



Partisan underestimation of the polarizing influence of group discussion☆



Jessica Keating*, Leaf Van Boven, Charles M. Judd

Department of Psychology and Neuroscience, University of Colorado Boulder, UCB 345, Boulder, CO 80309-0345, United States

HIGHLIGHTS

- In 2 experiments, groups of 4–6 like-minded participants discussed political topics.
- Group discussion polarized participants' attitudes.
- Participants misremembered pre-discussion attitudes as having been more extreme.
- Individuals underestimated the extent to which their attitudes had polarized.
- Group polarization pressures appear to be more potent than partisans realize.

ARTICLE INFO

Article history:

Received 3 August 2015

Revised 1 March 2016

Accepted 11 March 2016

Available online 18 March 2016

Keywords:

Group polarization

Political psychology

Conformity

Attitude consistency

ABSTRACT

Group polarization occurs when people's attitudes become more extreme following discussion with like-minded others. We hypothesized that people underestimate how much a relatively brief group discussion polarizes their own attitudes. People often perceive their own attitudes as unbiased and stable over time. Therefore, people's polarized post-discussion attitudes may cause them to misremember their pre-discussion attitudes as having been more extreme than they were. In two experiments, participants engaged in 15-minute discussions with 4–6 like-minded others regarding two political topics: whether Barack Obama or George W. Bush was the better president (Experiment 1) and whether they supported Barack Obama or Mitt Romney during the 2012 presidential election (Experiment 2). Group discussion polarized participants' attitudes, and participants misremembered their pre-discussion attitudes as having been more extreme than they actually were. Participants' polarized post-discussion attitudes significantly predicted their recalled pre-discussion attitudes, controlling for their actual pre-discussion attitudes, suggesting that their post-discussion attitudes guided reconstruction of their pre-discussion attitudes. These findings have implications for people's awareness of psychological biases and for the societal effects of partisan enclavement.

© 2016 Elsevier Inc. All rights reserved.

In a 2010 commencement address at the University of Michigan, President Obama issued a warning about the state of public discourse in America:

If we choose only to expose ourselves to opinions and viewpoints that are in line with our own, studies suggest that we become more polarized, more set in our ways. That will only reinforce and even deepen the political divides in this country (The White House, Office of the Press Secretary, 2010).

President Obama's warning about the effects of group polarization is particularly trenchant at a time when Americans have become ideologically sorted (Bishop, 2008). Americans tend to socialize with, live near, and work with others who are politically like-minded (McPherson,

Smith-Lovin, & Cook, 2001). In fact, people who believe their ideology to be out-of-sync with those living around them may actively seek out a new community that provides a better ideological match (i.e. the ideological migration hypothesis; Motyl, Iyer, Oishi, Trawalter, & Nosek, 2014). Liberals consequently have greater social contact with other liberals, conservatives have greater social contact with other conservatives, and this ideological enclavement may have contributed to increased political polarization in the United States (Schkade, Sunstein, & Hastie, 2007; Westfall, Van Boven, Chambers, & Judd, 2015).

Group polarization occurs when people's attitudes become more extreme following discussions with like-minded others (Isenberg, 1986; Moscovici & Zavalloni, 1969; Myers & Lamm, 1976; Stoner, 1968; Sunstein, 2009). For example, in a study that examined group polarization of contemporary partisan attitudes, researchers asked liberal and conservative community members in Boulder and Colorado Springs, Colo., respectively, to discuss three partisan topics for 15 min: environmental policy to reduce greenhouse gases, civil unions for gay and

☆ NSF Grants 0552120 and 1049125 supported this research.

* Corresponding author.

E-mail address: jessica.keating@colorado.edu (J. Keating).

lesbian Americans, and affirmative action (Schkade et al., 2007). Partisans' private attitudes became more extreme following the discussion: Boulder liberals held more liberal attitudes and Colorado Springs conservatives held more conservative stances. Such findings lend credence to President Obama's warnings about politician polarization.¹

Group polarization is determined, broadly, by two sets of processes (Isenberg, 1986; Myers & Lamm, 1976; Sunstein, 2002, 2009). Informationally, the ideological slant of information presented in group discussion tends to reflect the group's initial leanings (Sunstein, 2009). During discussion, the predominance of new and repeated persuasive arguments on one side will begin to overwhelm remaining—and likely unspoken (Isenberg, 1986)—arguments on the other side, steering group members toward more extreme attitudes (Kaplan, 1977; Vinokur & Burnstein, 1978). Socially, as the emerging group view becomes clear, partisans' concerns with social comparisons—their desire to fit the valued group's norm—can lead more moderate individuals to endorse extreme views (Hogg, Turner, & Davidson, 1990; McGarty, Turner, Hogg, David, & Wetherell, 1992). Social concerns may also lead partisans to espouse more extreme stances as a way of “one-upping” others in the group and adhering to dominant group norms (Pruitt, 1971; Sunstein, 2009). Both informational and social processes entail repeated exposure to corroborating information (Myers, 1978; Pruitt, 1971), and repeated attitude expression (Brauer, Judd, & Gliner, 1995), which can both polarize attitudes.

We examine how much people are aware that group discussion with like-minded others polarizes their own partisan attitudes. If people fully recognize how much group discussion polarizes their attitudes, they are well equipped to make informed choices, such as weighing the polarizing effects of enclavement against the potential benefits of an ideologically congruent community (Motyl et al., 2014). But, if people underestimate how much group discussion polarizes their attitudes, they are ill equipped to make informed decisions about the polarizing effects of enclavement.

Given the frequency with which people engage in discussions with like-minded others, and the blatant processes that polarize group members, people might be keenly aware that group discussion polarizes their own attitudes. For example, a Boulder liberal who is moderately supportive of carbon-reducing initiatives might learn more about climate change from other liberals and explicitly become more strident in support of environmental policies. Yet, we hypothesize that people underestimate how much group discussion polarizes their own partisan attitudes, even when their attitudes have just recently polarized.

We derive our hypothesis from two broad psychological principles. First, although people's political attitudes are highly constructive (Lord & Lepper, 1999; Lodge, McGraw, & Stroh, 1989; Schwarz, 2007; Zaller & Feldman, 1992)—reflecting the particular constellation of available information, social context, and motivational considerations that are present at the time—people regard their own attitudes as objective, unbiased, realistic evaluations of the world “as it is” (Griffin & Ross, 1991; Pronin, Lin, & Ross, 2002). For example, even though people's attitudes toward policies are actually strongly swayed by the partisan framing of a policy, people believe that their attitudes result from careful consideration of the available evidence (Cohen, 2003; Pronin, Berger, & Molouki, 2007). People's blindness to contextual influences on their own attitudes implies that they may underestimate how much group discussion polarizes their attitudes.

The second psychological principle is that people's reconstructive memory is strongly influenced by their current attitudes (Levine, 1997; Levine & Safer, 2002; Schwarz, 2007) and their implicit theories about attitude stability over time (Bem, 1972; Ross, 1989). Because people typically believe that attitudes are stable, people whose attitudes have changed tend to reconstruct their previous attitudes as having

been consistent with their current attitudes (Bem & McConnell, 1970; Goethals & Reckman, 1973; Levine, 1997; Nisbett & Wilson, 1977).

Previous research suggests that attitudes newly formed through group discussion may influence estimates of previous attitudes. Goethals and Reckman (1973) asked groups of high school students to discuss the issue of bussing to desegregate schools, a highly controversial topic. Discussion groups included a confederate who, armed with a set of compelling and confrontational arguments, reversed discussants' attitudes toward bussing. Afterward, discussants recalled their pre-discussion attitudes as having been more consistent with their (newly) reversed post-discussion attitudes than they actually were. Although consistent with our hypotheses, this study differed from our experiments in that discussants in the Goethals and Reckman study reversed the direction of their attitudes, which might have created particularly strong motives for people to misremember their initial attitudes. Further, our investigation concerns polarization resulting from emergent properties of group discussion, rather than attitude change resulting from a compelling, and well-informed, experimental confederate.

In two experiments, we tested whether partisans underestimate the effect of group polarization on their own attitudes. We asked groups of like-minded partisans to discuss whether Barack Obama was a better president than George W. Bush (Experiment 1) and whether they supported Barack Obama or Mitt Romney for president in 2012 (Experiment 2). We expected group discussion to polarize participants' attitudes. More importantly, we predicted that participants would underestimate how much their attitudes had changed, misremembering their pre-discussion attitudes as more polarized than they actually were. In both experiments, we also tested whether people's recollections of their pre-discussion attitudes were significantly predicted by their post-discussion attitudes, controlling for their actual pre-discussion attitudes, as might be expected based on the attitude reconstruction explanation.

1. Experiment 1: Obama vs. Bush

1.1. Method

Eighty undergraduates at the University of Colorado Boulder (74% female) participated in eight sessions yielding 16 groups ranging in size from 4 to 6 students. We asked people to discuss with like-minded others whether Barack Obama or George W. Bush was the better president. Eleven groups comprised partisans who preferred Obama; 5 groups comprised partisans who preferred Bush. We conducted the study between mid-September and October 2011, before the 2012 presidential campaign.²

Participants first completed a questionnaire that, in addition to basic demographic information, asked them to consider the presidencies of George W. Bush and Barack Obama and to indicate who they thought was the better president. Participants expressed their opinion by selecting which president they believed was the better one and by drawing a slash through a 15.8 cm line that ranged from *Barack Obama is much better* (0 cm) to *Neutral/neither* (7.9 cm) to *George W. Bush is much better* (15.8 cm). The experimenter used these responses to divide the larger group into two like-minded groups of 4 to 6 participants, usually into an “Obama” group and a “Bush” group.

The experimenter escorted groups to separate rooms, seated them around a table, and informed them that everyone in their group preferred the same president and the other group (or a group convened earlier in the day, for single groups) preferred the other president. The experimenter instructed groups to spend 15 min discussing why their preferred president performed better on the economy and foreign

¹ Although he did not use proper APA citation style, we suspect that Obama's remarks were informed by the results of the Colorado study, given that one of the co-authors, Cass Sunstein, was then a member of the Obama administration.

² Our sample size was determined by the available recruiting window, from the time the department research participant pool opened in mid-September through the end of the extra credit deadline in late October.

policy, then left the room. After 15 min of discussion (monitored from outside the room), participants had 5 min to write (as a group, with a randomly-selected, designated recorder) a consensual description of why their preferred president had performed better on the economy and foreign policy. Finally, participants were told to “take a final vote to ensure that each group member believes the final statement reflects the opinions of the group” and adjust the statement to reach consensus. Research on conformity pressures suggested that forcing groups to achieve consensus might exacerbate group polarization (Isenberg, 1986; Myers & Lamm, 1976; Sunstein, 2009).

Participants then completed a post-discussion questionnaire that included two key responses. First, in the course of estimating each group member's attitudes, participants reported their current, post-discussion attitude toward the two presidents using the same scale from the pre-discussion questionnaire. Second, participants used the same response scale to recall what their attitude had been earlier, before the discussion.

The questionnaire included several exploratory measures regarding participants' perceptions of the group discussion,³ which are included in Supplemental Materials. After completing the post-discussion questionnaire, participants were thanked and debriefed.

1.2. Results and discussion

In both Experiments, we analyzed data using linear mixed effects models with random intercepts for group to account for the interdependence among group members' attitudes (Judd, Westfall, & Kenny, 2012). We excluded nine participants who failed to provide post-discussion attitudes (resulting $N = 71$). For each of our three measures—pre-discussion attitudes toward the two presidents, post discussion attitudes, and recollection of pre-discussion attitudes—we computed extremity scores by taking the absolute value of each participant's deviation from the scale midpoint (range = 0, 7.9 cm). Bush supporters tended to be more moderate in their initial, pre-discussion support than Obama supporters (attitude extremity: $M_{Obama} = 3.87$ cm, $SD_{Obama} = 2.29$, $M_{Bush} = 2.49$ cm, $SD_{Bush} = 2.17$, $z = -2.31$, $p = .021$, 95% CI $[-2.54, -0.21]$).

Replicating group polarization, participants were more extreme after group discussion ($M = 4.30$ cm, $SD = 2.20$) than beforehand ($M = 3.48$ cm, $SD = 2.33$; compare the post- and pre-discussion columns in Table 1). We modeled the difference between participants' post-discussion minus pre-discussion attitudes, a measure of polarization, as a function of participants' presidential preference (Obama = $-.5$; Bush = $+.5$). A significant intercept indicated polarization ($b = 0.87$, $z = 3.20$, $p = .001$, 95% CI $[0.34, 1.41]$), controlling for presidential preference. Obama and Bush supporters did not differ in polarization ($b = 0.26$, $z = 0.48$, 95% CI $[-0.81, 1.33]$).

We next examined whether participants underestimated attitude polarization. Given that actual polarization is the difference between post-discussion and pre-discussion attitudes, and that perceived polarization is the difference between post-discussion and recalled attitudes, polarization underestimation is the difference between recalled and pre-discussion attitudes. That is, the difference between how much people actually polarized and how much they perceived themselves to have polarized is the difference between their actual pre-discussion attitude and their recalled pre-discussion attitudes.

³ None of the exploratory measures about the discussion correlated significantly with polarization underestimation, our primary dependent variable. A couple measures exhibited weak correlations with the extent of polarization (self-influenced by others, $r = .28$, $p = .02$, and by new information, $r = .24$, $p = .05$) and perceived polarization (self-influenced by others, $r = .29$, $p = .02$, and others by repeated expression, $r = .29$, $p = .01$). None were related to underestimation of polarization. The inability to accurately introspect about the influence of cognitive processes (Nisbett & Wilson, 1977) and conformity pressures (Pronin et al., 2004) is well documented. The lack of relationships between our outcomes and these exploratory measures is consistent with those findings. Therefore, we do not present further discussion of these measures.

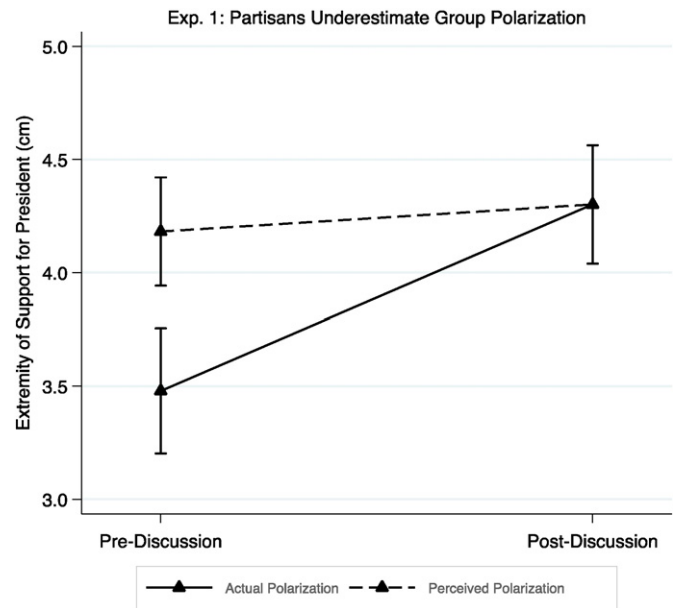


Fig. 1. Experiment 1: actual and perceived extremity shifts. Note: Attitude extremity was computed by taking the absolute value of each participant's deviation from the scale midpoint (range = 0, 7.9 cm) and collapsing across candidate preference. Standard errors are graphed for each point.

As predicted, participants recalled their attitudes as having been more extreme ($M = 4.18$ cm, $SD = 2.02$) than they actually were ($M = 3.48$ cm, $SD = 2.33$; compare the recalled and pre-discussion columns in Table 1 and see Fig. 1). We modeled the difference between recalled and pre-discussion attitudes from whether participants supported Bush or Obama. A significant intercept ($b = 0.78$, $z = 3.27$, $p < .001$, 95% CI $[0.31, 1.25]$) indicated that participants underestimated polarization. Obama and Bush supporters did not differ in their polarization underestimation ($b = 0.36$, $z = 0.76$, 95% CI $[-0.57, 1.30]$).

Not only did participants underestimate polarization, they did not recognize that their attitudes had significantly polarized. When we modeled the difference between participants' post-discussion attitudes and their recalled attitudes—that is, the measure of how much they perceived themselves to have polarized—the intercept was not significant ($b = 0.09$, $z = 0.56$, 95% CI $[-0.23, 0.42]$), and there was no difference between Obama and Bush supporters ($b = -0.10$, $z = -0.30$, 95% CI $[-0.75, 0.55]$).

We suggest that partisans underestimate polarization because their current, more extreme attitudes influence the reconstruction of what their attitudes were before the discussion. Consistent with this explanation, participants' recalled attitudes were predicted both by their post-discussion attitudes ($b = 0.63$, $z = 8.78$, $p < .001$, 95% CI $[0.49, 0.77]$) and by their pre-discussion attitudes ($b = 0.21$, $z = 3.12$, $p = .002$, 95% CI $[0.08, 0.35]$), controlling for which president they supported. These results demonstrate that partisans underestimate how much group discussion polarized their attitudes, and suggest that their recalled attitudes are shaped by their pre-discussion (moderate) and their post-discussion (extreme) attitudes.

2. Experiment 2: Obama vs. Romney

Experiment 2 took place just before the 2012 presidential election between Barack Obama and Mitt Romney. We asked partisans to discuss whether they supported the election of Obama or Romney, a salient and contentious partisan topic. We also wanted to test our hypothesis in a context with greater symmetry between the two political figures, both of whom were current presidential candidates, rather than asking partisans to compare a current (Obama) and past president (Bush), during

Table 1

Experiment 1: means and standard deviations by presidential preference, and total attitude extremity.

Preferred president	Attitude measures					
	Pre-discussion	Post-discussion	Recalled	Actual polarization	Perceived polarization	Polarization underestimation
Bush	10.39 cm (2.17)	11.40 cm (2.09)	11.36 cm (2.00)	1.01 cm (2.07)	0.04 cm (1.04)	0.96 cm (2.00)
Obama	4.03 cm (2.29)	3.28 cm (2.18)	3.43 cm (1.97)	0.75 cm (2.02)	0.15 cm (1.29)	0.60 cm (1.74)
Attitude extremity	3.48 cm (2.33)	4.30 cm (2.20)	4.18 cm (2.02)	0.82 cm (2.02)	0.12 cm (1.22)	0.70 cm (1.81)

Note: Attitude extremity was computed by taking the absolute value of each participant's deviation from the scale midpoint (range = 0, 7.9 cm) and, in Table 1, collapsing across presidential preference. Standard deviations are listed in parentheses.

whose administration many of our university student participants may not have been politically aware (Sears & Valentino, 1997).

Finally, we included a no-discussion control condition in which participants were placed into groups of like-minded others, asked to engage in a brief group introduction, and then instructed to write private descriptions explaining their presidential candidate preference. We expected that participants in the no-discussion control condition would not polarize as much as participants in the group discussion condition, and that control condition participants would therefore underestimate polarization less than group discussion participants.

2.1. Method

University of Colorado Boulder undergraduates ($N = 339$, 66% female) participated in 35 sessions in one of 66 groups of 4 to 6 students. Forty-six groups preferred Obama, and 20 groups preferred Romney. We conducted the study in September and October of 2012, directly before the presidential election.⁴ At the beginning of every session, participants reported whether they thought Barack Obama or Mitt Romney would make the better president and drew a slash through a line ranging from *Barack Obama is much better* (0 cm) to *Neutral/neither* (7.9 cm) to *Mitt Romney is much better* (15.8 cm). The experimenter assigned partisans to pro-Obama or pro-Romney groups based on these responses.

Participants briefly introduced themselves to the other group members to provide a sense of group identification in both the control and discussion conditions. Each participant then completed a second attitude measure of support for Obama or Romney, using the same scale above. We measured pre-task attitudes a second time to obtain an attitude rating prior to the task that distinguished the control and discussion conditions. Taking two measures of the initial attitudes also served as an indirect measure of attitude stability that we could use to predict the magnitude of the polarization effect.

At this point, the procedures in the discussion condition diverged from the no-discussion control condition. In the discussion condition, the experimenter chose a group member, ostensibly at random, to serve as the moderator. Actually, the experimenter chose the person with the most extreme attitude on the first questionnaire in order to strengthen the role of group discussion leader. The experimenter instructed the moderator to keep track of time and ensure a productive conversation. The group discussed for 15 min why their preferred candidate would make the better president and, then, had an additional 5 min to summarize the group's reasoning.

In the no-discussion control condition, the experimenter instructed participants to spend five minutes thinking about why their preferred candidate would make the better president and write those reasons down on their own.⁵ They were not allowed to speak with the others while writing. Note that, although they were not allowed to discuss

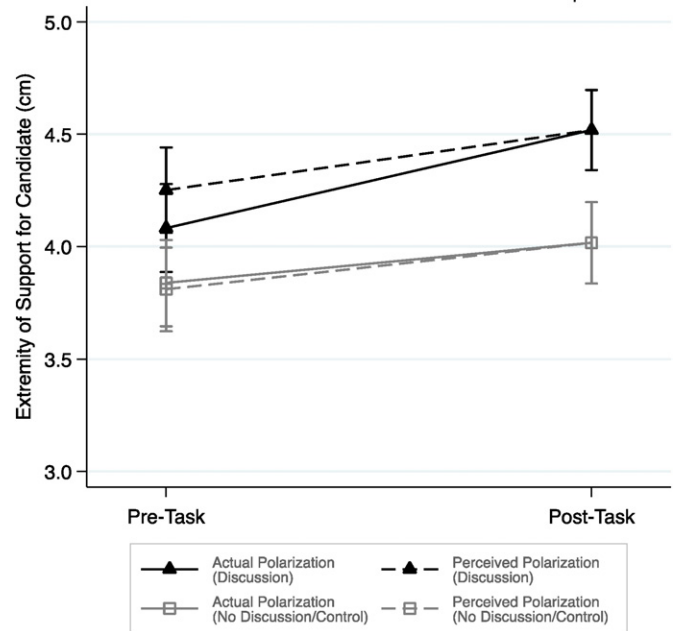
Exp. 2: Group Discussion Produces Greater Polarization and Greater Underestimation Than Private Contemplation

Fig. 2. Experiment 2: actual and perceived extremity shifts, by condition. Note: Attitude extremity was computed by taking the absolute value of each participant's deviation from the scale midpoint (range = 0, 7.9 cm) and collapsing across candidate preference. Standard errors are graphed for each point.

the topic, participants in the no-discussion control condition had been introduced to each other as supporters of the same candidate. Our intention was to create a control condition that was as similar as possible to the discussion procedure, but without the focused group discussion.

Finally, participants indicated, for a third time, their attitude toward Obama and Romney. They also recalled what their attitude had been "at the beginning of the study, when you were first separated into groups." As in Experiment 1, participants completed additional exploratory measures⁶ that are included in the Supplementary material.

The procedures for both the discussion and control conditions required that participants within the same session be assigned to the same condition in order to form discussion groups of like-minded others. We therefore randomly assigned each session to discussion (17 sessions with 32 groups) or control condition (18 sessions with 34 groups).

⁴ Our sample size was based on the available time window, from the time the department research participant pool opened in mid-September through the end of the extra credit deadline in late October.

⁵ Participants spent 5 min thinking and writing in the control condition because we feared their attention would wander from the task at 15 min.

⁶ Again, none of the exploratory measures about the discussion correlated significantly with underestimation of polarization. Two measures exhibited weak correlations with the extent of polarization (others influenced by group, $r = .18$, $p = .02$; and attitude changed during study, $r = .25$, $p = .002$) and perceived polarization (attitude changed during study, $r = .19$, $p = .02$). None were related to underestimation of polarization, our primary outcome. Therefore, the exploratory measures are not discussed further.

2.2. Results and discussion

We excluded 25 participants from analysis due to missing post-task responses (resulting $N = 314$). As in Experiment 1, we computed the absolute value of participants' deviation from the scale midpoint (range = 0 cm, 7.9 cm) for three attitude measures: pre-task attitudes (the average of participants' attitudes before and after the group introduction⁷), post-task attitudes, and recalled attitudes. Attitude ratings prior to and just after group assignment were extremely highly correlated ($r = 0.97$). However, 23 participants failed to provide attitude extremity ratings for themselves on the second questionnaire, leading us to average the two responses into an index of pre-task attitude. Pre-task attitude extremity did not differ significantly by condition ($b = .31, z = 1.03, p = .303, 95\% \text{ CI } [-.28, .91]$). Romney supporters had more moderate attitudes than did Obama supporters ($M_{Obama} = 4.20 \text{ cm}, SD_{Obama} = 2.42, M_{Bush} = 3.29 \text{ cm}, SD_{Bush} = 2.34, z = -2.99, p = .003, 95\% \text{ CI } [-1.51, -.31]$), but the interaction of condition and candidate preference was not significant ($b = .21, z = .35, 95\% \text{ CI } [-.98, 1.41]$).

We first compared polarization in the discussion and no-discussion control condition (Fig. 2). We modeled the difference between each participant's post-task attitude minus the average pre-task attitude score, our measure of group polarization, from participants' candidate preference (Obama = $-.5$, Romney = $+.5$), condition (no discussion = $-.5$, discussion = $+.5$), and their interaction. The model intercept was significantly positive ($b = 0.38, z = 4.50, p < .001, 95\% \text{ CI } [0.22, 0.55]$), indicating that participants' attitudes generally became more extreme (see Table 2 for raw means). The polarization effect was larger in the discussion condition ($b = 0.55, z = 4.55, p < .001, 95\% \text{ CI } [0.31, 0.78]$) than in the no-discussion condition ($b = 0.22, z = 1.82, p = .068, 95\% \text{ CI } [-0.02, 0.46]$), resulting in an overall effect of greater polarization in the discussion condition ($b = 0.33, z = 1.92, p = .055, 95\% \text{ CI } [-0.01, 0.66]$) controlling for choice of candidate. Finally, there was an effect of presidential preference ($b = 0.36, z = 2.09, p = .037, 95\% \text{ CI } [0.02, 0.69]$), indicating that Romney supporters polarized more than Obama supporters. The condition \times candidate preference interaction was not significant ($b = 0.37, z = 1.10, 95\% \text{ CI } [-0.29, 1.04]$). Thus, participants' attitudes polarized, and group discussion produced greater polarization than did writing about one's preferred presidential candidate.

We next examined polarization underestimation by modeling the difference between their post-task recollection of their initial attitudes and their actual initial attitudes from discussion condition, presidential candidate preference, and their interaction (see Table 2). There was an effect of condition ($b = 0.30, z = 2.18, p = .029, 95\% \text{ CI } [0.03, 0.57]$), reflecting that participants in the discussion condition underestimated polarization ($b = 0.26, z = 2.70, p = .007, 95\% \text{ CI } [0.07, 0.46]$) whereas participants in the no-discussion condition did not underestimate polarization ($b = -0.04, z = -.39, 95\% \text{ CI } [-0.23, 0.15]$). The condition \times candidate preference interaction was marginally significant ($b = 0.47, z = 1.70, p = .089, 95\% \text{ CI } [-0.07, 1.02]$), suggesting the greater tendency for group discussion to polarize participants who preferred Romney may have led to correspondingly greater underestimation of that polarization. Replicating and extending Experiment 1, participants underestimated how much their attitudes had polarized whereas participants who did not participate in group discussion, and whose attitudes were less polarized, did not underestimate polarization.

Participants did perceive that they had polarized somewhat, unlike Experiment 1 wherein participants did not perceive that they had become polarized. We modeled the difference between participants' post-task attitudes and their recalled attitudes from discussion condition, presidential candidate preference, and their interaction. The intercept was significant ($b = 0.27, z = 3.33, p = .001, 95\% \text{ CI } [0.11, 0.43]$), indicating that participants thought their attitudes were more extreme after the task than they were beforehand. Neither the effects of discussion condition nor presidential candidate preference nor their interaction was significant ($bs = 0.02, 0.17, \text{ and } -0.10$, respectively, all $ps = ns$). Thus, although people generally recognized that they had polarized, those in the discussion condition did not see themselves as having become more polarized than those in the no-discussion control condition, even though participants who engaged in group discussion had, in fact, become more polarized than those in the control condition.

We also examined whether participants' post-task attitudes predicted their recalled attitudes, consistent with our explanation that people's newly polarized attitudes influence reconstruction of their earlier attitudes. We modeled participants' recalled attitudes from their post-task attitudes, their pre-task attitudes, condition, presidential candidate preference, and the condition \times candidate preference interaction. Recalled attitudes were predicted by both post-task attitudes ($b = 0.40, z = 8.41, p < .001, 95\% \text{ CI } [0.30, 0.49]$) and by pre-task attitudes ($b = 0.55, z = 12.48, p < .001, 95\% \text{ CI } [0.46, 0.63]$). Neither condition nor candidate preference nor their interaction was significant (respectively, $bs = 0.17, 0.34, \text{ and } 0.38, zs = 1.38, 1.47, \text{ and } 1.51, ps = .169, .142, \text{ and } .131$). These results are consistent with our hypothesis that people's newly extreme post-discussion attitudes influence their reconstruction of pre-task attitudes.

Finally, we explored whether participants with less stable initial attitudes might be more prone to attitude polarization and, hence, to underestimating polarization. We computed the absolute difference between the two pre-discussion measures as an indicator of attitude instability. Pre-task attitude instability was not correlated with pre-task attitude extremity ($r(305) = -.08, p = .165$). Although imperfect, it seems reasonable that greater discrepancies between the two pre-task attitude measures reflected attitudes that were less certain and stable (Krosnick & Petty, 1995). Attitude instability significantly predicted attitude polarization ($b = 0.21, z = 2.38, p = .017, 95\% \text{ CI } [0.04, 0.38]$) and polarization underestimation ($b = 0.34, z = 4.52, p < .001, 95\% \text{ CI } [0.19, 0.49]$). Although exploratory, these results suggest that participants with relatively unstable attitudes are more prone to polarization and consequently more prone to underestimate polarization.^{8,9,10}

Taken together, these results suggest that group discussion causes more polarization than simply identifying with a group and contemplating one's own attitudes. And although people in this experiment generally recognized that their attitudes had polarized, those who had become more polarized following group discussion underestimated their attitude polarization more than people in the no-discussion control condition. Finally, an exploratory analysis indicated that those whose initial attitudes were more unstable, as measured by the difference between their two pre-task attitude measures, exhibited greater polarization and greater underestimation of that polarization.

⁸ The instability \times vote interaction was also significant ($b = 0.43, z = 2.31, p = .021, 95\% \text{ CI } [0.06, 0.79]$), suggesting that attitude instability was a stronger predictor of polarization in the Romney groups than in the Obama groups.

⁹ For underestimation, the instability \times condition interaction was non-significant ($p > .50$), and instability predicted polarization underestimation in the Romney groups more than in the Obama groups ($b = 0.33, z = 2.02, p = .043, 95\% \text{ CI } [0.01, 0.65]$).

¹⁰ Perceptions of polarization, the difference between post-task attitude and estimates of pre-task attitude, were negatively associated with attitude instability, although the effect was only marginally significant ($b = -0.14, z = -1.61, p = .107, 95\% \text{ CI } [-0.31, 0.03]$). That is, the more unstable participants' initial attitudes were, the less they perceived themselves to have polarized. The instability interactions with vote and condition were both non-significant ($p's > .50$).

⁷ Prior to conducting Experiment 2, we planned to use the second attitude rating as our pre-task measure because it occurred directly prior to when the discussion and control protocols diverged. The first attitude rating was intended only to help experimenters select the group leader and assign participants to the appropriate (Obama or Romney) group. However, the missing data on the second questionnaire and the strong correlation between the 2 pre-task ratings led us to decide, prior to analysis, to use the average of the two ratings.

Table 2

Experiment 2: means and standard deviations by condition and candidate support, and total attitude extremity by condition.

Attitude measures							
Preferred candidate		Pre-discussion	Post-discussion	Recalled	Actual polarization	Perceived polarization	Polarization underestimation
Discuss	Obama	3.59 cm (2.41)	3.30 cm (2.30)	3.54 cm (2.32)	0.29 cm (1.36)	−0.24 cm (1.31)	0.05 cm (1.07)
	Romney	11.40 cm (2.40)	12.22 cm (1.98)	11.88 cm (2.44)	0.82 cm (1.64)	0.34 cm (1.67)	0.48 cm (1.41)
Control	Obama	3.80 cm (2.42)	3.66 cm (2.36)	3.82 cm (2.36)	0.14 cm (0.91)	−0.15 cm (0.82)	−0.02 cm (0.99)
	Romney	10.98 cm (2.28)	11.28 cm (1.97)	10.92 cm (2.13)	0.30 cm (1.17)	0.36 cm (1.00)	−.06 cm (1.01)
Attitude extremity (discuss)		4.08 cm (2.43)	4.52 cm (2.21)	4.25 cm (2.36)	0.44 cm (1.46)	0.27 cm (1.42)	0.17 cm (1.89)
Attitude extremity (control)		3.84 cm (2.42)	4.02 cm (2.29)	3.81 cm (2.34)	0.18 cm (0.98)	0.21 cm (0.87)	−0.03 cm (0.99)

Note: Attitude extremity was computed by taking the absolute value of each participant's deviation from the scale midpoint (range = 0, 7.9 cm) and collapsing across candidate preference. Standard deviations are listed in parentheses.

3. General discussion

The results of two experiments demonstrate that people underestimate how much a brief group discussion polarizes their partisan attitudes. In Experiment 1, group discussion polarized partisans' attitudes about whether Barack Obama or George W. Bush was a better president. In Experiment 2, group discussion polarized partisans' support for the election of Obama or Romney in 2012—and group discussion polarized attitudes more than simply being identified with a group and writing privately about one's views. In both studies, participants underestimated how much their attitudes had polarized, misremembering their pre-discussion attitudes as more polarized than they actually were.

People's recalled attitudes were also partly predicted by their post-task attitudes, suggesting that their post-discussion attitudes guided reconstruction of their recalled pre-discussion attitudes. It seems unlikely that this influence simply reflects a self-presentation motive to appear consistent. Post-discussion attitude ratings were kept private from the rest of the discussion group. Participants also knew the experimenters possessed both pre-discussion and post-discussion measures and could easily compare the two. Such a design should, if anything, heighten accuracy concerns.

One question about these results concerns their generalizability beyond group discussions of like-minded university undergraduates in controlled laboratory settings. On the one hand, university undergraduates may be highly susceptible to group influences, may have particularly weak mental representations of political attitudes, and may therefore be particularly susceptible to biased reconstructive memory (Markus, 1986; Zaller & Feldman, 1992). On the other hand, community members, like university students, have been shown to polarize as a result of group discussion (Schkade et al., 2007). Furthermore, the group contexts in our experiments were constrained—certainly less so than partisans who live and work in ideologically homogeneous environments. Finally, the relatively short duration of the group discussions (15 min) may be a conservative estimate of how much people's attitudes would polarize over an extended period of everyday life. We strongly suspect that future research would find that community members in naturally-occurring like-minded contexts would similarly underestimate the effects of group polarization.

Our studies further demonstrate people's naiveté about the extent to which social factors shape their attitudes. Although people may acknowledge that social factors sometimes shape their views, they are less likely to perceive the influence of social factors in specific instances, and they generally underestimate their impact (Pronin, Gilovich, & Ross, 2004). People's blindness to personal bias is partly because they exert cognitive effort to perceive the world accurately and objectively (Pronin, 2009; Pronin et al., 2002; Van Boven, White, Kamada, & Gilovich, 2003). Because people are reluctant to view themselves as

politically biased, they take efforts to correct such biases to the extent they notice them (Wilson & Brekke, 1994). The present results suggest that people are unlikely to take such corrective efforts here because they underestimate how much group discussion polarizes their attitudes.

We believe that Obama was right to warn of the effects of exposing oneself only to like-minded others. Yet we suspect such warnings fall on deaf ears. People cannot correct biases they cannot see. Our results suggest that group polarization is seldom seen in the self.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jesp.2016.03.002>.

References

- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology*. Vol. 6. (pp. 1–62). New York: Academic Press. [http://dx.doi.org/10.1016/S0065-2601\(08\)60024-6](http://dx.doi.org/10.1016/S0065-2601(08)60024-6).
- Bem, D. J., & McConnell, H. K. (1970). Testing the self-perception explanation of dissonance phenomena: On the salience of premanipulation attitudes. *Journal of Personality and Social Psychology*, 14(1), 23–31. <http://dx.doi.org/10.1037/h0020916>.
- Bishop, B. (2008). *The big sort: Why the clustering of like-minded America is tearing us apart*. USA: Houghton Mifflin Harcourt.
- Brauer, M., Judd, C. M., & Gliner, M. D. (1995). The effects of repeated expressions on attitude polarization during group discussions. *Journal of Personality and Social Psychology*, 68(6), 1014–1029. <http://dx.doi.org/10.1037/0022-3514.68.6.1014>.
- Cohen, G. L. (2003). Party over policy: The dominating impact of group influence on political beliefs. *Journal of Personality and Social Psychology*, 85(5), 808–822. <http://dx.doi.org/10.1037/0022-3514.85.5.808>.
- Goethals, G. R., & Reckman, R. F. (1973). The perception of consistency in attitudes. *Journal of Experimental Social Psychology*, 9, 491–501. [http://dx.doi.org/10.1016/0022-1031\(73\)90030-9](http://dx.doi.org/10.1016/0022-1031(73)90030-9).
- Griffin, D. W., & Ross, L. (1991). Subjective construal, social inference, and human misunderstanding. In M. P. Zanna (Ed.), *Advances in experimental social psychology*. Vol. 24. (pp. 319–359). New York: Academic Press. [http://dx.doi.org/10.1016/S0065-2601\(08\)60333-0](http://dx.doi.org/10.1016/S0065-2601(08)60333-0).
- Hogg, M. A., Turner, J. C., & Davidson, B. (1990). Polarized norms and social frames of reference: A test of the self-categorization theory of group polarization. *Basic and Applied Social Psychology*, 11(1), 77–100. http://dx.doi.org/10.1207/s15324834basps1101_6.
- Isenberg, D. J. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology*, 50(6), 1141–1151. <http://dx.doi.org/10.1037/0022-3514.50.6.1141>.
- Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, 103(1), 54–69. <http://dx.doi.org/10.1037/a0028347>.
- Kaplan, M. F. (1977). Discussion polarization effects in a modified jury decision paradigm: informational influences. *Sociometry*, 40(3), 262–271. <http://dx.doi.org/10.2307/3033533>.
- Krosnick, J. A., & Petty, R. E. (1995). Attitude strength: An overview. In J. A. Krosnick, & R. E. Petty (Eds.), *Attitude strength: Antecedents and consequences* (pp. 1–24). New York: Lawrence Erlbaum Associates, Inc.
- Levine, L. J. (1997). Reconstructing memory for emotions. *Journal of Experimental Psychology: General*, 126(2), 165–177. <http://dx.doi.org/10.1037/0096-3445.126.2.165>.

- Levine, L. J., & Safer, M. A. (2002). Sources of bias in memory for emotions. *Current Directions in Psychological Science*, 11(5), 169–173. <http://dx.doi.org/10.1111/1467-8721.00193>.
- Lodge, M., McGraw, K. M., & Stroh, P. (1989). An impression-driven model of candidate evaluation. *American Political Science Review*, 83(2), 399–419. <http://dx.doi.org/10.2307/1962397>.
- Lord, C. G., & Lepper, M. R. (1999). Attitude representation theory. In M. P. Zanna (Ed.), *Advances in experimental social psychology*. Vol. 31. (pp. 265–343). New York: Academic Press. [http://dx.doi.org/10.1016/S0065-2601\(08\)60275-0](http://dx.doi.org/10.1016/S0065-2601(08)60275-0).
- Markus, G. B. (1986). Stability and change in political attitudes: Observed, recalled, and “explained”. *Political Behavior*, 8(1), 21–44. <http://dx.doi.org/10.1007/BF00987591>.
- McGarty, C., Turner, J. C., Hogg, M. A., David, B., & Wetherell, M. S. (1992). Group polarization as conformity to the prototypical group member. *British Journal of Social Psychology*, 31, 1–20. <http://dx.doi.org/10.1111/j.2044-8309.1992.tb00952.x>.
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27, 415–444. <http://dx.doi.org/10.1146/annurev.soc.27.1.415>.
- Moscovici, S., & Zavalloni, M. (1969). The group as a polarizer of attitudes. *Journal of Personality and Social Psychology*, 12(2), 125–135. <http://dx.doi.org/10.1037/h0027568>.
- Motyl, M., Iyer, R., Oishi, S., Trawalter, S., & Nosek, B. A. (2014). How ideological migration geographically segregates groups. *Journal of Experimental Social Psychology*, 51(C), 1–14. <http://doi.org/10.1016/j.jesp.2013.10.010>.
- Myers, D. G. (1978). Polarizing effects of social comparison. *Journal of Experimental Social Psychology*, 14(6), 554–563. [http://dx.doi.org/10.1016/0022-1031\(78\)90049-5](http://dx.doi.org/10.1016/0022-1031(78)90049-5).
- Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin*, 83(4), 602–627. <http://dx.doi.org/10.1037/0033-2909.83.4.602>.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review*, 84(3), 231–259. <http://dx.doi.org/10.1037/0033-295X.84.3.231>.
- Pronin, E. (2009). The introspection illusion. In M. P. Zanna (Ed.), *Advances in experimental social psychology*. Vol. 41. (pp. 1–67). Burlington, MA: Academic Press.
- Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, 28(3), 369–381. <http://dx.doi.org/10.1177/0146167202286008>.
- Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review*, 111(3), 781–799. <http://dx.doi.org/10.1037/0033-295X.111.3.781>.
- Pronin, E., Berger, J., & Molouki, S. (2007). Alone in a crowd of sheep: Asymmetric perceptions of conformity and their roots in an introspection illusion. *Journal of Personality and Social Psychology*, 92(4), 585–595. <http://dx.doi.org/10.1037/0022-3514.92.4.585>.
- Pruitt, D. G. (1971). Choice shifts in group discussion: An introductory review. *Journal of Personality and Social Psychology*, 20(3), 339–360. <http://dx.doi.org/10.1037/h0031922>.
- Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, 96(2), 341–357. <http://dx.doi.org/10.1037/0033-295X.96.2.341>.
- Schkade, D., Sunstein, C. R., & Hastie, R. (2007). What happened on deliberation day? *California Law Review*, 95, 915–940.
- Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition*, 25(5), 638–656. <http://dx.doi.org/10.1521/soco.2007.25.5.638>.
- Sears, D. O., & Valentino, N. A. (1997). Politics matters: Political events as catalysts for pre-adult socialization. *American Political Science Review*, 91(1), 45–65. <http://dx.doi.org/10.2307/2952258>.
- Stoner, J. A. F. (1968). Risky and cautious shifts in group decisions: The influence of widely held values. *Journal of Experimental Social Psychology*, 4, 442–459. [http://dx.doi.org/10.1016/0022-1031\(68\)90069-3](http://dx.doi.org/10.1016/0022-1031(68)90069-3).
- Sunstein, C. R. (2002). The law of group polarization. *The Journal of Political Philosophy*, 10(2), 175–195. <http://dx.doi.org/10.1111/1467-9760.00148>.
- Sunstein, C. R. (2009). *Going to extremes: How like minds unite and divide*. New York: Oxford University Press.
- The White House, Office of the Press Secretary. (2010, May 1). *Remarks by the President at University of Michigan Spring Commencement*. Retrieved from <https://www.whitehouse.gov/the-press-office/remarks-president-university-michigan-spring-commencement>.
- Van Boven, L., White, K., Kamada, A., & Gilovich, T. (2003). Intuitions about situational correction in self and others. *Journal of Personality and Social Psychology*, 85(2), 249–258. <http://dx.doi.org/10.1037/0022-3514.85.2.249>.
- Vinokur, A., & Burnstein, E. (1978). Novel argumentation and attitude change: The case of polarization following group discussion. *European Journal of Social Psychology*, 8(3), 335–348. <http://dx.doi.org/10.1002/ejsp.2420080306>.
- Westfall, J., Van Boven, L., Chambers, J. R., & Judd, C. M. (2015). Perceiving political polarization in the United States: Party identity strength and attitude extremity exacerbate the perceived partisan divide. *Perspectives on Psychological Science*, 10(2), 145–158. <http://dx.doi.org/10.1177/1745691615569849>.
- Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, 116(1), 117–142. <http://dx.doi.org/10.1037/0033-2909.116.1.117>.
- Zaller, J., & Feldman, S. (1992). A simple theory of the survey response: Answering questions versus revealing preferences. *American Journal of Political Science*, 36(3), 579–616. <http://dx.doi.org/10.2307/2111583>.