

Research Note: The Femin[ist] Mystique: How Group Perceptions Shape Affect and Identity

Elizabeth C. Connors, University of South Carolina
Katelyn E. Stauffer, University of South Carolina

April 6, 2023

Why do some individuals choose to identify as “feminist” while others choose to reject this label? In this note, we attempt to unpack this by proposing one particular driver behind the rejection of the feminist label: group perceptions. In particular, we argue that perceptions of feminists’ vocalicity should shape people’s affect towards and willingness to affiliate with feminists, and that the effect of these perceptions should differ by partisanship. We examine these expectations using open-ended responses and a survey experiment with 1,000 respondents from the 2021 Cooperative Election Study. We indeed find that perceptions of feminists’ vocalicity shape feminist affect and identity. Moreover, we find that changing these perceptions can even close the gap between Republicans and Democrats’ affect towards feminists. Our findings shed light on the currently extreme partisan differences in feminist identification as well as begin to uncover a sorely understudied, yet extremely important, political and social identity.

Key Words: American politics; group identity; political behavior; feminism

Paper prepared for the 2022 American Political Science Association Annual Conference in Montréal, Québec, Canada. Please do not cite or circulate without the permission of the authors. The abstract is 150 words, and the manuscript is approximately 4461 words.

Joy Behar: Can I read you the definition of feminism? I just looked it up ... the definition is that women and men should have equal rights and opportunities.

Deborah Roberts: Im not sure that I identify myself as a feminist even though by that definition I am ... it just feels like a very strong word. Although my daughter says you are a feminist Mom, you just dont like to wear the label.

– Excerpt from panel titled, Is ‘Feminism’ a Bad Word? on ABCs The View

Introduction

On a 2018 segment of the daytime talk show “The View,” the co-hosts participated in a segment titled “Is Feminism a Bad Word?” and discussed why some women embrace the feminist label while others are hesitant to adopt it. As the conversation went on, one of the co-hosts—Joy Behar—introduced a basic definition of feminism, the view that men and women should be equal.¹ Guest co-host Deborah Roberts acknowledged that she agreed with the sentiment that men and women should be equal in political, social, and economic life; yet, she still expressed a hesitancy to describe herself as feminist, calling it a “strong” word and suggesting that there is more wrapped up within this label than implied by the simple definition provided by Behar.

This exchange is emblematic of a broader paradox in American public opinion. Like Roberts, Americans are by-in-large supportive of gender equality (at least in the abstract), and are remarkably united on this point. In a 2020 Pew poll, 91% of Americans said it was “very important” women have the same rights as men.² This question taps the basic definition of feminism common in popular discourse—belief in gender equality. With this in mind, one unacquainted with American politics might expect to see a similarly large percentage of Americans identify as feminist. Yet, this is not what we observe in reality. In fact, public opinion polling consistently finds that

¹Of course, the definition of feminism is more complex than this. There are many variants of feminism and what it means to be “feminist” is a topic of much debate in scholarly and theoretical discourse. Nonetheless, the most simple definition that is often used in public communication is that of gender equality. And, when conceptualized in this simple way, people—like Deborah Roberts—are *still* hesitant to adopt the label despite agreeing with the sentiment.

²<https://www.pewresearch.org/global/2020/04/30/worldwide-optimism-about-future-of-gender-equality-even-as-many-see-advantages-for-men/>.

only a minority of Americans describe themselves as feminist,^{3,4,5,6} and more Americans describe feminists unfavorably than favorably (Huddy, Neely and Lafay, 2000). Further, for some, the term “feminist” is viewed as an insult (Huddy, Neely and Lafay, 2000). Even among those who believe most in women’s equality, feminists and feminism are not held in terribly high regard (Huddy, Neely and Lafay, 2000). What explains this contradiction? How is it that Americans are largely supportive of the basic principles of feminism, yet hostile towards the feminist label?

While a great deal of work in sociology has examined the antecedents of feminist identity, political science has been slower to unpack why some choose to identify as feminist yet others—despite holding similar gender equality views—do not. Further, even though canonical work in gender and politics identifies feminist consciousness as a consequential political identity that has policy implications (Conover, 1988), our understanding of how the term is currently used (or not used) by the public is limited. What work does exist finds that factors such as partisanship (Huddy, 2018), policy views (Conover, 1988; Cook and Wilcox, 1991), gender⁷, race, educational background, and class can influence feminist identification (Inglehart et al., 2003; Rhodebeck, 1996; Rinehart, 1992).

In this note, we explore how conceptions of feminists as a group influence the extent to which individuals are willing to use the term to describe themselves. To the degree that Americans hold unsavory views about feminists (e.g., their identities, tactics, demeanor, etc.), we argue that this influences feelings towards the group and thus one’s willingness to identify with it. Indeed, research on group politics tells us that individuals are hesitant to identify with groups they dislike or ascribe negative characteristics to (Klar and Krupnikov, 2016). If Americans have come to characterize feminism in a negative light, then it is little surprise that so few use the term to describe themselves. However, if these negative views were countered would individuals’ opinions towards feminists and willingness to affiliate change? In other words, would disrupting feminist stereotypes make individuals feel more warmly towards feminists? And would these shifts lead individuals to be more likely to call *themselves* feminist? In this note we examine one potential negative stereotype ascribed to feminists—that they are too vocal—and if disrupting this stereotype

³<https://www.ipsos.com/en-us/american-women-and-feminism>

⁴<https://www.washingtonpost.com/graphics/national/feminism-project/poll/>

⁵<https://www.pewresearch.org/fact-tank/2020/07/14/most-americans-support-gender-equality-even-if-they-dont-iden>

⁶<https://www.vox.com/2015/4/8/8372417/feminist-gender-equality-poll>

⁷Huddy (2018, but see) who note that gender is neither sufficient nor necessary to identify as feminist

influences affect towards and identification with feminists.

Using a survey experiment with a nationally representative sample of American adults, we find that many Americans view feminists as too vocal and that providing views of feminism that run counter to this view can drastically improve feelings towards and willingness to affiliate with feminists. This is especially pronounced among Republicans, who are currently much less likely to identify as feminist. This is likely because for Republicans, “vocal feminists” are seen as especially threatening, as feminists are generally viewed as part of the Democratic coalition (Mason, 2018).

The fact that so many Americans reject the feminist label presents an opportunity for political opponents to paint feminists as a political boogey-man—a radical minority attempting to subvert the preferences of the majority. In a moment when questions of women’s social and political equality have become particularly salient, understanding who identifies as feminist—and why—takes on renewed importance.

Group Perceptions, Affect, and Identity

Motivated by past research, we expect that how individuals perceive feminists—who they are and how they act—helps to drive: 1) positive (negative) feelings towards “feminists”; and 2) willingness to identify as feminist. Research finds that perceptions (even when incorrect) of social and political groups influence behavior and attitudes. For example, perceptions of American political parties—and the extent to which the average partisan embodies party stereotypes—can be a powerful driver of polarization, with those who see out-partisans as more “party stereotypical” judging them more harshly (Ahler and Sood 2018, see also Druckman et al. 2022). Similarly, negative perceptions about partisan behavior can lead people to not only view partisans negatively, but also be less likely to identify as partisan (Klar and Krupnikov, 2016).

Likewise, stereotypes about racial minorities, women, and other historically marginalized groups influence support for social welfare policies and policies aimed at counteracting racial and gender disparities (Cassese and Barnes, 2019; Kreitzer, Maltby and Smith, 2022). Indeed, many argue that Americans understand the political landscape in group-based terms and that this has powerful consequences for political attitudes and participation (Mason, 2018). Just as individuals hold beliefs about political parties and other groups, we expect that they also hold beliefs about

feminists—and that these beliefs can help us to understand why so many Americans are hesitant to wear the feminist label.

Not only do we expect that Americans hold stereotypes about feminists, but we also expect that Americans overestimate the extent to which feminists actually *fit* these stereotypes. As McCabe (2005) notes, “The backlash against feminism in the media and the relatively extreme positions taken by the more outspoken representatives of the feminist movement may have resulted in feminism being equated with ‘radical’ or ‘militant’ for much of the U.S. public.” Moreover, research suggests that individuals overinflate the number of group members who fit stereotypes. For example, Republicans overestimate the number of Democrats who are LGBT+ and very liberal, while Democrats overestimate the degree to which Republicans are very conservative and wealthy (Ahler and Sood, 2018). We likewise expect that people overestimate the number of feminists who embody negative stereotypes and that disrupting these stereotypes could lead to more favorable attitudes towards feminists and greater association with the group.

We focus on one stereotype in particular: that feminists are “loud” and/or “aggressive.” We focus on this stereotype for two reasons. First, it is one frequently ascribed to feminists in media and social discourse. Indeed, this stereotype came up quite frequently in open-ended responses of non-feminists (more details below). Second, this stereotype is likely perceived as negative, especially among those who also perceive feminists as politically incongruence. Research finds that individuals do not like people who talk about politics frequently (Klar, Krupnikov and Ryan, 2018) or those who view politics as central to their identity (Krupnikov and Ryan, 2022)—two characteristics that people likely assume feminists embody. Further, we expect this negative stereotype to be especially influential for Republicans, as research also finds that people especially do not like politically invested and vocal individuals when they assume they disagree with them (Klar, Krupnikov and Ryan, 2018; Krupnikov and Ryan, 2022). Because feminists are largely associated with the Democratic Party (e.g., Mason, 2018), Republicans likely assume feminists are extremely liberal Democrats. This perceived incongruence likely compounds the negative reaction to vocal feminists, making Republicans especially resistant to the group and label—yet also capable of changing their affect and affiliation once this stereotype is disrupted. Of course a perceived incongruence between Republican policies and feminism may be accurate; however, we expect disrupting stereotypes among Republicans may make them *relatively* more positive towards feminists.

Thus, we expect that many individuals hold a “vocal feminist” stereotype and that this makes Republicans in particular less warm towards, and less likely to identify with, feminists. This affect in turn should influence willingness to identify because research finds that how favorably one views a group influences their likelihood of affiliation (see, e.g., Klar and Krupnikov, 2016). Thus, group perceptions that shape affect should also shape affiliation.

Design & Analysis

To examine our expectations, we fielded a survey to a module of 1,000 respondents as part of the 2021 Cooperative Election Study (CES). The CES is an online, nationally representative sample of American adults and is administered by YouGov.⁸ As part of the module, we asked survey respondents whether they consider themselves to be “feminist,” “anti-feminist,” or “neither.” Consistent with prior public opinion polling with similarly worded questions (ANES 2021), only a minority identified as “feminist” (30.53%) and an even smaller minority identified as “anti-feminist” (5.61%). The modal response was “neither” (63.86%). Thus, the majority *did not* identify as feminist. Moreover, clear partisan differences emerge: while 49.09% of Democrats identified as feminist, just 6.78% of Republicans did.⁹

While this initial look at the data gives us an understanding of *who* identifies as feminist, it does little to explain *why* certain individuals reject (or accept) the label. To address this, we included an open-ended question asking respondents to, “Please explain why you do or do not identify as a feminist.”¹⁰ For the purposes of this note, we focus on respondents who indicated they did *not* identify as feminist. Because such a small percentage selected the anti-feminist option, we treat feminist identification as binary, capturing whether respondents identified as feminist (1, 30.53%) or not (0, 69.47%). Among those who did *not* identify as feminist, one of the most common themes to emerge in these responses related to feeling that feminists were “aggressive,” “overly assertive,” “obnoxious,” and “loud” in advocating for their beliefs.

For example, one respondent stated:

⁸For more information see (Ansolabehere and Rivers, 2013).

⁹Being a Democrat (as compared to Republican) predicts identifying as feminist in bivariate logistic regression ($p < .001$).

¹⁰To avoid priming effects, we randomized the order of the multiple choice and open-ended questions. Results are similar regardless of question ordering.

“I’m a woman but not a feminist - I think some feminists are too vocal - I think women should hav3 (sic) equal opportunities and equal pay, but I don’t like the term ‘feminist.’”

Another respondent noted:

“Feminists tend to be overly aggressive and take on make (sic) traits to try to force their way. They are distasteful people.”

The open-ended responses suggest that at least part of the hesitancy to identify as feminist among many individuals is indeed rooted in beliefs about who feminists are and how they act. To next examine if changing these stereotypes can lead individuals to adopt warmer feelings towards feminists, and perhaps identify with them, we turn to a survey experiment.

After answering the two aforementioned questions, we asked respondents to report feeling thermometer scores towards different types of feminists.¹¹ Respondents were randomly assigned to one of three conditions. The first condition used the same wording as the American National Election Studies (ANES) and asked respondents to rate “feminists” (*control*). The second condition asked respondents to rate “feminists who **rarely** talk about feminist issues” (*quiet*). The last condition asked respondents to rate “feminists who **frequently** talk about feminist issues” (*vocal*). These scores serve as the dependent variable for our analysis of how group perceptions shape affect.

The last piece of our survey tests how group perceptions shape identity. Thus, after the identity and feeling thermometer questions, we asked those in the control condition, “Does the way feminists act influence your likelihood of identifying as a feminist” (response options: yes, definitely; maybe; and no) and those in the treatment conditions, “If most feminists [*rarely* / *frequently*] talked about feminist issues, would you be more or less likely to identify as a feminist” (response options: more likely; less likely; the same). These questions serve as dependent variables for our analysis of how group perceptions shape identity.

Consequences for Affect

For our first set of analyses, we examine how group perceptions shape respondents’ feelings towards feminists. To do so, we compare how respondents feel about feminists in the control versus the

¹¹Four respondents did not provide answers to the feeling thermometer questions.

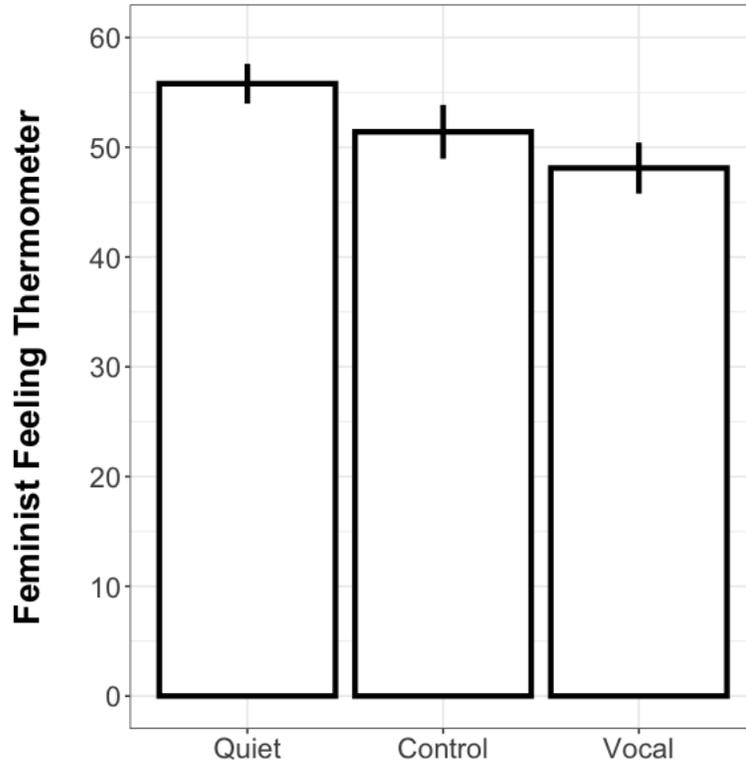


Figure 1: Feeling thermometer scores for “feminists” across conditions. Bars represent 84% confidence intervals. In cases where 84% confidence intervals do not overlap, we can conclude that means are significantly different at the $p < .05$ level

quiet and vocal conditions (see Table 1 and Figure 1). We see that, on average, people feel most positively towards quiet feminists and most negatively towards vocal feminists, with the control condition in between. While the quiet and control conditions are significantly different from one another ($p=.043$), the vocal and control conditions are not ($p=.171$). This suggests that without specificity, respondents assume feminists are vocal and thus view them more negatively—but that respondents feel more positively about feminists if they are “quiet.”

Next, we examine how the effect of these group perceptions changes by partisanship (see Figure 2). Recall that we predicted Republicans would be especially turned off by vocal feminists. Here we find that indeed Republicans rate vocal feminists slightly lower (-4.46 points, $p=.198$)—and quiet feminists much higher (21.32 points, $p<.001$)—than the control condition. That is a 92% increase in affect from *vocal* to *quiet*. Again, this suggests that: Republicans view vocal feminists negatively, they assume feminists *are* vocal, and their affect improves when asked about quiet

Table 1: Mean Feeling Thermometer Score Across Conditions

	<i>Dependent variable: Feminist Feeling Thermometer</i>								
	Overall Sample			Democrats			Republicans		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Control	51.40	31.66	329	69.94	26.38	149	32.52	24.35	106
Quiet	55.80	23.59	335	56.96	21.42	177	53.83	26.95	86
Vocal	48.11	30.17	332	62.64	26.17	168	28.06	25.28	101

feminists.

Notably, we do not change Republicans’ perception of feminists’ policy positions and beliefs, yet we still see a large shift just by changing perceptions of how often feminists *talk* about these positions and beliefs. In fact, Republicans view quiet feminists similarly to how *Democrats* view quiet feminists (means of 53.83 and 56.96, respectively). Among participants in the quiet condition, partisanship no longer predicts feelings towards feminists ($p=.308$), suggesting that updating perceptions of feminist behavior can actually close gaps between the two partisan groups’ affect towards feminists—gaps that in the control condition are 37.42 points ($p=.000$).

When we look at Democrats, we see quite a different story. Democrats slightly punish vocal feminists (-7.30, $p=.014$) and more severely punish quiet feminists (-12.98, $p<.001$). This puzzling finding makes sense, however, when we split the Democratic sample into feminists and non-feminists. Democratic *non-feminists* do not punish quiet feminists—in fact, the effect size is positive (1.13, $p=.759$), although insignificant—but *do* slightly punish vocal feminists (-6.63, $p=.078$), although this is only marginally significant. Democratic *feminists*, however, also slightly punish vocal feminists (-7.19, $p=.005$) and quite strongly punish quiet feminists (-27.80, $p<.001$). This quiet feminist result aligns with findings that people punish political in-group members who fail to act like “correct” group members by getting deeply involved (Krupnikov and Ryan 2022)—Democratic feminists might view feminists who fail to talk about (what they view as) important issues negatively because they are not fighting for “the cause.”¹²

¹²One might wonder, then, if this is just a story about feminists versus non-feminists rather than a story about partisanship and feminism. We do not find this to be the case. While feminists act similarly to Democratic feminists (punishing quiet feminists as well as slightly punishing vocal feminists) and non-feminists act similarly to Republican non-feminists (rewarding quiet feminists but treating vocal feminists the same), these trends are exaggerated among the partisan groups. Democratic feminists punish quiet feminists more than feminists as a whole do: an effect of

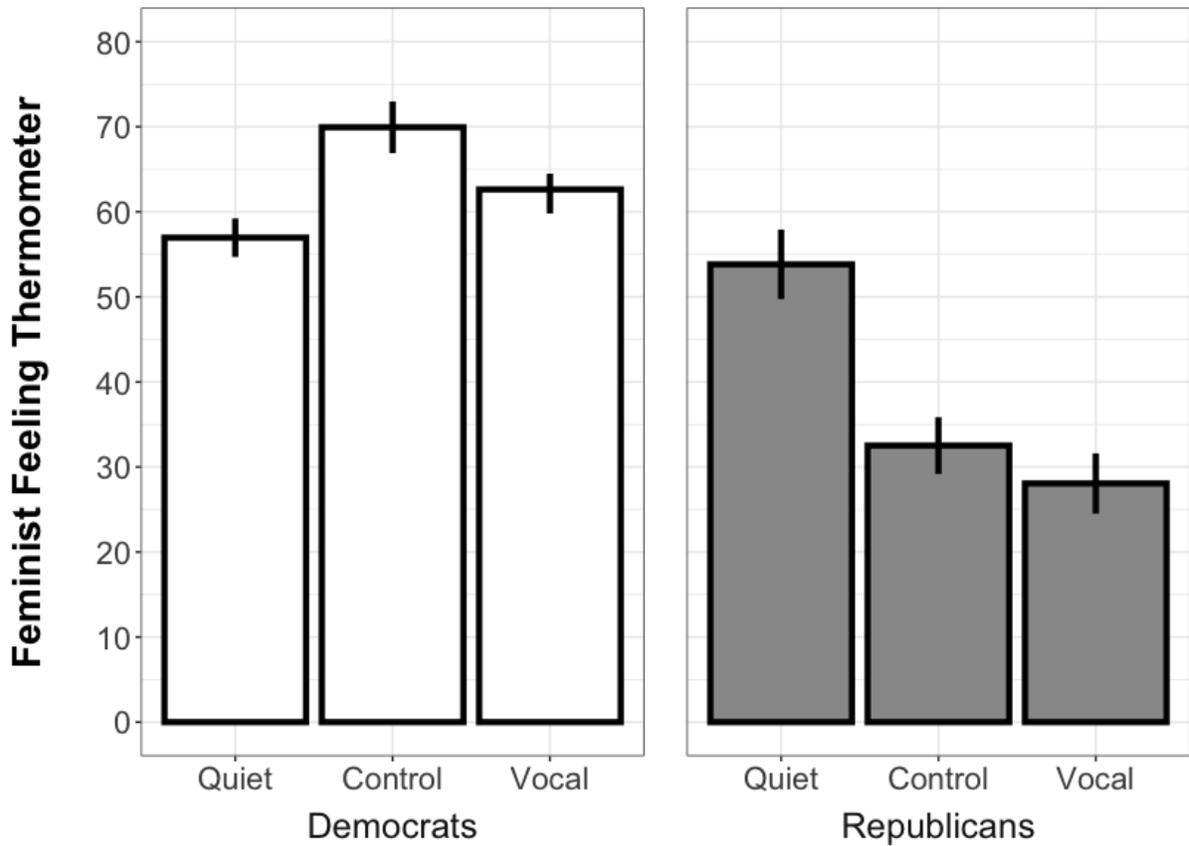


Figure 2: Feeling thermometer scores for “feminists” across condition by respondent party. Bars represent 84% confidence intervals. In cases where 84% confidence intervals do not overlap, we can conclude that means are significantly different at the $p < .05$ level

Thus, consistent with what we would expect given the open-ended responses, we find that how feminists act shapes respondents' group affect. Further, this changes by partisanship: while on average Democrats view quiet feminists negatively—an effect that is driven by Democratic *feminists* rather than non-feminists—Republicans on average view quiet feminists *positively*—an effect that seems to be driven by both feminists and non-feminists, although it is stronger among the latter. Next we turn to how this might translate to identification.

Consequences for Identity

As our last set of analyses, we examine how group perceptions shape individuals' willingness to identify as feminist. To do this, we first look at responses to the conditional identity question in the control condition. We find that over a quarter of respondents (27.96%) say their willingness to identify with the group is “definitely” shaped by how feminists act and another 25.23% say it “maybe” is. For Republicans, these respective percentages change to 35.85% and 19.81% and for Democrats these change to 22.92% and 31.64%.

While this high percentage (especially among Republicans) aligns with open-ended responses that suggest how (people perceive) feminists act shapes identity, to test our particular expectation that how vocal feminists are shapes who is willing to identify as feminist, we compare responses in the treatment groups. Here, we find that in both treatment conditions the modal response was “the same” (74.18% in *quiet* and 55.09% in *vocal*)—as in they would be just as likely to identify as feminist. Further, twice as many respondents said “less likely” in *vocal* (29.34%) than in *quiet* (15.43%), but about 5% more respondents said “more likely” in *vocal* (15.57%) than in *quiet* (10.39%), necessitating a deeper dive into the data.

In particular, given our theory and previous findings, we expect this masks partisan heterogeneity. This is indeed what we find. First, we find that Republicans say “less likely” *far* more often in *vocal* (51.46%) than in *quiet* (15.12%) and say “more likely” more often in *quiet* (12.79%) than in *vocal* (3.88%). Further, the gap between the more and less likely responses in *vocal* is almost half the scale: while a whopping 51.46% of Republicans said they would be *less* likely to identify as feminist if most feminists were vocal, only 3.88% of Republicans say they would be *more* likely

-27.80 ($p < .001$) as opposed to -20.76 ($p < .001$). Similarly, Republican non-feminists *reward* quiet feminists more than non-feminists as whole do: an effect of 24.99 ($p < .001$) as opposed to 15.33 ($p < .001$).

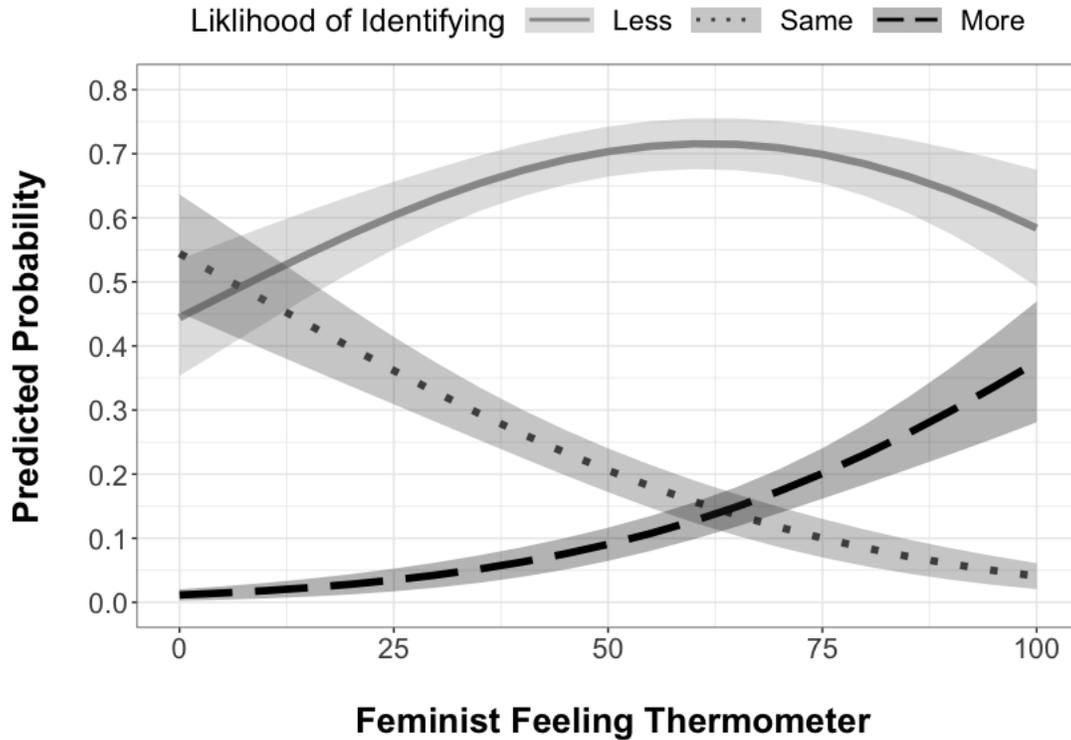


Figure 3: The effect of affect on willingness to identify as feminist.

to. Note that our previous findings suggest that the low affiliation among Republicans is partly due to their assumption that feminists are vocal and their dislike of this quality among feminists. Thus, their reporting that more vocal feminists would *further* decrease their likelihood of affiliation is notable.

Again, though, while findings for Republicans are clear, findings for Democrats are muddled and deserve further research. While Democrats say “less likely” more often in *vocal* (19.64%) than in *quiet* (15.64%), they also say “more likely” more often in *vocal* (23.81%) than in *quiet* (8.38%). This suggests that Democratic feminists might be reacting negatively to quiet feminists, like in the previous analysis. Indeed, we find that more Democratic feminists said they would be more likely to identify as feminist if most feminists were vocal (38.75%) than quiet (11.11%). Similarly, more Democratic feminists said they would be *less* likely to identify if most feminists were quiet (17.78%) than vocal (8.75%).

As a last step, we examine how individual affect towards feminists predicts their willingness to identify with the group. This should be the case because, as we note, research finds that the

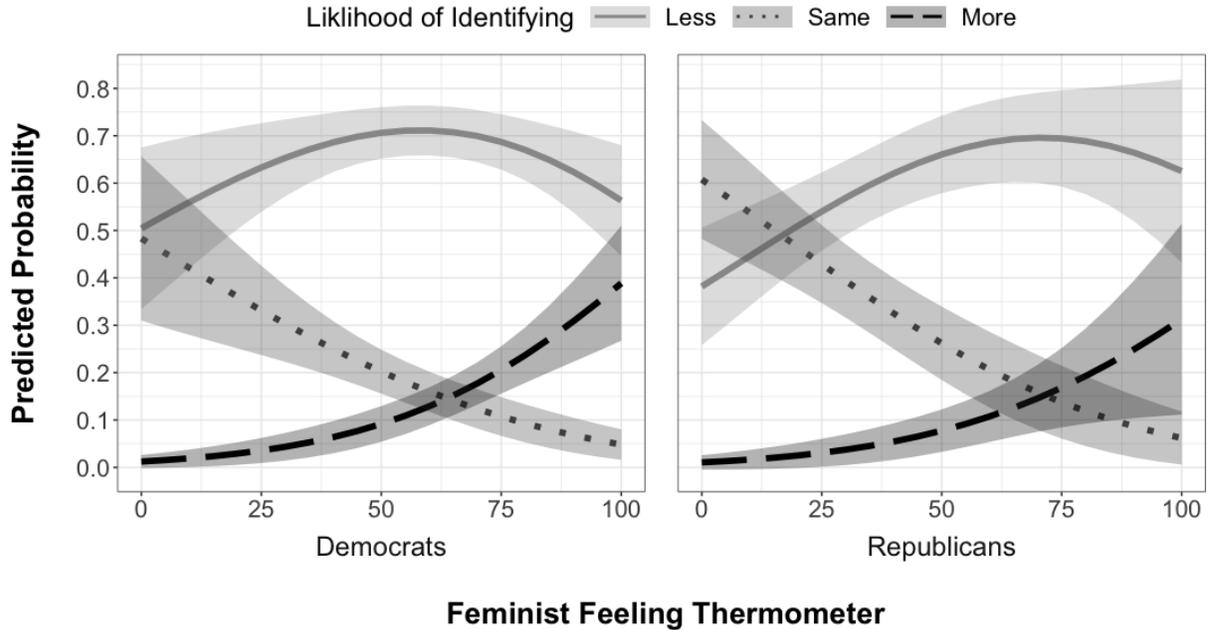


Figure 4: The effect of affect on willingness to identify as feminist by respondent party

more positively one views a group the more likely they should be to affiliate with it. Further, we expect that some of the low feminist affiliation among Republicans is due to their dislike of vocal feminists and thus hesitance to affiliate with them. Thus, we should see that how much one likes quiet or vocal feminists is correlated with respondents' likelihood of affiliating with that group based on its composition (i.e., if most of the group were quiet or vocal).

This is exactly what we find (see Figures 3 and 4): liking quiet (vocal) feminists more predicts respondents saying they would be more likely to identify if most feminists were quiet (vocal).¹³ Similarly, liking quiet (vocal) feminists *less* predicts respondents saying they would be *less* likely to identify if most feminists were quiet (vocal).¹⁴ Lastly, to examine this more directly, we run mediation analysis (see appendix). Here, we find that the *quiet* and *vocal* treatments influenced feminist identification (partly for Democrats but largely for Republicans) through changing respondents' affect towards feminists as a group. This aligns with both our argument and main findings.

¹³Coefficient of thermometer in a multinomial logit: .03, $p < .001$.

¹⁴Coefficient of thermometer in multinomial logit: -.03, $p < .001$.

Conclusion

In this piece we aimed to understand why so few Americans identify as feminist despite overwhelming endorsement of gender equality. We began unpacking this quandary by examining how group perceptions—in particular, whether the group is “vocal” or “quiet”—shape people’s affect towards and likelihood of affiliating with feminists. We found these group perceptions to be incredibly important, especially for Republicans. While Republicans dislike vocal feminists and assume feminists are vocal, they feel more warmly towards quiet feminists and are subsequently more likely to identify as feminist if most feminists fit this description. This aligns with findings that most people do not like those who are vocal about politics, especially when they disagree with them (Klar, Krupnikov and Ryan, 2018; Krupnikov and Ryan, 2022). Conversely, Democrats—in particular, feminist Democrats—feel more *negatively* towards quiet feminists and are subsequently *less* likely to identify if most feminists were quiet. This also aligns with existing work, as some group members can feel negatively towards ingroup members who do not express themselves “correctly” (Krupnikov and Ryan, 2022), although more research here is needed.

While these findings are illuminating and demonstrate a way to close the immense gap between Democrats’ and Republicans’ affect towards feminists and thus likelihood of identifying as feminist, there is still much to unpack. For example, how does policy matter here? And, how do other (mis)perceptions of group aspects shape affect and identity? Further, how does one’s social environment shape these perceptions and thus affect and affiliation? Lastly, and importantly, why do a shockingly significant number (50.91%) of *Democrats*—whose goals are largely aligned with feminists—*not* identify as feminist? In particular, how do the 22.82% and 31.54% of Democrats who said that how feminists act would “definitely” or “maybe” (respectively) change their willingness to identify actually want feminists to act, if the group’s vocality is not the story?¹⁵ We leave this and other important questions for future research and hope we are just the first word in more research aiming to understand feminist affiliation and goals of gender equality.

¹⁵Narrowing this to non-feminist Democrats, these percentages are 15.79% and 31.58%.

References

- Ahler, Douglas and Gaurav Sood. 2018. "The Parties in Our Heads: Misperceptions about Party Composition and their Consequences." *The Journal of Politics* 80(3):964–981.
- Ansolabehere, Stephen and Douglas Rivers. 2013. "Cooperative survey research." *Annual Review of Political Science* 16:307–329.
- Cassese, Erin C and Tiffany D Barnes. 2019. "Intersectional motherhood: Investigating public support for child care subsidies." *Politics, Groups, and Identities* 7(4):775–793.
- Conover, Pamela Johnston. 1988. "Feminists and the gender gap." *The Journal of Politics* 50(4):985–1010.
- Cook, Elizabeth Adell and Clyde Wilcox. 1991. "Feminism and the gender gap—a second look." *The Journal of Politics* 53(4):1111–1122.
- Druckman, James N. and, Samara Klar, Yanna Krupnikov, Matthew Levendusky and John Barry Ryan. 2022. "(Mis)estimating affective polarization." *The Journal of Politics* 84(2):1106–1117.
- Huddy, Leonie. 2018. "The Group Foundations of Democratic Political Behavior." *Critical Review* 30(1-2):71–86.
- Huddy, Leonie, Francis K. Neely and Marilyn R. Lafay. 2000. "Trends: Support for the Women's Movement." *Public Opinion Quarterly* 64(3):309–350.
- Imai, Kosuke, Luke Keele and Dustin Tingley. 2010. "A general approach to causal mediation analysis." *Psychological methods* 15(4):309.
- Inglehart, Ronald, Pippa Norris, Inglehart Ronald et al. 2003. *Rising tide: Gender equality and cultural change around the world*. Cambridge University Press.
- Klar, Samara and Yanna Krupnikov. 2016. *Independent Politics: How American Disdain for Parties Leads to Political Inaction*. New York: Cambridge University Press.
- Klar, Samara, Yanna Krupnikov and John Barry Ryan. 2018. "Untangling a Dislike for the Opposing Party from a Dislike of Partisanship." *Public Opinion Quarterly* 82(2):379–390.

- Kreitzer, Rebecca J, Elizabeth A Maltby and Candis Watts Smith. 2022. "Fifty shades of deservingness: an analysis of state-level variation and effect of social constructions on policy outcomes." *Journal of Public Policy* pp. 1–29.
- Krupnikov, Yanna and John Barry Ryan. 2022. *The Other Divide: Polarization and Disengagement in American Politics*. New York: Cambridge University Press.
- Mason, Lilliana. 2018. *Uncivil Agreement: How Politics Became Our Identity*. University of Chicago Press.
- McCabe, Janice. 2005. "Whats in a label? The relationship between feminist self-identification and feminist attitudes among US women and men." *Gender & Society* 19(4):480–505.
- Rhodebeck, Laurie A. 1996. "The structure of men's and women's feminist orientations: Feminist identity and feminist opinion." *Gender & Society* 10(4):386–403.
- Rinehart, Sue Tolleson. 1992. *Gender consciousness and politics*. Routledge.

Appendix

Survey

Randomly order the following two questions:

- *Feminist ID*: Do you consider yourself a feminist, an anti-feminist, or neither of these? [feminist / anti-feminist / neither]
- *Open-Ended*: Please explain why you do or do not identify as a feminist: [long open-ended]

Random assignment to one of the following three conditions:

- *Control Condition*: We'd like to get your feelings toward some people and groups who are in the news these days using something we call the feeling thermometer. Ratings between 50 and 100 degrees mean that you feel favorable and warm toward the person or group. Ratings between 0 and 50 mean that you don't feel favorable and warm toward the group. You would rate them at the 50 degree mark if you don't feel particularly warm or cold toward them. How would you rate feminists? [feeling thermometer]
- *Quiet Treatment*: We'd like to get your feelings toward some people and groups who are in the news these days using something we call the feeling thermometer. Ratings between 50 and 100 degrees mean that you feel favorable and warm toward the person or group. Ratings between 0 and 50 mean that you don't feel favorable and warm toward the group. You would rate them at the 50 degree mark if you don't feel particularly warm or cold toward them. How would you rate feminists who **rarely** talk about feminist issues? [feeling thermometer]
- *Vocal Treatment*: We'd like to get your feelings toward some people and groups who are in the news these days using something we call the feeling thermometer. Ratings between 50 and 100 degrees mean that you feel favorable and warm toward the person or group. Ratings between 0 and 50 mean that you don't feel favorable and warm toward the group. You would rate them at the 50 degree mark if you don't feel particularly warm or cold toward them. How would you rate feminists who **frequently** talk about feminist issues? [feeling thermometer]

For control group:

- *Conditional ID*: Does the way feminists act influence your likelihood of identifying as a feminist? [yes, definitely / maybe / no]

For quiet treatment group:

- *Conditional ID*: If most feminists **rarely** talked about feminist issues, would you be more or less likely to identify as a feminist? [more likely / less likely / the same]

For vocal treatment group:

- *Conditional ID*: If most feminists **frequently** talked about feminist issues, would you be more or less likely to identify as a feminist? [more likely / less likely / the same]

Mediation Analysis

In the main text, we found that when individuals were primed to think about “quiet” feminists (as opposed to “vocal” feminists or feminists generally) they expressed greater feelings of warmth towards—and willingness to identify with—feminists as a group. After examining partisan heterogeneity, we found that this effect was driven by Republican respondents, whereas results for Democratic respondents were more nuanced.

In the analyses in the main text, we treat affect towards feminists and feminist identification as distinct, analyzing each of these outcomes separately. However, we also argued that theoretically these concepts are linked, such that higher (lower) levels of affect towards feminists should lead to greater (less) willingness to adopt the label for oneself. In this respect, there are two potential avenues through which our treatments might influence individual willingness to identify as feminist. The first is directly. The second is indirectly by changing individuals’ feelings about feminists. Theoretically, this linkage makes intuitive sense. For ease of interpretation and readability, however, in the main text we present simplified analyses in which the path of treatment \rightarrow affect \rightarrow identification is left theoretical and affect and identification are analyzed independently. In this section, we more formally test the link between these two outcomes through mediation analysis (see Imai, Keele and Tingley 2010).

Mediation analysis requires the estimation of two models. In the first model, the mediator (in our case affect towards feminists) is regressed on the treatment. In the second model,

the outcome (in our case feminist identification) is regressed on the treatment *and* the mediator. Estimating these models as a system, researchers can then estimate both the direct and indirect effects of the treatment. In our case, an indirect effect would represent the extent to which our treatments shaped feminist identification due to changes in affect towards feminists. Direct effects would represent the extent to which our treatments influenced identification through any other means.

Here we present the results of four mediation analyses: 1) models where the outcome variable is a binary indicator capturing whether Democratic respondents were *more* likely to identify as feminist (as opposed to “less” or “about the same”); 2) models where the outcome variable is a binary indicator capturing whether Democratic respondents were *less* likely to identify as feminist (as opposed to “more” or “about the same”); 3) models where the outcome variable is a binary indicator capturing whether Republican respondents were *more* likely to identify as feminist (as opposed to “less” or “about the same”); and 4) models where the outcome variable is a binary indicator capturing whether Republican respondents were *less* likely to identify as feminist (as opposed to “more” or “about the same”).¹⁶ We estimate each set of models via 10,000 simulations and use bootstrapped standard errors.

Figure 5 presents the direct and indirect effects for each partisan group across outcomes. We turn first to the results where the outcome captures whether respondents reported being *more* likely to identify as feminist. Here, we find that the results generally support our conclusions from the main text. For Democratic respondents, we see a positive and significant indirect effect, suggesting that willingness to identify as feminist is at least partially attributable to more positive feelings associated with vocal feminists. However, we also see a positive and significant *direct* effect, suggesting that our treatment had an impact on identification outside of its effect on affect. For Republicans, we see the opposite pattern, also consistent with our findings from the main text. For Republicans, receiving the vocal treatment dampens affect towards feminists, which in turn decreases the likelihood that an individual reports being more likely to identify as feminist. Moreover, the absence of a significant direct effect suggests that, for Republicans, the effects of the treatments on identification are largely attributable to underlying feelings about the group.

Next we turn our attention to results when the outcome is being “less likely” to identify as

¹⁶For ease of interpretation we use linear probability models to estimate these effects.

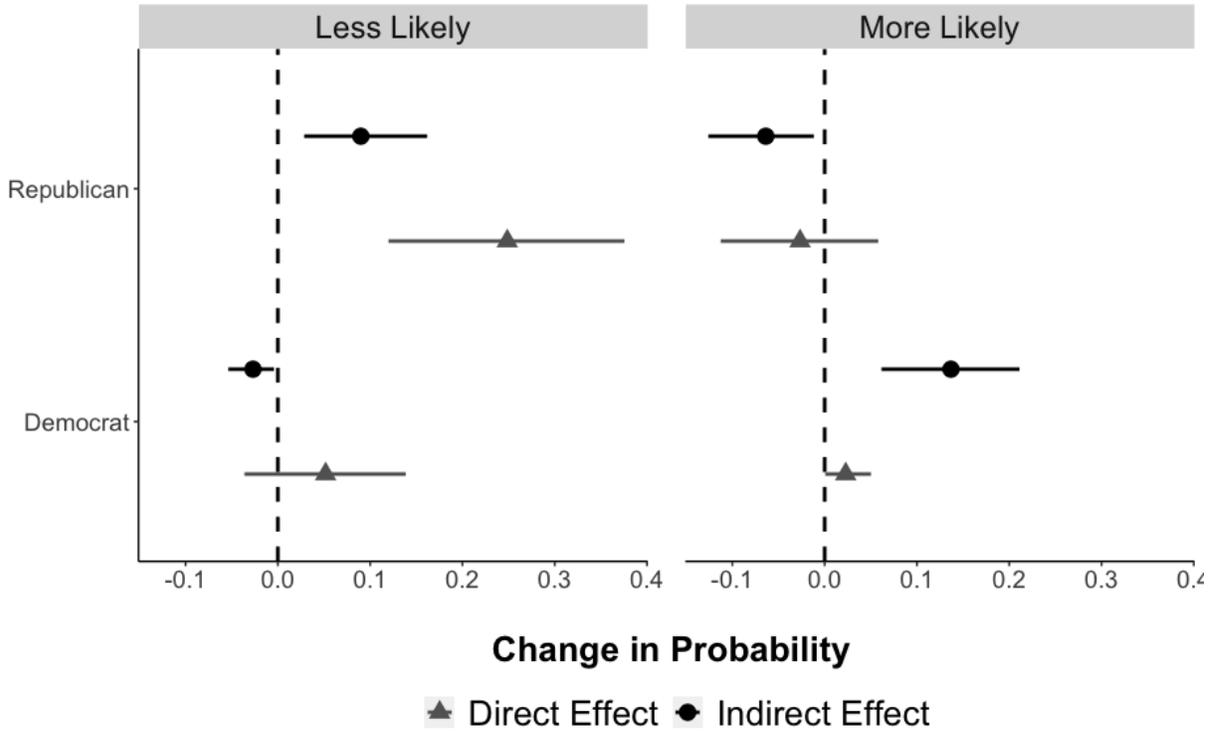


Figure 5: Mediation Results

feminist. Here again, the results lend themselves to similar conclusions. For Democrats, receiving the *vocal* treatment indirectly reduces the likelihood of individuals saying they would be less likely to identify as feminist. In contrast, for Republicans, receiving this treatment *increases* the likelihood of respondents saying they would be less likely to identify as feminist.

0.1 Mediation Sensitivity Analysis

The preceding discussion relied on a series of mediation analyses to more formally test the extent to which increasing (decreasing) positive affect towards feminists leads to an increased (decreased) willingness to adopt the feminist label. In other words, we tested how our treatments influenced self-identification *indirectly* through changes in affect towards feminists as a group. We estimated these models using the procedures outlined in Imai, Keele and Tingley (2010). Central to this endeavor, however, is the assumption of *sequential ignorability*, which requires that two conditions be met. First, there must be no unobserved pre-treatment covariates that influence *both* the treatment (the experimental conditions) and the mediator (feeling thermometer scores) or the outcome (self-identification). Second, there must be no unobserved pre-treatment covariates that influence *both*

affect towards feminists and individual willingness to identify as a feminist.

The assumption of sequential ignorability is quite a strong assumption. In our case, because respondents were randomly assigned to treatment conditions, our analysis should meet the first condition of sequential ignorability. However, because our mediator was *not* randomly assigned, it is possible that some unobserved factor is influencing both affect towards feminists and feminist self-identification, which would violate the second requirement of sequential ignorability. Unfortunately, there is no way to directly test whether this assumption is met in our analysis. However, we *can* assess the robustness of our results to the violation of this assumption using sensitivity analyses.

For each of our four outcomes of interest (models 1 through 4, from above: for both Democrats and Republicans, the propensity to be *more* or *less* likely to identify as a feminist), we calculate the estimated indirect effect of receiving the *vocal* condition (compared to the *quiet* condition) as a function of the correlation between error terms for: 1) models in which the outcome is feeling thermometer scores; and 2) models in which the outcome is being more (less) likely to identify as feminist. This correlation is labeled ρ .

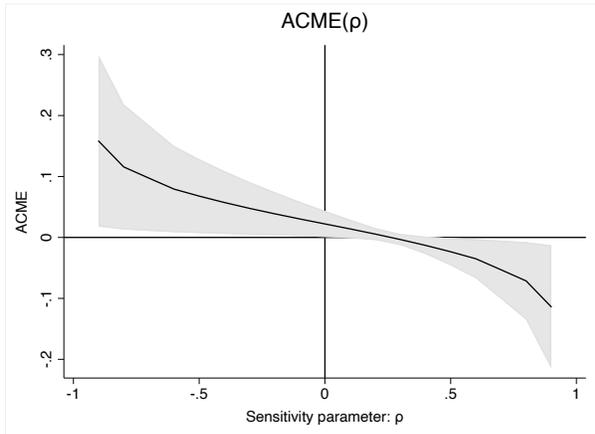
In Figures 6 and 7, we plot the estimated indirect effects across the range of ρ , which allows us to see how large ρ would need to be before we would expect an effect of 0 or an effect in the opposite direction of the one estimated in our models. Small values of ρ indicate small violations of sequential ignorability—in cases where ρ is small and the expected effect is 0, this would indicate that the relatively minor violations of sequential ignorability would invalidate the results. In contrast, higher values of ρ indicate findings that are more robust to the violation of sequential ignorability.

We begin by examining the sensitivity of our results predicting whether respondents reported being *more* likely to identify. These results are presented in Figures 6a and 6b. Again, in this case, the treatment is receiving the *vocal* feminist condition as opposed to the *quiet* feminist condition. Here we see that among Democrats, we would expect the ACME to be zero when ρ is .28. For Republicans, we would expect a zero effect at $\rho = .31$. This indicates that mediation results are more robust for Republicans, which is consistent with our expectations and findings in the main text.

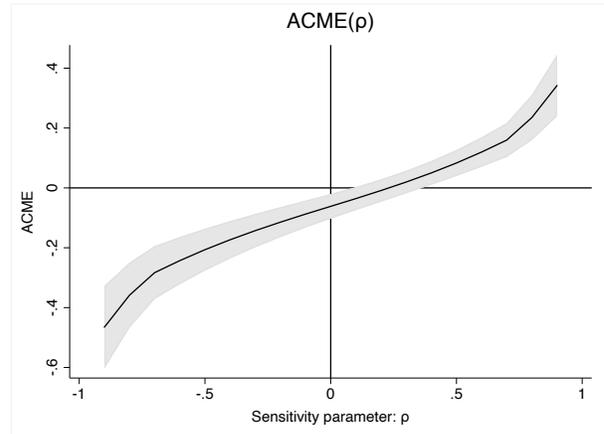
Figures 7a and 7b present the same sensitivity analyses, but for the mediation models where the outcome is whether respondents report being *less* likely to identify as feminist (as opposed

to *more* likely, like in Figures 6a and 6b. Like in the previous analyses, the results appear to be moderately robust to violations of the sequential ignorability assumption, though moreso for Republicans. This is again consistent with expectations and findings in the main text.

Thus, our findings suggest that the quiet and vocal feminist treatments influenced feminist identification (partly for Democrats and largely for Republicans) through changing respondents' affect towards feminists as a group. When respondents felt more warm towards feminists (i.e., when Democrats were responding to vocal feminists and Republicans were responding to quiet feminists), they were then subsequently more likely to identify as feminist. Conversely, when respondents felt more *cold* towards feminists (i.e., when Democrats were responding to quiet feminists and Republicans were responding to vocal feminists), they were then subsequently *less* likely to identify as feminist. This is consistent with our theoretical argument from the main text.

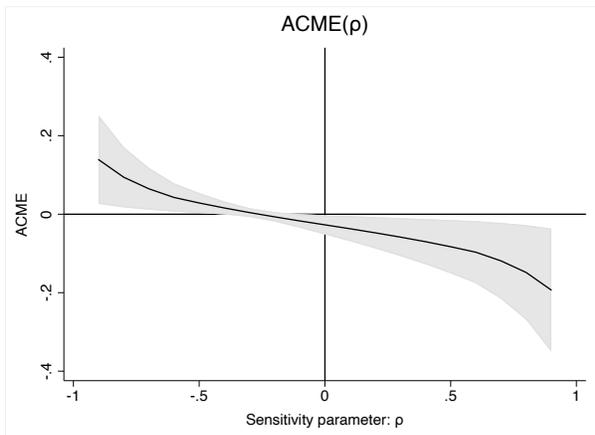


(a) Democrats

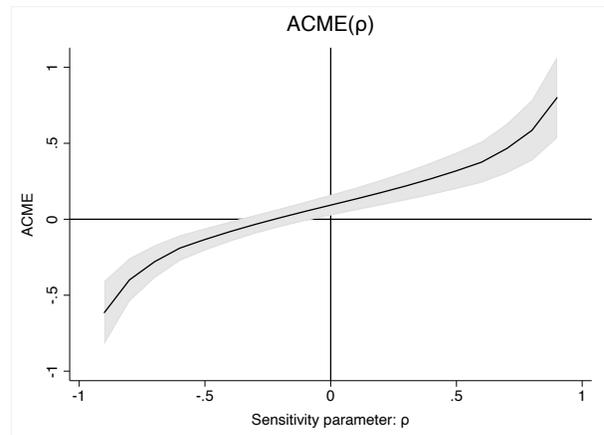


(b) Republicans

Figure 6: Sensitivity Analysis: More Likely to Identify



(a) Democrats



(b) Republicans

Figure 7: Sensitivity Analysis: Less Likely to Identify