Results**

This analysis plan was pre-registered on OSF, with data and reproducible analyses available on Github (links in Appendix C)[19]. Packages not used in Studies 1 and 2A include "doParallel" (Folashade et al., 2022) and "doRNG" (Gaujoux, 2020). Analyses are provided in the same order as the predictions (see above). Most of the analyses are multilevel models, as they are predicting ratings for the target from ratings for the self to capture projection/counter-projection. As in Study 1, Level 1 of all multilevel projection/counter-projection models in Study 2B consisted of responses for the self predicting those for the target (the item level). Along with the intercept, random slopes for responses for the self were also modeled. When stereotyping was included in model, it was also a Level 1 predictor and was allowed to vary randomly. The item level was nested within the participant level (Level 2). The variables included at the participant level vary depending on the model but may include target-related threat and/or the instruction variable. Predictors were always allowed to interact, including across levels.

[19] Amendments to the original plan are noted on OSF and throughout this paper.

If not specified, models used the BFI ratings to operationalize projection/counter-projection at the item level. This was partly due to there being zero variance (i.e., singularity) for the random intercept when modeling the ELI, while the BFI produced normal variance. Singularities often indicate the model is overfit; this occurred with the ELI even in an intercept-only model (intercept predicting target responses). However, the model was not singular if the random intercept was fixed and random slopes were allowed to vary. As the lack of random intercepts is somewhat concerning regarding the nature of the data (it can indicate there could be an error in the data, even though none was identified upon looking at the descriptives or data cleaning process, or when checking that participant responses matched after cleaning, etc.), the decision was made to use the ELI only when testing differences in residual counter-projection, and not in analyses testing the effect of the instruction variable (i.e., the mediation).

Results of the multilevel models should be interpreted like those in Study 1, meaning a significant interaction between ratings for the self and another predictor variable indicates a main effect for that predictor variable on projection/counter-projection (e.g., an interaction of responses for the self with target-related threat indicates a main effect of target-related threat on projection/counter-projection). Slopes were again used as a diagnostic measure to determine if people are projecting or counter-projecting, with significant negative slopes indicating counter-projection and significant positive slopes indicating projection. If the slope was not different from zero (i.e., flat), people were neither projecting nor counter-projecting. Again, "projection" will be used as the default terminology when describing main effects or interactions using ratings for the self predicting ratings for the target (i.e., the coefficients for the main effects produced from the multilevel analysis do not necessarily indicate if people are counter-projecting, just if people are projecting for certain or projecting less). After directionality and difference of slopes

from zero have been determined – typically through simple slopes – the terms "projection" and "counter-projection" will be used to indicate the correct process.

Level 1 continuous variables (i.e., responses on the ELI or the BFI – depending on the model – for the self, the target, and the stereotypical out-group member) were mean-centered within the subject to allow the mean to reflect variations around each individual participant's personal average on the items, as opposed to variation around the average of everyone who participated across all items[20]. Subject-level continuous variables (i.e., target-related threat) were grand mean-centered. Both the manipulated target variable and manipulated instruction variable were dummy-coded with their respective control conditions as the reference condition (i.e., the least threatening target was the reference condition for the target variable; the no instructions control was the reference condition for the instructions variable). In models in which the outcome was not projection (i.e., the model in which the manipulated target variable was used to predict target-related threat when testing if the analyses from Study 2A replicated), models other than multilevel models were employed (e.g., linear regression, ANOVA, EFA). These models are specified as they appear in their respective sections.

**Threat measure replication.** To verify the structure of target-related threat replicated from Study 2A, the following steps were taken: (1) a raw correlation matrix was calculated, (2) an EFA was conducted using the implicit and direct-but-not-explicit measures of threat, and (3) descriptive statistics were obtained for the subsequent composite measure (i.e., Cronbach's Alpha, mean, standard deviation). Following the verification of the structure of target-related

[20] Person mean-centering was initially employed to try to counteract the lack of variance in the random intercept for the ELI. Though this strategy did not solve the singularity issue, it was retained for Study 2B as the type of mean-centering did not appear to alter the results (i.e., the pattern of estimates were the same as grand mean centering) and seemed more representative of the data (i.e., it showed variation within person across items from the personal average, not the overall average – which may not be reflective of the individual's average).

threat, an ANOVA examined target-related threat broken down by target variable (comparing participants' reports of target-related threat for the most threatening target, the moderately threatening target, and the least threatening target – the targets determined by Study 2A).

The raw correlation matrix (available in Figure 7) demonstrated that symbolic and realistic threat (i.e., the direct-but-not-explicit measures) were, again, highly correlated with one another. As was the case for the results in Study 2A, the implicit distancing measures were weakly correlated with the other four measures of threat. However, unlike in Study 2A, the implicit distancing measures were also weakly correlated with one another, as opposed to being uncorrelated. When looking at the relationship of the direct-but-not-explicit measures with explicit threat, the pattern of results again indicated that symbolic and realistic threat were highly correlated with target-related explicit threat, but not group-level explicit threat. As such, the items replicate the pattern of intercorrelations among threat measures that was seen in Study 2A.

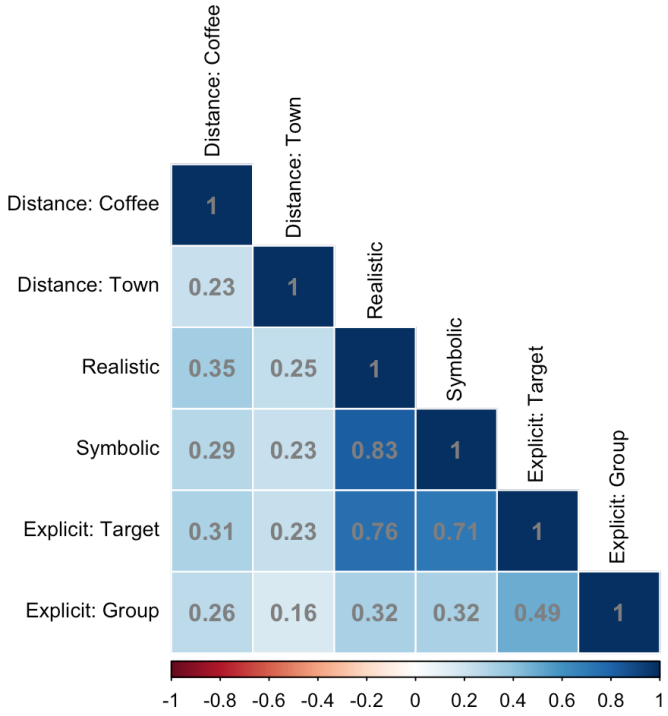



*Figure 7.* Raw correlation matrix for all measures of threat in Study 2B.

The four measures of indirect threat (i.e., the implicit distancing items and the direct-but-not-explicit items) were again subjected to a one-factor EFA[21] to identify whether there was a single latent factor of target-related threat. The factor loadings were rotated using oblimin rotation. The EFA results (available in Table 6) replicated those found in Study 2A, with the results supporting one latent factor on which symbolic and realistic threat had the highest loadings. The two distance measures (seating in the coffee shop; how far the target lived from the participant) had much lower factor loadings and higher uniqueness. The measures of symbolic and realistic threat were reliable ($\alpha = 0.90$, 95% CI [0.88, 0.92]), and thus were again collapsed into one composite measure, known as target-related threat ($M = 2.69$, $SD = 1.07$).

**Table 6**

*Factor Loadings for Exploratory Factor Analysis in Study 2B*

| Items | Factor Loadings | Uniqueness |
|---|---|---|
| Realistic Threat | 0.96 | 0.08 |
| Symbolic Threat | 0.84 | 0.29 |
| Distance: Coffee | 0.38 | 0.86 |
| Distance: Town | 0.29 | 0.92 |

An ANOVA looking at target-related threat broken down by the target variable was conducted in Study 2B to confirm the conditions evoked differing degrees of threat. The ANOVA was significant, $F(2, 421) = 200.00$, $p < 0.001$. Marginal means were calculated for target-related threat broken down by the target variable condition and can be seen in Table 7.

[21] It was initially pre-registered that the structure resulting from Study 2A (i.e., symbolic and realistic threat) predicting a latent factor of target-related threat would be analyzed in a confirmatory factor analysis (CFA). However, models with so few predictors are unidentifiable, so an EFA was again conducted.

**Table 7**

*Marginal Means for Target-related Threat Predicted by Target Variable in Study 2B*

| *Target Variable* | *Marginal Mean* | *SE* |
|---|---|---|
| Least threatening target | 1.70 | 0.06 |
| Moderately threatening target | 3.11 | 0.06 |
| Most threatening target | 3.39 | 0.07 |

Pairwise comparisons indicated that the least threatening target (i.e., the non-political control target) was perceived as less threatening than both the moderately threatening target (i.e., the warm target), $b = -1.69$, $p < 0.001$, and the highly threatening target (i.e., the political loss target), $b = -1.41$, $p < 0.001$. The moderately threatening target also was rated as significantly less threatening than the highly threatening target, $b = -0.28$, $p = 0.008$. These results replicated the findings of Study 2A, where the political loss target was also found to be the most threatening target, the non-political control target was found to be the least threatening, and the warm target was found to be moderately threatening. As such, these results indicate that the target variable successfully caused variation in participants' perceptions of target-related threat. Subsequently, the measure of target-related threat was employed as the measure of threat in all following analyses[22]. By employing target-related threat (versus the target variable), the following results are in line with Denning and Hodges' (2022) previous research that also tested the effect of an individual difference (i.e., in-group identification) on counter-projection.

***Perception in target differences on the BFI.*** As the target conditions were manipulated to differ in characteristics like warmth and competitiveness (i.e., to manipulate threat) that may

[22] As expected, analyses using just the target variable led to the same results. Few models were conducted using target-related threat and the target variable together as they captured the same variance and might have been collinear, as the target variable was developed to manipulate target-related threat in Study 2A (and that this relationship was confirmed through the manipulation check predicting target-related threat by target variable in Study 2B and the extension of the mediation testing the target variable's effect).

also lead to differences in the perceptions of target personality traits, separate analyses were calculated comparing participants' ratings of the three targets on each facet of personality (i.e., extraversion, open-mindedness, agreeableness, negative-emotionality, conscientiousness, and honesty) on the BFI. The items for each facet were collapsed into a single composite measure for each unique facet of personality (e.g., the three items corresponding to agreeableness were collapsed into one average for each participant on that facet). Separate ANOVAs were calculated predicting each facet by the target variable.

These analyses were done post-hoc and were exploratory. However, it was expected that more positive traits, such as agreeableness, conscientiousness, open-mindedness, and honesty, should be perceived as highest for the least threatening target (i.e., non-political, described warmly) and lowest for the most threatening target (described as competitive, wanting to take away the participants' resources). On the other hand, negative emotionality should show the reverse pattern, with the most threatening target rated as having the most. It was unclear how participants would perceive target extraversion, as they were only provided a brief vignette with little information about introversion/extraversion, a trait possibly clearer in-person or via longer narratives.

Target variable significantly predicted agreeableness, $F(2, 421) = 165.5$, $p < 0.001$. Marginal means (available in Table 8) and Tukey's post-hoc tests were calculated for this and all following models. Results indicated that people perceived the least threatening target to be most agreeable and to be more agreeable than the moderately threatening target, $b = 1.07$, $p < 0.001$, and the most threatening target, $b = 1.85$, $p < 0.001$. People also perceived more agreeableness in the moderately threatening target than the most threatening target, $b = 0.78$, $p < 0.001$.

**Table 8**

*Marginal Means for BFI Facets Predicted by Target Variable in Study 2B*

| | Marginal Means *(SE)* | | | | | |
|---|---|---|---|---|---|---|
| *Target Variable* | *Agreeable-ness* | *Open-mindedness* | *Conscientious-ness* | *Honesty* | *Negative emotionality* | *Extraversion* |
| Least threatening target | 3.86 *(0.07)* | 3.79 *(0.07)* | 3.96 *(0.06)* | 3.81 *(0.06)* | 2.57 *(0.06)* | 3.45 *(0.05)* |
| Moderately threatening target | 2.79 *(0.08)* | 3.10 *(0.07)* | 3.60 *(0.07)* | 3.21 *(0.06)* | 2.82 *(0.06)* | 3.58 *(0.06)* |
| Most threatening target | 2.01 *(0.07)* | 2.26 *(0.07)* | 2.96 *(0.06)* | 2.44 *(0.06)* | 3.07 *(0.06)* | 3.63 *(0.06)* |

Similar results were found for open-mindedness, $F(2, 421) = 121.4$, $p < 0.001$. Participants again perceived the least threatening target to be the most open-minded and also to be more open-minded than in the moderately threatening target ($b = 0.70$, $p < 0.001$) and the most threatening target ($b = 1.53$, $p < 0.001$). People also perceived more open-mindedness in the moderately threatening target than in the most threatening target, $b = 0.84$, $p < 0.001$.

Likewise, target variable significantly predicted conscientiousness, $F(2, 421) = 61.15$, $p < 0.001$. Participants again perceived the least threatening target to be most conscientious and more conscientious than both the moderately threatening target ($b = 0.36$, $p < 0.001$) and the most threatening target ($b = 1.00$, $p < 0.001$). People also perceived more conscientiousness in the moderately threatening target than in the most threatening target, $b = 0.64$, $p < 0.001$.

This same pattern persisted with honesty, which also was significantly predicted by the target variable, $F(2, 421) = 117.80$, $p < 0.001$. The least threatening target was again perceived to be the most honest and was perceived to be more honest than both the moderately threatening target ($b = 0.60$, $p < 0.001$) and the most threatening target ($b = 1.37$, $p < 0.001$). People also

perceived more honesty in the moderately threatening target than in the most threatening target, $b = 0.77$, $p < 0.001$.

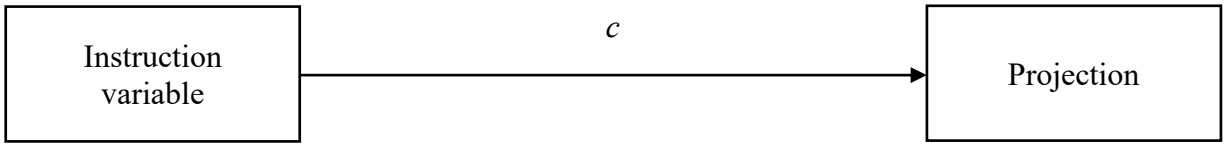
The opposite pattern was found regarding how participants perceived the targets' negative emotionality, $F(2, 421) = 17.49$, $p < 0.001$. Tukey's post-hoc tests revealed that participants perceived the most threatening target to have the most negative emotionality and more than both the least threatening target ($b = -0.51$, $p < 0.001$) and the moderately threatening target condition ($b = -0.25$, $p < 0.011$). The moderately threatening target also was perceived as having more negative emotionality than the least threatening target, $b = -0.26$, $p = 0.006$.

Finally, the model predicting perceptions of target extraversion by the target variable was only marginally significant, $F(2, 421) = 2.77$, $p = 0.063$. Marginal means are still available in Table 8 and post-hoc tests were conducted for completeness. Only the most threatening target was rated to have marginally higher extraversion than the least threatening target, $b = -0.18$, $p = 0.06$. All other comparisons were not significant, $ps > 0.21$.

**Effect of instruction variable mediated by target-related threat.** A mediational model with bootstrapping (10,000 iterations) was conducted. Samples for each analysis were randomly sampled, resulting in a total of 415 participants out of 424 that were selected per sample. (Subjects were not allowed to repeat within iterations because the models would not run due to differences in length of items across participants. For instance, if the same subject identification number [ID] was allowed to repeat within the model, a participant would have more than 19 BFI items attached to their ID, while other participants would only have 19). At the item level, items were randomly sampled with replacement within subject (the 19 items on the BFI were sampled with replacement per subject; items were allowed to repeat).

Three different models were calculated to determine if responses for target-related threat mediated the effect of the instruction variable on counter-projection. The first model calculated the direct effect of the instruction variable on counter-projection (i.e., without considering target-related threat; path *c* in Figure 8). Figure 8 uses the placeholder term of "projection" instead of "counter-projection" for interpretation with the mediational estimates. Specifically, since counter-projection is operationalized as a negative correlation between self and other traits (i.e., the opposite of projection), a negative estimate for projection indicates reduced projection, and likely counter-projection. A positive estimate would indicate projection. "Projection" will be used to described projection/counter-projection in the rest of the mediational model pathways. The multilevel model corresponding to path *c* included the instruction variable at Level 2. The coefficient for the fixed effect estimate corresponding to the interaction of the instruction variable with responses for the self was extracted each time the model was calculated. When the model was calculated, the instruction variable was mistakenly coded as the reference condition when dummy coding, though the reference was intended to be the no instruction control. To account for this, estimates including the instruction variable were multiplied by -1 for interpretability (i.e., with only two conditions, the estimate reflected the difference between the two. As such, misidentifying the reference only affected directionality, not the magnitude of the difference between the two).

Direct effect:




Indirect effect:

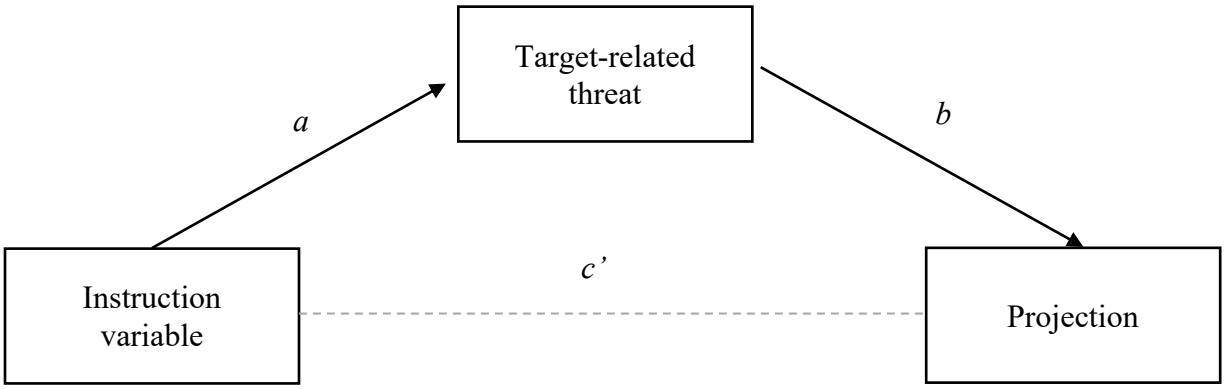



*Figure 8.* The direct and indirect pathways testing the mediational effect of target-related threat on the effect of instruction variable on projection in Study 2B.

The second model calculated the effect of the instruction variable on target-related threat (path *a* in Figure 8). Since this model was at the subject-level, an ANOVA was calculated predicting target-related threat by the instruction variable[23]. The coefficient corresponding to the main effect of the instruction variable was extracted. (As in the first model estimating the direct effect, the estimate was multiplied by -1 to reverse the reference condition to be the no instructions control condition for the instruction variable.)

The third model calculated both the effect of target-related threat on projection (path *b* in Figure 8) and the extent the direct effect of the instruction variable on projection was altered by

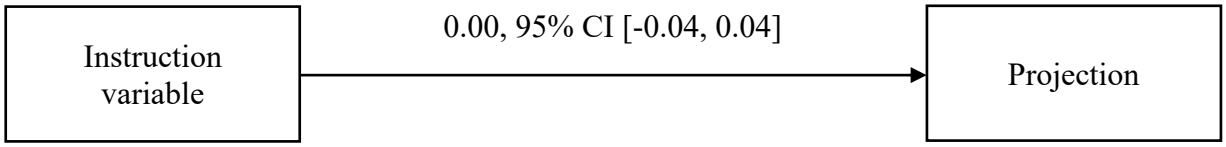
[23] Though this analysis is a one-way ANOVA because it only includes one categorical predictor, it reduced computer processing time to use the "lm" function in R (that corresponds to a linear model) because the "lm" method required one less function to obtain coefficients (i.e., important when iterating over 10,000 times). As an ANOVA is a type of linear model, this was not expected to alter results. However, this assumption was confirmed by testing the results outputted for a single sample using both "lm" and the traditional ANOVA approach ("aov" and "emmeans"). As the estimates calculated were the same using both methods, the model was calculated using the "lm" function.

including target-related threat in the model (path *c'* in Figure 8). The model employed was almost identical to the multilevel model used to calculate the direct effect of the instruction variable (path *c*). However, an additional Level 2 predictor of target-related threat was added. The coefficients for the fixed effects extracted from each iteration of the model included (1) the interaction of responses for the self and target-related threat (path *b*) and (2) the interaction of responses for the self with the instruction variable after accounting for target-related threat (path *c'*).

After all coefficients were extracted for each iteration, the total indirect effect was calculated by multiplying the *a* (i.e., the effect of the instruction variable on the target-related threat measure) and *b* (i.e., the effect of target-related threat on projection) pathways together for each model. Also, a total effect was calculated by adding the total indirect effect (i.e., *a* by *b*) to *c'* (i.e., the effect of the instruction variable on projection after accounting for target-related threat). This total effect was compared to the estimates calculated for *c* (i.e., the direct effect of the instruction variable on projection before accounting for target-related threat) for a subset of models, finding they were the same within rounding variation, indicating the mediation was calculated correctly. Mean and 95% confidence intervals[24] were calculated for each pathway in the mediation (i.e., across all estimates extracted for one effect) as well as for the indirect effect found from multiplying *a* by *b*. These estimates are available in Figure 9.

[24] The median was also calculated and nearly identical to the mean (*a:* -0.15, *b*: -0.15, *c:* 0.0, *c'*: -0.02, *a x b:* 0.02).

Direct effect:




Indirect effect:

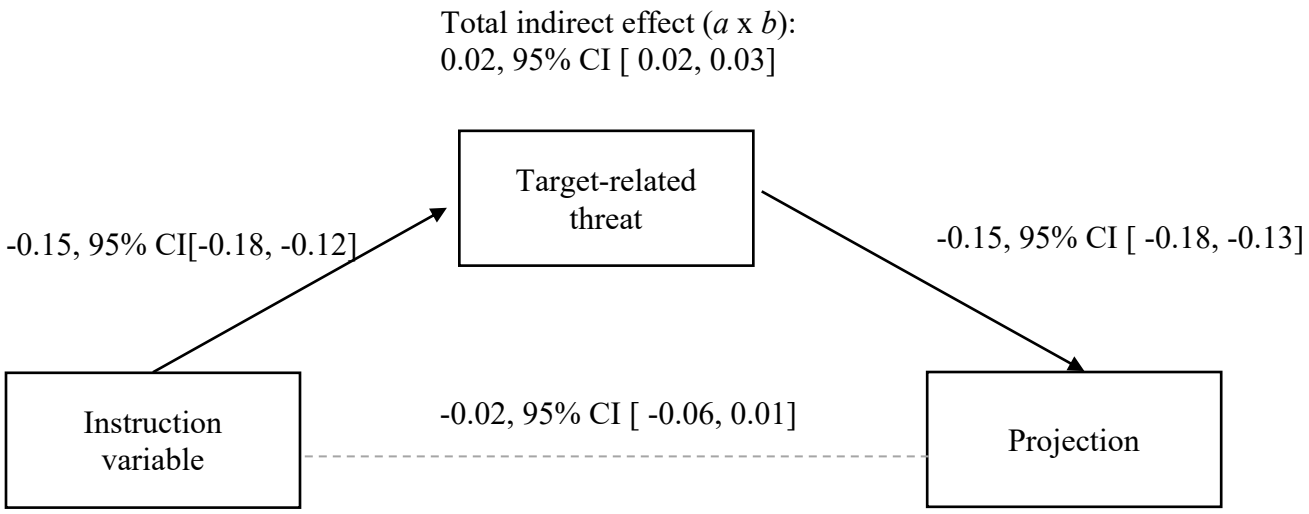



*Figure 9.* The estimates of the mediational model using the BFI in Study 2B, showing an indirect effect of the target variable on projection through target-related threat.

The results of the mediational model demonstrate there is no direct effect of the instructions variable on projection using a between-subjects design. However, there is an indirect effect of the instruction variable on projection through target-related threat. As can be seen in path *a*, the instruction variable significantly predicted target-related threat, with those in the analogous perspective-taking condition reporting less target-related threat than those in the no instructions control condition. Target-related threat also significantly predicted projection (path *b*), where an increase in target-related threat predicted less projection (and more counter-projection). The total indirect effect of the instruction variable on projection through target-related threat is positive, indicating that reductions in target-related threat due to analogous

perspective-taking instructions have a small effect on the extent people project (versus counter-project) with others. The residual effect of the instruction variable on projection after accounting for target-related threat was still not different from zero.

***Related simple slopes for target-related threat.*** As the mediation indicates an effect of target-related threat on counter-projection, this relationship was explored more deeply by conducting a single multilevel model on the entire dataset (i.e., predicting responses for target from responses for the self and perceptions of target-related threat) and examining the corresponding simple slopes. Unsurprisingly and as found in the mediation, there was a significant main effect of target-related threat on projection, $b = -0.13$, $p < 0.001$, 95% CI [-0.16, -0.11]. Simple slopes analyses were calculated, breaking target-related threat into the following perceived threat groups: high perceivers of threat (i.e., one standard deviation above the average), low perceivers of threat (i.e., one standard deviation below the average), and average perceivers of threat. As seen in Figure 10, the results revealed that high perceivers of threat counter-projected, $b = -0.13$, $p < 0.001$, 95% CI [-0.18, -0.10], average perceivers of threat neither projected nor counter-projected, $b = 0.00$, $p = 0.730$, 95% CI [-0.02, 0.03], and low perceivers of threat positively projected, $b = 0.15$, $p < 0.001$, 95% CI [0.11, 0.19]. Pairwise comparisons indicated that all the simple slopes significantly differed from one another, $bs$ = 0.14 – 0.28, $ps < 0.001$. Thus, as perceived threat increased, positive projection decreased, ultimately resulting in counter-projection for high perceivers of threat.

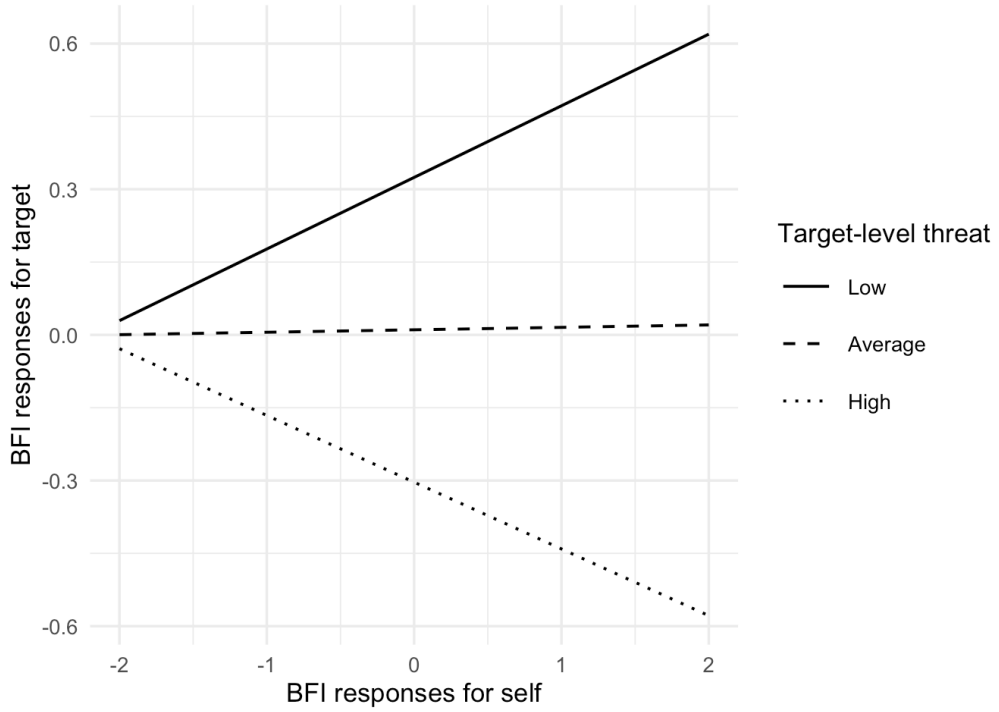



*Figure 10.* Main effect of target-related threat on projection using the BFI in Study 2B.

***Mediation extension exploring target variable.*** An additional model was conducted to explore if the mediating effect of the *manipulated* target variable was like that found using the *measure* of target-related threat, by examining the main effects of both the instruction variable and target variable on projection in the same model. As such, this analysis also tests the two routes by which target-related threat was manipulated and thought to affect projection – the instruction variable (indirectly demonstrated through target-level threat via the previous mediational model) and the target variable (previous manipulation checks demonstrated that target-related threat is predicted by the target variable). This model is predicated on the knowledge that analogous perspective-taking predicts target-related threat (path *a* in the mediational model) – suggesting that the two are related – and that target-related threat affects projection (path *b*). If the target variable and target-related threat occupy a similar conceptual

space, the manipulated target variable should have the same effect on projection as target-related threat (as compared by looking at coefficients and simple slopes).

A multilevel model was calculated – again using the BFI at the item-level – but including both the instruction variable and target variable as Level 2 predictors. Target variable and instruction variable were allowed to interact with responses for the self on the BFI, but not with one another. As in other models, both variables were dummy coded and their respective control conditions (i.e., the control target condition for the target variable and the no-instructions control condition/least threatening target condition for the instruction variable) were used as reference conditions. Results indicated that there was a significant main effect of the target variable using both contrasts (contrast comparing least threat threatening target to the moderately threatening target: $b$ = -0.21, $p$ < 0.001, 95% CI [-0.28, -0.14]; contrast comparing the least threatening target to the most threatening target: $b$ = -0.36, $p$ < 0.001, 95% CI [-0.43, -0.29]), but the main effect of the instruction variable was not significant, $b$ = -0.00, $p$ = 0.960, 95% CI [-0.06, 0.06].

Simple slopes were calculated decomposing the main effect of the target variable on projection. Results indicated people counter-projected in the most threatening target condition, $b$ = -0.18, $p$ < 0.001, 95% CI [-0.23, -0.12]; neither projected nor counter-projected in the moderately threatening target condition, $b$ = -0.02, $p$ = 0.330, 95% CI [-0.07, 0.02]: and (positively) projected in the least threatening target condition, $b$ = 0.19, $p$ < 0.001, 95% CI [0.14, 0.23]. Pairwise comparisons indicated all simple slopes were significantly different from one another, $bs$ = -0.15 – 0.23, $ps$ < 0.001.

These results indicate that, when included in the same model, the results are similar to those found in the mediation using target-related threat instead of the target variable. Specifically, it is threat – not analogous perspective taking-instructions – that predicts when

people project or counter-project. Though the instruction variable does have an indirect effect on projection through target-related threat in the mediation, it continues to not have a direct effect in these analyses and does not appear to reduce variance explained by the target variable. Instead, the effect on counter-projection using the manipulated target variable is similar to that previously found using measured target-level threat. Specifically, when comparing the simple slopes obtained for the target-related threat model decomposing path *b* to those found for this model including the target variable, the estimates for related effects are within the upper and lower bounds of each other's 95% confidence intervals (e.g., participants who perceived high threat had similar estimates for counter-projection as participants in the moderately threatening target condition). These results suggest we were successful in manipulating perceived threat via the target variable.

**Residual counter-projection.**

***Target-related threat with the BFI.*** The mediational model clearly demonstrated that target-related threat explains a significant portion of the variance in counter-projection, with counter-projection being prompted by perceptions of high threat. *However, is this effect of counter-projection the result of a distinct process of psychological distancing, or is it due to people having stereotypes that their out-group is opposite themselves?* To address this, an additional model was calculated with predicting ratings of the target from self ratings on the BFI at the item-level, but this time adding participants' stereotypes of the outgroup on the BFI as an additional item-level predictor. Stereotyping was also allowed to have a random slope.

Results indicated that the main effect of target-related threat on projection persisted after stereotyping was added to the model, $b = -0.07$, $p < 0.001$, 95% CI [-0.10, -0.05]. However, there was also a main effect of stereotyping, $b = 0.22$, $p < 0.001$, 95% CI [0.19, 0.26]. Moreover, a

model comparison indicated that adding stereotyping to the model improved model fit, $\chi^2$ = 1944.4, $p < 0.001$, Δ Conditional $R^2 = 0.14$.

When simple slopes were examined for the main effect of target-related threat on projection, results indicated that high perceivers of threat no longer counter-projected, $b = -0.01$, $p = 0.572$, 95% CI [-0.04, 0.02], though they also did not project. Perceivers of both average, $b = 0.07$, $p < 0.001$, 95% CI [0.04, 0.09], and low threat, $b = 0.14$, $p < 0.001$, 95% CI [0.11, 0.18], projected positively after accounting for stereotyping. Pairwise comparisons revealed that all simple slopes differed significantly from one another, $bs = 0.08 – 0.15$, $ps < 0.001$. This relationship is depicted in Figure 11. Counter to what was predicted, these results demonstrate that there is no residual counter-projection with the BFI, regardless of how much threat people perceive. Counter-projection of personality traits appears to be entirely explained by people's use of idiographic stereotypes of their out-group.

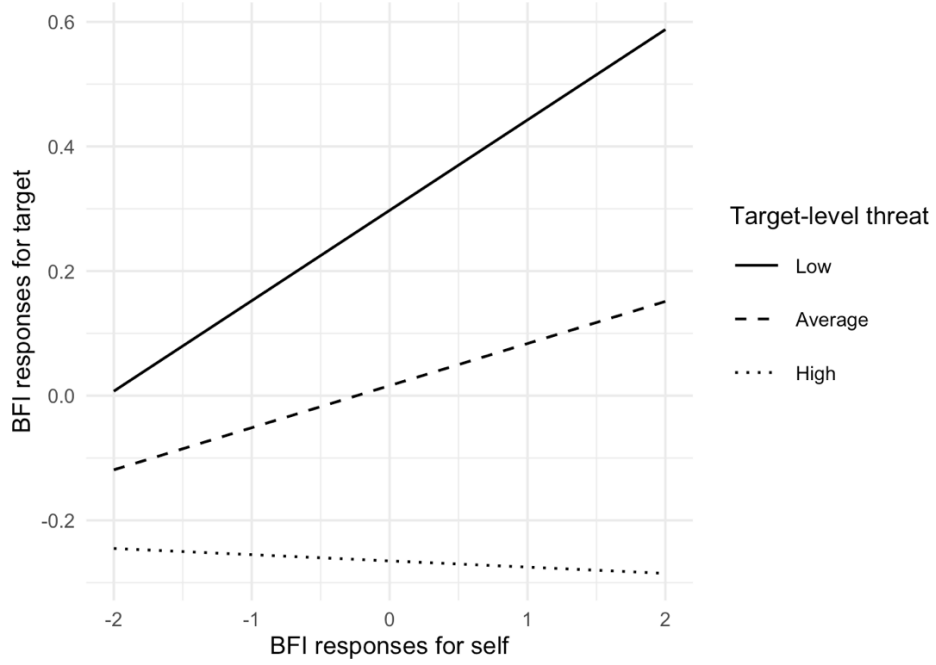



*Figure 11.* Main effect of target-related threat on projection using the BFI after accounting for stereotyping in Study 2B.

***Target-related threat with the ELI.*** Because counter-projection was previously found to depend on the type of information being projected (Denning & Hodges, 2022 found residual counter-projection after accounting for stereotyping with everyday preferences regardless of stereotyping, but residual counter-projection with personality traits and morals depended on in-group identification), analyses employing the ELI were also conducted. These analyses also explored whether residual counter-projection after accounting for stereotyping depended on target-related threat, only this time the analyses used self, target, and stereotypical ratings on the ELI instead of the BFI. As a reminder, it was expected that residual counter-projection would be found with the ELI because people do not have political stereotypes associated with everyday preferences. Target-related threat was expected to predict counter-projection (i.e., prior to including stereotyping in the model) with the ELI in a similar manner to when people counter-projected with the BFI (i.e., high perceivers of threat will counter-project the most), meaning those who perceived high threat would counter-project with everyday preferences. However, as Denning and Hodges (2022) previously found counter-projection across more levels of in-group identification with the ELI versus the BFI (both low and average identifiers counter-projected to some degree with the ELI), it was unclear if counter-projection would be found with those who perceived average or low target-related threat.

Stereotyping was not expected to change the effect of target-related threat, and thus, it was predicted there would still be *residual* counter-projection. In other words, we should see counter-projection across both models with and without stereotyping using the ELI (this prediction contrasts what was seen with the BFI). To test these predictions, two models were conducted: (1) one model predicted ratings for the target on the ELI from ratings for the self at Level 1 with target-related threat as a Level 2 predictor, and (2) another model added

stereotyping on the ELI (responses for the stereotypical out-group member on the ELI) as another Level 1 predictor. The first model without stereotyping was conducted for comparison purposes.

As mentioned earlier and to our surprise, when random intercepts and slopes at the item level were allowed to vary using the ELI, the model was singular due to zero variance in the intercept. Further testing revealed that even in an intercept-only model, the model was still singular[25]. To allow for comparison to the BFI, the random intercepts were removed, but the random slopes were retained for ratings for the self. Random slopes for stereotyping using the ELI were also employed.

The base model, without stereotyping, revealed a significant main effect of target-related threat, $b = -0.09$, $p < 0.001$, 95% CI [-0.11, -0.06]. Simple slopes were conducted to decompose the model. As seen in Figure 12, high perceivers of threat counter-projected, $b = -0.08$, $p < 0.001$, 95% CI [-0.12, -0.04]; average perceivers of threat neither projected nor counter-projected, $b = 0.01$, $p = 0.378$, 95% CI [-0.02, 0.04]; and low perceivers of threat projected positively, $b = 0.11$, $p < 0.001$, 95% CI [0.06, 0.15]. Pairwise comparisons revealed that all simple slopes differed significantly from one another, $bs = 0.09 – 0.18$, $ps < 0.001$.

[25] Tests were conducted to verify the data was normal, that there was variance in the responses, and that the responses were correctly attached to subject ID number. Additional analyses were conducted removing the new ELI items, but the random intercept remained singular using only the five original ELI items. The distribution of responses for each item from the original ELI items was also found to be consistent with that found in the Denning and Hodges (2022) previous study.

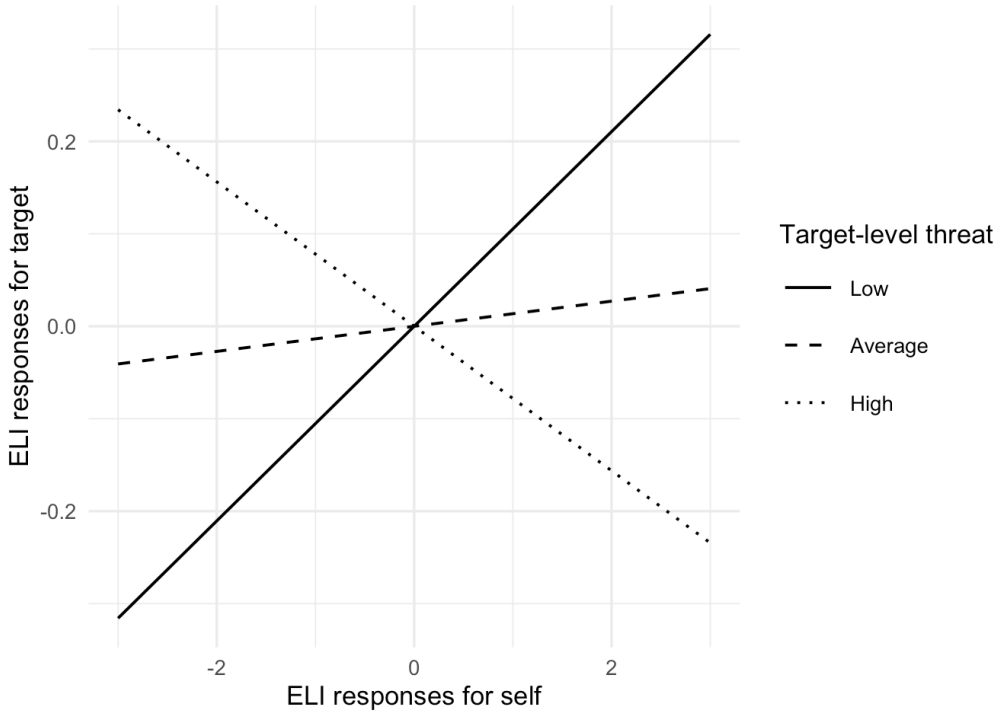



*Figure 12.* Main effect of target-related threat on projection using the ELI in Study 2B.

The second model was then conducted accounting for stereotyping. Results revealed that the main effect of target-related threat was still significant, $b = -0.06$, $p < 0.001$, 95% CI [-0.09, -0.04]. The main effect of stereotyping was also significant, $b = 0.27$, $p < 0.001$, 95% CI [0.24, 0.31], and a model comparison revealed the model with stereotyping fit the data better, $\chi^2 = 844.88$, $p < 0.001$, $\Delta$ Conditional $R^2 = 0.22$.

Simple slopes revealed that high perceivers of threat continued to counter-project, $b = -0.04$, $p = 0.044$, 95% CI [-0.08, -0.001], indicating *residual* counter-projection. As seen in Figure 13, simple slopes analyses revealed that average perceivers of threat now projected positively, $b = 0.03$, $p = 0.044$, 95% CI [0.001, 0.05], while low perceivers of threat also continued to project positively, $b = 0.09$, $p < 0.001$, 95% CI [0.05, 0.13]. Pairwise comparisons indicated that the simple slopes all differed from one another, $bs = 0.7 – 0.13$, $ps < 0.001$.

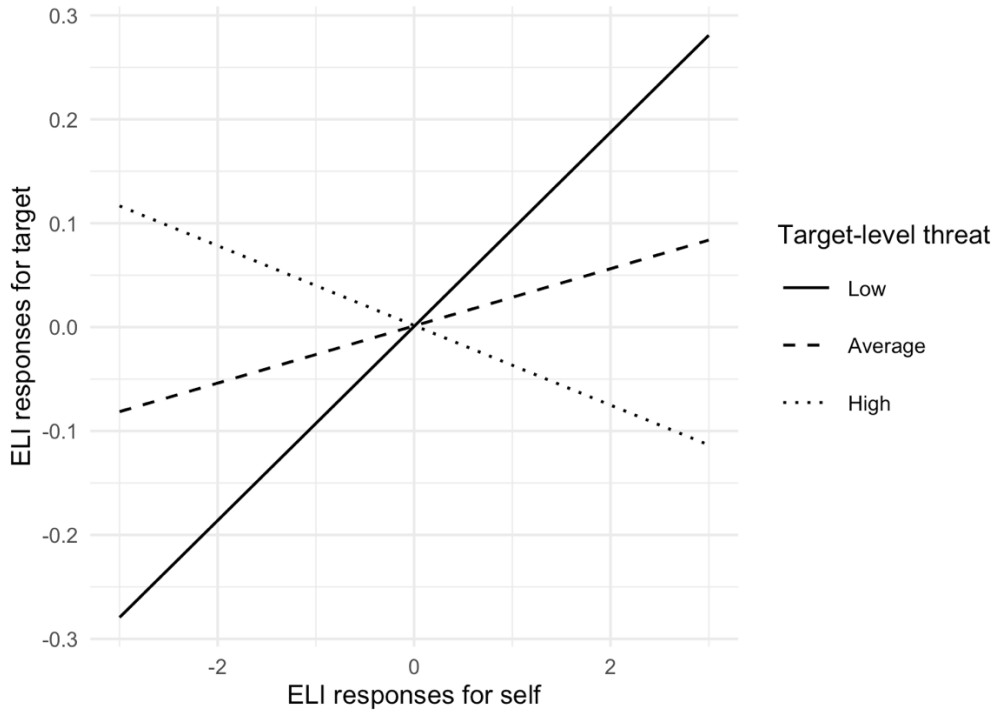



*Figure 13.* Main effect for target-related threat on projection using the ELI after accounting for stereotyping in Study 2B.

The comparison of the two models using the ELI revealed mixed support for predictions. Specifically, counter-projection persisted across both models, which supports the results found by Denning and Hodges (2022). Target-related threat predicts when people counter-project using the ELI, and counter-projection using the ELI is unaffected by stereotyping. However, the results for average perceivers of threat differed across models, a result not in line with Denning and Hodges (2022). Specifically, perceivers of average threat neither projected nor counter-projected when stereotyping was not included but they positively projected after stereotyping was accounted for. As such, projection with the ELI is affected by stereotyping in the current sample, with the responses that drove average perceivers to see the out-group as less like them on the ELI being explained by stereotyping.

These results also indicate that participants have a stereotype for typical out-group members using the ELI. However, this stereotype does not appear to be as strong as those held about personality traits, since we continue to see counter-projection with those who perceive high threat with the ELI and not the BFI after accounting for stereotyping.

***Reducing residual counter-projection with the ELI.*** One of the initial research questions regarding residual counter-projection focused on whether residual counter-projection can be reduced with analogous perspective-taking instructions. As residual counter-projection was only found with the ELI, this question will only be explored with this measure. Building upon the previous model, the instruction variable was added as an additional Level 2 predictor. The instruction variable was dummy-coded with the control condition as the reference condition.

Results indicated that there was a significant interaction of the instruction variable and target-related threat predicting projection after accounting for stereotyping, $b = 0.05$, $p = 0.049$, 95% CI [0.0003, 0.10]. To explore whether residual counter-projection was reduced by analogous perspective-taking instructions, the interaction was decomposed with simple slopes analyses. Results indicated that high perceivers of threat continued to counter-project after stereotyping was accounted for participants in the control condition, $b = -0.07$, $p = 0.006$, 95% CI [-0.12, -0.02], but not for participants in the analogous perspective-taking instruction condition, $b = -0.04$, $p = 0.216$, 95% CI [-0.09, 0.02], where they neither projected nor counter-projected. However, pairwise comparisons indicated these slopes were not significantly different from one another, $b = 0.03$, $p = 0.948$. As such, support for analogous perspective-taking reducing residual counter-projection is mixed – significant residual counter-projection is found in the control condition, not the analogous instruction condition, but the difference in slopes was not significant.

**Effects of residual counter-projection on behavioral intentions.** To test the effects of counter-projection on social behavior, Study 2B also included a measure of behavioral intentions – specifically, how long participants would be willing to have a conversation with the target. In order to conduct these analyses, coefficients that corresponded to ratings for the self predicting ratings for the target for each *individual participant* were extracted from a previous model. Due to the coefficients being at the individual level, individual differences in residual counter-projection could still be observed within the BFI, even though stereotyping accounted for the variance *across* individuals in the previous model. For example, one person may have projected the opposite of their own personality traits to such an extent that stereotyping did not entirely explain their individual counter-projection, even if stereotyping does explain counter-projection on the BFI overall across individuals.

The coefficients were extracted from both the ELI and the BFI models that included stereotyping as a covariate (the models also included target-related threat as a predictor). Coefficients for the effect of target-related threat on projection and the main effect of projection (i.e., ratings for the self predicting ratings for the target – no interactions) were extracted. However, models including coefficients for the effect of target-related threat as the predictor were singular, regardless of the type of measure (i.e., BFI or ELI). As such, the coefficients corresponding to the main effect of projection were used as the predictor in a linear regression predicting how long participants were willing to spend in a conversation with the target. Separate models using the coefficients for the ELI and the BFI were run. The coefficients were not centered as the scale was already meaningful (i.e., a negative coefficient represents counter-projection, a coefficient around zero represents neither projection nor counter-projection, and a positive coefficient represents projection), nor was the outcome variable centered.

***Using the BFI.*** Results indicated that projection coefficients from the BFI model positively predict willingness to have a conversation, $b = 1.32$, $t = 5.02$, $p < 0.001$. As can be seen in Figure 14, as coefficients increase, time willing to have a conversation also increases. Importantly, those with negative coefficients – representing *individual* residual counter-projection – were the least willing to spend time having a conversation with the out-group target.

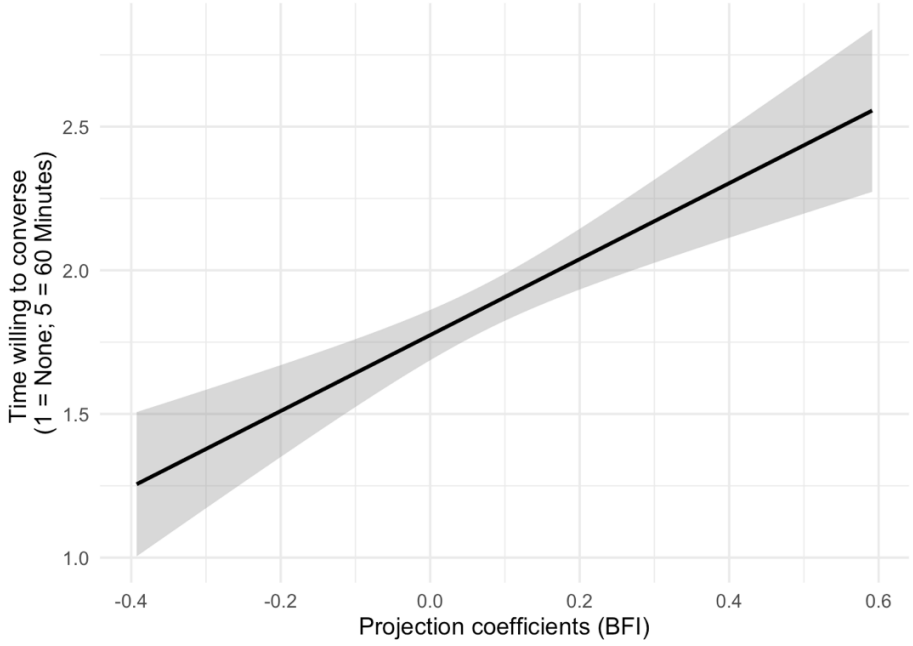



*Figure 14.* Effect of projection on willingness to have a conversation with the target using BFI model coefficients in Study 2B. Negative coefficients indicate counter-projection.

The model was expanded upon by exploring if the instruction variable reduced residual counter-projection's effect on willingness to have a conversation. The two predictors – projection coefficients and instruction variable – were allowed to interact. Results indicated there was no main effect $b = -0.14$, $t = -1.61$, $p = 0.101$, nor interaction including the instruction variable $b = 0.44$, $t = 0.83$, $p = 0.405$. The main effect of the projection coefficients persisted, $b = 1.06$, $t = 2.61$, $p = 0.009$.

***Using the ELI.*** Results using the ELI were similar, indicating that projection coefficients also positively predicted willingness to have a conversation, $b = 1.93$, $t = 4.34$, $p < 0.001$. As can be seen in Figure 15, as coefficients increase, time willing to have a conversation also increases. Again, those with negative coefficients were the least willing to spend time having a conversation with the out-group target.

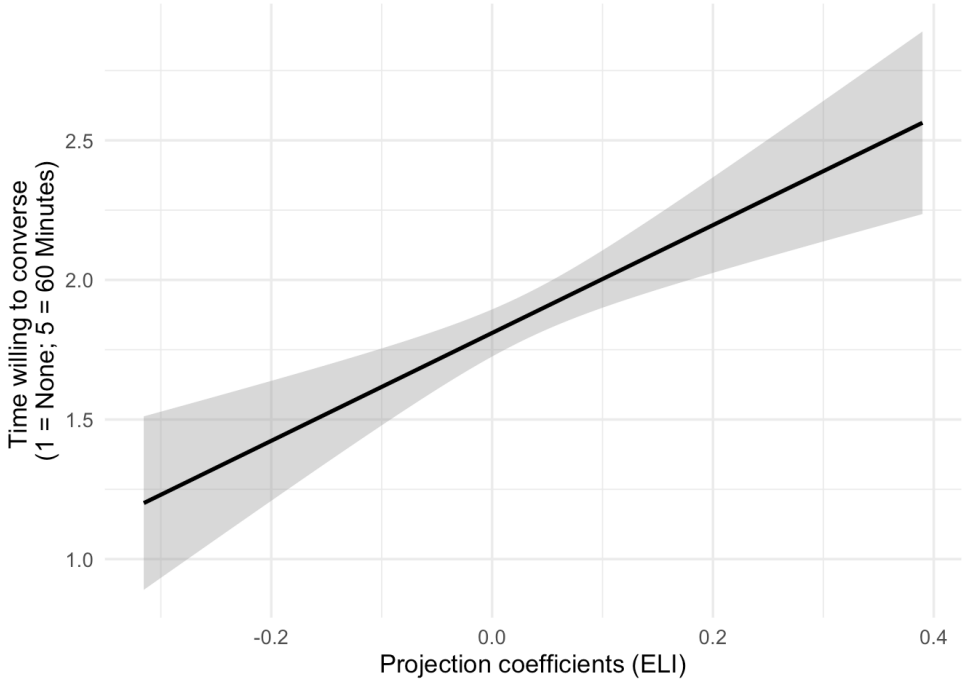



*Figure 15.* Effect of projection on willingness to have a conversation with the target using ELI model coefficients in Study 2B. Negative coefficients indicate counter-projection.

As with the ELI, the model was again expanded upon by exploring if the instruction variable reduced residual counter-projection's effect on willingness to have a conversation. Results indicated there was no main effect $b = -0.11$, $t = -1.32$, $p = 0.187$, nor interaction including the instruction variable $b = -0.03$, $t = -0.04$, $p = 0.970$. The main effect of the projection coefficients persisted, $b = 1.95$, $t = 3.23$, $p = 0.001$.

**Discussion**

Across this series of analyses – and as predicted – target-related threat was demonstrated to be a predictor of counter-projection. Specifically, as perceptions of target-related threat increased, projection decreased, until people ultimately counter-projected when they perceived high threat. In fact, the *only* circumstance under which people counter-projected was when they perceived high target-related threat. The finding that variation in perceptions of target-related threat triggers counter-projection corroborates Denning and Hodges' (2022) previous results because they also found counter-projection to be dependent on an individual difference variable – how highly people identified with their political in-group.

However, the current studies also found that some of counter-projection's variance was explained by stereotyping, particularly when looking at counter-projection of personality traits (i.e., the BFI). Specifically, after stereotyping was accounted for in the model, high perceivers of threat no longer counter-projected – though they did not begin to project positively, either. When exploring counter-projection with everyday preferences (i.e., the ELI), results indicated that counter-projection was again explained by threat (i.e., people projected more as threat increased), and that counter-projection was still present after including stereotyping as a predictor. These differing results between the BFI and ELI replicated prior findings by Denning and Hodges (2022), who also found that differences in residual counter-projection (i.e., after accounting for stereotyping) are dependent on information type. These results were also consistent with predictions that there would be little effect of political stereotypes on results for the everyday preferences, with people still projecting the opposite of their preferences. This result was expected because people likely do not have *strong* political stereotypes, especially since the items were designed to be orthogonal to politics (Denning & Hodges, 2022), even though in the

current study, people did have some stereotypes of the outgroup for everyday preferences. As such, it appears that stereotyping may explain counter-projection if strong pre-existing stereotypes are available, but in their absence, people appear to project opposite preferences to create distance between themselves and threatening others. However, when strong pre-existing stereotypes *are* available, people use those instead of psychological distancing, indicating that some out-group stereotypes are opposite one's in-group. Of note, projection with the ELI was not entirely unaffected by stereotyping – participants who initially perceived an average level of threat projected positively after stereotyping was accounted for. However, participants who perceived high threat who initially counter-projected still counter-projected.

Counter-projection may not only contribute to polarizing people's perceptions of out-group members, but also by affecting people's behavioral intentions towards out-group members. Specifically, using both personality traits and everyday preferences, the results indicated that people who continued to counter-project after stereotyping was accounted for were the least willing to spend time having a conversation with the political out-group target. Consequently, even though residual counter-projection was not found overall with the BFI when looking at groupings of perceptions of threat, there were some individuals who showed counter-projection above and beyond stereotyping, and this effect appears to negatively impact intentions to socially interact with outgroup members.

Though the current study did not test the association between political polarization and counter-projection, the results are consistent with the idea that counter-projection would increase polarization between political groups by making people see the other side as even more different from themselves than members of the other side actually are. Further, given residual counter-projection has the ability to make people reluctant to talk with people from the other side, the

distance between polarized groups would be unlikely to change. Thus, one would hope a method to reduce counter-projection could be developed, particularly given the current highly politically polarized climate. However, in contrast to our predictions, counter-projection was not reduced via analogous perspective-taking instructions (and this was the case for both scales – the BFI and ELI – that were tested). Moreover, we found only mixed support for residual counter-projection being reduced by the analogous perspective-taking instructions. Specifically, only those in the control condition who perceived high threat continued to counter-project after stereotyping was accounted for. However, in this group, the slope for counter-projection was not significantly different from the flat line found for those in the analogous perspective-taking instruction condition who also perceived high threat, indicating whatever reduction in counter-projection that occurred for those in the analogous perspective-taking instruction condition was not significant. Considering the rest of the findings of Study 2B, these inconclusive findings regarding residual counter-projection do not lend promising support for analogous perspective taking as an intervention.

This result for analogous perspective taking in Study 2B conflicted with the results for analogous perspective taking in Study 1 (where counter-projection was reduced for the post-instruction target, but only when the target was male). The different pattern of results may stem from differences in study design. Specifically, Study 1 used a pre-post design (instead of the between-subjects design used in Study 2B) and the pre-post design may have led to a reduction in counter-projection because participants were comparing the post-instruction target to the pre-instruction target. In other words, participants in Study 1 may have found the post-instruction target less threatening than the pre-instruction target, independent of the analogous perspective-taking manipulation.

The possible "comparison effect" found in Study 1 may also have stemmed from a difference in the degree the targets appeared threatening. Specifically, counter-projection was only reduced by the analogous perspective-taking instructions when the post-intervention target was male. The male target in Study 1 was the same as the moderately threatening target (i.e., the warm target) used in Studies 2A and 2B – that is, except for being male. In Study 2A, this target was reported to be less threatening – though not significantly – than the competent target description, which was the same description as the female target in Study 1. Thus, it is possible that it wasn't the use of the analogous perspective-taking instructions that reduced counter-projection in Study 1, but the use of a less threatening target after a previously-shown more threatening target that reduced counter-projection (e.g., "this guy isn't as bad as that woman, I guess he's more like me then!"). The implications of this comparison effect on political polarization will be discussed further in the General Discussion.

# CHAPTER V

# GENERAL DISCUSSION

This series of studies explored the role of counter-projection in polarized contexts by answering four questions: (1) Is counter-projection driven by threat?; (2) Does analogous perspective-taking reduce counter-projection?; (3) Is the effect of analogous perspective-taking on counter-projection mediated by reducing threat?; and (4) When counter-projection persists after accounting for stereotyping, does threat predict the residual counter-projection variance? If so, can the residual counter-projection explained by threat be reduced by analogous perspective-taking? Though the resulting answers to these questions were not entirely expected, they were enlightening about counter-projection's role in polarized relations.

*Is counter-projection driven by threat?* Both Studies 1 and 2B explored whether counter-projection was driven by an individual's perception of threat. Study 1 measured baseline perceptions of threat triggered by the out-group (i.e., group-level threat), finding that as perceptions of group-level threat increased, so did counter-projection. However, people may vary in the degree they attribute threatening characteristics of the out-group, possibly affecting counter-projection. Thus, a measure of target-related threat was developed in Study 2A (i.e., a composite measure of realistic and symbolic threat based on Stephan et al.'s 2002 and 2016 work was adapted to mention the target and be measured after the target was introduced) and variations in counter-projection due to perceptions of target-related threat were explored in Study 2B. Study 2B produced similar findings to Study 1 – participants counter-projected when they perceived high target-related threat. Taken together, it appears counter-projection occurs when people perceive high threat, whether that threat is related to a target or stems from beliefs more generally about the participant's out-group.

Viewed through another lens, these results reveal that not everyone counter-projects with political out-group targets. For example, low and average perceivers of threat either did not counter-project or positively projected. This indicates people *can* see polarized out-group individuals as similar to themselves. Assuming polarization is related to counter-projection (though this was not directly tested in these experiments), it is promising that most people do not counter-project, particularly because it suggests that polarization is not inevitable in intergroup interactions (i.e., most people see the outgroup as polar opposite of themselves, but few extend this to individuals). This result may hint at the ability to reduce future polarization by focusing on individual relationships across political lines. Nevertheless, the results do indicate that there are times when people consistently see out-group individuals as the opposite of themselves. Specifically, it appears likely that counter-projection occurs when people either have more extreme political views (e.g., those who identify as right or left-wing may perceive more threat and may increasingly counter-project) or are exposed to more threatening political circumstances (e.g., everyone may counter-project when the target is highly threatening, regardless of whether they have moderate or more polarized political views). As variation in perceptions of threat (i.e., target-related threat) were manipulated by threatening circumstances (i.e., the manipulated target threat variable) in Study 2B, the two were unable to be parsed apart; However, future research should examine the individual effect of both on counter-projection.

*Does analogous perspective-taking reduce counter-projection?* The analogous perspective-taking results were not conclusive and varied across studies. Specifically, the results found in the pre-post design study (Study 1) were not consistent with those found in Study 2B, which had a between-subjects design. Study 1 found that people counter-projected to the target they saw prior to the analogous perspective-taking instructions, but not with the target they saw

after. However, after exploring if there was an effect of target order, it appeared the instructions only reduced counter-projection for one target – when the male target who was described in a warm manner was shown after the female target who was described in a competent manner (i.e., when the reverse order occurred, counter-projection was unaffected by the instructions). Study 2B did not find an effect of the analogous perspective-taking instructions on counter-projection (though a small, but inconclusive effect was found with residual counter-projection), even though the moderately threatening (female) target in Study 2B was described in the same manner as the male target from Study 1. The differing results suggest that study design (i.e., pre-post versus between-subjects) may have contributed to the inconsistent findings. However, given that other subtle factors were varied in Study 1 (i.e., the order in which the targets were presented; how the targets were described), it is also possible the analogous perspective-taking instructions only work when targets appear deviant from their in-group (as the male target did in Study 1). These explanations, along with others, will be explored more later in this chapter.

*Is the effect of analogous perspective-taking on counter-projection mediated by reducing threat?* This question was only explored in Study 2B, where the instructions were found not to influence counter-projection. As such, there was not any support for a mediation. However, there was support for an indirect effect of analogous perspective-taking instructions through perceptions of target-related threat (though this effect was small).

Regardless of the lack of mediation, the theory behind this question may still apply to other interventions. Specifically, increased perceptions of threat were found to increase counter-projection, while those who perceived less threat did not counter-project. As such, it appears that reducing threat might reduce counter-projection, making a future intervention that reduces counter-projection likely to be mediated by perceptions in threat.

*When counter-projection persists after accounting for stereotyping, does threat predict the residual counter-projection variance?* Study 2B explored the question of residual counter-projection – the occurrence, after accounting for stereotyping, of people continuing to attribute their opposite traits, beliefs, and preferences to the out-group target. Results indicated that stereotyping entirely accounted for counter-projection of personality traits, a result consistent with findings by Denning and Hodges (2022), who also found idiographic stereotyping explained most counter-projection of personality traits. In other words, people who perceived high threat were using a stereotype that their out-group was their opposite when responding about the target individual's personality traits. However, stereotyping did not explain all of people's counter-projection of everyday preferences. This finding supported predictions that counter-projection with everyday preferences would be unaffected by stereotyping. This result was expected because the items used to measure everyday preferences (i.e., the everyday likes items, or ELI; Denning & Hodges, 2022) were designed to be orthogonal to politics, meaning people should not have strong previously established stereotypes of everyday preferences for their political out-group. This result also replicated prior work by Denning and Hodges (2022), who found stereotyping to explain counter-projection for personality traits but not everyday preferences.

There was also a second component to the fourth research question: *Can the residual counter-projection explained by threat be reduced by analogous perspective-taking?* Since residual counter-projection was not found with personality traits, this question was only explored with everyday preferences, and the results were inconclusive. Specifically, those who exhibited residual counter-projection with everyday preferences – high perceivers of threat – were only found to counter-project (i.e., have a negative slope) in the control condition. High perceivers of threat in the analogous perspective-taking instruction condition did not exhibit residual counter-

projection (nor did they positively project). As such, the analogous perspective-taking instruction appeared to reduce the extent high perceivers of threat saw their out-group as their opposite. Yet, this reduction did not produce results that were significantly different from the residual counter-projection exhibited by the high perceivers in the control condition. Moreover, analogous perspective taking did not reduce the predictive effect of counter-projection or residual counter-projection on the behavioral intention outcome – how long people were willing to spend in a conversation with the target. When considering the other lackluster findings regarding analogous perspective taking in Study 2B, these results provide further unfavorable support for analogous perspective taking as a standalone intervention.

**Counter-projection can be reliably manipulated**

Let's take a moment to consider a common scenario: Two strangers are sitting next to each other on a plane, having a conversation. The conversation is pleasant enough, but then one of them makes a joke about politics. From this joke – and the other's reflexive reaction – it is clear they align with opposing political parties. Does the knowledge the other person is an out-group member change how they are going to *perceive* each other? Based on accumulating evidence in support of counter-projection (e.g., this dissertation; Denning & Hodges, 2022), the answer would be "yes" – people appear to switch from projection to counter-projection depending on the target's association with a contentious or threatening out-group. However, this claim differs from previous literature on *projection* (i.e., when people positively correlate self and other characteristics), which argued projection was a default process employed with everyone, because it was cognitively less effortful to see people as like oneself (Davis, 2017; Krueger, 2007). This belief led many past findings of counter-projection to be regarded as flukes or surprises (e.g., Davis, 2017; Machunsky et al., 2014), instead of seeing counter-projection as a

reliable social cognitive phenomenon. Over time, instances of counter-projection consistently appeared, suggesting there may be more substance to counter-projection than previously given merit (see Denning & Hodges, 2022 for review).

When combined with past findings of counter-projection, the current results (that counter-projection was triggered when people perceived high threat, both when threat was target-related and at the group-level) reveal an underlying trend – targets need to be contentious or threatening to evoke counter-projection. For instance, counter-projection was previously found with targets who were idiographically disliked (Davis, 2017), negatively valenced (Machunsky et al., 2014), members of political out-groups (Castelli et al., 2009; Denning & Hodges, 2022), and on opposite political sides of a specific issue (Gromet & Van Boven, 2020; Mullen et al., 1992). Counter-projection was also found when people were highly identified with their in-group (Denning & Hodges, 2022; Riketta, 2005), which was replicated in Study 1 of this dissertation. Finally, counter-projection was found exclusively with out-group targets – people did not counter-project to their in-group (Denning & Hodges, 2022).

Past theoretical explanations for counter-projection also point to threat as a trigger, arguing that perceivers counter-project to create distance between themselves and others who threaten their – the perceivers' – sense of self (Cadinu & Rothbart, 1996; Clement & Krueger, 2002), often until that distance places the perceiver and target at opposite ends. In other words, counter-projection may serve as a protective mechanism for the self. In terms of the current findings, it is likely that those who perceived their out-group (or a specific out-group target) as threatening were specifically experiencing threats to their sense of self. This theory is consistent with work on shared reality, or lack thereof (Hodges et al., 2018). Specifically, Hodges and colleagues theorized that perceptions of threat prevent people from perceiving when others share

their same viewpoint. As such, people may counter-project to maintain the distinctly separate realities that threat prompts.

Together it appears that accentuated opposing views (i.e., voting differently) make conditions ripe for regarding someone as a threat to the self. This may explain how past contexts (e.g., targets that disagree about specific political issues; Gromet & Van Boven, 2020) and individual differences (e.g., high in-group identification; Denning & Hodges, 2022; Riketta, 2005) led to counter-projection, as these past studies highlighted group differences that likely posed a threat to the self. One consequence of these results is that there may be additional circumstances in which people counter-project, so long as there is a difference they find threatening (e.g., religious differences). Yet, as previously mentioned, not everyone is prone to finding circumstances threatening enough to counter-project. Thus, the current studies demonstrated that counter-projection is a phenomenon that can be consistently predicted to occur, but one that is not found in everyone or in all circumstances. As such, though the current findings expand the projection literature by demonstrating that counter-projection can be manipulated, these results also reinforce previous findings that (positive) projection is more the default process of the two.

**Stereotyping – explanation or mechanism?**

Returning to the example on the plane: When the two out-group strangers are determining how to perceive one another, projection and counter-projection are not the only two social cognitive processes available. Instead, the strangers may use the out-group's stereotype (Park & Rothbart, 1982), assuming their neighbor on the plane possesses stereotypical characteristics of that group. However, if people employ oppositional stereotypes (i.e.,

stereotypes that the out-group is opposite them) to create psychological distance, *are they still counter-projecting?* Study 2B was conducted in part to answer this question.

In replicating prior findings by Denning and Hodges (2022), Study 2B demonstrated that there are circumstances in which stereotyping explains counter-projection. Specifically, stereotyping explained when high perceivers of target-related threat counter-projected personality traits. The finding that people have stereotypes of their political out-group's personality traits is consistent with past work (Gerber et al., 2011). Given that responses for both the target and the stereotypical out-group member overlap, it is possible oppositional stereotyping entirely explains counter-projection: People have pre-existing stereotypes of the out-group that are the opposite of themselves, and they use those stereotypes when responding about the target individual.

Stereotyping may even be extended to (positive) projection, with people using stereotypes that their out-group shares similar traits to them instead of using the self as their basis for positive projection. Let's unpack this by considering a different example – the stereotype of people who like to cook. If you are a baker, your stereotype of cooks may overlap with your own experience because the groups are similar. Thus, if you are interacting with someone who likes to cook and you attribute self traits to them, those traits may be indiscernible from stereotypes. In other words, projection and stereotyping would overlap, and it is unclear which one is being applied in perceiving another person (i.e., which comes first?). Unlike counter-projection, the possible overlap of (positive) projection and in-group stereotyping is mostly innocuous (e.g., seeing others as similar probably leads to positive social outcomes), making it previously unnecessary to distinguish between the two. Yet, their potential overlap changes the interpretation of projection by operationalizing it as a possible result of stereotyping. Through

the lens of stereotyping, both projection and counter-projection are explained by cognitive heuristics – people employ their generalizations about a target's associated group to form quick perceptions about the relationship between the self and the target.

Yet, people may *want* to see themselves more like a positive target, or *desire* to put distance between themselves and a threatening target. These affective responses imply people are *motivated* to project or counter-project. Given that counter-projection with everyday preferences was unaffected by stereotyping in Study 2B, this motivational lens has some credence: The items on the everyday likes measure were designed to be orthogonal to political stereotypes (i.e., people likely did not have strong pre-existing stereotypes available to employ; Denning & Hodges, 2022), meaning people may have been motivated to counter-project as an alternative method to create distance when using cognitive heuristics (stereotyping) was not enough.

In this lens, counter-projection is both cognitive and affective: when readily available oppositional stereotypes exist, people employ those stereotypes to create the necessary distance between themselves and the out-group target because it is cognitively less effortful. Thus, stereotyping is a *mechanism to counter-projection,* as distancing is still achieved. However, when strong oppositional stereotypes do not exist, people may still be motivated to create distance – they use themselves as an anchor and then distance away from the threatening target until an ideal distance has been achieved, thus counter-projecting (Machunsky et al., 2014). As such, stereotyping may be a cognitive mechanism behind distancing, engaged to reduce effortful processing when there is a strong and pre-existing stereotype available. However, counter-projection may still occur to create the *desired* distance between the self and a threatening other when only weak stereotypes about the out-group exist (or possibly no stereotypes about the out-

group exist). The view that both cognitive and motivational processes could be at play is consistent with Machunsky and colleagues' (2014) explanation of counter-projection.

However, the current studies may have underestimated residual counter-projection, meaning it is possible residual counter-projection occurs more often in more ecologically valid circumstances. For instance, the targets used were often described using impoverished descriptions that largely relied on group stereotypes (i.e., voting preference). For the most threatening target in Study 2B, the entire description was centered on Jen's voting preference and political desires. These descriptions may have led participants to employ stereotypes in their perceptions more than they would have otherwise – largely because they had little other information to go on.

To address this point, an exploratory analysis was conducted examining the extent participants' idiographic stereotype on each ELI item differed from the average stereotype for that ELI item (i.e., the averaged response across all participants for the stereotypical out-group member in Study 2B for each item). Results indicated that residual counter-projection for those who perceived high target-related threat (i.e., the result previously found with the ELI) increased as people's idiographic stereotypes became increasingly disparate from the average (results available in the "supplemental_analyses.Rmd" for Study 2B on Github, link available in Appendix C). As such, it appears that people counter-projected with the ELI above and beyond stereotyping because they did not have a consistent stereotype of the out-group. It is possible that, had there been targets in the current study described with information that did not trigger established out-group stereotypes, more residual counter-projection may have occurred.

Likewise, residual counter-projection may occur more among some people, particularly those who experience increased threat daily. For instance, minority groups for whom a 2024

Trump Presidency would pose a potential threat to their rights or well-being (e.g., increased violence towards black and brown individuals; decreased voting accessibility) may be more likely to counter-project above and beyond stereotyping because the threat they experience is so great they need to distance themselves further than their out-group stereotypes allow. This is also possible with female participants (e.g., women may experience more threat due to lack of control over their bodies, through changed abortion laws), though additional exploratory analyses examining gender of participants in the current studies demonstrated that gender did not moderate the effect of target-related threat on counter-projection or residual counter-projection (with either personality traits or everyday preferences; available in the "supplemental_analyses.Rmd" on Github for Study 2B, available in Appendix C).

The effect of intersectionality (combinations of personal identities) on counter-projection may be more prominent in in-person interactions. For instance, interacting face to face with an out-group target may pose a *greater* threat than someone participants have never met and never will meet (e.g., an impoverished online survey target). On the other hand, participants may be more likely to find common ground with someone they meet in-person, leading them to *individuate* the target and see them as more willing to compromise (see Niella, 2020). Consequently, participants may not feel the need to distance themselves as much, if at all, from in-person targets. Future research should explore the extent people counter-project with targets introduced through a variety of different mediums, including video and in-person conversations.

These theories regarding counter-projection should be considered along with the limitations of the current studies. First, it is worth noting that the data for everyday preferences (i.e., the ELI) had some unexplainable characteristics (i.e., everyday likes items had zero variance with the random intercept). However, the data appeared normal in other regards (e.g.,

normality) and the findings mostly replicated those by Denning and Hodges (2022), thus providing more confidence in these results and their implications. Yet, the ELI items dealt with largely trivial topics, so the extent the results generalize to other circumstances and information types is up for debate. For instance, it is likely people do not often consider their out-group's everyday preferences (which is why they did not have strong pre-existing stereotypes about these preferences to explain counter-projection), which may have made the task of answering questions about these preferences somewhat unnatural for them. For this result to generalize and be meaningful, there would need to be other types of information that people do not have pre-existing stereotypes about *but* still use regularly when perceiving others. However, using a piece of information regularly enough may lead to the development of a stereotype, making the results with everyday preferences possibly artificial.

Second, these studies were not designed to explicitly test differences between counter-projection and stereotyping. Future research should parse apart the two processes, including exploring the relationship between (positive) projection and in-group stereotyping.

Third, the studies in this dissertation as well as those in Denning and Hodges' (2022) article used predominantly liberal samples (e.g., those who voted for Biden in 2020). The use of liberal samples is both a pro and con of this experimental design. Specifically, using liberal samples allowed for replication of results found with the Study 1 sample; added knowledge to the field of political threat about when liberals experience threat (i.e., in situations that threaten political loss); and demonstrated that the Intergroup Threat Theory items (Stephan's et al., 2002, 2016) can measure target-related threat. However, this limitation prevents knowledge about whether residual counter-projection is exclusive to *liberals* who experience high threat or also occurs with conservatives.

Though recent work has demonstrated that liberals and conservatives both experience political threat equally (Brandt et al., 2014; Brandt et al., 2021; Crawford, 2017), the types of circumstances that prompt their perceptions of threat differ (e.g., conservatives are more threatened by threats to safety – such as war – while liberals are more threatened by threats to healthcare and poverty; Brandt et al., 2021; Crawford, 2017; Eadeh & Chang, 2020). It seems likely that with issues that are threatening to conservatives, residual counter-projection would still be found. However, it is possible that there are differences between liberals and conservatives when it comes to counter-projection. For instance, conservatives may use stereotypes more, thereby showing less support for a motivational account of counter-projection. Future studies should explore differences in counter-projection between liberals and conservatives.

**The meaning of counter-projection**

Determining whether counter-projection is explained by stereotyping may be moot if the effect of counter-projection on real-world behaviors is minimal. When considering the meaningfulness of counter-projection at both a societal and interpersonal level, it is important to first consider how counter-projection and polarization may be related. Though counter-projection was not directly tied to political polarization in these studies, perceiving one's out-group as opposite oneself likely fans the fires of polarization. In some ways, the relationship of polarization and counter-projection poses a "chicken and an egg" conundrum – do people first perceive distance and then become polarized, or does polarization lead to distancing? Though this was not explored in the current studies, we suggest that polarization likely occurs first, but that counter-projection *reinforces* it. Specifically, there presumably is a circular pattern in which perceived polarization from out-group members causes the out-group to appear increasingly

threatening, ultimately increasing counter-projection and its negative consequences, which in turn leads to more polarization and threat.

This cyclical relationship poses a question for how to design an intervention to reduce polarization: *Should interventions focus on the psychological mechanisms that reinforce polarization (e.g., counter-projection) or the perceptions of polarization in general? Or do reductions in both occur simultaneously?* For instance, if we help people realize their beliefs are not as polarized as they previously thought, it is likely people would counter-project less as well. Within studies examining a reduction in mechanisms thought to support polarization (such as counter-projection), participants may need to perceive reduced polarization for changes in the supporting mechanisms to also occur. Future research should explore if interventions designed to reduce polarization also reduce counter-projection, and vice versa.

When considering counter-projection's effect on downstream effects, it should also be considered whether counter-projection's effect size on real-world effects is meaningful. Study 2B demonstrated that residual counter-projection has implications on behavioral intentions – people's willingness to have a conversation with the out-group target – with those who exhibited residual counter-projection being the least willing to have a conversation with the out-group target. This effect may extend to other downstream behaviors, as those who are motivated to create psychological distance probably also engage in behaviors that promote *physical* distancing (e.g., they may live in politically segregated areas, thus, surrounding themselves with more like-minded individuals). However, is this self-reported reduction in willingness to have a conversation with an out-group individual meaningful, especially when it was only found to be predicted by residual counter-projection using everyday preferences? Counter-projection of everyday preferences in a survey may lead to negative intentions, but the actual outcomes

associated with these trivial differences may be easy to overcome in the real-world (e.g., our difference in preference for cats versus dogs may not stand in the way of working together on bi-partisan policies). People may also *say* they are less willing to have a conversation but *behave* differently in-person, possibly out of a desire not to appear rude or biased.

Yet, if differences on mundane issues like everyday preferences indicate an inclination to disagree about more important issues (issues that we, as a society, may need to agree on) *and* these differences affect in-person behavior, counter-projection may contribute to meaningful harm. For example, people often agree that education should be improved for children and that opioid overdoses need to be decreased. However, views on how to accomplish these tasks may be polarized when the individual overseeing the task is associated with one political group, even if political differences are unrelated to the solutions being suggested. Distancing in response to this person's threatening political association may impede progress on issues everyone would otherwise agree upon. This example demonstrates that *any* negative real-world social outcome resulting from counter-projection is potentially problematic.

Counter-projection's effects – and by extension, interventions to reduce counter-projection – could also be considered increasingly impactful if they last *longitudinally*. Though not examined in the current studies, if an intervention leads people to counter-project less with their out-group over time and this increases willingness to interact across the aisle, that would be more meaningful than one singular instance of reduced counter-projection that may only momentarily improve prosocial behavior. In addition to considerations of an intervention, the concept of longitudinal counter-projection is worthwhile to consider by itself, especially in the extent it occurs naturally. For instance, if one perceives their neighbor as politically threatening based on one conversation, they may counter-project about a variety of unrelated issues for

months afterward (e.g., “My neighbor does not take care of their yard because they do not care about our community, which I care a lot about!”). Longitudinal occurrences of counter-projection may also extend outside of politics when people experience anxiety, mental health issues, or other disadvantages. Specifically, when people experience high anxiety or trauma, it may be a luxury (that they do not have) to see others as similar. Instead, they may perceive everyone as posing some sort of threat and, thus, they may counter-project consistently with multiple different types of people. Future studies should explore variations in the pervasiveness of counter-projection among individuals and whether individuals most affected can be helped using a long-term intervention.

With this repeated discussion of an “intervention,” it is important to consider whether a reduction in counter-projection is needed to improve social outcomes. For instance, it may not matter that a person distances themselves from an out-group target, if they are still able to have a cordial and productive conversation with that out-group individual. Because residual counter-projection was found to lead to more negative behavioral intentions in Study 2B, positive social outcomes seem less likely to occur while people continue to counter-project. However, Study 2B explored only one behavioral intention and did not employ an in-person paradigm, where participants could behave differently. Future research should explore the extent to which positive social outcomes can occur when people counter-project.

It is also necessary to consider how much of a reduction in counter-projection is necessary to justify an intervention. This can be considered statistically when considering the slope of counter-projection, as simply achieving a flat line (i.e., neither projecting nor counter-projecting) may not be enough to justify an intervention. Though people are no longer counter-projecting, they still may not perceive the out-group target as similar to themselves. Justifications

for an intervention could also be operationalized through finding changes in social outcomes caused by reduced counter-projection that impactfully improve people's lives. For instance, if an intervention reduces counter-projection and leads people to increasingly work across the aisle on a community issue (e.g., children's education), the improved outcome would indicate support that the intervention led to meaningful change. Future studies should explore if interventions can be found that lead to meaningful effects tied to reduced counter-projection.

**The inconsistent intervention**

Unfortunately, the intervention tested in these studies – analogous perspective taking –did not provide promising results for a meaningful intervention. Study 1 – that used a pre-post design – found analogous perspective taking reduced counter-projection, but Study 2B – that employed a between-subjects design – found no effect on counter-projection (i.e., before stereotyping was accounted for; there was an inconclusive effect on residual counter-projection, but this effect did not translate to changes in behavioral intentions). These inconsistent findings are especially striking because some target descriptions that were used in both studies were nearly the same. Specifically, Study 1 found the instructions to be effective only when the post-instruction target was less threatening (i.e., the male target described in a warm manner) than the pre-instruction target (i.e., the female target described in a competent manner). However, when the same less threatening target description was employed in the between-subjects study (although this time that target was female), there was no effect. As such, it is possible differences in results stemmed from study design. For instance, the pre-post design may have reduced counter-projection in Study 1 due to participants' comparing the less threatening target (i.e., the male target described with traditional female gender roles) to the more threatening target. In comparing the two targets, participants may have chosen to extend more similar traits to the less threatening target (e.g., the

second target may not have seemed as "bad" in comparison, and maybe even more like them than they initially thought). As such, contrasting comparison as an explanation for the difference in results should be explored in future research. One possible direction would be to replicate Study 1 but add a "control" comparison condition (i.e., a condition with no instructions between the first and second targets). If the same results are found, it would suggest that analogous perspective taking did not drive the reduction in counter-projection. Regardless, an "intervention" based on comparison is likely not a fruitful future direction – though it may reduce counter-projection with one target, it could reinforce differences with more threatening others.

The inconsistency in results between Study 1 and Study 2B may have implications for prior analogous perspective-taking studies, particularly those whose methodologies employed multi-pronged interventions (e.g., Brockman & Kalla, 2016). Specifically, Brockman and Kalla (2016) employed an *in-person* canvassing study where canvassers not only asked participants to discuss an analogous personal scenario, but also engaged in *non-judgmental listening* while the participants responded and provided their own personal experiences as means to connect them with the target person's experience – known as *narrative persuasion*. Though the authors attempted to parse apart the different effects of non-judgmental listening and narrative persuasion in follow-up experiments, they did not isolate the effect of analogous perspective taking (Kalla & Brockman, 2020). Since the results of Study 2B did not find an effect of analogous perspective taking when other possibly confounding explanations were removed (e.g., comparison in the pre-post design), it is also possible that analogous perspective taking may not have been the driving factor in Brockman and Kalla's (2016; Kalla & Brockman, 2020) work.

Before entirely writing-off analogous perspective taking, it is also necessary to consider that analogous perspective taking *is* a social process that occurs and *is* an effective intervention but prompting it to occur requires different instructions than the current studies employed. For instance, most previous canvassing studies – such as Brockman and Kalla (2016) – introduced the target of analogous perspective taking at the beginning of the conversation. Though the analogous perspective taking task has participants consider a situation that happened to them, they are directed to consider a situation similar to the target individual's circumstance. By already having an idea of the target they are connecting their experience with, participants connect the self to the target – initiating perspective taking – consequently perceiving more similarity and extending more understanding towards the target. Without a relevant target or target-group in mind, participants may not perspective take at all (i.e., they are not considering how their experience connects with someone else's).

This explanation is consistent with a study by Saveski and colleagues (2021), who asked participants to think about a time they and a friend had a disagreement (i.e., analogous perspective-taking) every time they saw a Twitter post they disagreed with on a politically opposite Twitter feed. Participants knew the target was the political out-group owner of the twitter feed and were able to think about the target in light of considering past disagreements with a friend. The findings by Saveski and colleagues may also have been bolstered by the extension of positive traits associated with their friend onto the target (most participants in the current studies wrote about acquaintances or strangers). Further, Saveski and colleagues' instructions prompted participants to analogously perspective-take *every time* they disagreed. By participating in analogous perspective taking multiple times throughout the study, the instructions may have had a cumulative effect.

The need for the target to have been introduced to participants before analogously perspective taking could also explain why the instructions were effective in Study 1 but not Study 2B. In Study 1, participants were asked questions about political identification and shown an out-group target *before* the analogous perspective-taking instructions were introduced. As such, they may have been primed to think about a time they felt misunderstood politically, making them more understanding of a political out-group target. It may be possible to examine this explanation in the future by coding the responses of participants in both Studies 1 and 2B to see if those who wrote about a relevant situation (i.e., a time they felt politically misunderstood) were less likely to counter-project than those who wrote about non-analogous circumstances. Further coding should also be conducted to verify the content validity of the intervention, particularly focusing on how successfully participants wrote about a time they felt misunderstood (e.g., did they write about a personal situation with another individual or just general feelings of misunderstanding?).

An alternative explanation for the results in Study 1 may have stemmed from the male target being perceived as *deviant* from his in-group. This deviance may have created the opportune circumstance for the analogous perspective-taking instructions to be effective. For instance, participants may have expected a male Trump supporter to be typically "masculine" (e.g., the breadwinner who has a wife who does the cooking, and thus not interested in family-focused hobbies and career paths, as the male target was). When the target was not consistent with participants' stereotype of a male Trump supporter, participants may have perceived that he "wasn't like other" Trump supporters. This perception of the target as deviant from their stereotype may have been necessary for analogous perspective taking to be effective, as deviant

targets are probably more likely to be perceived as politically misunderstood than targets who exhibit stereotype-typical characteristics.

This deviance explanation may also explain the lack of effect with the female targets in both Studies 1 and 2B. In Study 1, the female target may have been described too stereotypically for a Trump supporter (i.e., she was a business manager, and Trump was known for support of businesses) for her to be deviant. In Study 2B, the most threatening target was described by her stereotypical voting preference, while the moderately threatening target was a *female* target described *with traditional female* characteristics. Though described in a similar manner to the male target in Study 1, the fact that the female target now was consistent with traditional female characteristics likely made her appear stereotype consistent (e.g., the stereotype of a women who supports Trump may be that she is a homemaker). As such, it is possible the results of Study 2B do not indicate that analogous perspective taking is ineffective in general, but that the intervention does not work with targets whose descriptions are minimal and largely based on stereotypical content.

**The success of the threat measure**

The wide variety of measures of threat employed in Studies 2A and 2B (e.g., explicit threat, target-related direct-but-not explicit threat, and indirect measures of physical distancing) demonstrate there are several different methods to measure threat. By employing a measure of threat consisting only of symbolic and realistic threat, the measure in Studies 2A and 2B was very specific – it measured how much people perceived the target and the target's related group as posing *harm* to them and/or their way of life. This definition makes harm critical: participants must perceive intended harm on behalf of the target or group to perceive threat. If measures other

than symbolic and realistic threat (i.e., the target-related direct-but-not explicit threat items) were used in the present studies, threat may have operated differently.

However, the physical distancing measures were not related to one another nor were they related to any other measure of threat. This raises a question of ecological validity for the physical distancing measures. For instance, it is possible participants reacted to this measure because they guessed the purpose of the question (i.e., that we expected them to want to sit further from their political out-group target)[26]. This interpretation is consistent with additional exploratory follow-up analyses that the explicit threat items in Study 2B did not predict counter-projection (see "supplemental_analyses.Rmd" for Study 2B on Github in Appendix C).[27] It appears that participants may have responded in a reactionary manner to questions asking explicitly about "threat" and physical distancing, prompting them to answer inconsistently with their possible actual perceptions of threat (e.g., those seen on their responses on symbolic and realistic threat in the target-related threat measure). This reactionary interpretation is reinforced by the fact that group-level threat measure had a different pattern of results in Study 1 when it was collected before the target was introduced.

Additionally, participants may have had different responses to the two measures of physical distancing. Some researchers have found people *perceive less* distance between themselves and a threatening target to keep the threat close (e.g., "Keep your enemies closer"; e.g., Xiao & Van Bavel, 2012). This differs from other studies that find people *create more* distance (e.g., Cook & Cottrell, 2021; Wyner & Calvini, 2011) between themselves and a

[26] The responses on this measure were not normally distributed nor were they skewed but showed more of a plateau.

[27] Participants responses on the explicit target-related threat item in Study 2B were positively skewed, indicating they may have lied about their responses to appear less biased. This may have explained the lack of counter-projection.

threatening target when given the opportunity. These discrepant results suggest there is a difference between measures based on whether the distance already exists or if participants are creating it, which is consistent with the current two measures: the distance already existed in the distance in town measure (where the least distance was perceived from the highly threatening target who wanted participants' candidate to lose the upcoming election), but participants created the distance in the coffee shop measure (where the most distance was reported with this same target who wanted their candidate to lose the upcoming election). However, the responses on the physical distancing measure in town were not entirely consistent with this theory: participants perceived the most distance with the warm target, not the control target. Though participants did appear to keep the most threatening target close, they should have allowed the least threatening target (the non-political control target) to be the farthest away. As such, the possibility that participants were reactive and their responses were not valid is still likely, and this explanation may extend to other measures of physical distancing employed in survey studies (e.g., participants' responses in these past studies may have stemmed from biased reactions as well).

Yet, this isn't to say physical distancing measures for threat never work. Studies using in-person paradigms (e.g., Wyner & Calvini, 2011) may be more successful because participants are less likely to realize what is being asked of them. For instance, if there are multiple places for a research participant to sit in a room where a study is being conducted, the experimenter doesn't have to ask participants where they would sit to measure physical distancing (a move that may prompt reactance or bias). The experimenter could simply watch where the participant sits and measure how far that was from the threatening target, all without the participant's knowledge that this information was being collected. In this case, placing physical distance between the self and another person may be a more valid measure of threat.

When discussing the measures of threat employed in Study 2B, it is important to note that threat was both manipulated by and measured using target conditions. These target conditions were used to ensure more variability in perceived target-related threat (with no variability in threat, it may not have been possible to see threat's effect on counter-projection). However, this use of a manipulated and a measured variable both tapping the same construct in one study is unusual. It also raises the question of whether the target-related threat measure captures any individual differences related to threat, or if people's scores on this measure entirely followed the pattern of threat created by the manipulated threat variable (i.e., the different targets). Results using target condition as the predictor instead of target-related threat were almost identical, indicating individual differences in target-related threat are likely not highly consequential.

**Final conclusions**

Before we conclude, let us zoom in on counter-projection's relationship with political polarization. Depending on which results are emphasized, the implications of counter-projection on political polarization can be interpreted as either disheartening *or* optimistic. On one hand, people counter-project the opposite of their traits and everyday preferences, which indicates people not only see others as their *political* opposites, but their opposites in mundane everyday characteristics as well. These results illustrate the depressing pervasiveness with which polarization has crept into some people's daily lives. Yet, on the other hand, there is also reason to hope. The fact that counter-projection to out-group targets was not always present suggests that polarization may be reduced. Counter-projection was only consistently found in specific circumstances (i.e., when people perceive high threat) with a somewhat trivial type of information, implying many people do not see their political out-group as significantly differing from themselves. This latter finding is in contrast with the portrayal of extreme polarization

presented in the news and on social media. Thus, knowing when people do and don't counter-project could help reduce polarization by showing people they aren't as different from another as they thought.

## APPENDIX A

## TRAIT MEASURES

*Instructions for self:* Please select a number for each statement to indicate the extent to which you agree or disagree. Please use the scale in the table below to respond to the statements.

*Instructions for target/stereotypical target group:* Please select a number for each statement to indicate the extent to which you think [Jen/ those who supported Trump in the 2020 U.S. Presidential Election] agree[s] or disagree[s] with that statement. Please use the scale in the table below to respond to the statements.

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Disagree strongly | Disagree a little | Neutral | Agree a little | Agree strongly |

Each statement begins with the phrase:

***[I am/Jen is] someone who… OR People who supported Trump in the 2020 U.S. Presidential Election…***

1. Tend[s] to be quiet.
2. [Is/are] compassionate, has a soft heart.
3. Tend[s] to be disorganized.
4. [Worry/Worries] a lot.
5. [Is/Are] fascinated by art, music, or literature.
6. [Is/Are] dominant, acts as a leader.
7. [Is/Are] sometimes rude to others.
8. [Has/Have] difficulty getting started on tasks.
9. Tend[s] to feel depressed, blue.
10. [Has/Have] little interest in abstract ideas.
11. [Is/Are] full of energy.
12. Assume[s] the best about people.
13. [Is/Are] reliable, can always be counted on.
14. [Is/Are] emotionally stable, not easily upset.
15. [Is/Are] original, comes up with new ideas.
16. [Is/Are] not good at deceiving other people.*
17. Take[s] risks that could cause trouble for [me/her/him/them].*
18. Use[s] others for [my/her/his/their] own ends.*
19. Stick[s] to the rules.*

*Sixth factor of personality – honesty – in addition to the BFI2-XS.

## APPENDIX B

## TARGET IMAGES

*Image 1.* Mike from Study 1.

*Image 2.* Jen from Study 1 and Studies 2A and 2B

## APPENDIX C

## GITHUB AND OSF LINKS

This appendix contains all external links referenced in this proposal. OSF refers to the Open Science Foundation, where analysis plans were pre-registered for Study 1 and Study 2B. Study 2A was a pilot study and, thus, considered exploratory. Github houses the data and code for all analyses, making them reproducible.

1. Link to OSF pre-registration for Study 1:

   https://osf.io/mkwfu/?view_only=e9f6827259cb448080a16727b4f81251

2. Github link for Study 1:

   https://github.com/kdenning/analogous-perspective-taking-2

3. Github link for Studies 2A and 2B power analysis and simulations:

   https://github.com/kdenning/dissertation_simulations

4. Github link for Study 2A (labeled as "pilot" on Github):

   https://github.com/kdenning/diss_threat_pilot

5. Github link for Study 2B (labeled as "main" study on Github):

   https://github.com/kdenning/diss_main_study

## APPENDIX D

## TARGET THREAT MANIPULATION DESCRIPTIONS

The target descriptions accompanied an image ("Jen" in Appendix B). They are listed in order of what was expected to be the most threatening to least threatening (i.e., non-political control target condition).

- *Political loss target condition*: "Jen wants liberals to lose the Presidency and the House in 2024. She hopes President Trump will run again, so that she can vote for him again."
- *Competent target condition*: "Jen enjoys running, works as a business manager, and voted for Trump in the 2020 U.S. Presidential Election."
- *COVID target condition*: "Jen does not believe in getting the Covid-19 vaccine. She thinks the severity of Covid-19 has been exaggerated to push vaccines for a liberal agenda, at the cost of personal freedom."
- *Warm target condition*: Jen enjoys cooking, works as a high school teacher, and voted for Trump in the 2020 U.S. Presidential Election."
- *Non-political control target condition* "Jen likes to cook and works as a high-school teacher."

## APPENDIX E

## COFFEE SHOP DISTANCE MEASURE

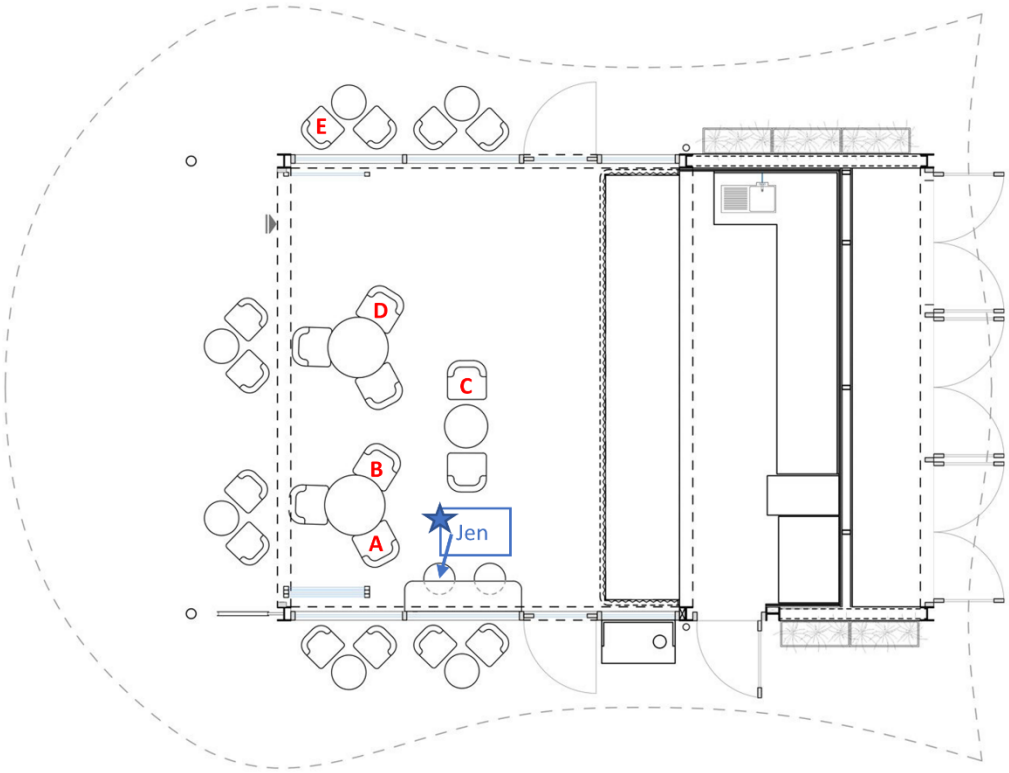



*Image 3.* Image used in one of the measures of threat measuring physical distance in Study 2, in which participants report where they would sit in relation to the target in a coffee shop. Image from Ravenscroft (2019).

## APPENDIX F

## EVERYDAY LIKES ITEMS

1. How do you/[stereotypical out-group or target individual] feel about stand-up comedy shows:
    a. 1 = Dislike a great deal
    b. 2 = Dislike a moderate amount
    c. 3 = Neither like nor dislike
    d. 4 = Like a moderate amount
    e. 5 = Like a great deal

2. How do you/[stereotypical out-group or target individual] feel about condiments (e.g., ketchup, mayo, mustard, pickles)?
    a. 1 = Dislike a great deal
    b. 2 = Dislike a moderate amount
    c. 3 = Neither like nor dislike
    d. 4 = Like a moderate amount
    e. 5 = Like a great deal

3. Do you/[stereotypical out-group or target individual] prefer dogs or cats?
    a. 1 = Greatly prefer dogs
    b. 2 = Moderately prefer dogs to cats
    c. 3 = No preference
    d. 4 = Moderately prefer cats to dogs
    e. 5 = Greatly prefer cats to dogs

4. Do you/[stereotypical out-group or target individual] prefer Winter versus Summer?
    a. 1 = Greatly prefer Winter to Summer
    b. 2 =Moderately prefer Winter to Summer
    c. 3 = No preference

d. 4 = Moderately prefer Summer to Winter

e. 5 = Greatly prefer Summer to Winter

5. How much do you/[ stereotypical out-group or target individual] like the color yellow?

a. 1 = Dislike a great deal

b. 2 = Dislike a moderate amount

c. 3 = Neither like nor dislike

d. 4 = Like a moderate amount

e. 5 = Like a great deal

6. Do you [stereotypical out-group or target individual] like pineapple as a pizza topping?

a. 1 = Dislike a great deal

b. 2 = Dislike a moderate amount

c. 3 = Neither like nor dislike

d. 4 = Like a moderate amount

e. 5 = Like a great deal

7. Do you [stereotypical out-group or target individual] prefer coffee or tea?

a. 1 = Greatly prefer coffee to tea

b. 2 = Moderately prefer coffee to tea

c. 3 = No preference

d. 4 = Moderately prefer tea to coffee

e. 5 = Greatly prefer tea to coffee

8. Do you [stereotypical out-group or target individual] like playing video games?

a. 1 = Dislike a great deal

b. 2 = Dislike a moderate amount

c. 3 = Neither like nor dislike

d. 4 = Like a moderate amount

e. 5 = Like a great deal

9. Do you [stereotypical out-group or target individual] prefer soccer or basketball?

    a. 1 = Greatly prefer soccer to basketball

    b. 2 = Moderately prefer soccer to basketball

    c. 3 = No preference

    d. 4 = Moderately prefer basketball to soccer

    e. 5 = Greatly prefer basketball to soccer

10. Do you [stereotypical out-group or target individual] like reading books for pleasure?

    a. 1 = Dislike a great deal

    b. 2 = Dislike a moderate amount

    c. 3 = Neither like nor dislike

    d. 4 = Like a moderate amount

    e. 5 = Like a great deal

## REFERENCES CITED


Audette, N., Horowitz, J., & Michelitch, K. (2020). *Personal narratives reduce negative attitudes toward refugees and immigrant outgroups: Evidence from Kenya.* [Unpublished Manuscript]. https://www.vanderbilt.edu/csdi/includes/WP 1 2020 FINAL.pdf.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi:10.18637/jss.v067.i01.

Brambilla, M., & Butz, D. A. (2013). Intergroup threat and outgroup attitudes: Macro-level symbolic threat increases prejudice against gay men. *Social Psychology, 44*(5), 311–319. doi: 10.1027/1864-9335/a000127

Brandt, M. J., Reyna, C., Chambers, J. R., Crawford, J. T., & Wetherell, G. (2014). The ideological-conflict hypothesis: Intolerance among both liberals and conservatives. *Current Directions in Psychological Science, 23*(1), 27–34. doi: 10.1177/0963721413510932

Brandt, M. J., Turner-Zwinkels, F. M., Karapirinler, B., Van Leeuwen, F., Bender, M., van Osch, Y., & Adams, B. (2021). The association between threat and politics depends on the type of threat, the political domain, and the country. *Personality and Social Psychology Bulletin*, *47*(2), 324–343. doi: 10.1177/0146167220946187

Brockman, D., & Kalla, J. (2016). Durably reducing transphobia: A field experiment on door-to-door canvassing. *Science*, *352*(6282), 220–224. doi: 10.1126/science.aad9713

Brown, J.R. & Enos, R.D. (2021) The measurement of partisan sorting for 180 million voters. *Nature Human Behavior,* 5, 998–1008. doi: 10.1038/s41562-021-01066-z

Cadinu, M. R., & Rothbart, M. (1996). Self-anchoring and differentiation processes in the minimal group setting. *Journal of Personality and Social Psychology*, *70*(4), 661-677. doi: 10.1037/0022-3514.70.4.661

Castelli, L., Arcuri, L., & Carraro, L. (2009). Projection processes in the perception of political leaders. *Basic and Applied Social Psychology*, *31*(3), 189-196. doi: 10.1080/01973530903058151

Carney, D. R., Jost, J. T., Gosling, S. D., & Potter, J. (2008). The secret lives of liberals and conservatives: Personality profiles, interaction styles, and the things they leave behind. *Political Psychology*, *29*(6), 807-840. doi: 10.1111/j.1467-9221.2008.00668.x

Cesario, J., & Navarrete, C. D. (2014). Perceptual bias in threat distance: The critical roles of in-group support and target evaluations in defensive threat regulation. *Social Psychological and Personality Science, 5*(1), 12–17. doi: 10.1177/1948550613485605

Chambers, J., Baron, R., & Inman, M. (2006). Misperceptions in intergroup conflict: Disagreeing about what we disagree about. *Psychological Science, 17*(1), 38-45. doi:http://dx.doi.org/10.1111/j.1467-9280.2005.01662.x

Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl., S. Anandkumar, A., Ford, A., Volcic, R., & De Rosario., H. (2020). *Pwr: Basic functions for power analysis* (R package version 1.3.0). https://CRAN.R-project.org/package=pwr

Chan, C., Chan, G. C., Leeper, T. J., & Becker, J. (2021). *rio: A Swiss-army knife for data file I/O* (R package version 0.5.26). https://cran.r-project.org/web/packages/rio/index.html

Clement, R. W., & Krueger, J. (2002). Social categorization moderates social projection. *Journal of Experimental Social Psychology, 38*(3), 219–231. doi: 10.1006/jesp.2001.1503

Cook, C. L., & Cottrell, C. A. (2021). You don't know where he's been: Sexual promiscuity negatively affects responses toward both gay and straight men. *Psychology of Men & Masculinities, 22*(1), 63–76. doi: 10.1037/men0000270

Crawford, J. T. (2014). Ideological symmetries and asymmetries in political intolerance and prejudice toward political activist groups. *Journal of Experimental Social Psychology*, *55*, 284-298. doi: 10.1016/j.jesp.2014.08.002

Crawford, J. T. (2017). Are conservatives more sensitive to threat than liberals? It depends on how we define threat and conservatism. *Social Cognition, 35*(4), 354–373. doi: 10.1521/soco.2017.35.4.354

Davis, M. H. (2017). Social projection to liked and disliked targets: The role of perceived similarity. *Journal of Experimental Social Psychology*, *70*, 286-293. doi: 10.1016/j.jesp.2016.11.012

Deane, C. & Gramlich, J. (2020, November 6). *America is exceptional in the nature of its political divide.* Pew Research Center. https://www.pewresearch.org/fact-tank/2020/11/13/america-is-exceptional-in-the-nature-of-its-political-divide/

Denning, K. R., & Hodges, S. D. (2022). When polarization triggers out-group "counter-projection" across the political divide. *Personality and Social Psychology Bulletin, 48*(4), 638-656. doi: 10.1177/01461672211021211

DiDonato, T. E., Ullrich, J., & Krueger, J. I. (2011). Social perception as induction and inference: An integrative model of intergroup differentiation, ingroup favoritism, and differential accuracy. *Journal of Personality and Social Psychology*, *100*(1), 66-83. doi:http://dx.doi.org/10.1037/a0021051

Eadeh, F. R., & Chang, K. K. (2020). Can threat increase support for liberalism? New insights into the relationship between threat and political attitudes. *Social Psychological and Personality Science*, *11*(1), 88–96. doi: 10.1177/1948550618815919

Folashade, D., Microsoft Corporation, Weston, S. & Tenenbaum, D. (2022). *doParallel: Foreach parallel adaptor for the 'praallel' package* (R package version 1.0.17). https://cran.r-project.org/web/packages/doParallel/index.html

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878–902. doi: doi.org/10.1037/0022-3514.82.6.878

Fox, J. W. S. (2019). *An R companion to applied regression*, *3rd edition*. Sage, Thousand Oaks CA. https://socialsciences.mcmaster.ca/jfox/Books/Companion/index.html.

Gaujoux, R. (2020). *doRNG: Generic reproducible parallel backend for 'foreach' loops* (R package version 1.0.17). https://cran.r-project.org/web/packages/doRNG/index.html

Gerber, A. S., Huber, G. A., Doherty, D., & Dowling, C. M. (2011). The big five personality traits in the political arena. *Annual Review of Political Science*, *14*, 265-287. doi:10.1146/annurev-polisci-051010-111659

Gehlbach, H., & Brinkworth, M. E. (2012). The social perspective taking process: Strategies and sources of evidence in taking another's perspective. *Teachers College Record*, *114*(1), 1–29. doi: 10.1177/016146811211400107

Gramlich, J. (2016, November 7). *America's political divisions in 5 charts.* Pew Research Center. http://www.pewresearch.org/fact-tank/2016/11/07/americas-political-divisions-in-5-charts

Gromet, D. M., & Van Boven, L. (2020). *Oppositional projection in perceived support for cap and trade* [Unpublished manuscript]. Department of Psychology, University of Colorado, Boulder, Colorado.

Hodges, S. D., Denning, K. R., & Lieber, S. (2018). Perspective taking: Motivation and impediment to shared reality. *Current Opinion in Psychology.* https://doi.org/10.1016/j.copsyc.2018.02.007

Hughes, B. T., Flournoy, J. C., & Srivastava, S. (2021). Is perceived similarity more than assumed similarity? An interpersonal path to seeing similarity between self and others. *Journal of Personality and Social Psychology, 121*(1), 184–200. https://doi.org/10.1037/pspp0000369

Jilani, Z. & Smith, J. A. (2018, November 7). *What's driving political violence in America?* Greater Good Magazine. https://greatergood.berkeley.edu/article/item/whats_driving_political_violence_in_america

Jost, J. T., Stern, C., Rule, N. O., & Sterling, J. (2017). The politics of fear: Is there an ideological asymmetry in existential motivation? *Social Cognition, 35*(4), 324–353. doi: 10.1521/soco.2017.35.4.324

Kalla, J., & Brockman, D. (2016). 060615 Trans Eq Ally Script Final. *Replication Data for: Durably reducing transphobia: a field experiment on door-to-door canvassing*. Harvard Dataverse, V3. doi: 10.7910/DVN/WKR39N/IGBEHY

Kalla, J., & Brockman, D. (2020). Reducing exclusionary attitudes through interpersonal conversation: Evidence from three field experiments. *American Political Science Review, 114*(2), 410-425. doi: 10.1017/S0003055419000923

Kassambra, A. & Mundt, F. (2020). *Factoextra: Extract and visualize the results of multivariate data analyses* (R package version 1.0.7). https://CRAN.R-project.org/package=factoextra

Kennedy, R., Clifford, S., Burleigh, T., Waggoner, P. D., Jewell, R., & Winter, N. J. G. (2020). The shape of and solutions to the mturk quality crisis. *Political Science Research and Methods*, *8*(4), 614-619. doi:https://doi.org/10.1017/psrm.2020.6

Krueger, J. I. (2007). From social projection to social behaviour. *European Review of Social Psychology*, *18*(1), 1-35. doi: 10.1080/10463280701284645

Ku, G., Wang, C. S., & Galinsky, A. D. (2015). The promise and perversity of perspective-taking in organizations. *Research in Organizational Behavior*, *35*, 79-102. doi: 10.1016/j.riob.2015.07.003

Lenth, R. (2021). *emmeans: Estimated marginal means, aka least-squares means* (R package version 1.5.4). https://CRAN.R-project.org/package=emmeans

Levitsky, S. & Way, L. (2022, January 10). *America's coming age of instability: Why constitutional crises and political violence could soon be the norm.* Foreign Affairs. https://www.foreignaffairs.com/articles/united-states/2022-01-20/americas-coming-age-instability

Lüdecke, D. (2021). sjPlot: *Data visualization for statistics in social science* (R package version 2.8.7). https://CRAN.R-project.org/package=sjPlot.

Lüdecke, D., Patil, I., & Makowski, D. (2020). Parameters: Extracting, computing and exploring the parameters of statistical models using R. *Journal of Open Source Software*, *5*(53), 2445-2449. doi:https://joss.theoj.org/papers/10.21105/joss.02445

Machunsky, M., Toma, C., Yzerbyt, V., & Corneille, O. (2014). Social projection increases for positive targets: Ascertaining the effect and exploring its antecedents. *Personality and Social Psychology Bulletin, 40*(10), 1373-1388. doi: 10.1177/0146167214545039

Mackie, D. M., Devos, T., & Smith, E. R. (2000). Intergroup emotions: Explaining offensive action tendencies in an intergroup context. *Journal of Personality and Social Psychology, 79*(4), 602-616. doi: 10.1037/0022-3514.79.4.602

Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 630-633. doi: 10.3758/BF03206543

Mooijman, M., & Stern, C. (2016). When perspective taking creates a motivational threat: The case of conservatism, same-sex sexual behavior, and anti-gay attitudes. *Personality and Social Psychology Bulletin*, *42*(6), 738-754. doi: 10.1177/0146167216636633

Mullen, B., Dovidio, J. F., Johnson, C., & Copper, C. (1992). In-group-out-group differences in social projection. *Journal of Experimental Social Psychology*, *28*(5), 422-440. doi: 10.1016/0022-1031(92)90040-Q

Niella, T. (2020). *Moral and political polarization: An intervention to increase consensus in face-to-face discussions and how outcomes compare to our expectations* [Unpublished thesis]. Department of Psychology, University of Oregon, Eugene, Oregon

Obeidallah, D. (2022, January 10). *Even Canadians fear US democracy could end soon.* CNN. https://www.cnn.com/2022/01/09/opinions/canadians-fear-us-democracy-collapse-obeidallah/index.html

Park, B., & Rothbart, M. (1982). Perception of out-group homogeneity and levels of social categorization: Memory for the subordinate attributes of in-group and out-group members. *Journal of Personality and Social Psychology, 42*(6), 1051–1068. doi: 10.1037/0022-3514.42.6.1051

Pew Research Center. (2012, June 4). *Partisan polarization surges in Bush, Obama years.* https://www.people-press.org/2012/06/04/partisan-polarization-surges-in-bush-obama-years/

Pew Research Center (2019, October 10). *Partisan antipathy: More intense, more personal.* https://www.pewresearch.org/politics/2019/10/10/partisan-antipathy-more-intense-more-personal

Pierce, J. R., Kilduff, G. J., Galinsky, A. D., & Sivanathan, N. (2013). From glue to gasoline: How competition turns perspective takers unethical. *Psychological Science*, *24*(10), 1986-1994. doi: 10.1177/0956797613482144

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team (2022). *nlme: Linear and nonlinear mixed effects models* (R package version 3.1-155). https://CRAN.R-project.org/package=nlme.

Pyszczynski, T., Greenberg, J., Solomon, S., Cather, C., Gat, I., & Sideris, J. (1995). Defensive distancing from victims of serious illness: The role of delay. *Personality and Social Psychology Bulletin, 21*(1), 13–20. doi: 10.1177/0146167295211003

Pyszczynski, T., Greenberg, J., Solomon, S., Sideris, J., & Stubing, M. J. (1993). Emotional expression and the reduction of motivated cognitive bias: Evidence from cognitive dissonance and distancing from victims' paradigms. *Journal of Personality and Social Psychology*, *64*(2), 177. doi: 10.1037/0022-3514.64.2.177

R Core Team (2020). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.* https://www.R-project.org/.

Ravenscroft, T (2019, July 31). *Mizzi Studio completes stingray café alongside the Serpetine.* Dezeen. https://www.dezeen.com/2019/07/31/serpentine-coffee-house-mizzi-studio-hyde-park/

Revelle, W. (2021). *psych: Procedures for psychological, psychometric, and personality research* (R package version 2.1.3). https://CRAN.R-project.org/package=psych.

Rice, S. E. (2020. September 22). *A divided America is a national security threat.* The New York Times. https://www.nytimes.com/2020/09/22/opinion/trump-national-security.html

Robbins, J. M., & Krueger, J. I. (2005). Social projection to ingroups and outgroups: A review and meta-analysis. *Personality and Social Psychology Review*, *9*(1), 32–47. doi: 10.1207/s15327957pspr0901_3

Sassenrath, C., Hodges, S. D., & Pfattheicher, S. (2016). It's all about the self: When perspective taking backfires. *Current Directions in Psychological Science*, *25*(6), 405-410. doi: 10.1177/0963721416659253

Saveski, M., Gillani, N., Yuan, A., Vijayaraghavan, P., & Roy, D. Perspective-taking to reduce affective polarization on social media. (2021). *arXiv preprint.* https://arxiv.org/pdf/2110.05596.pdf

Soto, C. J., & John, O. P. (2017). Short and extra-short forms of the Big Five Inventory–2: The BFI-2-S and BFI2-XS. *Journal of Research in Personality*, *68*, 69-81. doi: http://dx.doi.org/10.1016/j.jrp.2017.02.004

Stephan, W. G., Boniecki, K. A., Ybarra, O., Bettencourt, A., Ervin, K. S., Jackson, L. A., McNatt, P. S., & Renfro, C. L. (2002). The role of threats in the racial attitudes of blacks and whites. *Personality and Social Psychology Bulletin*, *28*(9), 1242–1254. doi: 10.1177/01461672022812009

Stephan, W. G., Ybarra, O., & Morrison, K. R. (2016). Intergroup threat theory. In T. D. Nelson (Ed.), *Handbook of Prejudice, Stereotyping, and Discrimination* (pp. 43–59). Psychology Press.

Talley, A. E., & Bettencourt, B. A. (2008). Evaluations and aggression directed at a gay male target: The role of threat and antigay prejudice. *Journal of Applied Social Psychology, 38*(3), 647–683. doi: 10.1111/j.1559-1816.2007.00321.x

Tarrant, M., Calitri, R., & Weston, D. (2012). Social identification structures the effects of perspective taking. *Psychological Science*, *23*(9), 973-978. doi: 10.1177/0956797612441221

Thalmayer, A.G., Saucier, G., & Eigenhuis, A. (2011). Comparative validity of brief to medium length Big Five and Big Six personality questionnaires. *Psychological Assessment*, 23, 995-1009. doi: 10.1037/a0024165

United States Census Bureau (2021). *QuickFacts United States* (Table of Census data, vintage version July 1, 2021). https://www.census.gov/quickfacts/fact/table/US/PST045221

Xiao, Y. J., & Van Bavel, J. J. (2012). See your friends close and your enemies closer: Social identity and identity threat shape the representation of physical distance. *Personality and Social Psychology Bulletin*, *38*(7), 959–972. https://doi.org/10.1177/0146167212442228

Wei, T. & Simko, V. (2021). *R package 'corrplot': Visualization of a correlation matrix* (R package version 0.92). https://github.com/taiyun/corrplot.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, LD., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., … & Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. doi:10.21105/joss.01686.

Wyer, N. A., & Calvini, G. (2011). Don't sit so close to me: Unconsciously elicited affect automatically provokes social avoidance. *Emotion, 11*(5), 1230–1234. doi: 10.1037/a0023981