

Measuring Norm Pluralism and Perceived Polarization in U.S. Politics

Folco Panizza¹, Eugen Dimant², Erik O. Kimbrough³, and Alexander Vostroknutov⁴

¹Molecular Mind Laboratory, IMT School for Advanced Studies Lucca, Piazza San Francesco, 19, 55100, Italy

²Center for Social Norms and Behavioral Dynamics, University of Pennsylvania

³Smith Institute for Political Economy and Philosophy, Chapman University

⁴Department of Microeconomics and Public Economics, Maastricht University

ABSTRACT

Recent research has shown how norms shape political and economic decision-making. Much of this work assumes that a single norm influences the behavior of all people, but in fact, many situations are characterized by the existence of competing normative viewpoints. We apply a method for measuring belief in the simultaneous existence of multiple norms. Such multiplicity arises naturally when norms are associated with distinctive groups, and thus political polarization can be characterized, in part, as a product of diverging norms between groups. We thus assess the validity of our measurement technique by testing whether it can recover polarized norms about seven salient political issues on which U.S. Democrats and Republicans tend to hold different views. We apply our method to recover beliefs about multiple norms and we compare it to beliefs elicited separately for Democrats and Republicans. Our study uncovers a wide range of views on social norms between and within political groups. Partisans understand their group's norms and hold personal beliefs that align with them. They can also recognize the diversity and polarization in U.S. public opinion by identifying norms specific to political parties and acknowledge the variety of opinions within their own parties, which may indicate internal divisions. This research underscores the importance of nuanced approaches to political norms that go beyond party lines. By acknowledging the plurality of norms, we can encourage productive discussions and bridge ideological divides.

Keywords: social norms, polarization, belief elicitation

1 INTRODUCTION

Political polarization poses a credible threat to democratic values. The political divide runs deep, affecting our perceptions of and interactions with others, which can be difficult to correct (Baldassarri and Bearman, 2007; Bar-On et al., 2024; Braley et al., 2023; Dimant et al., 2024; Druckman, 2023; Graham and Svolik, 2020; Lees and Cikara, 2020). Crucially, polarization is not just a reflection of latent differences between people but is exacerbated by the strong identities that people derive from their membership in social groups. A key aspect of membership in a social group (e.g. Republicans or Democrats in U.S. politics) is conformity to that group's norms. People who identify with one group or another take cues from others in their group about what they ought to do and what they ought to believe (Bicchieri et al., 2022; Dimant et al., 2023). In oppositional contexts like political competition, group members are expected to support their own group's agenda and differentiate their group from others. Consequently, to the extent that individuals are motivated to adhere to group norms, this can lead to divided societies characterized by racial disparities, conflicts based on distinctive identities, and deep-seated partisan hostility (Bursztyn et al., 2020; Dimant, 2023b; Iyengar et al., 2019; Levy, 2021).

Social scientists have long recognized the importance of norm-following motivations in shaping social behavior (e.g. Bicchieri, 2006; Cialdini and Trost, 1998; Sherif, 1936; Tomasello, 2019). Recent work has adopted a simple conceptual framework for understanding norm-driven decision-making in which individuals' choices reflect both a desire to satisfy idiosyncratic self-interest and a desire to live up to the (perceived) expectations of members of their social groups (e.g. Bénabou and Tirole, 2011; Kimbrough

and Vostroknutov, 2016; Krupka and Weber, 2013). When these motivations conflict, individuals face trade-offs, which they resolve differently depending on the strength of their identity with the group and the strength of group norms (McConnell et al., 2018; Pickup et al., 2021; White et al., 2014). However, when individuals' actions are only loosely linked to outcomes, as in politics (Downs, 1957), norms play an especially important role in shaping behavior and beliefs (Groenendyk et al., 2023). Since individual donations or votes are unlikely to sway an election or change the discourse on an issue, the influence of self-interest on decisions becomes negligible, and the influence of norms comes to dominate.

Thus, to understand the consequences of political polarization, we need to understand what group members think one another ought to do and ought to believe, and we also need to understand how members of one group perceive other groups. In other words, we need tools that can measure how people perceive groups' injunctive norms.¹

Conceptually, injunctive norms are understood as sets of shared beliefs about what is appropriate or inappropriate in a given setting (Bicchieri, 2006; Bicchieri and Dimant, 2019). *Personal normative beliefs* are an individual's first-order beliefs about what is appropriate or inappropriate. *Normative expectations* - the key beliefs from which norms get their injunctive force - are second-order beliefs among group members regarding what other group members do and do not approve. Thus, social scientists have recognized that measuring norms requires measuring these normative expectations (Aycinena et al., 2024; Bicchieri and Xiao, 2009; Dimant, 2023a; Krupka and Weber, 2013). To measure normative expectations, most methods require researchers to specify *ex ante* the reference groups within which norms are shared. But sometimes we may want to measure normative beliefs without prior knowledge of what groups exist in the population. This may be useful, for example, when it is unclear which members of the population share normative beliefs, when groups are dynamically changing, or when researchers suspect that there will be within-group heterogeneity.

With this in mind, ref. Panizza et al. (2023) introduces a method of measuring normative expectations when beliefs in the population are heterogeneous and groups are unknown. The key innovation of the "Norm-Drawing" task is to incentive-compatibly measure *third-order* beliefs about the distribution of normative expectations in the population, i.e., what I believe other people in my group understand to be the prevailing norms. In the task, respondents are asked to draw a set of "views" that might be held by others and to estimate what share of the group holds each view that they draw. Respondents' payments depend on how closely their own guess corresponds to the average guesses made by other respondents who were asked about the same reference group.

The authors have shown that the method elicits a multiplicity of beliefs in simple, abstract laboratory tasks used to study prosociality (e.g. heterogeneous views about the appropriate amount of windfall money to give to a stranger in a dictator game; see also Kimbrough et al. (2024)) and that the average belief closely resembles the average normative expectations (second-order beliefs) elicited for similar tasks in the literature. However, their initial test of the method examines contexts in which researchers have no strong *ex ante* reason to believe there are distinct groups with separate norms. Here, we apply the method to the U.S. political landscape, a context in which groups are known *ex ante* to hold distinctive normative views, and we assess the extent to which the method introduced by Panizza et al. (2023) can recover those beliefs.

Diversity of normative views is the very basis of politics. Political identity groups (e.g. parties) are arguably defined by their norms - that is, by their shared expectations that members in good standing ought to support (or oppose) particular policies and candidates. However, the extent to which group members actually share beliefs and understand the heterogeneity of beliefs in the population, especially across partisan lines, remains a point of debate (Dias et al., 2023). Thus, our first goal is to test whether it is possible to elicit the different normative beliefs in the U.S. population without specifying *ex ante* the political groups that compose it. To test this, we ask Democrat-leaning and Republican-leaning partisans to report the norms of their respective groups, and then we compare these beliefs to those elicited using the Norm-Drawing method of Panizza et al. (2023) with a separate sample of U.S. respondents who were asked to report normative beliefs among Americans in general. We predict that this latter exercise should be able to capture the normative beliefs of both Democrats and Republicans, without naming them explicitly.

¹ Social psychologists distinguish injunctive and descriptive norms, where the latter refers to norms of behavior that are followed simply because they are useful guides to action in a particular context and which are learned from observing what others do (Cialdini et al., 1991). We focus on injunctive norms as key motivators to political action.

Empirically our approach is analogous to statistical tests assessing the ability of a model to recover known parameters that are used to simulate data. The modeler knows the data-generating process and seeks to test whether the statistical procedure can extract the underlying signal from noise. Here we have strong priors that Democrats and Republicans differ in their views about a variety of social issues, and we assess the extent to which a new elicitation technique can successfully recover this known heterogeneity when subjects are asked to report their third-order normative beliefs about Americans in general. However, rather than rely on simulation or introspection to define the norms of Democrats and Republicans, we employ existing methods for measuring norms *within a reference group* to provide a basis for comparison to our method.

We compare the output of the Norm-Drawing task to another norm measurement technique due to Krupka and Weber (Krupka and Weber, 2013) that measures second-order beliefs for a particular reference group. In this task, subjects are asked to report a single normative viewpoint and their payment depends on whether their report matches the most common report made by others in their reference group. For clearly defined groups with closely shared norms, the Krupka-Weber task and the Norm-Drawing task should return similar data, but when there is more than one normative view within a reference group (e.g., the U.S. population), the Norm-Drawing task should better capture that heterogeneity.

As a second goal of this study, we also investigate the perceived political divide *within* political groups. We use the Norm-Drawing method to ask partisans about normative beliefs in their political groups, and measure how heterogeneous this third-order perception is.

We thus asked a politically representative sample of U.S. respondents to report their personal and normative beliefs (either second-order using Krupka Weber or third-order using Norm-Drawing) about seven political issues. We find that: (1) when asked to report a single normative view using the Krupka-Weber task, Democrats and Republicans generally agree with group members on a set of norms that distinguish their group from the other and they hold personal views that, on average, are close to their group's norms; (2) when allowed to report multiple views among Americans using the Norm-Drawing task, respondents are generally able to identify party-specific norms, as well as additional heterogeneity that is not captured when we ask just for a single view among Americans (via Krupka-Weber); (3) when allowed to report multiple views *within* their own parties via the Norm-Drawing task, both Democrats and Republicans identify diverse views, some of which differ only slightly from one another and some of which arguably reflect even within-party polarization about certain issues.

Our evidence suggests that the normative landscape of American politics is quite complex and that it is risky to treat partisans' views as monolithic. We suggest that within-party disagreement may provide fruitful ground for political actors who seek to unify views among party members or to attract defectors to other parties whose views on certain key issues are out of alignment with the typical member of their current party.

2 METHODS

The experimental design, original hypotheses, and analyses were pre-registered on the website of the Open Science Framework. Pre-registration and data are available at: osf.io/yh6gd. The study received ethical approval from the University of Maastricht (IRB number: ERCIC_457_31_05_2023).

We adopt a standard 5% significance level to test against the null hypotheses. Post-hoc tests and multiple analyses were corrected for multiple comparisons using a Benjamini-Hochberg correction (Benjamini and Hochberg, 1995).

2.1 Experimental Design

We selected seven issues around which there is an active public debate in the US: abortion, same-sex rights, marijuana legalization, federal spending, gun violence, immigration, and energy investments. The issues were selected based on their currency in the political discussion and the disagreement they generate between as well as within the Democratic and Republican parties. We included and adapted questions from prior surveys (full list available at osf.io/xg2sm) about specific policies (e.g., the legalization of medical marijuana) and asked respondents to evaluate their appropriateness. A policy's appropriateness was evaluated on a three-point scale: appropriate, neither, or inappropriate. The number of policies varied by issue. The full list of policies is included on the online repository (osf.io/2f5mh) and shown in Figures 2 and 6-11.

To understand how respondents perceive public opinion, they were asked questions about the views of other U.S. citizens or sympathizers with either the Republican or Democratic party. The reference group was determined based on respondents' self-reported political identity; we treated as Democrats (Republicans) both subjects who explicitly identified as Democrats (Republicans) and those who identified as Independents but lean Democrat (Republican). Self-identified Democrats and Democratic-leaning respondents answered questions about the beliefs of either "other Democrats" or "Americans". Republicans and Republican-leaning respondents answered questions about the beliefs of either "other Republicans" or "Americans". Finally, respondents who did not identify with either party answered questions about Americans.

The set of questions asked was randomized across participants. Half of the respondents were asked to guess a single normative view among people in the reference group regarding each policy presented. The responses were incentivized to elicit accurate ratings using the method developed by Krupka and Weber (2013), in which respondents are paid a small bonus if their guess on a randomly chosen policy corresponds to the most common guess made by other subjects. We call this the Krupka-Weber task. The other half of the respondents instead answered questions about the distribution of normative views within the reference group. Respondents could report up to nine different views (e.g., the view that it is appropriate to legalize marijuana for medical use but inappropriate to legalize its recreational use or the view that any legalization is inappropriate) and were asked to report the prevalence of that view (in percentage points) within the reference group. This task was also incentivized for accurate responses, based on the method developed by Panizza et al. (2023) in which the overlap between the views drawn by the respondent and the set of views drawn by other participants determines their probability of receiving a small bonus payment. We call this the Norm-Drawing task.

Therefore, the survey was conducted using a 3 (reference group: Democrats versus Republicans versus Americans) by 2 (task: Krupka-Weber versus Norm-Drawing) between-subjects design. The survey proceeded as follows. Participants first answered a series of demographic and political identity questions (including questions about the strength of their partisan identity). They then answered questions about three of the issues, which were randomly selected. For each issue, participants first expressed their personal views and then guessed the views of the assigned reference group using one of the two tasks described above. Participants responded using a graphical interface where they mapped each policy into a rating.² Lastly, participants completed a six-item version of the Cognitive Reflection Test (Frederick, 2005). As part of the instructions, participants had to practice using the interface to make normative judgments. This ensured that they understood how to register their views. Instructions also included detailed rules about how bonus payments were determined. The survey took approximately 9 minutes to complete in the Krupka-Weber version and 16 minutes in the Norm-Drawing version, and participants were paid \$1.60 for completing the survey and could earn an additional \$1.00 bonus based on their response to the Krupka-Weber or Norm-Drawing task.

3 RESULTS

We recruited a politically representative sample of U.S. respondents ($N = 996$, 49% female, 75% college educated) stratified by age, gender, and political orientation from the online platform Prolific. 50% of respondents identified with the Democratic Party, 36% identified with the Republican Party, and 14% reported other or no affiliation.

3.1 Do Democrats and Republicans Have Distinct Norms?

We begin by reporting the personal normative beliefs and normative expectations (i.e. second-order beliefs elicited using the Krupka-Weber task) among reference groups of Democrats and Republicans on each of the issues. This gives us a basis for comparison with the data from the Norm-Drawing task.

Figure 1 displays average personal normative beliefs (1st order beliefs, in grey) and normative expectations elicited in the Krupka-Weber task (2nd order beliefs, in black) for each reference group. Here we focus on the bottom two rows of the figure which show the data for the Democrat and Republican reference groups.

The figure shows a fairly close correspondence between 1st and 2nd order beliefs among both reference groups, which suggests that partisans' own stated views are similar, on average, to the most common view

²Watch a video of the interface here.

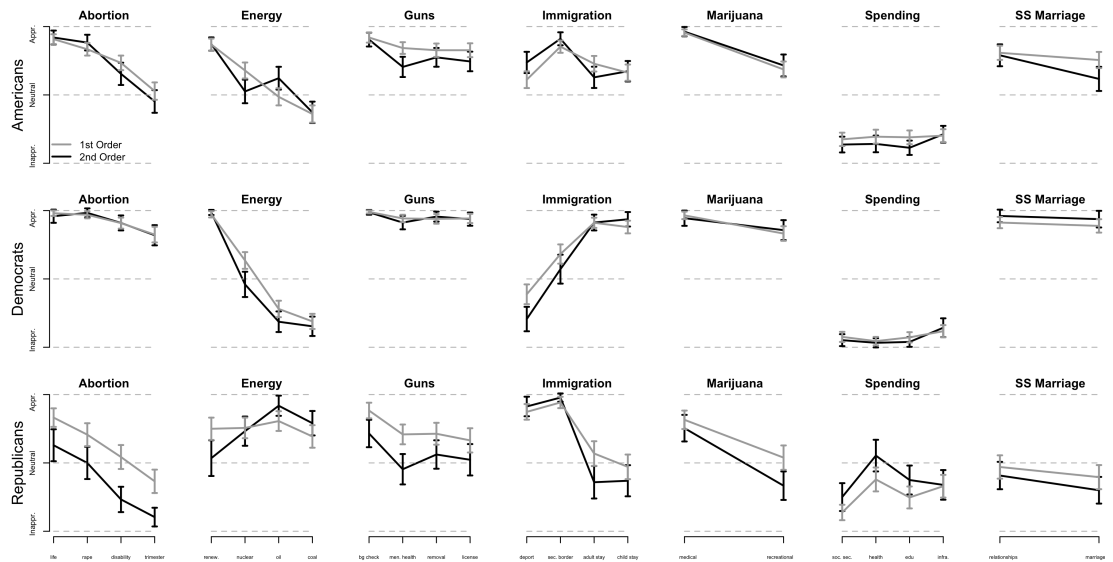


Figure 1. Correspondence between self-reported personal normative beliefs (1st order) and normative expectations elicited via the Krupka-Weber task (2nd order). Means \pm 2 SEs. Each row of the figure corresponds to a single matching group, and each column to a single issue.

they expect co-partisans to hold. For each of the 48 reference group \times issue \times policy combinations, we conduct a χ^2 test of the null hypothesis that 1st and 2nd order beliefs are drawn from the same distribution. For 39 of 48 comparisons, we fail to reject the null (multiple hypothesis corrected p -values > 0.05). That is, on a large majority of issues, among respondents who identify as partisans, the distribution of personal normative beliefs is statistically indistinguishable from the distribution of normative expectations about co-partisans. This close correspondence between personal views and norms comports with previous evidence that partisans are frequently aware of and prone to conform to group norms (Groenendyk et al., 2023).³

The exceptions are intriguing: Republicans report more personal approval of abortion, gun background checks, mental health support, and recreational marijuana than they expect to receive from their co-partisans. Democrats report more personal approval of deporting illegal immigrants and of investments in nuclear power than they expect from their co-partisans. These points of disagreement suggest that there may actually be more heterogeneity in beliefs within-group than is assumed by the Krupka-Weber task, a possibility to which we will return below.

By contrast, χ^2 tests can sharply reject the null hypothesis that the normative expectations of Democrats and Republicans are drawn from the same distribution for all 24 issue \times policy combinations (all p -values < 0.01).⁴ This suggests that not only do members of each group hold coherent normative views on the issues we selected, they also hold views that differ from those held by members of the other party.

Finding 1: *The Krupka-Weber task reveals distinct norms among partisan groups.*

3.2 Can We Recover Democrats' and Republicans' Distinct Norms with the Norm-Drawing Task?

The Norm-Drawing task with Americans as the reference group offers us an opportunity to test to what extent the norms of salient subgroups can be recovered from a sample that contains multiple such groups. As such, we ask whether the norms reported in the Norm-Drawing task include the views elicited with the Krupka-Weber method in the Democrat and Republican reference groups.

To illustrate the output of the Norm-Drawing task in finer detail, consider participants' predictions about normative views regarding immigration policy in the U.S. population (Figure 2, left panel; see

³In Supplementary Analysis 12.5 we compare how close beliefs elicited using Krupka-Weber are to personal views across reference groups. We show that, on average, Democrats' reported norms are closer to their personal views than are Republicans'.

⁴We can similarly reject the null that Democrats and Republicans report the same distribution of personal normative beliefs for 23 out of 24 items.

Supplementary Figures 6-11 for analogous figures for the other issues). Figure 2 shows the three views thought to be most prevalent in the population: one view that strongly favors increasing restrictions on immigration, another view that favors making it easier for immigrants to integrate into the country, and another view that favors both types of policies. The view to which respondents assign the highest prevalence in the Norm-Drawing task (left panel) is similar—though by no means identical—in shape to the average normative expectations of Democrats in the Krupka-Weber task (right panel, blue line), and the view with the second highest weight is quite similar to the average normative expectations of Republicans (right panel, red line). This provides qualitative support for the idea that the Norm-Drawing task can recover the norms of each reference group.⁵

However, it is worth noting that respondents expect only 35% of Americans to hold one of these two views. Thus, the Norm-Drawing task reveals that Americans perceive considerable heterogeneity in normative expectations that goes beyond the distinct norms identified by the Krupka-Weber method. On average, participants in the Norm-Drawing task reported 3 to 4 views on each issue about which they were surveyed.

Despite this multiplicity, participants still reported only a small subset of all the possible views, regardless of the issue or of the group for which the prediction was made (χ^2 tests sharply reject the null hypothesis that participants report random views, all $p < .001$). This finding suggests that respondents focused on a set of salient views that represent real views held by some part of the population.

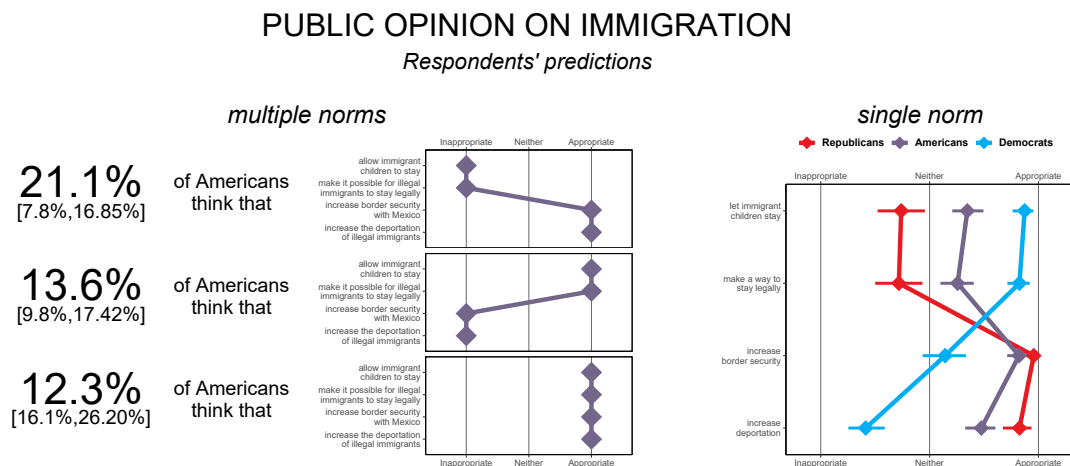


Figure 2. Perceived polarization in the U.S. population on immigration policy. Left: the three views believed to be most common among Americans, as measured by the Norm-Drawing task. Percentages indicate the predicted average share of the population believed to hold that view. Percentages in square brackets indicate the 95% confidence interval. Right: the average of responses in the Krupka-Weber task about the same issue, by reference group. Error bars indicate bootstrap-estimated 95% confidence intervals.

Another subset of respondents were instead asked to guess the single most common normative view among *Americans* for each policy using the Krupka-Weber task; see the top row of Figure 1. This offers an opportunity to test whether, when views are polarized, the Krupka-Weber task will struggle to capture the diversity of opinion in the population. Unsurprisingly, when asked about a single view among all Americans on issues where Democrats and Republicans are polarized, the responses were not very informative. The right panel of Figure 2 exemplifies this problem: while respondents were able to express the presence of contrasting views on immigration in the Norm-Drawing task, the Krupka-Weber data is unable to summarize this diversity of views when the reference group does not correspond to a group with clear normative expectations.⁶

⁵Predictions for the U.S. population do not vary significantly among respondents with different political preferences, Supplementary Analysis 12.4.

⁶In Supplementary Analysis 12.5 we show that personal views and normative expectations in Krupka-Weber are more divergent when the reference group is Americans than when it is Democrats; perhaps surprisingly, there is no significant difference in the gap

More broadly, the ability of respondents to coordinate in the Krupka-Weber task is lower on those issues and in those reference groups where, on average, more diverse views are reported in the Norm-Drawing task. The relation between these two variables is confirmed by a logistic regression with diversity as a predictor of coordination in the Krupka-Weber task ($\beta = -8.75$, $z = -6.78$, $p < .001$, see also Supplementary Analysis 12.3).

Finding 2: *The Norm-Drawing task with Americans as a reference group can recover distinct norms corresponding to the norms of partisan groups. The task reveals more within-group diversity of views than the Krupka-Weber task allows.*

3.3 Can We Use the Norm-Drawing Task within Each Reference Group to Enrich Our Picture of Group Differences?

Next, we ask what can be learned from the two versions of the Norm-Drawing task that used Democrats and Republicans as the reference group.

3.3.1 Polarization between Partisan Groups

One sanity check on the data is whether the views reported in the Norm-Drawing task reflect the same kind of polarization between partisan groups that was identified in the Krupka-Weber task. This is indeed the case, as the Norm-Drawing task confirms that Democrats and Republicans perceive their respective parties to hold very different views: for instance, the share of Democrats believed to find abortion always appropriate is much larger than the share of Republicans holding the same view (Dirichlet regression, $\beta = 1.82$, $z = 10.86$, $p < .001$), and vice versa; the share of Republicans believed to find abortion always inappropriate is much larger than the share of Democrats holding the same view ($\beta = 0.80$, $z = 4.43$, $p < .001$). Similar differences are observed across issues (Supplementary Table 2).

3.3.2 Polarization within Partisan Groups

If Republicans' and Democrats' self-perceived norms are so different, is it also true that they perceive their groups to hold a single, group-specific norm about each issue? If so, then we would expect respondents in the Norm-Drawing task with a partisan reference group to agree upon a single norm, corresponding to the one elicited in the Krupka-Weber task.

Contrary to this prediction, we found evidence of multiple views within each partisan group (Figure 4, central and right panel; Supplementary Figures 12-16). Agreement among co-partisans varied considerably by group and issue. For example, the highest level of agreement was measured among Democrats predicting their own party's views on same-sex rights: the most commonly reported view (that same-sex couples should have the right to both date and marry) was reported by 95% of respondents, and a mean share (the proportion of members believed to hold that view) of 76% (Figure 3, top left). The lowest level of agreement was measured among Republicans predicting their own party's views on gun control: the most frequently reported view (that all proposed policies to reduce gun violence are appropriate) was reported by only 50% of respondents and a mean share of 13% (Supplementary Figure 16).

As a result of this disagreement, Democrats and Republicans differ in how well they are able to coordinate with group members (Figure 4, central panel). Democrats are more likely to agree on what the most common views are within their group than Republicans are (mixed-effects linear regression, $\beta = 0.15$, $t(478) = 9.07$, $p < .001$). If we overlay a participant's guess with the average guess of other participants, the overlap would be on average 15 percentage points (95% CI = [11.8pp, 18.7pp]) larger if the participant was a Democrat rather than a Republican. Respondents in the Americans reference group, composed of respondents across the entire political spectrum, agree less than Democrats ($\beta = -0.98$, $t(476) = -7.30$, $p < .001$) but more than Republicans ($\beta = 0.55$, $t(478) = 3.41$, $p < .001$). Interestingly, there appears to be one issue on which Democrats tend to agree less on their views than Republicans: immigration, where the average agreement is 23.5% (95% CI = [20.7%, 26.9%]), compared to an average agreement of 29.5% ([25.1%, 33.9%]) among Republicans and 29.0% ([25.8%, 32.3%]) in the Americans reference group (see also Supplementary Figure 12).

To estimate agreement, we proceeded as follows. For each participant, we computed the average share of each view reported by other participants within the same reference group, then for each view we computed the minimum of the share reported by the participant and the average share, according to

for the Republican reference group.

the formula: $\sum_{i=1} \min\{s_i, E[s_i]\}$, where s_i is the share of people assumed to hold view i and $E[s_i]$ is the average of the shares assigned to view i by all other participants (see also Panizza et al., 2023). If the participant did not report view i , or if only the participant reported view i , then $\min\{s_i, E[s_i]\} = 0$. Once we computed agreement, we ran a linear mixed-effects regression with view agreement as the dependent variable, reference group as the predictor, and issue and participant id as random intercepts. To test for differences between reference groups, we performed pairwise contrasts.

In the Appendix, we also report regression analysis examining how well subjects coordinate their reported views across *tasks*, by reference group and issue. We show that coordination rates are highest in the Krupka-Weber among Democrats; coordination is lower in the Norm-Drawing task and generally lower when the reference group is Americans or Republicans. Subjects with higher CRT scores are better able to coordinate.

Issues and groups differed not only in the extent to which respondents shared a common perception of their group's norms but also in the extent to which group members' *own* views were perceived as polarized or fragmented (Figure 4, right panel).

Consider the views reported by respondents about same-sex rights and marijuana legalization (Figure 3). In the case of same-sex rights, views are clearly divided along party lines, with Democrats thought to rally behind the view that same-sex relationships and marriages are appropriate, whereas Republicans think that a vast majority of them would disapprove both. When dealing with the issue of marijuana legalization, however, views are much more sparse, especially among Republicans, who see themselves as much more divided. Fragmentation of views is also visible on gun violence (Supplementary Figure 16): the three views that are perceived to be most common among Republicans are that none of the listed policies are appropriate (predicted $M = 14.4\%$ 95% CI = [8.4%, 20.4%]), that all listed policies are appropriate (predicted $M = 13.0\%$ 95% CI = [6.0%, 19.9%]), or that all policies are neither appropriate nor inappropriate (predicted $M = 8.0\%$ 95% CI = [2.1%, 13.9%]). Conversely, Democrats see themselves as much more united on this issue: the two most common views are that all policies are appropriate (predicted $M = 62.0\%$ 95% CI = [54.8%, 69.2%]), or that all policies are appropriate except for increasing funding for mental health screening and treatment, which is rated as neither appropriate nor inappropriate (predicted $M = 9.9\%$ 95% CI = [4.4%, 15.4%]).

At the same time, even within the same issue, respondents may perceive consensus on some policies but disagreement on others (Figure 5). For instance, Democrats overwhelmingly report views agreeing that investing in renewable energy is appropriate, whereas they perceive disagreement on the appropriateness of investing in nuclear energy. Similarly, Republicans' perceive views that coincide more on the inappropriateness of aborting before the end of the second trimester, than they do on whether abortion is appropriate when the baby risks severe health problems.

It is noteworthy that perceived polarization and pluralism are measured not only at the group level but also at the individual level: indeed, a large majority of respondents (53-92%, depending on the issue and reference group, Table 1) report at least two views that disagree on what policy is appropriate (e.g., it is appropriate to legalize all uses of marijuana versus it is appropriate to legalize only medical use). The highest proportions of respondents reporting no conflicting views are on the issue of cutting the U.S. government budget, where Democrats' most frequently reported view is that it is inappropriate to cut spending at all, and same-sex rights, where Republicans mostly perceive themselves as not supporting marriages and relationships.

Our data also offers an opportunity to compare the fragmentation of reported views in the Norm-Drawing task across reference groups. We estimate linear mixed-effects regressions with view diversity as the predicted variable, reference group as the predictor, and issue and participant ID as random intercepts. To test for differences between reference groups, we performed pairwise contrasts. A test across reference groups reveals that respondents in the American reference group report views about Americans that are much more divided than the separate views of Republicans or Democrats about themselves (Figure 4, right panel; Americans > Republicans: $\beta = 0.13$, $t(477) = 4.92$, $p < .001$; Americans > Democrats: $\beta = 0.28$, $t(477) = 12.35$, $p < .001$). In addition, Republicans tend to see themselves as more divided than Democrats do ($\beta = 0.14$, $t(478) = 5.07$, $p < .001$).

Finding 3: *The Norm-Drawing task reveals both polarization between partisan groups and varying degrees of diversity within partisan groups.*

Issue	Americans			Democrats			Republicans		
	ON	SA	C	ON	SA	C	ON	SA	C
Abortion	14.7	0.0	85.3	17.1	0.0	82.9	13.0	4.3	82.6
Energy	12.3	3.7	84.0	11.0	1.4	87.7	18.9	0.0	81.1
Gun violence	12.5	0.0	87.5	23.7	0.0	76.3	7.9	0.0	92.1
Immigration	10.8	8.4	80.7	16.7	1.5	81.8	10.0	5.0	85.0
Marijuana	7.1	1.0	91.8	15.0	0.0	85.0	17.8	0.0	82.2
Spending cuts	13.4	5.2	81.4	20.6	26.5	52.9	20.0	5.7	74.3
Same-sex rights	12.1	2.0	85.9	31.6	0.0	68.4	16.7	13.9	69.4

Table 1. Normative views reported in the Norm-Drawing task, by issue, by reference group. ON: proportion of respondents who reported only one norm; SA: proportion of respondents who reported multiple normative views, but with the same policies rated as appropriate; C: proportion of respondents who reported multiple normative views with different policies rated as appropriate (i.e. conflicting norms).

4 CONCLUSION

This study explores how norm pluralism and polarization shape the perception of several political issues in the United States. To do so, we use a new elicitation method as introduced by (Panizza et al., 2023) and apply it to the context of contentious topics in U.S. politics. Our research reveals that U.S. Democrats and Republicans hold distinct normative views on political issues, and that there is nevertheless significant within-party diversity in beliefs. Our analyses show how the Norm-Drawing task is able to recover this nuanced political landscape by capturing the multiplicity of norms both between and within political groups.

Our analysis in this Research Report only scratches the surface of the richness of our data, but we believe it highlights the value of exploring the whole normative landscape of American politics through the eyes of participants. On the one hand, we see clear evidence of polarization in the fact that partisan groups are generally able to coordinate on a single norm when asked to do so. But our Norm-Drawing task raises questions about the extent to which that apparent polarization masks real within-group heterogeneity. Previous experimental work has shown that being asked to think about political group norms can increase conformity and encourage individuals to engage in forms of costly political expression (Groenendyk et al., 2023; Pickup et al., 2021, 2023; White et al., 2014), but we might ask whether that is partly driven by the fact that these designs treat partisan views as singletons. Perhaps when asked to reflect on heterogeneity of belief among members of their groups partisans will be more comfortable acting in ways that don't just seek to please co-partisans.

5 ACKNOWLEDGMENTS

The authors would like to thank Marco de Carlo, Eric Groenendyk and Mark Pickup for valuable advice on the design of the survey.

6 SUPPLEMENTARY MATERIAL

Supplementary material is available below.

7 FUNDING

This work is supported in part by funds from the International Foundation for Research in Experimental Economics.

8 AUTHOR CONTRIBUTIONS STATEMENT

F.P., E.D., E.K., and A.V. developed the experimental design; F.P. designed the survey; F.P. and A.V. conducted the experiment; F.P., E.K., and A.V. analyzed the data, E.K. and F.P. wrote the original draft; and F.P., E.D., E.K., and A.V. wrote the manuscript.

9 PREVIOUS PRESENTATION

These results were previously presented at Princeton University and the 2024 Polarization Research Lab (PRL) Annual Meeting on Partisan Animosity in America.

10 PREPRINTS

A preprint of this article is published at this link.

11 DATA AVAILABILITY

The data underlying this article are available at: osf.io/yh6gd.

REFERENCES

- Aycinena, D., Bogliacino, F., and Kimbrough, E. O. (2024). Measuring norms: A comparison of the descriptive and predictive power of three methods. Mimeo.
- Baldassarri, D. and Bearman, P. (2007). Dynamics of political polarization. *American Sociological Review*, 72(5):784–811.
- Bar-On, K. K., Dimant, E., Lelkes, Y., and Rand, D. G. (2024). Navigating polarization: Insights from behavioral science. Working paper.
- Bénabou, R. and Tirole, J. (2011). Identity, morals, and taboos: Beliefs as assets. *The Quarterly Journal of Economics*, 126(2):805–855.
- Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, pages 289–300.
- Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press.
- Bicchieri, C. and Dimant, E. (2019). Nudging with care: The risks and benefits of social information. *Public choice*, 191(3):443–464.
- Bicchieri, C., Dimant, E., Gächter, S., and Nosenzo, D. (2022). Social proximity and the erosion of norm compliance. *Games and Economic Behavior*, 132:59–72.
- Bicchieri, C. and Xiao, E. (2009). Do the right thing: but only if others do so. *Journal of Behavioral Decision Making*, 22(2):191–208.
- Braley, A., Lenz, G. S., Adjodah, D., Rahnama, H., and Pentland, A. (2023). Why voters who value democracy participate in democratic backsliding. *Nature human behaviour*, 7(8):1282–1293.
- Bursztyn, L., Egorov, G., and Fiorin, S. (2020). From extreme to mainstream: The erosion of social norms. *American Economic Review*, 110(11):3522–48.
- Cialdini, R. B., Kallgren, C. A., and Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. In *Advances in experimental social psychology*, volume 24, pages 201–234. Elsevier.
- Cialdini, R. B. and Trost, M. R. (1998). Social influence: Social norms, conformity and compliance.
- Dias, N. C., Pearl, J., and Lelkes, Y. (2023). American partisans vastly underestimate the diversity of other partisans’ attitudes.
- Dimant, E. (2023a). Beyond average: A method for measuring the tightness, looseness, and polarization of social norms. *Economics Letters*, 233:111417.
- Dimant, E. (2023b). Hate trumps love: The impact of political polarization on social preferences. *Management Science*.
- Dimant, E., Galeotti, F., and Villeval, M. C. (2023). Motivated information acquisition and social norm formation. Working Paper Available at SSRN: <https://dx.doi.org/10.2139/ssrn.4525398>.
- Dimant, E., Gelfand, M., Hochleitner, A., and Sonderegger, S. (2024). Strategic behavior with tight, loose, and polarized norms. *Management Science*.
- Downs, A. (1957). *An economic theory of democracy*. Harper and Row.
- Druckman, J. N. (2023). Correcting misperceptions of the other political party does not robustly reduce support for undemocratic practices or partisan violence. *Proceedings of the National Academy of Sciences*, 120(37):e2308938120.

- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives*, 19(4):25–42.
- Graham, M. H. and Svolik, M. W. (2020). Democracy in America? Partisanship, Polarization, and the Robustness of Support for Democracy in the United States. *American Political Science Review*, 114(2):392–409.
- Groenendyk, E., Kimbrough, E. O., and Pickup, M. (2023). How norms shape the nature of belief systems in mass publics. *American Journal of Political Science*, 67(3):623–638.
- Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., and Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, 22:129–146.
- Kimbrough, E. O., Krupka, E. L., Kumar, R., Murray, J., Ramalingam, A., Sánchez-Franco, S., Sarmiento, O. L., Kee, F., and Hunter, R. (2024). On the stability of norms and norm-following propensity: A cross-cultural panel study with adolescents. *Experimental Economics*.
- Kimbrough, E. O. and Vostroknutov, A. (2016). Norms make preferences social. *Journal of the European Economic Association*, 14(3):608–638.
- Krupka, E. L. and Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, 11(3):495–524.
- Lees, J. and Cikara, M. (2020). Inaccurate group meta-perceptions drive negative out-group attributions in competitive contexts. *Nature Human Behaviour*, 4(3):279–286.
- Levy, R. (2021). Social media, news consumption, and polarization: evidence from a field experiment. *American Economic Review*, 111(3):831–70.
- McConnell, C., Margalit, Y., Malhotra, N., and Levendusky, M. (2018). The economic consequences of partisanship in a polarized era. *American Journal of Political Science*, 62(1):5–18.
- Panizza, F., Dimant, E., Kimbrough, E. O., and Vostroknutov, A. (2023). Measuring norm pluralism and tolerance. *Available at SSRN*.
- Pickup, M., Kimbrough, E. O., and de Rooij, E. A. (2021). Expressive politics as (costly) norm following. *Political behavior*, pages 1–21.
- Pickup, M., Kimbrough, E. O., and de Rooij, E. A. (2023). Crossing the line: Evidence for the categorization theory of spatial voting. *British Journal of Political Science*, page 1–11.
- Sherif, M. (1936). The psychology of social norms.
- Tomasello, M. (2019). *Becoming human: A theory of ontogeny*. Harvard University Press.
- Volgenant, T. and Jonker, R. (1987). On some generalizations of the travelling-salesman problem. *Journal of the Operational Research Society*, 38(11):1073–1079.
- White, I. K., Laird, C. N., and Allen, T. D. (2014). Selling out?: The politics of navigating conflicts between racial group interest and self-interest. *American Political Science Review*, 108(4):783–800.

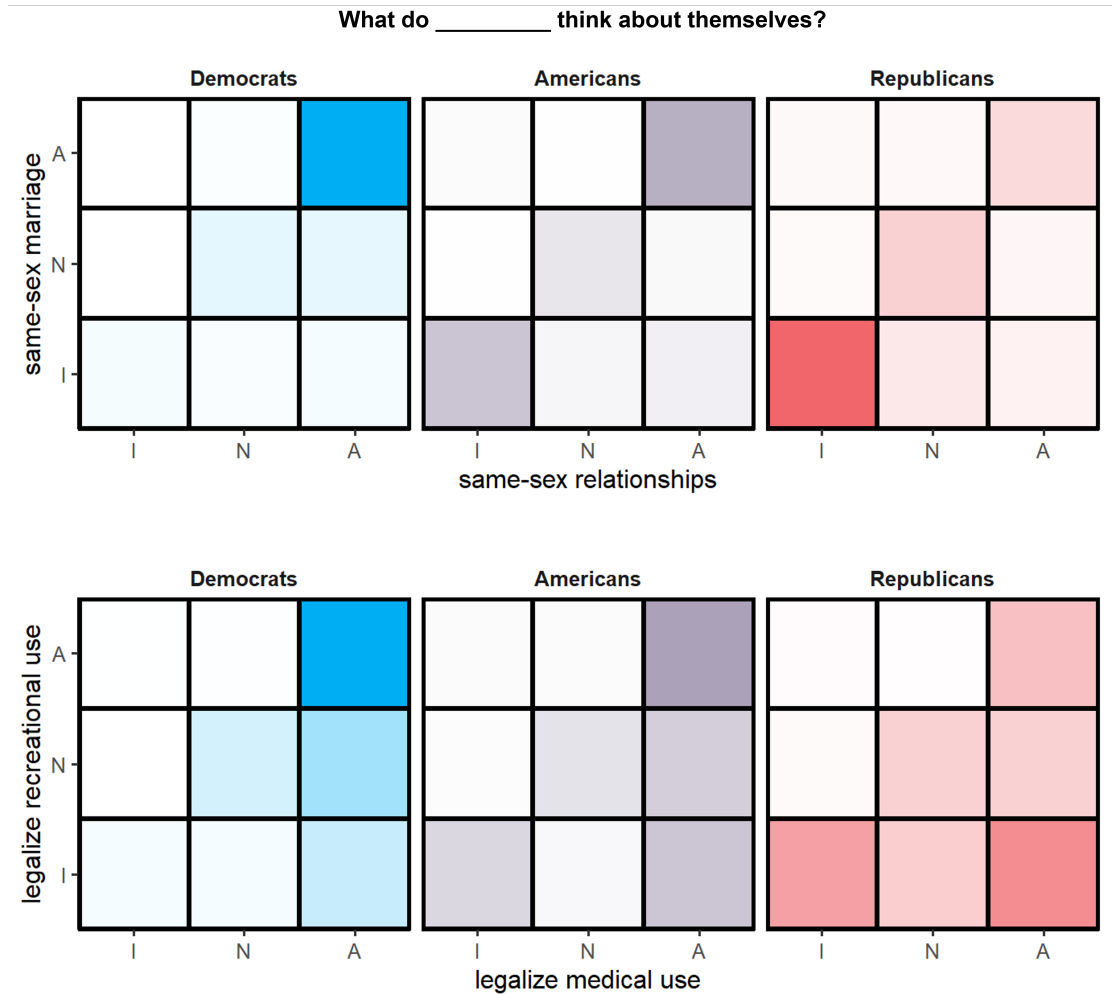


Figure 3. The share of the different possible views reported in the Norm-Drawing task on same-sex rights (top) and marijuana legalization (bottom), by reference group. Each tile is shaded by the average share of people believed to hold that normative view: the darker the shade, the larger the share of people. Views on marijuana legalization are more heterogeneous than views on same-sex rights. In addition, respondents in the Americans reference group report a joint distribution of the Republican and Democratic views. A: appropriate; N: neither appropriate nor inappropriate; I: inappropriate.

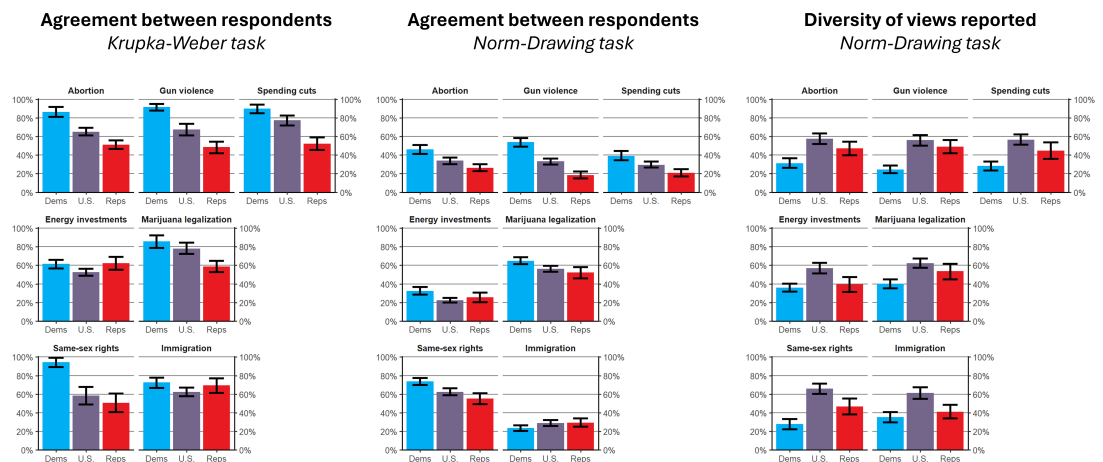


Figure 4. Agreement in responses and diversity of views. Left: the probability that a respondent rating matches the most common rating among other respondents in the same reference group, by issue in the Krupka-Weber task. A higher likelihood of a match indicates greater agreement on the most common normative view in that group. Center: the overlap of respondents' predictions with the average of other respondents in the same reference group, by issue in the Norm-Drawing task. Higher overlap indicates greater agreement in the views reported. Democrats tend to agree much more than Republicans on all issues, except for immigration policy. Right: diversity of views reported by respondents in each reference group, by issue in the Norm-Drawing task. Respondents predicting U.S. public opinion consistently report a higher diversity of views than respondents predicting views within Democrats or Republicans. Error bars indicate bootstrap-estimated 95% confidence intervals.

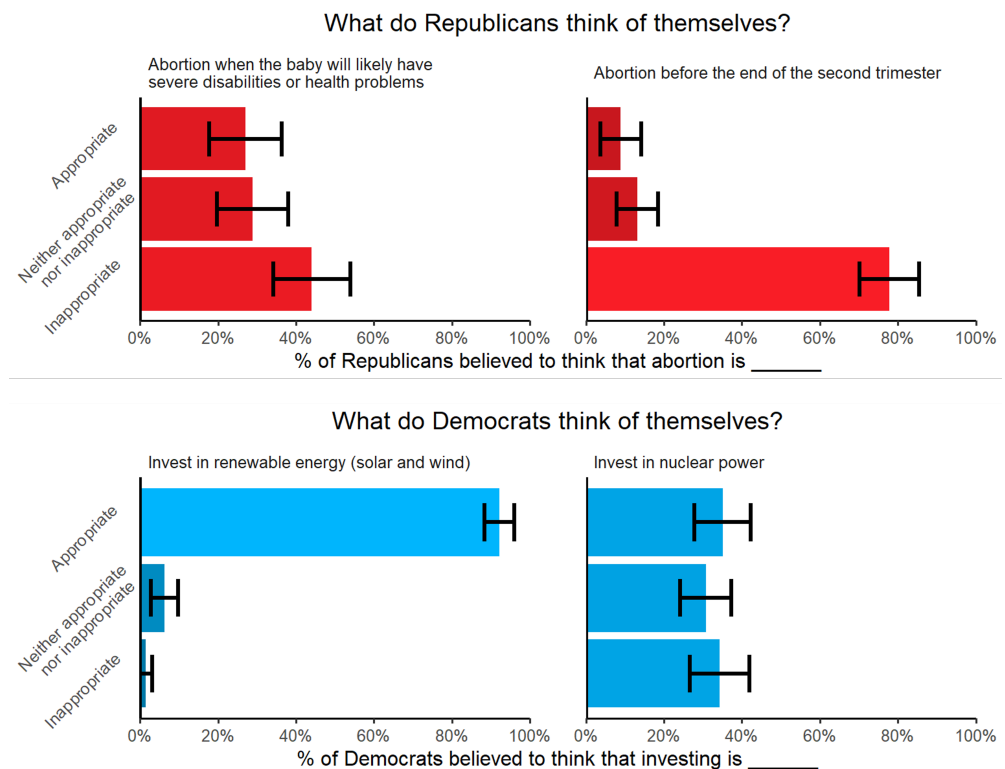


Figure 5. Policy agreement and disagreement. Top row: Republicans perceive agreement that abortion is inappropriate before the end of the second trimester, but perceive disagreement when the baby is at risk of severe health problems. Bottom row: Democrats perceive agreement that it is appropriate to invest in renewable energy sources, but they perceive disagreement when it comes to investing in nuclear energy. Error bars indicate 95% confidence intervals.

12 APPENDIX: SUPPLEMENTARY ANALYSES

12.1 Perceived normative landscape, by reference group

Figures 6-11 show the 3 most frequently drawn views in the Norm-Drawing task for the Americans reference group (left panel) alongside the average view reported in the Krupka-Weber for each reference group (right panel).

Figures 12-16 show the distribution of views reported in the Norm-Drawing task for each reference group regarding immigration, spending, energy, abortion, and gun control.

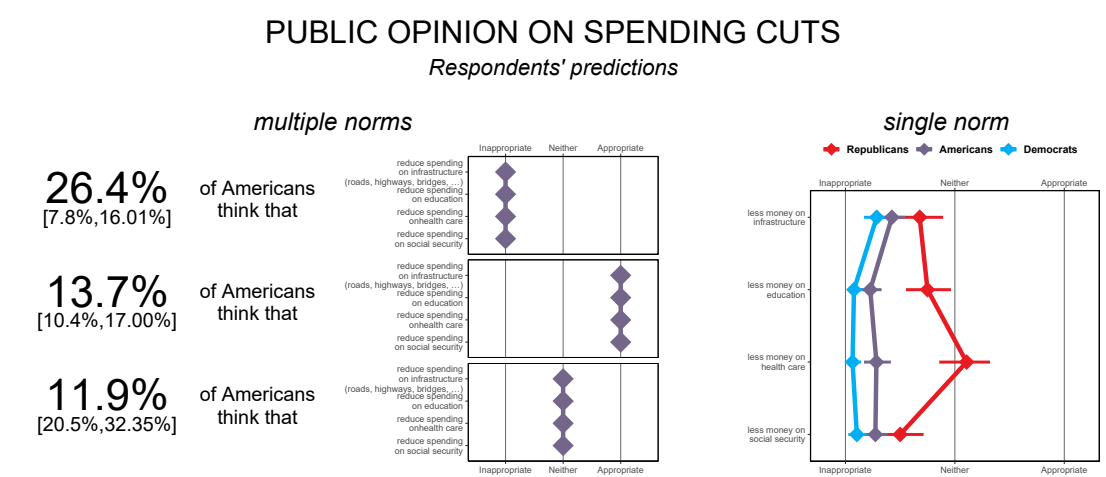


Figure 6. Perceived polarization in the U.S. population on government spending. Left: the three views believed to be most common among Americans, as measured by the Norm-Drawing task. Percentages indicate the predicted average share of the population believed to hold that view. Percentages in square brackets indicate the 95% confidence interval. Right: the average of responses in the Krupka-Weber task about the same issue, by reference group. Error bars indicate bootstrap-estimated 95% confidence intervals.

PUBLIC OPINION ON SAME SEX RIGHTS

Respondents' predictions



Figure 7. Perceived polarization in the U.S. population on same sex relationships. Left: the three views believed to be most common among Americans, as measured by the Norm-Drawing task. Percentages indicate the predicted average share of the population believed to hold that view. Percentages in square brackets indicate the 95% confidence interval. Right: the average of responses in the Krupka-Weber task about the same issue, by reference group. Error bars indicate bootstrap-estimated 95% confidence intervals.

PUBLIC OPINION ON ABORTION

Respondents' predictions

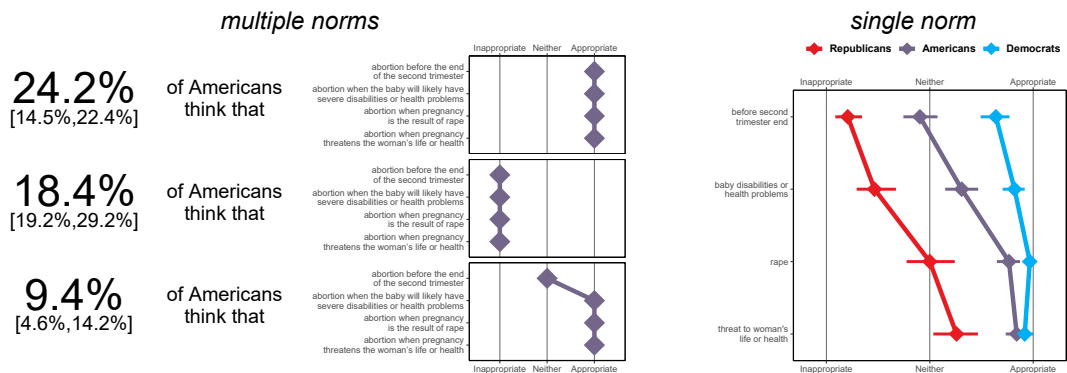


Figure 8. Perceived polarization in the U.S. population on abortion policy. Left: the three views believed to be most common among Americans, as measured by the Norm-Drawing task. Percentages indicate the predicted average share of the population believed to hold that view. Percentages in square brackets indicate the 95% confidence interval. Right: the average of responses in the Krupka-Weber task about the same issue, by reference group. Error bars indicate bootstrap-estimated 95% confidence intervals.

PUBLIC OPINION ON ENERGY INVESTMENTS

Respondents' predictions

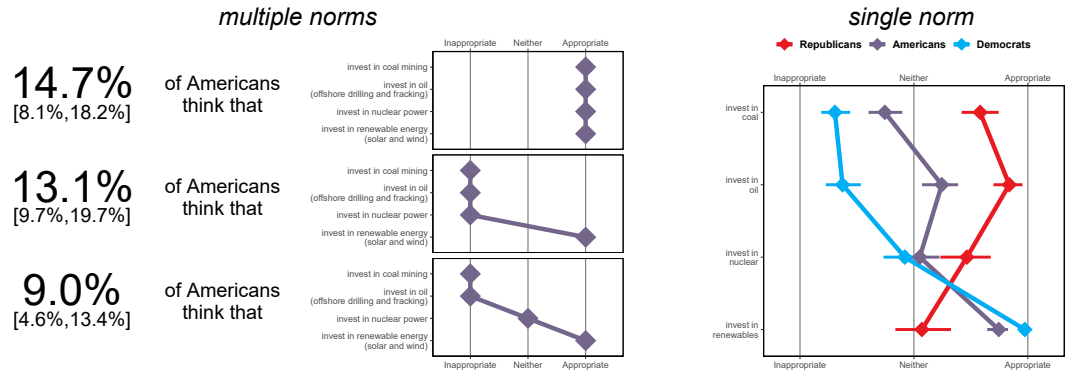


Figure 9. Perceived polarization in the U.S. population on energy investments. Left: the three views believed to be most common among Americans, as measured by the Norm-Drawing task. Percentages indicate the predicted average share of the population believed to hold that view. Percentages in square brackets indicate the 95% confidence interval. Right: the average of responses in the Krupka-Weber task about the same issue, by reference group. Error bars indicate bootstrap-estimated 95% confidence intervals.

PUBLIC OPINION ON MARIJUANA LEGALIZATION

Respondents' predictions

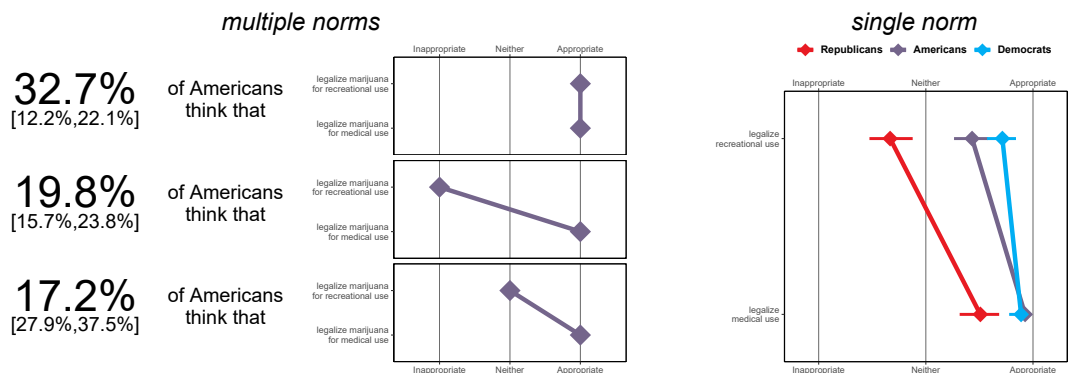


Figure 10. Perceived polarization in the U.S. population on marijuana legalization. Left: the three views believed to be most common among Americans, as measured by the Norm-Drawing task. Percentages indicate the predicted average share of the population believed to hold that view. Percentages in square brackets indicate the 95% confidence interval. Right: the average of responses in the Krupka-Weber task about the same issue, by reference group. Error bars indicate bootstrap-estimated 95% confidence intervals.

PUBLIC OPINION ON GUN VIOLENCE

Respondents' predictions

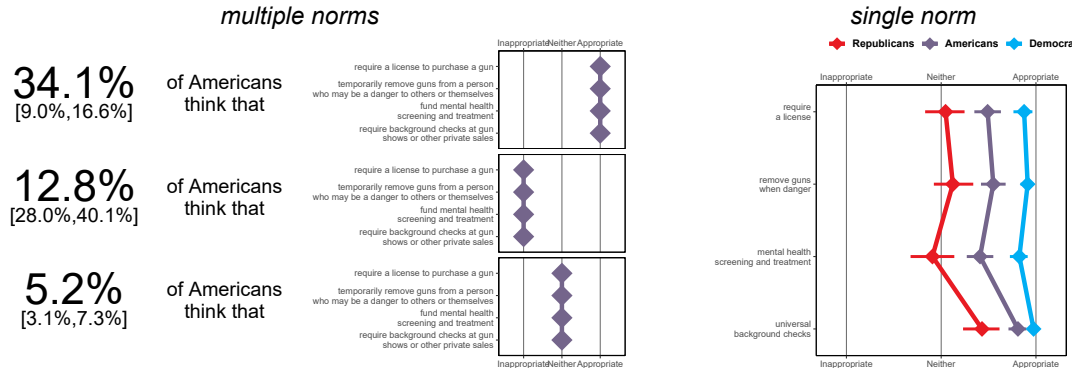


Figure 11. Perceived polarization in the U.S. population on gun control. Left: the three views believed to be most common among Americans, as measured by the Norm-Drawing task. Percentages indicate the predicted average share of the population believed to hold that view. Percentages in square brackets indicate the 95% confidence interval. Right: the average of responses in the Krupka-Weber task about the same issue, by reference group. Error bars indicate bootstrap-estimated 95% confidence intervals.

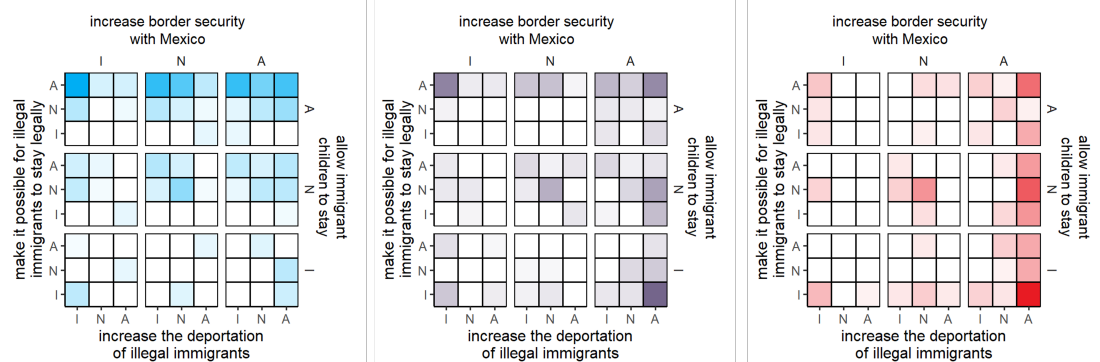


Figure 12. The share of the different possible views reported in the Norm-Drawing task on immigration, by reference group. Each tile is shaded by the average share of people believed to hold that normative view: the darker the shade, the larger the share of people. A: appropriate; N: neither appropriate nor inappropriate; I: inappropriate.

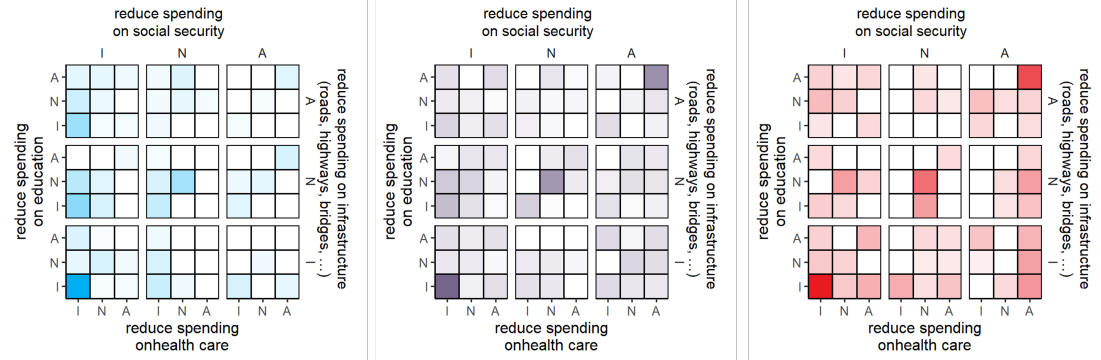


Figure 13. The share of the different possible views reported in the Norm-Drawing task on spending cuts, by reference group. Each tile is shaded by the average share of people believed to hold that normative view: the darker the shade, the larger the share of people. A: appropriate; N: neither appropriate nor inappropriate; I: inappropriate.



Figure 14. The share of the different possible views reported in the Norm-Drawing task on abortion, by reference group. Each tile is shaded by the average share of people believed to hold that normative view: the darker the shade, the larger the share of people. A: appropriate; N: neither appropriate nor inappropriate; I: inappropriate.

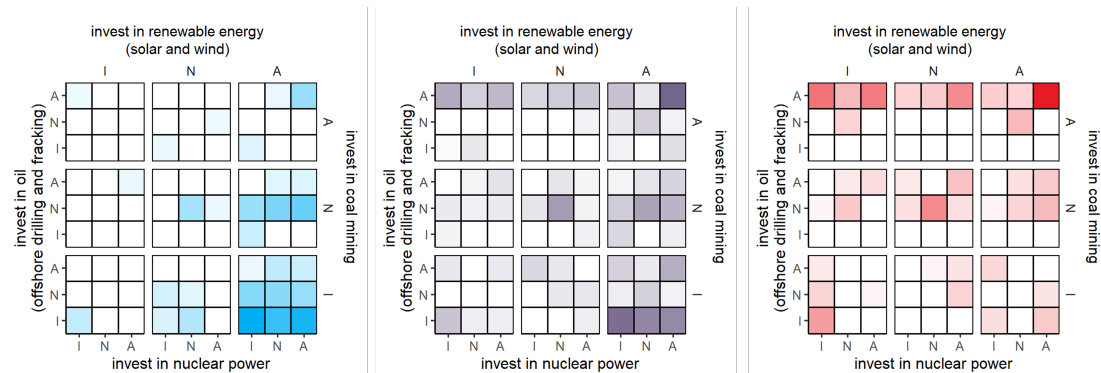


Figure 15. The share of the different possible views reported in the Norm-Drawing task on energy investments, by reference group. Each tile is shaded by the average share of people believed to hold that normative view: the darker the shade, the larger the share of people. A: appropriate; N: neither appropriate nor inappropriate; I: inappropriate.

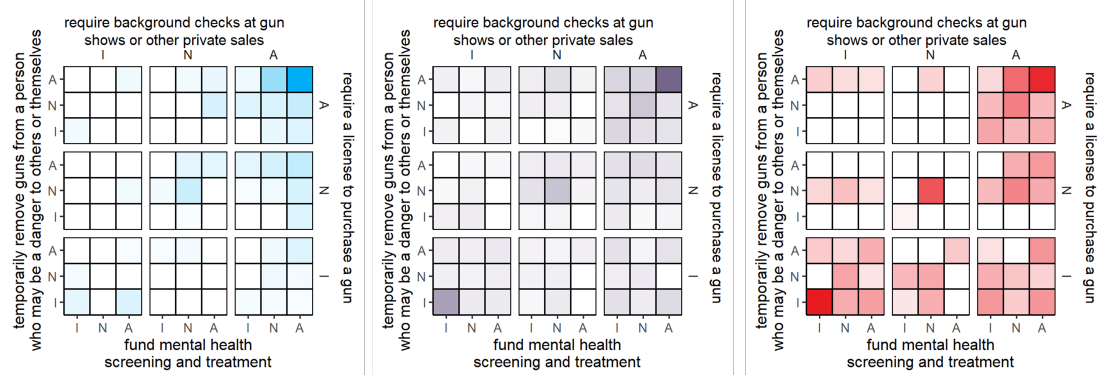


Figure 16. The share of the different possible views reported in the Norm-Drawing task on gun violence, by reference group. Each tile is shaded by the average share of people believed to hold that normative view: the darker the shade, the larger the share of people. A: appropriate; N: neither appropriate nor inappropriate; I: inappropriate.

12.2 Differences between the Republican and Democratic reference group in the Norm-Drawing task

Topic	View	Estimate	Std. Error	Z-value	Pr
Abortion	IIII	0.799	0.180	4.434	< 0.001
Abortion	AAAA	-1.818	0.167	-10.861	< 0.001
Marijuana	II	0.701	0.181	3.871	< 0.001
Marijuana	AI	0.467	0.179	2.614	0.020
Marijuana	AN	-0.536	0.182	-2.950	0.010
Marijuana	AA	-1.263	0.177	-7.130	< 0.001
Spending	IIII	-0.879	0.180	-4.885	< 0.001
Energy	AIII	-1.071	0.194	-5.520	< 0.001
Energy	ANII	-0.689	0.197	-3.503	0.006
Energy	AAII	-0.738	0.196	-3.776	0.003
Energy	AAAA	1.174	0.185	6.352	< 0.001
Guns	IIII	0.630	0.190	3.309	0.027
Guns	AAAA	-2.368	0.167	-14.143	< 0.001
Immigration	IIAA	-0.590	0.194	-3.040	0.046
Immigration	AAII	1.213	0.186	6.526	< 0.001
Immigration	AANN	0.685	0.193	3.556	0.011
Same-Sex	II	1.584	0.189	8.376	< 0.001
Same-Sex	AA	-2.140	0.190	-11.275	< 0.001

Table 2. Views for each topic that differ between the Republican and Democratic reference groups according to Dirichlet regressions run by issue, filtered for significant differences (corrected for multiple comparisons). A positive estimate means that the view receives higher shares within the Republican reference group, whereas a negative estimate means that the view receives higher shares within the Democratic reference group. Each letter represents the rating for a policy: A: Appropriate, I: Inappropriate N: Neither.

12.3 Diversity of views in the Norm-Drawing task and coordination ability in the Krupka-Weber task

To measure how responses in the Norm-Drawing task predict coordination in the Krupka-Weber, we proceed as follows. We define coordination as whether a participant's rating of the appropriateness of a policy matches the most frequent rating given by other respondents. For example, if the participant guessed that most Democrats would find same-sex relationships appropriate, and the majority of other respondents in the Democrat reference group also rated same-sex relationships as appropriate, then there is a match. As a predictor of coordination, we look at the diversity of normative views for each issue and for each reference group in the Norm-Drawing task. To measure the diversity of views, we first compute the average set of views reported in the Norm-Drawing task. We do this by calculating the average proportion of each view reported. For example, if a view was reported by one out of three participants, and that participant thought that 15% of people held that view, then since the other two participants did not report that view, the average share is $(0\% + 0\% + 15\%)/3 = 5\%$. After calculating the average of the multiple views, we calculate their diversity. To measure the diversity of views on a given issue for a given reference group, we compute the following formula:

$$d = 2 \times \sqrt{\sum_i s_i \delta_i} \quad (1)$$

Where s_i is the proportion of people holding the i th view and δ_i is the normalized Euclidean distance between view i , and the weighted average of all views reported:

$$\delta = \sqrt{\sum_j (p_j - \bar{p}_j)^2 / N_p} \quad (2)$$

Where p_j is the appropriateness rating for the j th policy of the view (appropriate = 1; neither = 0.5; inappropriate = 0), \bar{p}_j is the weighted average rating of the j th policy among all the views reported, and N_p is the number of policies for the given issue. The diversity index d ranges from 0 to 1. It is 0 when only one view is reported and 1 when all views are maximally distant from the average view.

To test whether the diversity of views in the Norm-Drawing task predicts response matching in the Krupka-Weber, we perform a logistic regression with robust standard errors clustered at the issue and reference group level. We include fixed effects for each participant. Diversity negatively predicts response matching ($\beta = -8.75, z = -6.78, p < .001$); that is, as diversity increases, the likelihood of matching the rating of other respondents in the Krupka-Weber decreases.

12.4 Partisan differences in predicting U.S. public opinion in the Norm-Drawing task

Since we have the partisan identity of all respondents, we can also check whether Democrats and Republicans differ in their ability to predict norms among Americans: statistical tests find no significant differences in predictions except that Democrats report the view that policies against gun violence are appropriate slightly less frequently ($\beta = -1.00, z = 5.27, p < .001$) but report more frequently views that it is appropriate to legalize medical marijuana or both recreational and medical use (all $p < .001$).

12.5 Supplementary Hypothesis Test I (pre-registered hypothesis H2B "Distortion"):

If public opinion is split on an issue, then beliefs about the most common normative view elicited via the Krupka-Weber task should not be a good predictor of what people think. We test this hypothesis by measuring the distance between a participant's guess of the most common view and the same participant's personal normative beliefs. We hypothesize that participants predicting Americans' views will report on average personal views that are more distant from their guess about the most common view than participants predicting the most common view among those who identify with their political party. We use 3 to compute the distance between personal and predicted most common view. We then ran a fixed effects linear regression with distance as the predicted variable, reference group as the predictor, and issue and participant ID as random intercepts. Pairwise contrasts reveal that Democrats' predictions about the most common view in the Krupka-Weber task are closer to their personal beliefs than the predictions in the American reference groups ($\beta = -0.19, z = -9.35, p < .001$). However, contrary to expectations, predictions in the Americans reference group are not significantly closer to their personal views than Republicans' are ($\beta = -0.02, z = -0.73, p = .464$).

12.6 Correspondence between Personal Normative Beliefs (1st order) and the Perceived Distribution of Normative Expectations (3rd order)

Since we also collected the personal normative beliefs of respondents in the Norm-Drawing task, it is possible to compare the predicted distribution of views with the reported distribution of personal beliefs in the sample. In this regard, Democrats' guesses correspond more closely to Democrats' personal beliefs than Republicans do to theirs ($\beta = 0.15, z = 12.73, p < .001$). Perhaps surprisingly, respondents in the Americans reference group seem better at predicting the sample's personal beliefs than Republicans are at predicting their own group's personal beliefs ($\beta = 0.03, z = 2.76, p = .006$), although they perform less well than Democrats do at predicting their group's personal beliefs ($\beta = -0.12, z = -12.73, p < .001$).

To estimate the distance between the distribution of (1st order) personal normative beliefs and the (3rd order) distribution of views predicted by participants, we computed the minimum distance between the two using a linear assignment algorithm (Volgenant and Jonker, 1987). To account for the fact that participants made their predictions in percentages, while personal beliefs did not sum to 100 (but rather had an N equal to the size of that group), we changed the number of observations proportionally so that both would sum to the same number ($N = 300$). In other words, 300 "representatives" were "chosen" to proportionally represent the predicted views of the subject, and 300 more represented the personal views in the group. Representatives were selected using a Hamilton quota, and ties were randomly assigned. To determine the minimum distance between the two distributions, each of the 300 representatives of the participant's prediction was matched with the closest of the 300 representatives of the personal beliefs. The distance metric adopted was a variation of Equation 2:

$$\delta_p = \sqrt{\sum_j (p_j^{\text{predicted}} - p_j^{\text{personal}})^2 / N_p} \quad (3)$$

Where $p_j^{\text{predicted}}$ is the appropriateness rating to the j th policy given by the representative of the predicted view, and p_j^{personal} is the appropriateness rating to the j th policy given by the representative

of the personal view. δ_p took values between zero (the predicted view and the personal view were the same) and one (the predicted view and the personal view were the opposite). Once all representatives were matched, we computed the average distance which, like δ_p was bounded between zero and one. After computing the distance between each participant's prediction and the distribution of personal views, we ran a linear mixed-effects regression with view distance as the predicted variable, reference group as the predictor, and issue and participant ID as random intercepts. To test for differences between reference groups, we performed pairwise contrasts.

Moreover, if the Norm-Drawing task captures the diversity of personal views within the reference group, then the diversity of personal views on a given issue for a given reference group should predict the diversity of views reported in the Norm-Drawing task (pre-registered hypothesis 1). The diversity of personal views was computed using the same procedure as in Equation 1. Since the diversity of predicted views was distributed in the form of a zero-inflated (semi-)continuous distribution, we followed a (two-stage) hurdle test of the hypothesis: first, we tested the association using a logistic regression with a dummy that takes value 1 if the diversity is greater than zero, and zero otherwise; second, we ran a linear regression with all predictions with diversity greater than zero. Since the diversity of personal views on an issue is the same for all respondents predicting that reference group, the errors are clustered around the issue and reference group. While diversity in personal views significantly predicted diversity in predictions in the logistic regression ($\beta = 1.59, z = 3.91, p < .001$), the association was only marginally significant for the linear regression ($\beta = 0.55, t(1435) = 3.81, p = .062$). Given the differences between reference groups, we explored whether this lack of significance may have been driven by this factor. We thus re-run the linear regression including the reference group and its interaction with the diversity of personal views as a covariate. When accounting for the reference group in the linear regression, the relationship was significant in the Democrats reference group ($\beta = 0.28, t(510) = 4.72, p = .010$), but not significant in the Republicans ($\beta = 0.27, t(275) = 1.57, p = .214$) and Americans ($\beta = 0.19, t(646) = 1.39, p = .214$) reference groups.

12.7 How Does Coordination Vary by Task and Reference Group?

In Table 3 we report estimates of regression models in which the dependent variable is the overlap between one's own response and the responses of other respondents on a given policy proposal in a given reference group. In the Krupka-Weber task, this is the percentage of respondents who offered the same normative evaluation of that policy, and in the Norm-Drawing task, this is the overlap between the respondent's distribution of views and the average distribution reported by others. We regress these measures on dummy variables for the reference group, the task, and their interaction, with the Americans reference group in the Krupka-Weber captured by the Constant term. We include a second specification with demographic variables, including Age, Sex and Cognitive Reflection Test score. Our estimates report heteroskedasticity robust standard errors.

The table reveals that coordination rates are generally highest among Democrats in the Krupka-Weber and lowest among Republicans in the Norm-Drawing task. Age and sex are not significantly associated with successful coordination, but we see that higher CRT scores are predictive of coordination for all but one issue (Gun Control).

Table 3. Match Rate by Task and Reference Group

	Immigration (1)	(2)	Spending (3)	(4)	SS Marriage (5)	(6)	Abortion (7)	(8)	Energy (9)	(10)	Marijuana (11)	(12)	Gun Control (13)	(14)
Democrats	0.12*** (0.02)	0.12*** (0.02)	0.18*** (0.03)	0.18*** (0.02)	0.44*** (0.03)	0.44*** (0.03)	0.23*** (0.02)	0.22*** (0.02)	0.19*** (0.02)	0.19*** (0.02)	0.08** (0.03)	0.07* (0.03)	0.32*** (0.02)	0.31*** (0.02)
Republicans	0.10*** (0.02)	0.11*** (0.02)	-0.24*** (0.02)	-0.24*** (0.02)	-0.07** (0.02)	-0.07** (0.02)	-0.06** (0.02)	-0.07*** (0.02)	0.09*** (0.02)	0.08*** (0.02)	-0.24*** (0.02)	-0.24*** (0.02)	-0.15*** (0.02)	-0.15*** (0.02)
Mult. Views Task	-0.21*** (0.02)	-0.20*** (0.02)	-0.34*** (0.02)	-0.34*** (0.02)	0.17*** (0.03)	0.17*** (0.03)	-0.21*** (0.02)	-0.22*** (0.02)	-0.20*** (0.01)	-0.20*** (0.01)	-0.11*** (0.02)	-0.10*** (0.02)	-0.20*** (0.02)	-0.20*** (0.02)
Dem. x Mult. Views	-0.17*** (0.03)	-0.17*** (0.03)	-0.09* (0.04)	-0.09* (0.04)	-0.33*** (0.04)	-0.32*** (0.04)	-0.11** (0.04)	-0.10** (0.04)	-0.09** (0.03)	-0.09** (0.03)	0.00 (0.04)	-0.00 (0.04)	-0.11** (0.04)	-0.11** (0.04)
Rep. x Mult. Views	-0.10** (0.04)	-0.10** (0.04)	0.16*** (0.03)	0.15*** (0.03)	-0.00 (0.04)	-0.00 (0.04)	-0.01 (0.03)	-0.01 (0.03)	-0.05 (0.03)	-0.06 (0.03)	0.20*** (0.04)	0.19*** (0.04)	-0.00 (0.03)	0.00 (0.03)
CRT Score		0.01** (0.00)		0.01*** (0.00)		0.01** (0.00)		0.01* (0.00)		0.01** (0.00)		0.02*** (0.00)		0.00 (0.00)
Male		-0.00 (0.01)		-0.01 (0.01)		-0.00 (0.02)		-0.01 (0.01)		-0.01 (0.01)		-0.01 (0.01)		-0.03* (0.01)
Age		-0.00 (0.00)		-0.00 (0.00)		-0.00 (0.00)		-0.00 (0.00)		0.00 (0.00)		-0.00 (0.00)		0.00 (0.00)
Constant (Amer.)	0.50*** (0.01)	0.47*** (0.02)	0.63*** (0.02)	0.60*** (0.03)	0.45*** (0.02)	0.44*** (0.03)	0.55*** (0.01)	0.57*** (0.03)	0.42*** (0.01)	0.39*** (0.02)	0.67*** (0.02)	0.61*** (0.03)	0.53*** (0.02)	0.53*** (0.03)
Observations	399	399	439	439	424	424	431	431	407	407	438	438	447	447
R Sq.	0.541	0.551	0.605	0.617	0.500	0.511	0.487	0.496	0.550	0.562	0.238	0.296	0.622	0.627