

# Testing a Civil Discourse Norms Campaign in Michigan

Wave 3 Experimental Report: Michigan State of the State Survey 94

National Social Norms Center, Michigan State University | September 2, 2026

---

## Executive Summary

Waves 1 and 2 of this project established that Michigan residents hold inaccurate beliefs about people on the other side of the political spectrum. Wave 3 tested whether a campaign can change those beliefs. It can, and it did so at a dose far below what any real campaign would deliver.

The headline for planning purposes is the ratio of dose to effect. Respondents saw four static images once, embedded in an unrelated survey, framed as a design evaluation rather than as a message about politics, with a minimum exposure of 20 seconds and no repetition, no reinforcement, and no reason to attend to the content as persuasion. That single incidental exposure produced statistically reliable movement on the overall misperception index, on the specific beliefs the messages addressed, and on respondents' own stated attitudes. Effects of this kind are normally difficult to detect at all in a one-shot design.

- The overall misperception index fell 4.05 points, from 28.68 to 23.88 (SE = 1.89,  $p = .032$ ,  $d = -0.19$ ), closing 14% of the total misperception present in the control group. The average absolute size of respondents' errors fell 2.67 points ( $p = .049$ ).
- On the three beliefs the billboards addressed, misperception fell 5.80 points, from 34.47 to 27.98 (SE = 1.95,  $p = .003$ ,  $d = -0.26$ ), closing 17% of the gap on those norms.
- Effects concentrated on what the messages actually said. The four unaddressed misperceptions also moved 2.8 points as a set ( $p = .215$ ).
- The largest single result concerns bipartisan lawmaking, the subject of two of the four billboards—indicate a dose-response effect. Michiganders in the control group believed only 36% of the other side thinks it is important for political leaders to work across the aisle, when the true figure is 94%. Among those who saw the messages, the estimate rose to 46%, closing roughly one sixth of a 58-point gap from one exposure.
- The campaign worked equally well on the left, the right, and the middle. On bipartisan lawmaking, the improvement was 8.5 points among conservatives, 8.3 among liberals, and 8.8 among moderates, with no statistical difference between groups.
- Two of the four unaddressed misperceptions improved at borderline significance, but the four cannot be reliably distinguished from one another, so we do not draw conclusions about which beliefs spillover reaches.
- Beliefs about out-group goodwill moved. Beliefs about out-group hostility did not. Underestimation of the other side's good faith fell 5.9 points ( $p = .002$ ), while overestimation of the other side's hostility was unchanged ( $p = .293$ ).
- The messages also changed personal attitudes, not just perceptions. Willingness to say it is acceptable to insult the other side fell from 48% to 37% ( $p = .013$ ).
- The baseline held steady between waves. The Wave 3 control group reproduced the Wave 2 gaps almost exactly, confirming the problem is stable and that the experimental comparison is sound.

The practical implication is that the campaign moves the specific beliefs it names, so message selection matters a great deal. Creative should lead with cooperation and common ground, which is where beliefs proved changeable, rather than with corrections to how hostile the other side is. The first message tested was preferred by 38% of respondents, roughly double either of the two lowest performers.

## I. Project Purpose and Study Design

This project measures and seeks to reduce the civil discourse perception gap in Michigan: the difference between what residents believe people on the other side think and do, and what those people actually report about themselves. Waves 1 and 2 documented the size of the gap. Wave 3 tested an intervention.

YouGov interviewed 1,196 Michigan residents between August 10 and August 20, 2026, matched to a final sample of 1,000 adults and weighted to the state population on gender, age, race, education, and 2024 presidential vote. Respondents were randomly assigned to one of two conditions. Those in the message condition (n = 482) viewed four social norms billboards and were asked which design they preferred. Those in the control condition (n = 518) went straight through the survey without seeing the messages. All perception questions came later in the survey, after the exposure, so any difference between the groups is attributable to the messages.

The exposure was deliberately incidental. Respondents were told they were evaluating billboard designs, not receiving a message about politics. This is a conservative test: it approximates the passive, low-attention exposure a real billboard receives, rather than the focused attention of a lab study.

### What the messages said, and what they did not say

The four ONE Michigan billboards each stated a fact about Michiganders drawn from the March 2026 survey. Two addressed the importance of bipartisan lawmaking, one statewide and one broken out by ideology. One addressed the acceptability of offensive political language. One addressed regularly talking with people who hold different political views. Figure 1 reproduces all four.

Respondents in the message condition were told we were designing billboards for a campaign about how Michiganders perceive and communicate with one another. All four designs appeared together in a grid, and after 20 seconds a prompt asked which one should go on a billboard. Again, the framing was design evaluation, not persuasion.

This means the campaign spoke to three of the seven misperceptions we measure and stayed silent on the other four. That is not a flaw in the study; it creates the most useful test available. We can ask whether the messages moved the beliefs they addressed, and separately whether they moved related beliefs they never mentioned. The first question tells you whether the creative works. The second tells you how much ground each message covers, which determines how many messages a full campaign needs.

**Figure 1.** The four campaign messages tested



**Note.** Every percentage on the creative is an accurate figure from the March 2026 Michigan State of the State Survey. Addressed: the importance of bipartisan lawmaking (Messages 1 and 2), the acceptability of hostile language toward the other side (Message 3), and regularly talking with the other side (Message 4). Not addressed: wanting to understand the other side, whether insulting their arguments is acceptable, harmful language on social media, and political violence.

**Table 1. Wave 3 study design at a glance**

| Element           | Detail                                                                          |
|-------------------|---------------------------------------------------------------------------------|
| Survey            | Michigan State of the State Survey 94, fielded by YouGov                        |
| Field period      | August 10 to August 20, 2026                                                    |
| Sample            | 1,000 Michigan adults, weighted to state population                             |
| Design            | Randomized, two conditions, between subjects                                    |
| Message condition | 482 respondents viewed four norms billboards                                    |
| Control condition | 518 respondents saw no messages                                                 |
| Outcomes          | Seven paired perception and self-report items on civil discourse                |
| Analysis          | Weighted regression with robust standard errors, controlling for age and gender |

## II. How Small the Dose Was, and What It Still Produced

Because this is the result most relevant to whether the approach scales, it is worth stating the dose precisely before reporting the effects. Respondents in the message condition received a single presentation of four static images, viewed once, inside a survey about unrelated topics. The block was framed as a request for design feedback, so respondents had no reason to process the content as a claim about politics. Minimum exposure was 20 seconds. There was no repetition, no second wave of exposure, no accompanying media, and no measurement of whether respondents read the statistics at all. This is a weaker manipulation than a person would receive from driving past an actual billboard twice a week for a month.

Against that dose, the effects below are not marginal curiosities. They are reliable at conventional thresholds, they appear on multiple independent outcome families, and several close a meaningful share of a gap that has been stable across two survey waves six months apart.

**Table 2. What one incidental exposure produced**

| Outcome                                     | Control | Effect | SE   | p     | d     | Share of gap closed |
|---------------------------------------------|---------|--------|------|-------|-------|---------------------|
| Overall misperception index                 | 28.7    | -4.05  | 1.89 | .032  | -0.19 | 14%                 |
| Average size of error                       | 38.0    | -2.67  | 1.36 | .049  | -0.17 | 7%                  |
| Misperception of the three addressed norms  | 34.7    | -5.80  | 1.95 | .003  | -0.26 | 17%                 |
| Bipartisan lawmaking gap                    | -58.0   | +9.13  | 2.61 | <.001 | 0.29  | 16%                 |
| Talks with the other side gap               | -19.2   | +4.55  | 2.36 | .054  | 0.16  | 24%                 |
| Wants to understand the other side gap      | -22.0   | +3.94  | 2.27 | .082  | 0.15  | 18%                 |
| Says insulting the other side is acceptable | 48.1%   | -10.30 | 4.14 | .013  | -0.21 | 21%                 |
| Wants to understand the other side          | 47.7%   | +7.28  | 4.17 | .081  | 0.15  | 15%                 |

*Note.* All estimates are weighted least squares with HC3 robust standard errors, controlling for age and gender; n = 993 to 1,000. Effect is the message condition minus the control condition. For the last two rows the outcome is the respondent's own stated attitude in percentage points, and the final column is the relative change from the control percentage. For the gap rows the final column is the effect as a proportion of the control-group gap.

Three features of this table matter for judging whether the approach is worth scaling. First, the effects are not confined to a single item: they appear on a seven-item composite, on individual gaps, and on respondents'

own attitudes, which are three logically distinct outcomes. Second, the effect sizes are small in standardized terms, between  $d = 0.15$  and  $d = 0.29$ , but that is the expected magnitude for a single incidental exposure, and the relevant comparison is not to a laboratory intervention with forced attention. Third, the proportions in the last column are substantial: one viewing closed between 14% and 24% of gaps that had not moved at all on their own over the preceding six months.

The self-report results deserve particular emphasis because they were not the design's target. The campaign was built to correct beliefs about other people. It also changed what respondents said about themselves. Willingness to say it is acceptable to insult the other side fell from 48.1% to 36.7%, a 21% relative reduction ( $b = -10.30$ ,  $SE = 4.14$ ,  $p = .013$ ), and stated interest in understanding the other side rose from 47.7% to 55.0% ( $b = 7.28$ ,  $p = .081$ ). This is the mechanism the theory of normative social behavior predicts: correcting a descriptive norm changes personal endorsement of the behavior. Demonstrating that link in a general population sample, from one exposure, is the strongest evidence in the study that the approach reaches beyond perception and toward conduct.

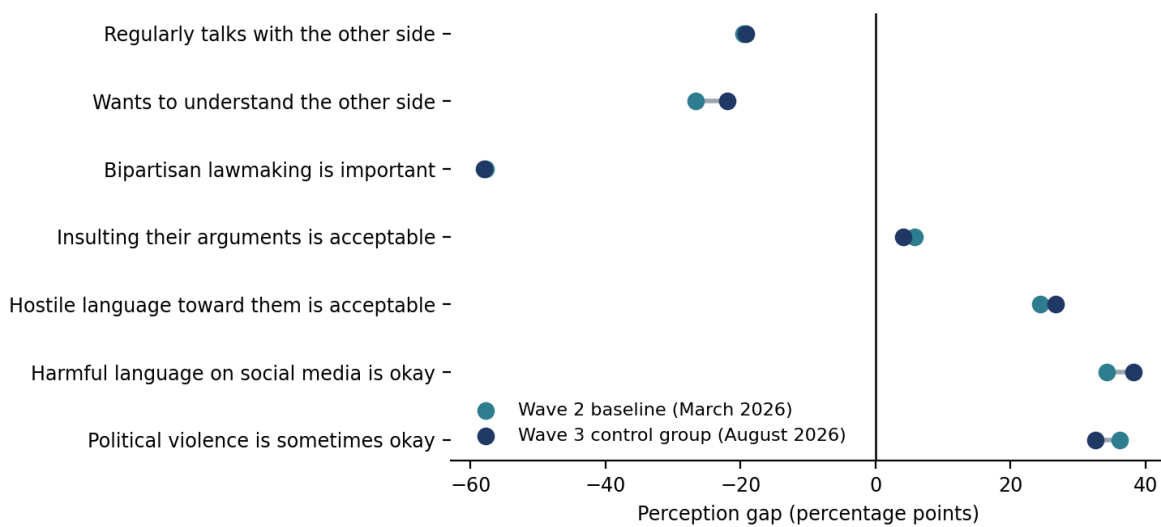
One boundary on the dose argument should be stated plainly, since a larger dose is an assumption rather than a finding. This study establishes that a minimal exposure produces a detectable effect. It does not establish the shape of the dose-response curve, and effects could plausibly saturate, accumulate, or decay. Estimating that curve requires a design with randomized exposure frequency, which is the natural next study.

### III. The Baseline Held Between Waves

Before evaluating the campaign, we confirmed that the problem had not shifted on its own. Comparing the Wave 2 baseline sample from March 2026 to the Wave 3 control group from August 2026, the gaps are essentially identical. None of the seven topics differed once we account for testing seven at once, and the overall misperception index was 29.2 at Wave 2 versus 28.7 in the Wave 3 control group, a difference well within sampling noise ( $p = .694$ ).

This matters for two reasons. It confirms the perception gap is a stable feature of the Michigan public rather than a momentary reaction to the news cycle, and it establishes that the control group is a valid comparison for judging the campaign's effect.

**Figure 2.** *The perception gap is unchanged between March and August 2026*



*Note.* Each line connects the Wave 2 baseline estimate to the Wave 3 control group estimate for the same topic. Values are percentage points; negative means residents underestimate the other side, positive means they overestimate.

#### IV. Did the Norms Message Exposure Work?

Yes, and the pattern lines up with the billboard exposure. The two strongest results are both on topics the campaign addressed. Michiganders who saw the messages gave significantly more accurate estimates of how much the other side values bipartisan lawmaking, the subject of two billboards, and moved in the right direction on regularly talking with the other side, the subject of a third. Table 3 marks which topics the campaign targeted.

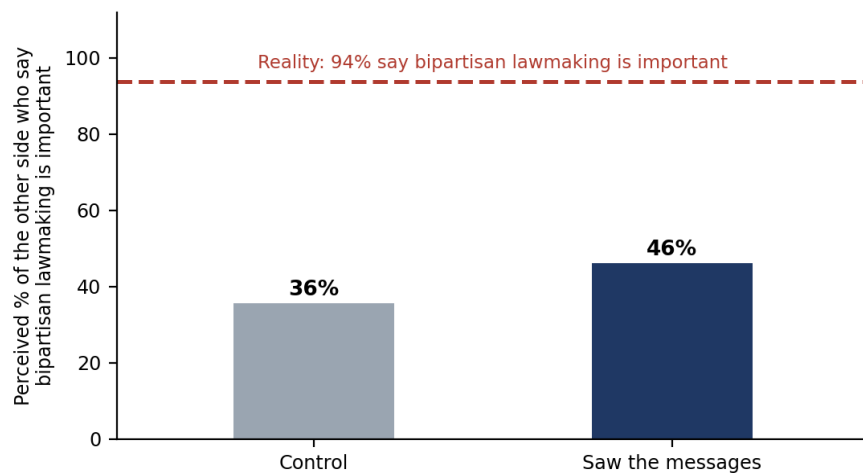
**Table 3.** *What Michiganders believe about the other side, by condition*

| Topic                                                          | Targeted in billboard? | Actual | Control believes | Message group believes | Change |
|----------------------------------------------------------------|------------------------|--------|------------------|------------------------|--------|
| Regularly talks with the other side                            | Yes                    | 59%    | 40%              | 45%                    | +4.2   |
| Bipartisan lawmaking is important                              | Yes                    | 94%    | 36%              | 46%                    | +9.1   |
| Hostile language toward the other side is sometimes acceptable | Yes                    | 29%    | 56%              | 52%                    | -3.3   |
| Wants to understand the other side's views                     | No                     | 46%    | 23%              | 29%                    | +3.7   |
| Insulting the other side's arguments is sometimes acceptable   | No                     | 54%    | 59%              | 59%                    | +0.1   |
| Harmful or mean language on social media is sometimes okay     | No                     | 20%    | 59%              | 53%                    | -4.8   |
| Political violence is sometimes okay                           | No                     | 10%    | 43%              | 40%                    | -1.5   |

*Note.* Topics addressed by a billboard are listed first. Actual is the average of the true conservative and liberal rates. Change is the difference attributable to message exposure after controlling for age and gender.

The clearest result is bipartisan lawmaking. Ninety four percent of Michiganders on both sides say it is important for political leaders to work with the other party to pass new laws. Residents in the control group believed only 36% of the other side thinks so. Seeing the messages raised that estimate to 46%.

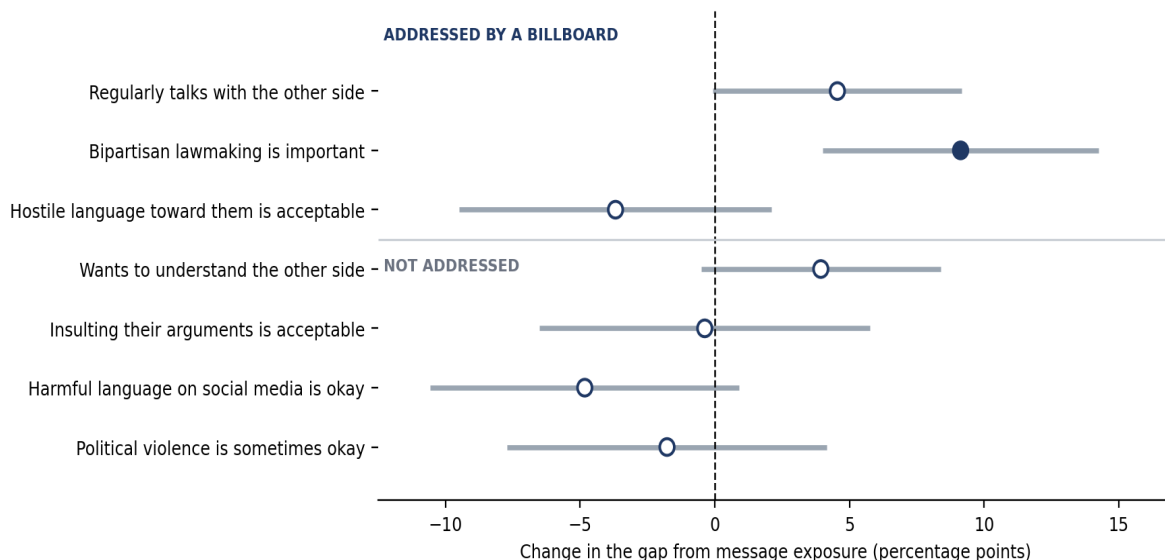
**Figure 3.** *What Michiganders think the other side believes about bipartisan lawmaking, against reality*



*Note.* The dashed line is the actual weighted percentage of Michiganders who say bipartisan lawmaking is important. Even after exposure, residents underestimate the other side by roughly 48 points.

Figure 4 shows the effect on each of the seven gaps with its margin of error. Bars crossing the zero line indicate topics where the campaign produced no detectable change.

**Figure 4.** *Effect of message exposure on each perception gap*



*Note.* Points are the estimated change from exposure; lines are 95% confidence intervals. The filled point is statistically significant; open points are not. Topics are grouped by whether a billboard addressed them.

## V. Did the Messages Only Move What They Named?

Grouping the seven topics by whether a billboard addressed them gives the clearest read on how the campaign works. The three targeted misperceptions improved by 5.8 points as a set, a reliable effect. The four untargeted misperceptions improved by 2.8 points, which is within the range of chance.

**Table 4.** *Improvement in misperception, by whether the campaign addressed the topic*

| Set of beliefs                             | Control | Saw messages | Improvement | p    |
|--------------------------------------------|---------|--------------|-------------|------|
| Three topics the billboards addressed      | 34.5    | 28.0         | +5.8        | .003 |
| Four topics the billboards did not address | 24.0    | 20.8         | +2.8        | .215 |

*Note.* Values are the average misperception across the topics in each set, in percentage points. Improvement is movement toward accuracy. Addressed topics: bipartisan lawmaking, regularly talking with the other side, and the acceptability of hostile language toward them.

One caution belongs here, because it affects how far the finding can be pushed. The targeted set moved reliably and the untargeted set did not, but a direct statistical comparison of the two effects is not itself significant ( $p = .117$ ). The honest reading is that the campaign clearly moves what it names, and may also move related beliefs somewhat less strongly. It would take a larger sample or a stronger dose to separate those two possibilities.

Two of the four unaddressed topics did improve at borderline significance: wanting to understand the other side ( $p = .082$ ) and whether harmful language on social media is okay ( $p = .099$ ). The other two, whether insulting the other side's arguments is acceptable and whether political violence is acceptable, did not move ( $p = .906$  and  $p = .558$ ).

It is tempting to read a pattern into which two moved, and we tested that directly rather than assuming it. Among control-group respondents, the two unaddressed hostility beliefs are statistically indistinguishable in how closely they track the addressed hostile-language belief ( $r = .68$  and  $r = .66$ ; bootstrapped difference  $+ .02$ , 95% CI  $[-.07, +.11]$ ), so neither is meaningfully nearer to the message. Nor do their treatment effects differ reliably from each other ( $p = .080$ ). With only four unaddressed beliefs, the study can establish that effects concentrate on addressed norms but cannot say which unaddressed beliefs spillover reaches. We would not build a campaign plan on that distinction until a larger design tests it.

For campaign planning, this is the most actionable result in the study. It means a campaign cannot expect a few messages about cooperation to repair the whole picture Michiganders hold of one another. Each cluster of beliefs likely needs its own creative.

## VI. Where the Campaign Moved People, and Where It Did Not

Grouping the seven topics reveals a clean pattern. Beliefs about the other side's goodwill proved changeable. Beliefs about the other side's hostility did not.

**Table 5.** Overall misperception, by condition

| Measure                               | Control | Saw messages | Change | p    |
|---------------------------------------|---------|--------------|--------|------|
| Pessimistic misperception index       | 28.7    | 23.9         | -4.0   | .032 |
| Underestimation of out-group goodwill | 33.1    | 25.9         | -5.9   | .002 |
| Overestimation of out-group hostility | 25.4    | 22.3         | -2.7   | .293 |
| Absolute gap magnitude                | 38.0    | 34.7         | -2.7   | .049 |

*Note.* Higher numbers mean more misperception. Absolute gap magnitude measures the size of error regardless of direction.

Underestimation of the other side's good faith fell by 5.9 points, a reliable effect. Overestimation of the other side's hostility fell by 2.7 points, which is not distinguishable from zero. Two explanations are plausible and they carry different implications for the campaign. Beliefs about hostility may be defended because they serve a purpose for people, or they may simply be reinforced daily by media coverage that a billboard cannot outweigh. A sustained campaign with repeated exposure would help distinguish these.

## VII. The Campaign Reached Across the Spectrum

Our central practical question is whether a norms campaign can work statewide without preaching to one side. On the strongest outcome, it does. The improvement in perceived support for bipartisan lawmaking was almost the same size for conservatives, liberals, and those in the middle, and the difference between groups was statistically nil.

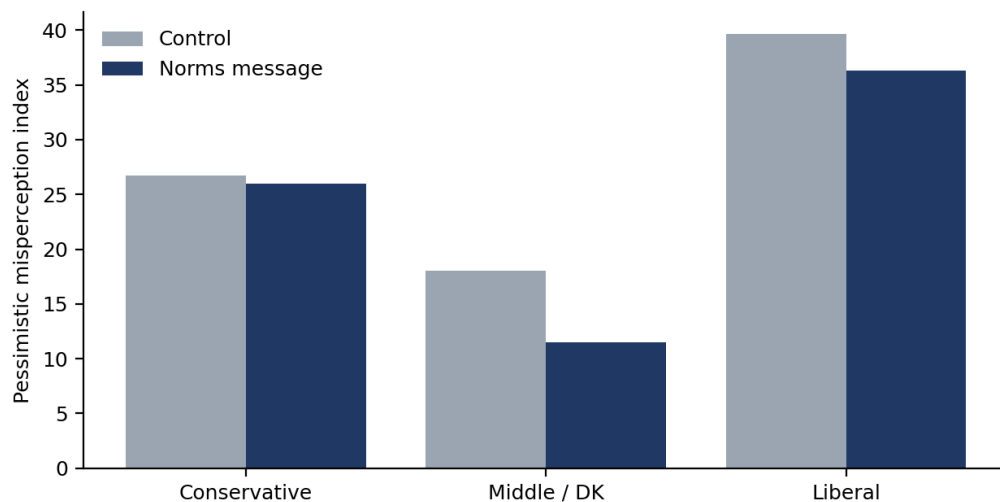
**Table 6.** *Change in perceived out-group support for bipartisan lawmaking, by ideology*

| Ideology         | Control gap | Message group gap | Improvement | p    | Group size |
|------------------|-------------|-------------------|-------------|------|------------|
| Conservative     | -61.4       | -51.3             | +8.5        | .036 | 332        |
| Middle or unsure | -48.7       | -38.9             | +8.8        | .070 | 302        |
| Liberal          | -61.4       | -53.4             | +8.3        | .041 | 360        |

*Note.* Gap values are negative because all three groups underestimate the other side. Improvement is movement toward accuracy.

Baseline misperception does differ sharply by ideology. Liberals hold the largest misperceptions of conservatives, conservatives hold intermediate misperceptions of liberals, and those in the middle are the most accurate. The campaign did not eliminate that difference; it shifted all three groups by a similar amount.

**Figure 5.** *Overall misperception by ideology and condition*



*Note.* Higher bars indicate more inaccurate, more pessimistic beliefs about the other side.

**Table 7.** *Overall misperception index by ideology and condition*

| Ideology         | Control | Saw messages | Change |
|------------------|---------|--------------|--------|
| Conservative     | 26.7    | 26.0         | -0.7   |
| Middle or unsure | 18.2    | 11.5         | -6.6   |

| Ideology | Control | Saw messages | Change |
|----------|---------|--------------|--------|
| Liberal  | 40.1    | 36.3         | -3.8   |

## VIII. The Messages Changed Personal Attitudes Too

We did not expect this and it is arguably the most encouraging result in the study. Respondents who saw the messages were substantially less willing to say it is acceptable to insult people on the other side, dropping from 48% in the control group to 37%. Interest in understanding the other side's views rose from 48% to 55%, a marginal effect.

This is what norms theory predicts. When people learn that a behavior is less common than they assumed, their own endorsement of it declines. It suggests the campaign's reach extends beyond correcting beliefs about others to shifting personal conduct, which is ultimately what the project aims to change.

**Table 8.** *Respondents' own attitudes, by condition*

| Statement about the respondent                                 | Control | Saw messages | Change | p    |
|----------------------------------------------------------------|---------|--------------|--------|------|
| Regularly talks with the other side                            | 57%     | 54%          | -1.6   | .707 |
| Wants to understand the other side's views                     | 48%     | 55%          | +7.3   | .081 |
| Bipartisan lawmaking is important                              | 92%     | 89%          | -1.6   | .518 |
| Insulting the other side's arguments is sometimes acceptable   | 48%     | 37%          | -10.3  | .013 |
| Hostile language toward the other side is sometimes acceptable | 25%     | 23%          | -0.4   | .909 |
| Harmful or mean language on social media is sometimes okay     | 19%     | 16%          | -1.8   | .569 |
| Political violence is sometimes okay                           | 8%      | 6%           | -3.2   | .153 |

## IX. Message Preference

Among the 482 respondents who viewed the billboards, the first message was clearly preferred, chosen by 38%. The second drew 23%, and the third and fourth drew about 20% each. Because all four were shown together and preference was asked afterward, this indicates design appeal rather than persuasive power, and the two should not be assumed to coincide. Still, it gives a starting point for creative development.

**Table 9.** *Billboard design preference among respondents who saw the messages*

| Message   | Share preferring it |
|-----------|---------------------|
| Message 1 | 38%                 |
| Message 2 | 23%                 |
| Message 3 | 20%                 |
| Message 4 | 20%                 |

## **X. What This Means for the Campaign**

- Choose message topics deliberately, because the campaign moves what it names. The three beliefs the billboards addressed improved; the four they ignored largely did not. Coverage is a creative decision, not a side effect.
- Lead with cooperation and common ground. That is where beliefs proved changeable and where the effect held across the political spectrum.
- Do not expect hostility corrections to work the same way. The offensive-language billboard was the one targeted message that did not produce a reliable shift, and the untargeted hostility beliefs did not move either.
- Budget for breadth. If the campaign needs to move beliefs about insults or political violence, it will likely need messages that address them directly rather than relying on spillover from cooperation creative.
- The per-exposure effect is small and the implied return on a real campaign is not. One viewing closed 14% of the overall gap and 17% of the gap on addressed norms. Prior work on the U Celebrate alcohol harm reduction campaign shows effects accumulate with repeated exposure, though the dose-response curve for this campaign has not yet been estimated.
- Fund the dose-response study. The single most valuable next design randomizes exposure frequency, which would convert the present finding into an estimate of how much campaign is needed to close how much gap.
- The statewide reach is genuine. A message that moves conservatives, liberals, and moderates by the same amount can run everywhere without separate creative for each audience.

## **XI. Limitations**

Five points bound these conclusions. All four billboards were shown together, so the effect belongs to the message set rather than to any single design, and the targeted versus untargeted comparison is a comparison of belief topics rather than randomly assigned messages. The difference between the targeted and untargeted effects is not itself statistically significant, so the specificity finding is suggestive rather than settled. There was no measurement before exposure, so the comparison rests on random assignment; the message group did contain more women than the control group, and all reported figures adjust for age and gender accordingly. Respondents who placed themselves in the middle of the ideology scale were asked about a different target group in the billboard block than in the perception block, though results hold under either assignment. Finally, the questionnaire specification called for the liberal and conservative fill to be keyed to party affiliation, but the delivered data show it was implemented on the ideology question; the analyses above follow the data as fielded.