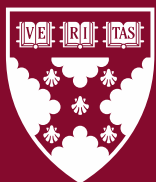


Working Paper 23-049

Change We Can't Believe In: Distrust of Political Converts

Julian J. Zlatev
Amos Schurr
Nir Halevy



**Harvard
Business
School**

Change We Can't Believe In: Distrust of Political Converts

Julian J. Zlatev
Harvard Business School

Amos Schurr
Ben-Gurion University of the Negev

Nir Halevy
Stanford University

Working Paper 23-049

Copyright © 2023 by Julian J. Zlatev, Amos Schurr, Nir Halevy.

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

Funding for this research was provided in part by Harvard Business School.

Change We Can't Believe In: Distrust of Political Converts

Julian J. Zlatev¹

Harvard University

Amos Schurr

Ben-Gurion University of the Negev

Nir Halevy

Stanford University

¹ Corresponding Author: Julian J. Zlatev (jzlatev@hbs.edu)

Abstract

We propose and test three hypotheses regarding how people respond to political converts—individuals who switch their voting from one political party to another. Across two experiments, using behavioral and attitudinal measures of trust in two different countries, we find evidence that people simultaneously trust in-party members more than out-party members and party maintainers more than party switchers (i.e., political converts). We additionally find that this increased trust in party maintainers holds when evaluating just out-party members or just in-party members independently. Finally, we find inconsistent evidence that these results are moderated by ideology. We discuss implications of this work for intergroup relations, polarization, and impression formation.

In the U.S. presidential election of 1884, a group of well-known Republican party members switched their political allegiance and voted for the Democratic candidate, Grover Cleveland, who they viewed as less corrupt than the Republican candidate at the time. In response to this, the editor of the New York Sun labeled these party switchers “mugwumps,” a derogatory term meant to imply that they were self-important, impudent status seekers. The hostility toward these mugwumps for seemingly abandoning their beliefs was summed up by a contemporary of the time, who wrote: ““I would rather be in the claws of the 'Tammany tiger', than in the kidgloved hands of the hypocritical, unprincipled, unscrupulous, vainglorious Republican renegade who goes to another party tooting his gilded trumpet of 'Reform'” (Blodgett, 1962).

Though perhaps not surprising, these criticisms raise questions about the way political converts—defined as individuals who switch their voting from one party to another—are viewed, not only by members of their former in-group but their new one as well. The possibility of changing political allegiances is critical for the proper functioning of representative democracy. This system is built on the premise that political parties represent their constituents, and act according to their will and opinion. Once a constituent changes her beliefs or finds a party that better represents her ideological worldview, she should be able to change her voting behavior accordingly. Consequently, political parties are highly permeable and the technical procedure of changing one’s party affiliation is easy. Indeed, in the U.S. alone, millions of people switch their political affiliation each year (Pew Research Center, 2020).²

However, the ideals underlying representative democracy often clash with political and psychological realities. Particularly, political affiliation plays a central role in people’s self-

² Because party affiliation is public but actual voting behavior is private, the former is easier to measure. However, it is possible that using political affiliation as a proxy may over- or under-estimate changes in actual voting behavior.

concept (Green, Palmquist, & Schickler, 2004; Greene, 2002) and can be a powerful social identity on par with nationality and religion (Cameron, 2004; Deaux, Mizrahi, & Ethier, 1995; Ysseldyk, Matheson, & Anisman, 2010). In fact, previous research has shown that citizens often prioritize party affiliation over ideological content in supporting policies (Cohen, 2003), speaking to the gravitational pull of political parties on their members. These psychological processes have become even more pronounced given the recent rise in political polarization and animus toward members of political out-groups (Finkel et al., 2020; Iyengar et al., 2019).

The significance of party affiliation as a core aspect of an individual's social identity highlights a number of important questions: How do people judge political converts? Do people evaluate political converts who left "them" and joined "us" differently than those who left "us" and joined "them?" Do people with different ideological backgrounds evaluate political converts differently? Finally, in light of the contempt expressed toward the historical "mugwumps," what are the behavioral consequences of perceiving political converts as immoral? Integrating the literatures on intergroup relations, social perception, and trust, the current research addresses these important questions at the intersection of social, moral, and political psychology.

From the perspective of current members of a political party, political converts who join that party can represent both a potential threat and a welcome opportunity. On the one hand, previous out-group members may have opinions or stances outside the norm of the party. Therefore, their addition can threaten group cohesiveness and consensus and muddy the party's ideological message. On the other hand, political power in a democratic system requires popular support; as such, new people joining a party strengthens that party's ability to gain or hold positions of power.

Beyond any costs and benefits to the parties themselves, the act itself of switching parties may be seen as indicating something about the individual doing so, as demonstrated by the opening historical anecdote. This may be especially true if most voters do not switch parties often, such that the act of switching carries information about the individual rather than about the prevailing norm (Kelley, 1973). Switching may reflect positively on the switcher, signaling that they are pragmatic, adaptable, and willing to update their beliefs with new information. However, switching may instead reflect negatively on the switcher, signaling a lack of consistency or conviction in their moral principles.

In this paper we examine how individuals respond to political converts. While previous work has examined perceptions of newcomers in the domains of immigration (e.g., Dinesen, 2012; van der Linden, Hooghe, de Vroome, & Van Laar, 2017) and organizational teams (e.g., Gruenfeld & Tan, 1999; Phillips, Liljenquist, and Neale, 2009), we do not know of previous work examining newcomers in the domain of political parties. This is important for three reasons. First, as discussed earlier, millions of Americans (and millions more people across the world) switch party affiliations each year. As such, it is important to know how people respond to political these converts. Second, whereas other work studying newcomers has primarily examined contexts where movement into a social group is effortful, costly, or risky (e.g., immigrating to a new country), switching one's political affiliation is often assumed to be costless, a hitherto untested assumption that the current research challenges. Although the boundaries of political parties are highly permeable, and political parties seemingly welcome newcomers of all kinds to their ranks, the act of converting may be moralized and therefore sanctioned by individuals. Examining how people perceive political converts in such a permeable yet moralized context can help elucidate when and why people may choose to switch

their party affiliation and when they may be hesitant to do so. Finally, understanding the relational and reputational consequences of political conversion can advance knowledge on a previously overlooked threat to representative democracy, which is built on the premise that individuals should be able to switch their political allegiances with little or no cost (Pitkin, 1967). Uncovering the psychological consequences of switching is of considerable interest also to practitioners and academics engaged in political persuasion, those who seek to bridge the political divide in society, and those who aim to strengthen democracy (Voelkel et al., 2022).

In the following sections we outline our three main hypotheses, including the logic behind each one, prior relevant work, and specific testable predictions.

In-group Preference Hypothesis

A large and robust stream of research suggests that people prefer to interact with and trust members of their own group over those of an out-group (Balliet et al., 2014; Brewer & Silver, 1978; McAllister, 1995). In the political domain, this can take many forms, such as preferring legislation endorsed by political allies (Cohen, 2003) and preferring to hire and date members of one's own political party (Gift & Gift, 2015; Huber & Malhotra, 2017).

When focusing on political converts, there is also reason to believe that people will trust targets who switch into their in-party over targets who switch into their out-party. From a purely pragmatic perspective, political parties are more powerful the more members they have. When people join a political party it strengthens that party's ability to win elections, pass legislation, and advance their agenda. The benefits of having political converts join the in-party are particularly valuable when these newcomers were previously associated with a rival political party, as this represents not just a gain to one's own numbers but also a loss to those of one's opponent. Political converts who join one's political in-party not only share instrumental goals

with its existing members (i.e., strengthening the group's political power); their defection from the rival ideological camp may also be viewed as a symbolic win.

Taken together, the logic outlined above represents what we call the in-group preference hypothesis:

H1a. People will trust current in-party members over current out-party members.

H1b. People will trust political converts who joined one's in-party more than political converts who joined one's out-party.

Principled Loyalty Hypothesis

It is possible that people also value consistency and loyalty not only when they benefit one's in-group, but for their own sake. Behaving in a predictable or consistent manner is widely viewed as a key component of trust (Mayer et al., 1995). A requirement of integrity in particular is that one displays coherence, or a consistency between one's beliefs and actions (McFall, 1987). Thus, it would appear that people would prefer to trust someone whose beliefs are more consistent than someone who is viewed as changing their beliefs over time.

Prior work has demonstrated the extent to which people value consistency and loyalty in intergroup contexts. For example, leaders who change their mind about moralized issues are viewed more negatively than those who stick to their beliefs (Kreps, Laurin, & Merritt, 2017). Additionally, people trust others more when they profess to care about moralized social issues, because caring about these social issues signals that one is principled and consistent in one's beliefs (Zlatev, 2019). Indeed, consistency is viewed as a key aspect of trust and moral character (Cohen & Morse, 2014; McFall, 1987). Taken together, this leads to what we call the principled loyalty hypothesis:

H2a. People will trust consistent voters more than voters who switch their party support (i.e., political converts).

Beyond this overall relationship between party switching and trust, the research cited above also supports more specific subgroup hypotheses:

H2b: People will trust party maintainers over party switchers when evaluating both in-party targets and out-party targets.

In support of Hypothesis 2b, Kreps and colleagues (2017) found that people negatively evaluated leaders who changed their opinion on moralized issues even when they agreed with the leader's new belief. Similarly, Zlatev (2019) found that people trusted others they disagreed with to a greater extent when those individuals cared more about a social issue (compared to when they cared relatively less about it). This provides some evidence that the principle of loyalty appears to be separable from the specific content of the beliefs one holds; people value loyalty even among those whom they disagree with. We examine whether this preference for loyalty extends to evaluations of people based on their political behavior.

Ideology Moderation Hypothesis

Finally, there is reason to believe that the hypotheses outlined above may differ as a function of the evaluator's political affiliation. Specifically, previous work has argued that political conservatism is associated with a number of individual differences that would point to lowered acceptance toward outsiders, including decreased tolerance of ambiguity and decreased integrative complexity (Jost, Glaser, Kruglanski, & Sulloway, 2003; Wilson, 1973).

Conservatives have also been found to be more distrustful of newcomers in other domains, such as immigration (Brooks, Manza, & Cohen, 2016; Hainmueller & Hopkins, 2014). This suggests that conservatives will distrust political converts to a greater extent than liberals

will (i.e., moderation of the in-group preference hypothesis). Other work has found evidence that conservatives value loyalty as a moral principle to a greater extent than liberals do (Graham, Haidt, & Nosek, 2009). This additionally suggests that members of more conservative political parties may view political converts (i.e., party switchers) more negatively than members of liberal political parties (i.e., moderation of the principled loyalty hypothesis). In sum, this leads to the ideology moderation hypothesis and its three components:

H3a: Conservatives will show a stronger in-party preference than liberals.

H3b: Conservatives will show a stronger consistency preference than liberals.

H3c: Conservatives will show a stronger consistency preference for the in-party than liberals.

The Present Research

We test the three overarching hypotheses outlined above in two large-scale experiments run with voters in two countries with two different political systems (a two-party system and a multi-party system). Table 1 provides an overview of each of these hypotheses.

It is important to note that none of the hypotheses outlined above are mutually exclusive. In other words, support for one of these hypotheses does not invalidate any of the other hypotheses. As such, the goal of this work isn't to adjudicate between conflicting theories, but rather to examine the viability of a series of related theories using large-scale, incentive-compatible studies from two different countries. For each study, we report all conditions, measures, and data exclusions.

Experiment 1

In Experiment 1 we examined the degree to which individuals trust others as a function both of who they were voting for in an upcoming election and who they voted for in a previous

election. This study was run in the United States within the context of a two-party political system.

Participants and Methods

A total of 989 (54.2% male, 44.5% female, 1.3% other; $M_{\text{age}} = 32.40$, $SD_{\text{age}} = 11.51$) participants on Prolific completed this study.³ This study was run approximately two weeks before the 2020 general election in the United States. After completing demographic questions, participants read instructions about the Rely or Verify game (Levine & Schweitzer, 2015). Specifically, participants were told that they would be participating in an incentivized two-player decision-making task in which they were assigned the role of Green Player. The Green Player's goal was to decide whether the amount of money in a jar was odd or even. The other player in this task, the Orange Player, knew the answer and would send them a message that the amount of money was either odd or even. Participants could then decide to rely on this information or verify it. The payoff structure, outlined below, is such that the Orange Player has a monetary incentive to lie. The Green Player has a monetary incentive to rely on the Orange Player's message if she believes the Orange Player is being truthful, and to verify the message if she believes the Orange Player is lying. We present the complete payoff structure below.

Participants in the role of the Green Player read detailed instructions for the decision making task and completed two comprehension questions. They subsequently received information about the Orange Player they would be paired with.⁴ Participants were randomly assigned to receive information about one of nine types of Orange Players, in a 3 (2016 election:

³ This sample size provided 80% power to detect a difference in proportions between groups of .50 and .58 with a 5% false-positive rate.

⁴ To avoid deception, we ran a smaller number of Orange Players (N=101) in a separate earlier study who were matched, with replacement, with Green Players.

Didn't Vote vs. Voted Republican vs. Voted Democrat) x 3 (2020 election: Won't Vote vs. Voting Republican vs. Voting Democrat) factorial design.

After reading about which Orange Player they would be paired with, participants then received a message from them that the amount of money in the jar was either odd or even. Participants next decided whether they would rely on this information or verify it. Finally, participants answered five Likert-type questions about the Orange Player's integrity, benevolence, competence, morality, and overall trustworthiness. Participants received payment based on the outcomes of their Rely-or-Verify games.

Results

In-group preference hypothesis

To test the in-group preference hypothesis, we first examined whether participants trusted current in-group members more than current out-group members (H1a). We found evidence in support of this hypothesis, such that participants trusted (i.e., chose to rely on information from) targets who were planning to vote the same as they were in the 2020 election (48.7%) to a greater extent than targets who were not planning to vote for the same party they were (32.8%; $b = 0.67, z = 4.41, p < .001$). This was true both for targets who were planning to vote for the opposing party ($b = 1.00, z = 5.28, p < .001$) and for targets who were planning not to vote at all ($b = 0.41, z = 2.37, p = .018$).

To further explore the in-group preference hypothesis, we additionally examined participants' perceptions of former out-group members who were currently in-group members compared to former in-group members who were currently out-group members (H1b).

Consistent with the in-group preference hypothesis, we found that targets who changed their vote

to be consistent with the participant's vote were trusted more (47.0%) than targets who changed their vote to be inconsistent with the participant's vote (24.4%; $b = 1.01, z = 4.35, p < .001$).

Principled loyalty hypothesis

To test the principled loyalty hypothesis, we first examined whether participants viewed targets who stayed members of the same political group more positively than targets who switched groups. Consistent with H2a, we found that participants trusted targets who were consistent in their voting patterns across 2016 and 2020 (44.9%) significantly more than targets who were inconsistent (36.0%; $b = 0.37, z = 2.57, p = .01$). This was true for targets who specifically switched from one political party to another (i.e., excluding targets who didn't vote in at least one election; $b = 0.44, z = 2.15, p = .03$), but not for targets who went from not voting to voting (compared to targets who didn't vote in either election; $b = 0.27, z = 1.08, p = .28$) or targets who went from voting to not voting (compared to targets who didn't vote in either election; $b = 0.32, z = 1.27, p = .20$).

We next examined whether this preference for principled loyalty was evident for both in-group members and out-group members (H2b). There was no interaction between target consistency and participant-target party match ($b = 0.15, z = 0.50, p = .62$). However, while the difference in behavioral trust was significant between consistent (40.7%) and inconsistent (31.0%) targets for participants' evaluations of out-group members ($b = 0.42, z = 2.35, p = .02$), this difference was not significant for participants' evaluations of in-group members (consistent: 52.9%, inconsistent: 46.0%; $b = 0.27, z = 1.13, p = .26$). Thus, in this setting it seems that in-group members do not receive 'trust credits' for being consistent (consistent with the idea that loyalty is expected), whereas out-group members who behave consistently are trusted more than those who behave inconsistently.

Ideology moderation hypothesis

Finally, we examined whether ideology moderated the key results described above. First, we examined whether the previously-identified preference for in-group members over out-group members was stronger for political conservatives than for political liberals (H3a). We did not find support for this prediction. Specifically, there was no interaction between political affiliation and preference for the in-group (Voted Democrat: 50%, Voted Republican: 46.5%) over the out-group (Voted Democrat: 30.6%, Voted Republican: 36.9%; $b = 0.42$, $z = 1.33$, $p = .18$), suggesting that conservatives and liberals trusted members of their respective in-groups to similar degrees.

Next, we explored whether the previously-identified preference for consistent voters over party switchers was stronger for conservatives than for liberals (H3b). Again, we did not find support for this prediction. Specifically, there was no interaction between political affiliation and preference for consistency ($b = 0.01$, $z = 0.03$, $p = .98$), suggesting that conservatives and liberals trusted consistent voters (Voted Democrat: 41.6%, Voted Republican: 44.1%) over inconsistent (Voted Democrat: 33.1%, Voted Republican: 35.6%) voters to a similar degree. Finally, we explored whether conservatives preferred consistent voters over inconsistent voters within their in-group specifically to a greater extent than liberals (H3c). Once again, this prediction was not supported, with no significant interaction between political affiliation and preference for consistency (Voted Democrat: 54.8%, Voted Republican: 47.2%) over inconsistency (Voted Democrat: 43.5%, Voted Republican: 48.4%) when focusing solely on participants' in-group members ($b = 0.50$, $z = 0.82$, $p = .41$)

Discussion

Experiment 1 finds support for both the in-group preference hypothesis and the principled loyalty hypothesis. Consistent with the in-group preference hypothesis, people trusted targets who were currently part of their in-group more than targets who were currently part of their out-group, in line with research on parochialism (De Dreu, Balliet, & Halevy, 2014). When judging others who changed their political mind between 2016 and 2020, participants trusted targets who left the out-group to join their in-group more than targets who left the in-group to join the out-group. Lending support to the principled loyalty hypothesis, people trusted consistent targets more than inconsistent targets, although, in partial support of H2b, this difference was significant when evaluating out-group members but non-significant when evaluating in-group members. Overall, no support was found for the ideology moderation hypothesis, such that none of the predicted interactions with political affiliation were significant.

Experiment 2

In Experiment 2 we collected additional data to further test the hypotheses outlined above. Specifically, we sought to replicate the results of Experiment 1 in a higher-powered, within-subjects design using a different population of respondents who participate in a multiparty democratic system, as opposed to the two-party democratic system examined in Experiment 1.

Participant and Methods

A total of 1026 (54.2% male, 44.5% female, 1.3% other; $M_{\text{age}} = 32.40$, $SD_{\text{age}} = 11.51$) participants on an Israeli political polling platform completed this study.⁵ This study was run approximately 8 weeks before the Israeli legislative election on March 2, 2020. After completing demographic questions, participants were presented with and asked to evaluate four hypothetical individuals in a random order: (1) A voter who always identified with a right-wing party, (2) a

⁵ This sample size provided 80% power to detect an effect size of Cohen's $d = .087$ or greater in a paired t-test with a 5% false-positive rate.

voter who always identified with a left-wing party, (3) a voter who had previously identified with a left-wing party but currently identifies with a right-wing party, and (4) a voter who had previously identified with a right-wing party but currently identifies with a left-wing party. Thus, this experiment used a fully within-subjects design. For each of these voters, participants answered a single, face-valid question (with wording matched to the condition) meant to assess perceived trustworthiness (translated from Hebrew): “I find people who have always identified politically with the right [left] trustworthy.” Or “I find people who have change their political identification from left to right [right to left] trustworthy.”

Results

Because each participant answered the perceived trustworthiness item about each of the four voters, we ended up with 4104 data points for use in our analyses. All analyses reported below were run using mixed-effects models with participant as a random effect.

In-group preference hypothesis

Similar to Experiment 1, we tested the principled loyalty hypothesis by examining whether participants viewed current in-group members more positively than current out-group members (H1a). We found evidence in support of this hypothesis, such that participants trusted current in-party targets ($M = 4.33$, $SD = 1.70$) more than current out-party targets ($M = 3.11$, $SD = 1.65$; $b = 1.21$, $t(3031) = 28.20$, $p < .001$).

We additionally examined participants’ perceptions of former out-group members who were currently in-group members compared to former in-group members who were currently out-group members (H1b). Consistent with the in-group preference hypothesis, we found that targets who changed their vote to be consistent with the participant’s vote ($M = 3.83$, $SD = 1.64$)

were trusted more than targets who changed their vote to be inconsistent with the participant's vote ($M = 3.23$, $SD = 1.7$; $b = .59$, $t(1010) = 9.65$, $p < .001$).

Principled loyalty hypothesis

Again similar to Experiment 1, we explored the principled loyalty hypothesis by examining whether participants viewed targets who stayed members of the same political wing more positively than targets who switched wings. Consistent with H2a, we found that participants trusted targets who were consistent in their ideological party support ($M = 3.91$, $SD = 1.84$) significantly more than targets who were inconsistent ($M = 3.53$, $SD = 1.70$; $b = 0.38$, $t(3031) = 7.92$, $p < .001$).

We next examined whether this preference for principled loyalty was evident for both in-group members and out-group members (H2b). There was a significant interaction between target consistency and participant-target party match ($b = 0.50$, $t(3029) = 5.98$, $p < .001$). Specifically, this interaction indicates that the difference in perceived trustworthiness between consistent and inconsistent targets was considerably larger for in-group members ($M_{\text{consistent}} = 4.64$, $SD_{\text{consistent}} = 1.70$, $M_{\text{inconsistent}} = 4.01$, $SD_{\text{inconsistent}} = 1.63$; $b = 0.63$, $t(1008) = 12.31$, $p < .001$) than it was for out-group members ($M_{\text{consistent}} = 3.18$, $SD_{\text{consistent}} = 1.67$, $M_{\text{inconsistent}} = 3.05$, $SD_{\text{inconsistent}} = 1.63$; $b = 0.13$, $z = 2.90$, $p = .004$).

Ideology moderation hypothesis

Finally, we examined whether ideology moderated the results described above. First, we examined whether the previously-identified preference for in-group members over out-group members was stronger for political conservatives than for political liberals (H3a). Unlike in Experiment 1, in Experiment 2 we find strong support for this prediction. Specifically, there was a significant interaction between political affiliation and preference for the in-group ($b = 0.70$,

$t(3030) = 7.28, p < .001$), such that right-wing individuals trusted their own in-group ($M_{\text{in-group}} = 4.40, SD_{\text{in-group}} = 1.76, M_{\text{out-group}} = 3.00, SD_{\text{out-group}} = 1.69; b = 1.40, t(2220) = 26.30, p < .001$) to a larger extent than did left-wing individuals ($M_{\text{in-group}} = 4.13, SD_{\text{in-group}} = 1.49, M_{\text{out-group}} = 3.43, SD_{\text{out-group}} = 1.49; b = 0.70, t(810) = 10.84, p < .001$).

Next, we explored whether the previously-identified preference for consistent voters over party switchers was stronger for conservatives than for liberals (H3b). Again, in contrast with Experiment 1, we find strong support for this prediction. Specifically, there was a significant interaction between political affiliation and preference for consistency ($b = 0.42, t(3030) = 3.91, p < .001$), such that right-wing individuals trusted consistent targets ($M_{\text{consistent}} = 3.94, SD_{\text{consistent}} = 1.93, M_{\text{inconsistent}} = 3.45, SD_{\text{inconsistent}} = 1.76; b = 0.49, t(2220) = 8.18, p < .001$) to a larger extent than did left-wing individuals ($M_{\text{consistent}} = 3.82, SD_{\text{consistent}} = 1.57, M_{\text{inconsistent}} = 3.75, SD_{\text{inconsistent}} = 1.49; b = 0.07, t(810) = 1.01, p = .32$). Lastly, we explored whether conservatives preferred consistent voters over inconsistent voters within their in-group specifically to a greater extent than liberals (H3c). Once again, this prediction was strongly supported, with a significant interaction between political affiliation and preference for consistency when focusing solely on participants' in-group members (Right-Wing: $M_{\text{consistent}} = 4.81, SD_{\text{consistent}} = 1.74, M_{\text{inconsistent}} = 3.99, SD_{\text{inconsistent}} = 1.68$; Left-Wing: $M_{\text{consistent}} = 4.19, SD_{\text{consistent}} = 1.52, M_{\text{inconsistent}} = 4.08, SD_{\text{inconsistent}} = 1.46; b = 0.71, t(1008) = 6.24, p < .001$)

Discussion

In Experiment 2 we found consistent support for the in-group preference and principled loyalty hypotheses. This is consistent with findings from Experiment 1, which examined these hypotheses in a different country using a different experimental paradigm. In addition, unlike in Experiment 1, we also found support for the ideology moderation hypotheses. A few possibilities

exist for why evidence for this last hypothesis was absent in Experiment 1 and present in Experiment 2, which we discuss in greater detail in the General Discussion.

General Discussion

Across two studies examining voters in two different countries comprising different governing structures, we examined three overarching hypotheses regarding how people perceive political converts—that is, individuals who switched their allegiance from one political party to another (see Table 2 and Figures 1-3 for a summary of the results). Overall, we found consistent support for both the in-group preference hypothesis and the principled loyalty hypothesis using an incentive-compatible behavioral measure (Experiment 1) and a self-reported attitudinal measure (Experiment 2). In support of the in-group preference hypothesis, people trusted current members of their in-group more than current members of their out-group (H1a) and also trusted people who switched into their political in-group more than people who switched out of it (H1b). In support of the principled loyalty hypothesis, people trusted consistent voters who maintained their party support over political converts who switched their party support (H2a) both for current out-group members (in Studies 1 and 2) and for current in-group members (in Experiment 2 but not in Experiment 1; H2b).

We also found clear support for the ideology moderation hypothesis in Experiment 2, but no support for this hypothesis in Experiment 1. Specifically, in Experiment 2 (but not Experiment 1) conservatives trusted current in-group members more than current out-group members (H3a), consistent voters more than party switchers (H3b) and consistently-voting in-group members more than party-switching in-group members (H2c) to a greater extent than liberals did. There are potential empirical and theoretical reasons for these discrepant findings. Empirically, Experiment 2 included more participants, employed a within-subject design, and

used a continuous measure of perceived trustworthiness rather than a binary measure of behavioral trust; as such, it was better-powered and thus Experiment 1 may have been underpowered to detect these effects. Theoretically, there may be important contextual differences between the two countries examined that may contribute to the divergent effects. Cultural differences in country-level collectivism, in the structure (two-party versus multiparty) and stability of each political system (Israel had five elections in the past four years⁶) and in political polarization and norms regarding voting could have all potentially contributed to the pattern of findings.

Theoretical Implications

This work contributes to a number of different literatures. First, we contribute to prior research examining political party identification and polarization. In line with this work, we find a strong preference for one's own party members (Chambers, Schlenker, & Collisson, 2013; Cohen 2003). However, the support for the principled loyalty hypothesis in both studies adds additional nuance to this picture. These findings suggest that people do not view the out-group as a homogenized whole, but rather differentiate between different out-group members whom they trust more versus less. In particular, people perceived consistent members of the opposing party to be more trustworthy than inconsistent ones, even though consistent members likely hold the strongest opposing views on political topics (Zlatev, 2019). This suggests that people's perceptions of out-party members are more complex than previously identified, for example in research on the out-group homogeneity effect (Ostrom & Sedikides, 1992). Future research should examine whether these findings extend to other social groups people can choose to join or leave, such as converting to another religion or switching sports team allegiances.

⁶ <https://www.cnn.com/2022/10/30/middleeast/israel-elections-explainer-intl>

Second, we contribute to a growing literature on signals of trustworthiness in interpersonal perceptions. Recent work has identified a number of factors that play a role in people's evaluation of others' integrity or benevolence—and thus their trustworthiness—including escalating commitment to a failing course of action (Dorison, Umphres, & Lerner, 2021), feigning emotions (Levine & Wald, 2020) or acknowledging emotions in others (Yu, Berg, & Zlatev 2021) and brokering social ties (Halevy, Halali, & Cohen, 2020). We provide an additional behavioral signal of perceived lower integrity: switching political parties. In doing so we add to the developing picture of how people form impressions of others and manage their own impressions, particularly in the domain of trust.

Finally, we contribute to the literature on ideology and social cognition (Jost, Federico, & Napier, 2009; Jost, Nosek, & Gosling, 2008). Specifically, we find mixed evidence that conservatives prefer in-group members and consistent or loyal others to a greater extent than liberals do (Graham, Haidt, & Nosek, 2009; Jost, Glaser, Kruglanski, & Sulloway, 2003). As mentioned above, it is unclear whether these discrepant findings across our two experiments are due to methodological or theoretical reasons; however, they suggest the possibility that either differences in national culture, governance structure, or political context may moderate some of these prior findings. Future work should explore these possibilities.

Strengths, Limitations, and Future Directions

Integrating the literatures on intergroup relations, social perception, and trust, we derived multiple hypotheses concerning reactions to political converts. Although our experiments used large samples and employed different methodologies to test our hypotheses, the current research has certain limitations that point to promising directions for future research. First, the current research is limited in its focus on just two types of political newcomers, namely those who have

switched allegiances from the out-party to the in-party or vice versa. Future research may compare reactions to political converts and reactions to other kinds of political newcomers, including immigrants who have gained citizenship and teenagers who have just become eligible to vote. Future research may also compare reactions to individuals who switch their political allegiances with reactions to individuals who change group membership in other contexts, including both moralized (e.g., nations, social movements, ideological groups) and non-moralized (e.g., work organizations, sport teams, consumer brands) domains.

A second question that awaits future research concerns the extent to which different motivations for switching political allegiances moderate reactions to political converts. For example, are political converts trusted more when the process that led to their switching involved thoughtful deliberation rather than an affective reaction (e.g., disgust with a candidate's scandalous behavior)? Third, the current research focused on the second-person perspective of individuals who made decisions to trust or not to trust others. Future research extend this to look at the first-person perspective of actors and examine the extent to which they are aware of the potential relational and reputational costs associated with switching political affiliation, as well as how such awareness shapes voting behavior. Finally, future research may go beyond capturing a snapshot of people's (un)willingness to trust political converts at a given moment in time by using repeated interaction paradigms to study the dynamics of trust over time (e.g., Bornstein, Kugler, & Ziegelmeyer, 2004).

Conclusion

Switching political parties is critical for the proper functioning of the democratic system. It is technically easy, incurs minimal financial costs, and is beneficial to the exchange of ideas within a society. We identify an important social cost to switching, namely that people distrust

political newcomers, whether they are switching into one's own party or an opposing party. Taken together with people's preference for trusting their own party members, as well as possible ideological differences in these perceptions, these results provide a more complete picture of how people interact with each other in politicized settings.

References

- Balliet, D., Wu, J., & De Dreu, C. K. (2014). In-group favoritism in cooperation: A meta-analysis. *Psychological Bulletin, 140*(6), 1556.
- Blodgett, G. T. (1962). The Mind of the Boston Mugwump. *The Mississippi Valley Historical Review, 48*(4), 614-634.
- Bornstein, G., Kugler, T., & Ziegelmeyer, A. (2004). Individual and group decisions in the centipede game: Are groups more “rational” players?. *Journal of Experimental Social Psychology, 40*(5), 599-605.
- Brewer, M. B., & Silver, M. (1978). In-group bias as a function of task characteristics. *European Journal of Social Psychology.*
- Brooks, C., Manza, J., & Cohen, E. D. (2016). Political ideology and immigrant acceptance. *Socius, 2*, 1-12.
- Cameron, J. E. (2004). A three-factor model of social identity. *Self and Identity, 3*(3), 239-262.
- Chambers, J. R., Schlenker, B. R., & Collisson, B. (2013). Ideology and prejudice: The role of value conflicts. *Psychological Science, 24*(2), 140-149.
- 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.
- Cohen, T. R., & Morse, L. (2014). Moral character: What it is and what it does. *Research in Organizational Behavior, 34*, 43-61.
- Deaux, K., Reid, A., Mizrahi, K., & Ethier, K. A. (1995). Parameters of social identity. *Journal of personality and social psychology, 68*(2), 280.
- De Dreu, C. K., Balliet, D., & Halevy, N. (2014). Parochial cooperation in humans: Forms and functions of self-sacrifice in intergroup conflict. *Advances in Motivation Science, 1*, 1-47.

- Dinesen, P. T. (2012). Does generalized (dis) trust travel? Examining the impact of cultural heritage and destination-country environment on trust of immigrants. *Political Psychology, 33*(4), 495-511.
- Dorison, C. A., Umphres, C. K., & Lerner, J. S. (2022). Staying the course: Decision makers who escalate commitment are trusted and trustworthy. *Journal of Experimental Psychology: General, 151*(4), 960.
- Finkel, E. J., Bail, C. A., Cikara, M., Ditto, P. H., Iyengar, S., Klar, S., ... & Druckman, J. N. (2020). Political sectarianism in America. *Science, 370*(6516), 533-536.
- Gift, K., & Gift, T. (2015). Does politics influence hiring? Evidence from a randomized experiment. *Political Behavior, 37*(3), 653-675.
- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*(5), 1029.
- Green, D. P., Palmquist, B., & Schickler, E. (2004). *Partisan hearts and minds: Political parties and the social identities of voters*. Yale University Press.
- Greene, S. (2004). Social identity theory and party identification. *Social Science Quarterly, 85*(1), 136-153.
- Gruenfeld, D. H., & Fan, E. T. (1999). What newcomers see and what oldtimers say: Discontinuities in knowledge exchange. In *Shared cognition in organizations* (pp. 245-266). Psychology Press.
- Hainmueller, J., & Hopkins, D. J. (2014). Public attitudes toward immigration. *Annual Review of Political Science, 17*.
- Halevy, N., Halali, E., & Cohen, T. R. (2020). Brokering orientations and social capital:

- Influencing others' relationships shapes status and trust. *Journal of Personality and Social Psychology*, 119(2), 293.
- Huber, G. A., & Malhotra, N. (2017). Political homophily in social relationships: Evidence from online dating behavior. *The Journal of Politics*, 79(1), 269-283.
- Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, 22(1), 129-146.
- Jost, J. T., Federico, C. M., & Napier, J. L. (2009). Political ideology: Its structure, functions, and elective affinities. *Annual Review of Psychology*, 60(1), 307-337.
- Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, 129(3), 339-375.
<https://doi.org/10.1037/0033-2909.129.3.339>
- Jost, J. T., Nosek, B. A., & Gosling, S. D. (2008). Ideology: Its resurgence in social, personality, and political psychology. *Perspectives on Psychological Science*, 3(2), 126-136.
- Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, 28(2), 107.
- Kreps, T. A., Laurin, K., & Merritt, A. C. (2017). Hypocritical flip-flop, or courageous evolution? When leaders change their moral minds. *Journal of Personality and Social Psychology*, 113(5), 730.
- Levine, E. E., & Schweitzer, M. E. (2015). Prosocial lies: When deception breeds trust. *Organizational Behavior and Human Decision Processes*, 126, 88-106.
- Levine, E. E., & Wald, K. A. (2020). Fibbing about your feelings: How feigning happiness in the face of personal hardship affects trust. *Organizational Behavior and Human Decision Processes*, 156, 135-154.

- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709-734.
- McAllister, D. J. (1995). Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal*, 38(1), 24-59.
- McFall, L. (1987). Integrity. *Ethics*, 98(1), 5-20.
- Ostrom, T. M., & Sedikides, C. (1992). Out-group homogeneity effects in natural and minimal groups. *Psychological Bulletin*, 112(3), 536.
- Pew Research Center. (2020). *Voters rarely switch parties, but recent shifts further educational, racial divergence*. <https://www.pewresearch.org/politics/2020/08/04/voters-rarely-switch-parties-but-recent-shifts-further-educational-racial-divergence/>
- Phillips, K. W., Liljenquist, K. A., & Neale, M. A. (2009). Is the pain worth the gain? The advantages and liabilities of agreeing with socially distinct newcomers. *Personality and Social Psychology Bulletin*, 35(3), 336-350.
- Pitkin, H. F. (1967). *The concept of representation*. Univ of California Press.
- van der Linden, M., Hooghe, M., de Vroome, T., & Van Laar, C. (2017). Extending trust to immigrants: Generalized trust, cross-group friendship and anti-immigrant sentiments in 21 European societies. *PloS one*, 12(5), e0177369.
- Voelkel, J. G., Stagnaro, M. N., Chu, J., Pink, S., Mernyk, J. S., Redekopp, C., ... & Willer, R. (2022). Megastudy identifying successful interventions to strengthen Americans' democratic attitudes. Preprint.
- Wilson, G. D. (1973). A dynamic theory of conservatism. In G. D. Wilson, *The Psychology of Conservatism*. Academic Press.
- Ysseldyk, R., Matheson, K., & Anisman, H. (2010). Religiosity as identity: Toward an

understanding of religion from a social identity perspective. *Personality and Social Psychology Review*, 14(1), 60-71.

Yu, A., Berg, J. M., & Zlatev, J. J. (2021). Emotional acknowledgment: How verbalizing others' emotions fosters interpersonal trust. *Organizational Behavior and Human Decision Processes*, 164, 116-135.

Zlatev, J. J. (2019). I may not agree with you, but I trust you: Caring about social issues signals integrity. *Psychological Science*, 30(6), 880-892.

Table 1. Overview of hypotheses.

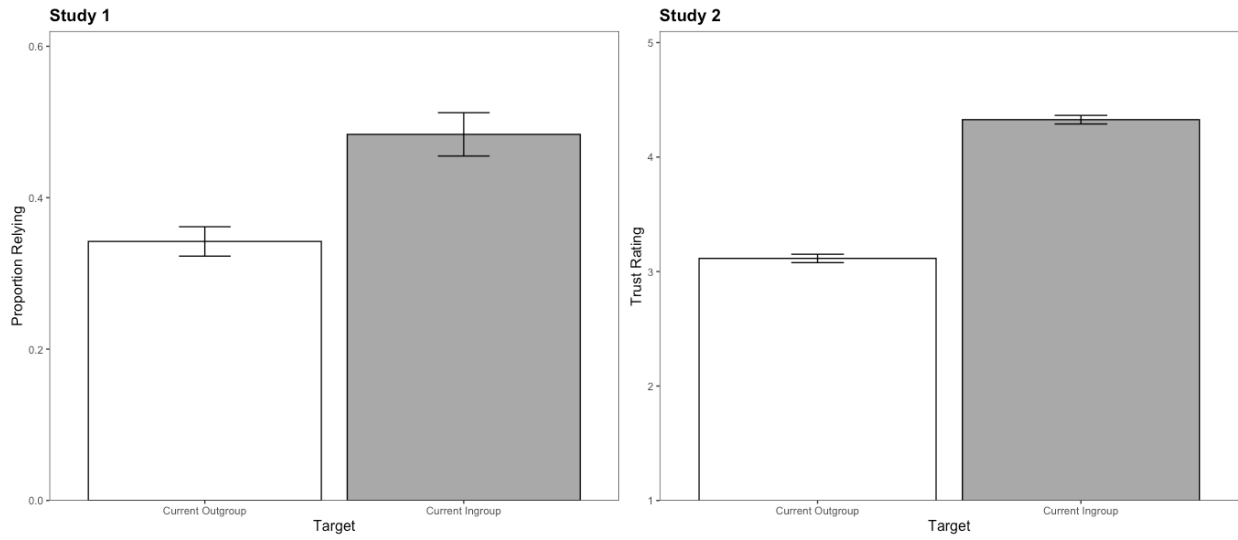
In-group Preference Hypothesis	H1a. People will trust current in-party members over current out-party members.
	H1b. People will trust party switchers who end up a member of one's in-party over party switchers who end up a member of one's out-party.
Principled Loyalty Hypothesis	H2a. People who maintain the same party affiliation over time (i.e., party maintainers) will be trusted more than people who switch their party affiliation (i.e., party switchers).
	H2b. People will trust party maintainers over party switchers when evaluating both in-party targets and out-party targets.
Ideology Moderation Hypothesis	H3a. Conservatives will show a stronger in-party preference than liberals.
	H3b. Conservatives will show a stronger consistency preference than liberals.
	H3c. Conservatives will show a stronger consistency preference for the in-party than liberals.

Table 2. Support for hypotheses by study.

		Experiment 1	Experiment 2
In-group Preference Hypothesis			
H1a.	Current in-party members > Current out-party members	Yes	Yes
H1b.	Switched into party > Switched out of party	Yes	Yes
Principled Loyalty Hypothesis			
H2a.	Party maintainers > Party switchers	Yes	Yes
H2b.	Party maintainers > Party switchers for both in-party and out-party members	Mixed	Yes
Ideology Moderation Hypothesis			
H3a.	Conservative in-party preference > Liberal in-party preference	No	Yes
H3b.	Conservative consistency preference > Liberal consistency preference	No	Yes
H3c.	Conservative consistency preference > Liberal consistency preference for in-party	No	Yes

Figure 1. Results for in-group preference hypothesis for Studies 1 and 2.

H1a.



H1b.

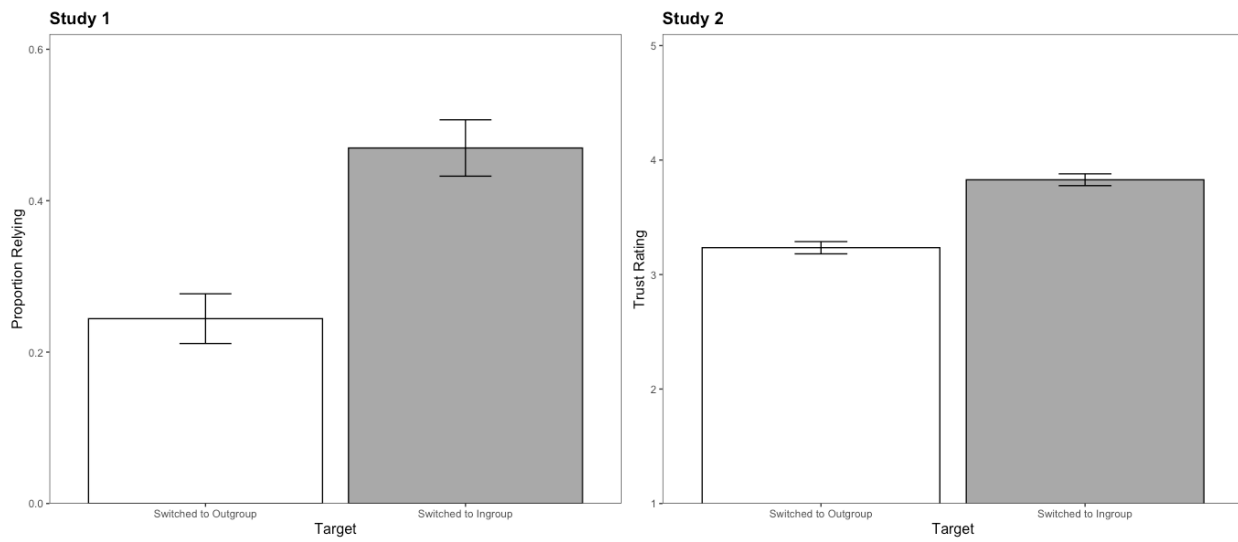
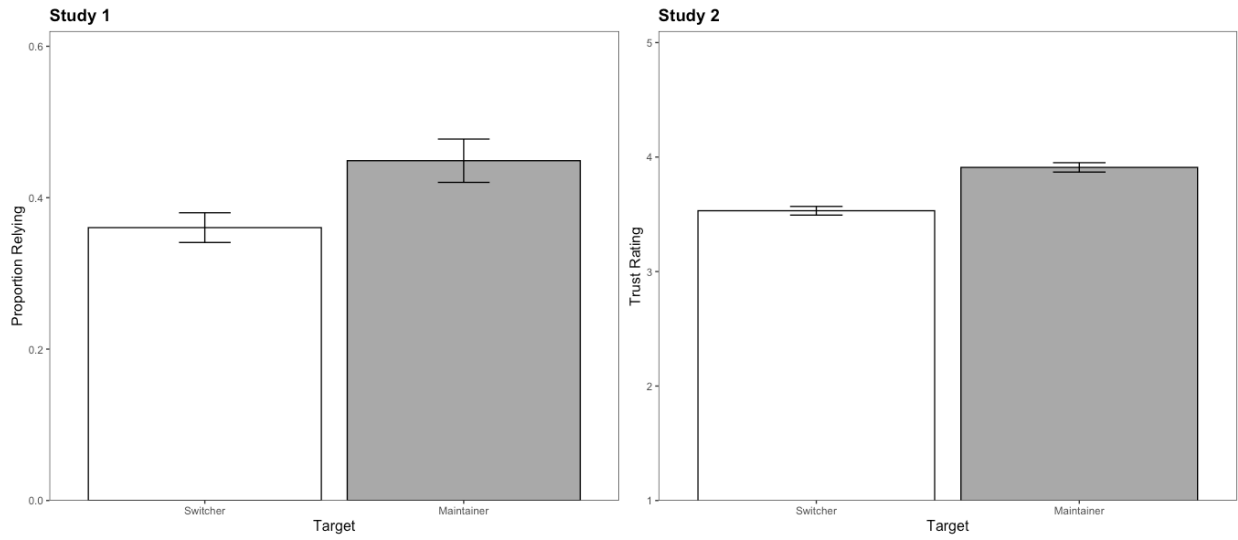


Figure 2. Results for principled loyalty hypothesis for Studies 1 and 2.

H2a.



H2b.

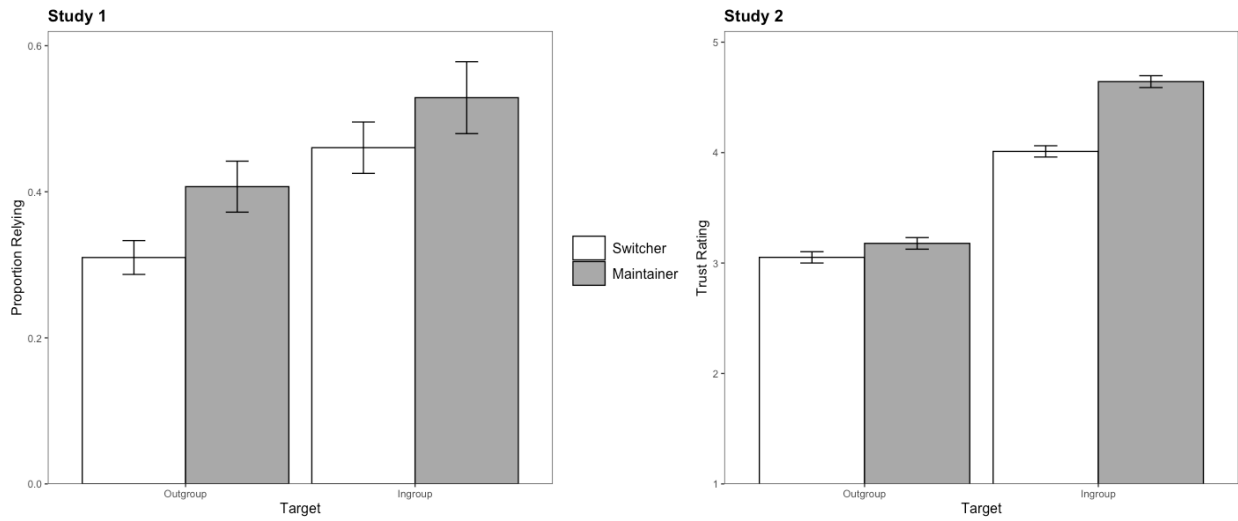
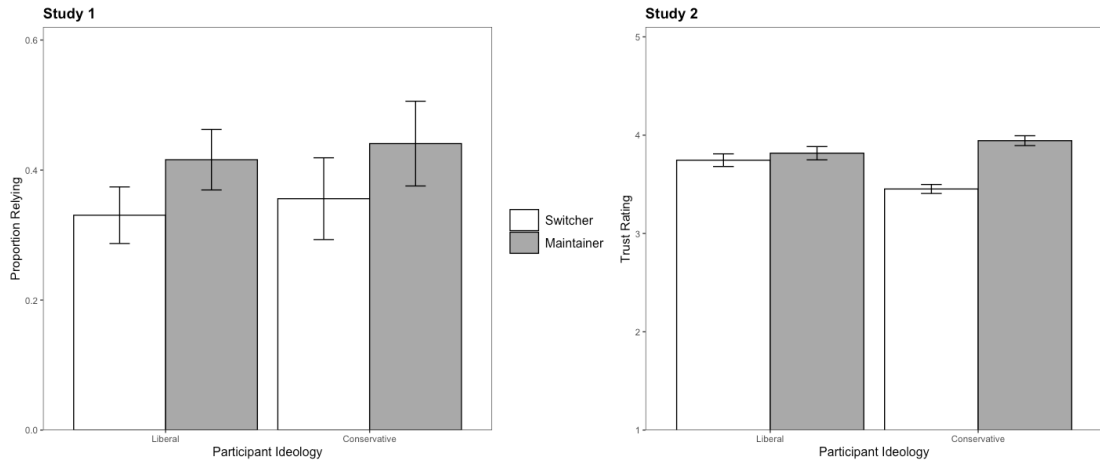


Figure 3. Results for ideology moderation hypothesis for Studies 1 and 2.

H3a.



H3b.



H3c.

