Learning from Twins: Can marketers learn about consumers across the political divide by interacting with AI

Benjamin Lira

2025-12-17

Table of contents

<div align="center">Sample Composition</div>

Learner Party	N
D→R	115
R→D	96

Overview

Political polarization creates challenges for marketers who must communicate effectively with ideologically diverse consumers. Misperceptions about political outgroups are widespread, with partisans often holding exaggerated and inaccurate views of those across the political divide. These misperceptions can lead to ineffective marketing strategies when firms attempt to appeal to broad consumer bases. This study tests whether interacting with an AI chatbot prompted to represent a political outgroup can improve marketers' understanding of that outgroup and reduce affective polarization.

Hypotheses

We hypothesized that after interacting with an AI chatbot representing their political outgroup, participants would:

1. **Report more accurate beliefs about the outgroup's attitudes** toward environmentally responsible consumption (Primary Outcome 1).
2. **Feel warmer toward the political outgroup** (Primary Outcome 2).

Data

Data Collection

Data were collected via Qualtrics from 211 participants. Participants self-reported their political orientation on a 0-100 slider scale, where values below 50 indicate Democratic leaning and values 50 and above indicate Republican leaning.

Measures

Primary Outcomes

- **Belief accuracy**: Absolute difference between participants' estimates of outgroup green attitudes and actual outgroup means (lower = more accurate)
- **Outgroup warmth**: Feeling thermometer (0-100 scale) toward the political outgroup

Secondary Outcomes

- Confidence in outgroup judgments (5-point scale)
- Bot informativeness and empathy ratings (5-point scale)

Pre-Registered Analyses

For each primary outcome, we estimated mixed-effects models of the form:

$$y \sim \text{time} \times \text{learner\_party} + (1|\text{id})$$

where `time` indexes pre- versus post-interaction measurement and `learner_party` indicates D→R versus R→D learning direction.
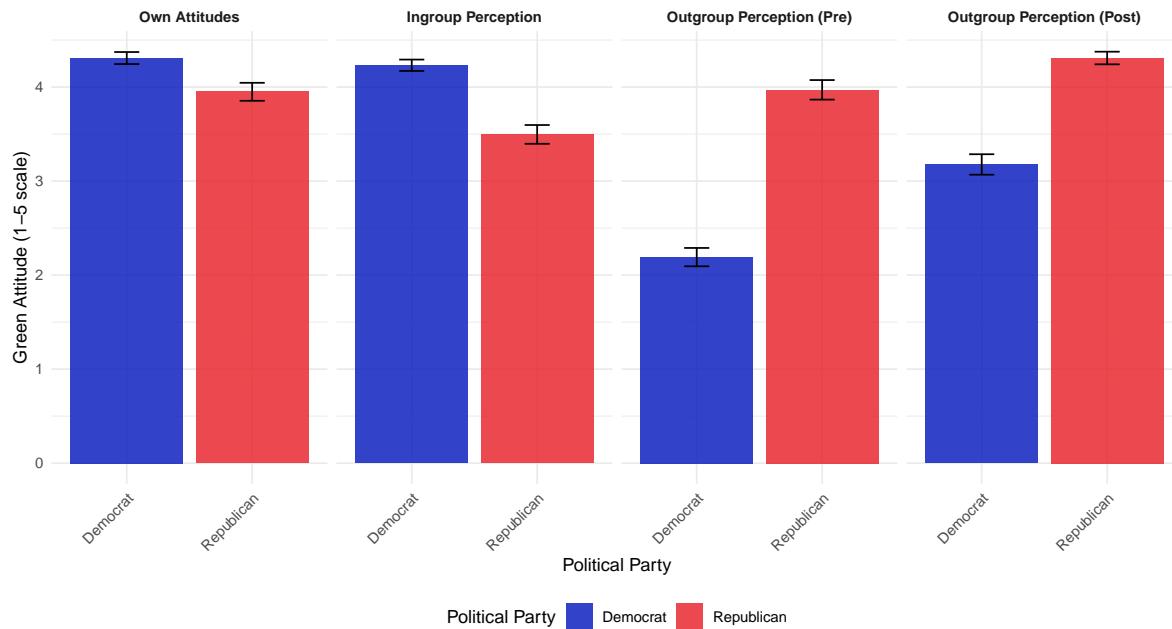
Results

Preliminaries

Democrats' and Republicans' actual green values and their perceptions of ingroup and outgroup attitudes.
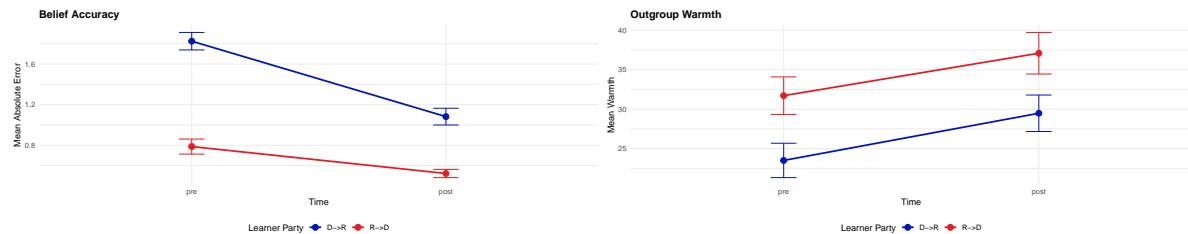
Green Attitudes: Actual and Perceived

Own attitudes, ingroup perceptions, and outgroup perceptions before and after interaction

Q1. Do people update their beliefs after interacting with the chatbot?

We examine whether participants changed their beliefs about the outgroup's environmental attitudes from pre- to post-interaction, analyzing both accuracy and warmth.

Q2. Is belief updating symmetric across political groups?

The primary analyses above test whether Democrats and Republicans show different patterns of belief updating. The key test is the `time × learner_party` interaction in the mixed-effects models. A significant interaction would indicate asymmetric updating across groups.

	Accuracy	Warmth
(Intercept)	1.353***	27.237***
	(0.060)	(1.692)
timepost	−0.525***	5.701***
	(0.062)	(0.957)
SD (Intercept ResponseId)	0.600	22.520
SD (Observations)	0.641	9.835
Num.Obs.	422	422

+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001
Note: Accuracy is absolute error (lower = more accurate). * p<0.05, ** p<0.01, *** p<0.001

	Accuracy	Warmth
(Intercept)	1.825***	23.504***
	(0.072)	(2.267)
timepost	−0.742***	5.974***
	(0.082)	(1.300)
learner_partyR→D	−1.037***	8.204*
	(0.107)	(3.361)
timepost × learner_partyR→D	0.477***	−0.599
	(0.121)	(1.927)
SD (Intercept ResponseId)	0.466	22.225
SD (Observations)	0.620	9.856
Num.Obs.	422	422

+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001
Note: Accuracy is absolute error (lower = more accurate). * p<0.05, ** p<0.01, *** p<0.001

	Accuracy	Warmth
timepost	−0.628*	2.920
	(0.255)	(4.018)
learner_partyR→D	−0.919**	15.267+
	(0.296)	(8.994)
political_extremism	0.011*	−0.394*
	(0.005)	(0.162)
timepost × learner_partyR→D	0.621+	−4.195
	(0.337)	(5.315)
timepost × political_extremism	−0.003	0.077
	(0.006)	(0.096)
learner_partyR→D × political_extremism	−0.002	−0.250
	(0.007)	(0.222)
timepost × learner_partyR→D × political_extremism	−0.004	0.111
	(0.008)	(0.131)
SD (Observations)	0.620	9.785
Num.Obs.	422	422

+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001
Note: Three-way interaction tests whether extremism moderates differential updating. * p<0.05, ** p<0.01, *** p<0.001

Q3. Does political extremity predict how much people update?

We explore whether participants' political extremism (distance from the political center) moderates the treatment effect.

Q4. Is belief updating associated with how the chatbot is perceived?
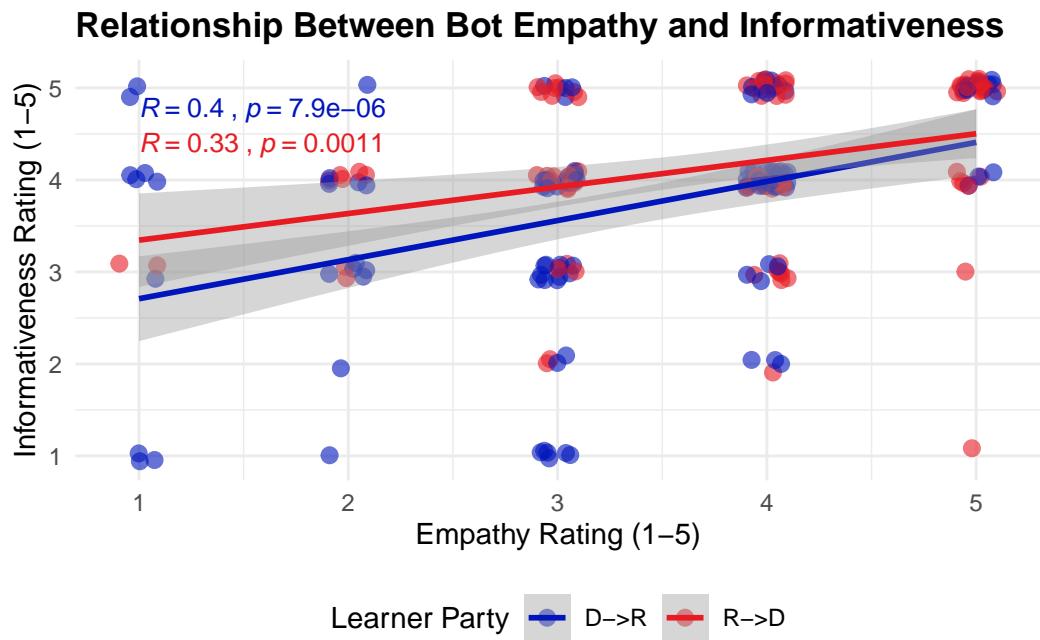
We examine whether participants who found the chatbot more informative or empathic showed greater belief updating, controlling for pre-interaction values.

Do perceptions of informativeness and empathy correlate with each other?

	Accuracy	Warmth
(Intercept)	1.174***	−6.583+
	(0.220)	(3.960)
accuracy_pre	0.326***	
	(0.047)	
info	−0.116*	2.621**
	(0.047)	(0.955)
empathy	−0.095*	1.543
	(0.045)	(0.945)
warmth_outgroup_pre		0.875***
		(0.040)
Num.Obs.	211	211
R2	0.292	0.730
R2 Adj.	0.282	0.726

+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001
Note: Regressions control for pre-interaction
values. * p<0.05, ** p<0.01, *** p<0.001

Relationship Between Bot Empathy and Informativeness

$R = 0.4$, $p = 7.9e{-}06$
$R = 0.33$, $p = 0.0011$

Informativeness Rating (1–5)

Empathy Rating (1–5)

Learner Party — D–>R — R–>D

Q5. Does engagement ("dose") matter for belief updating?

The pre-registration specified examining whether engagement with the chatbot—operationalized as the number of words typed by participants during the chat—is associated with post-interaction outcomes. We test whether greater engagement predicts more belief updating and warmth change.

```
Engagement metrics:

  Turn count: M = 10.06, SD = 4.21

  Word count: M = 182.85, SD = 4.21
```

Does engagement predict belief accuracy improvement?

Greater engagement—measured by turn count—was not significantly associated with accuracy change ($b = 0.0149$, $p = 0.169$). Using log-transformed word count, the effect was not significant ($b = $ -0.0297, $p = 0.799$).

	Turn Count	Log(Word Count)
(Intercept)	0.373*	0.672
	(0.158)	(0.606)
Turn Count	0.015	
	(0.011)	
Baseline Accuracy	−0.696***	−0.691***
	(0.054)	(0.054)
Learner: R→D	−0.234*	−0.241*
	(0.107)	(0.107)
Log(Word Count)		−0.030
		(0.116)
Num.Obs.	211	211
R2	0.482	0.478
R2 Adj.	0.475	0.470
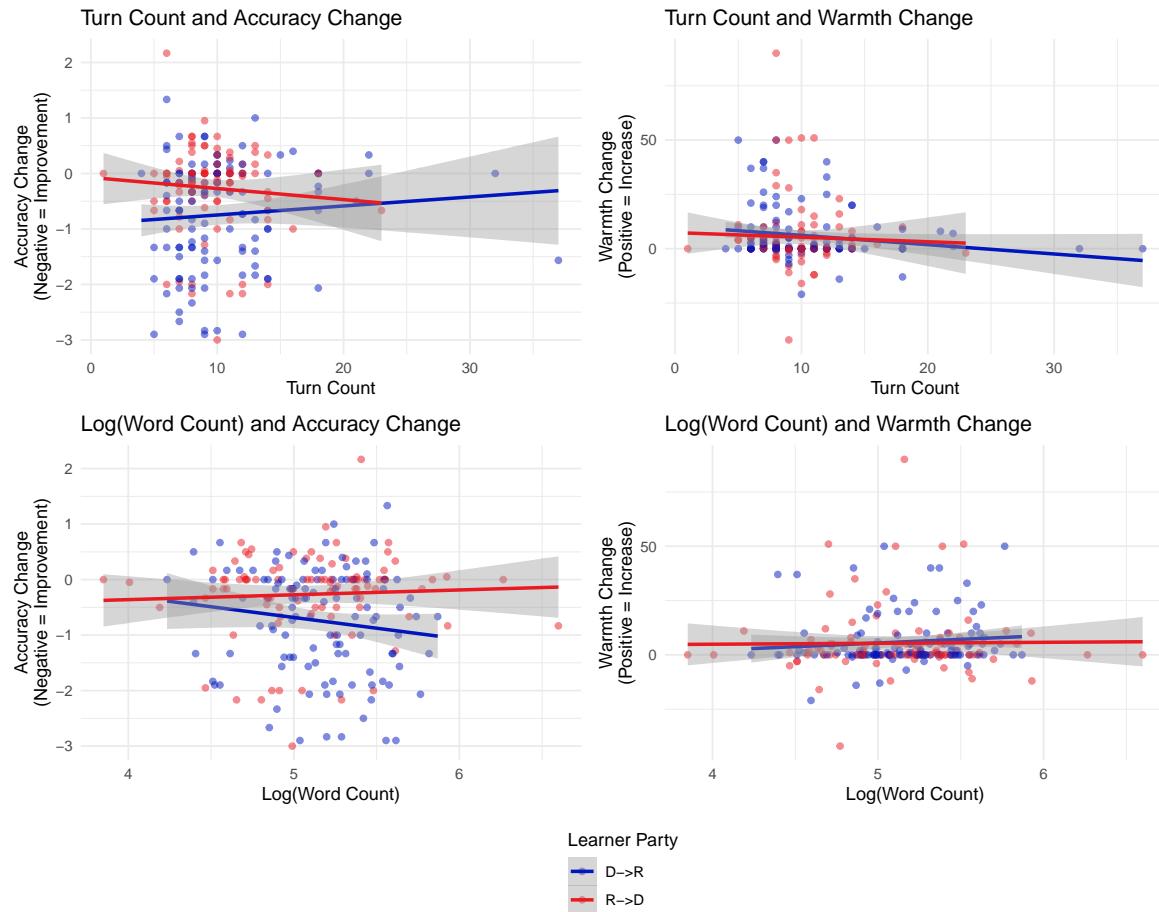
+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001

	Turn Count	Log(Word Count)
(Intercept)	12.035***	2.140
	(2.847)	(12.817)
Turn Count	−0.376+	
	(0.226)	
Baseline Warmth	−0.092*	−0.089*
	(0.041)	(0.041)
Learner: R→D	−0.125	0.194
	(1.937)	(1.944)
Log(Word Count)		1.151
		(2.448)
Num.Obs.	211	211
R2	0.037	0.025
R2 Adj.	0.023	0.011

+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001

Does engagement predict warmth change?

Greater engagement—measured by turn count—was not significantly associated with warmth change (b = -0.3762, p = 0.098). Using log-transformed word count, the effect was not significant (b = 1.1508, p = 0.639).

Q6. Does interacting with the chatbot improve applied judgment?

Participants wrote slogans for marketing energy-saving appliances to the political outgroup.
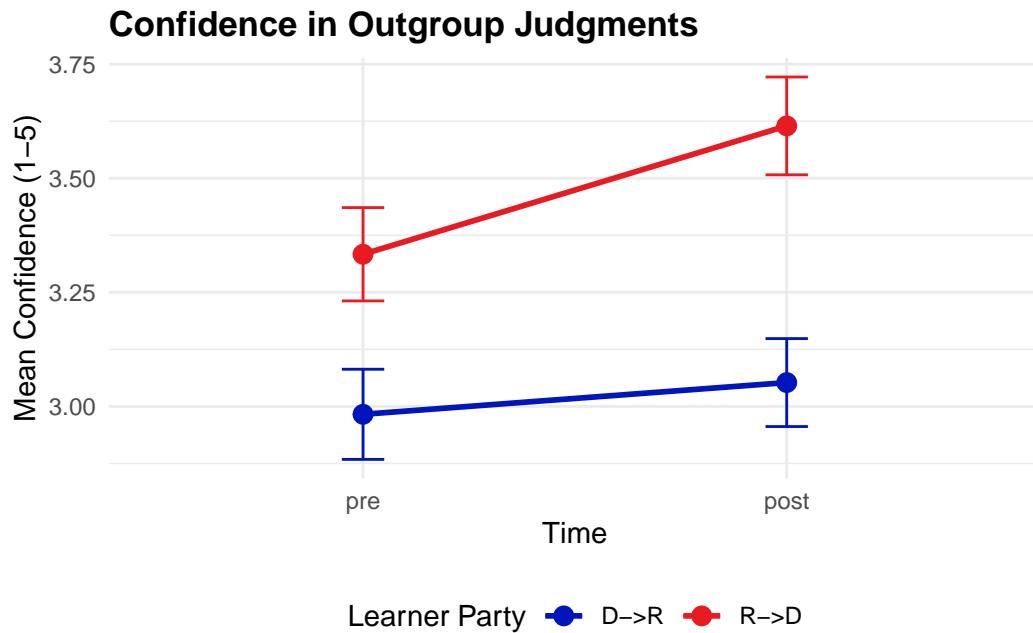
DATA AVAILABLE: Raw slogan text in variable `Q23` (n = 211 responses).

DATA NEEDED: 1. Semantic similarity scores (compare to AI-generated responses to detect copying) 2. Quality ratings (human or LLM ratings of slogan fit to target outgroup)

Q7. Does confidence track accuracy—or diverge from it?

We examine participants' metacognitive awareness by testing (1) whether confidence changes from pre to post, and (2) whether confidence changes track with actual accuracy improvements.

Model: Does confidence change pre-to-post?

Correlation: Does confidence change track accuracy improvement?
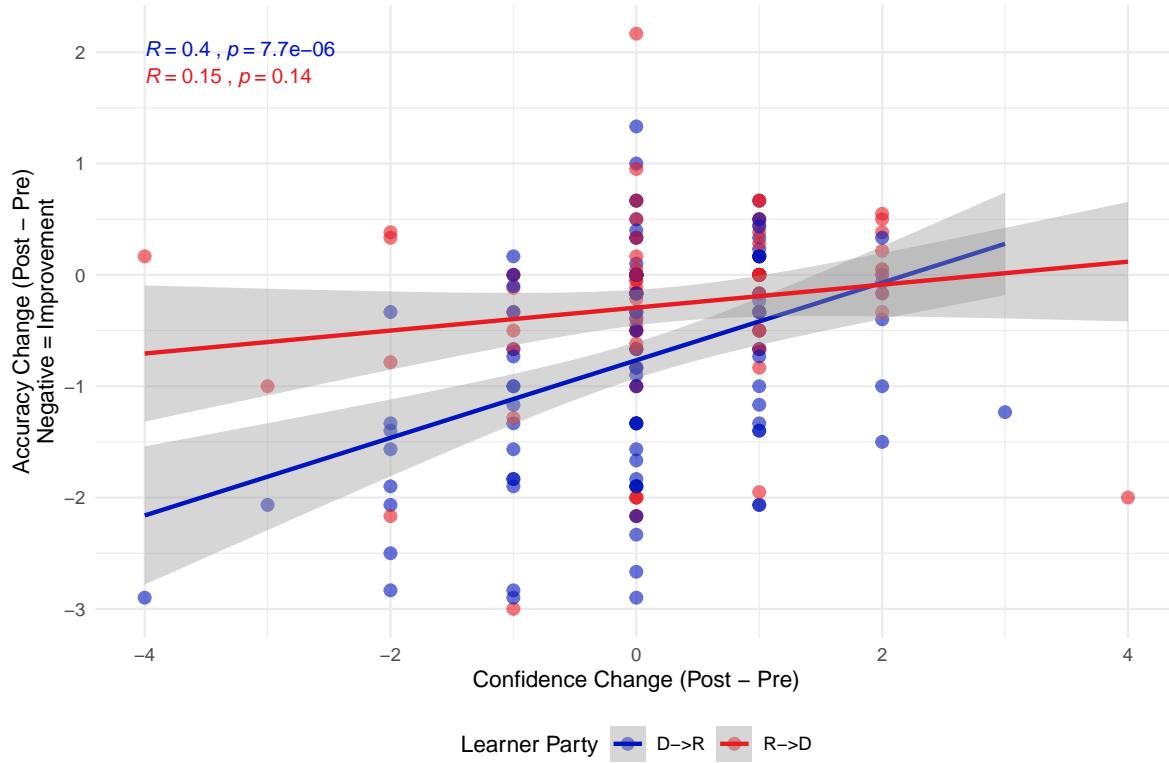
	Confidence (Main Effect)	Confidence (Interaction)
(Intercept)	3.142***	2.983***
	(0.073)	(0.097)
timepost	0.166*	0.070
	(0.077)	(0.104)
learner_partyR→D		0.351*
		(0.143)
timepost × learner_partyR→D		0.212
		(0.154)
SD (Intercept ResponseId)	0.706	0.672
SD (Observations)	0.792	0.790
Num.Obs.	422	422

+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001
Note: Confidence rated on 1-5 scale. * p<0.05, ** p<0.01, *** p<0.001

Confidence–Accuracy Calibration

Does confidence change track with accuracy improvement?

Note: Negative correlation would indicate increased confidence associated with improved accuracy

	Warmth Difference
(Intercept)	−52.487***
	(3.020)
timepost	7.600***
	(1.649)
learner_partyR→D	3.727
	(4.478)
timepost × learner_partyR→D	0.108
	(2.444)
SD (Intercept ResponseId)	29.879
SD (Observations)	12.500
Num.Obs.	422

+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001
Note: Warmth difference = outgroup - ingroup. Positive time coefficient indicates reduced polarization. * p<0.05, ** p<0.01, *** p<0.001

Robustness Checks

Warmth Difference Score

Following common practice in the affective polarization literature, we examine warmth using a difference score (outgroup - ingroup).

Discussion

Summary of Findings

This pilot study tested whether interacting with an AI chatbot representing a political outgroup could improve participants' understanding of and warmth toward that outgroup. Using a within-subjects pre-post design, we examined changes in outgroup belief accuracy and outgroup warmth following an AI-mediated interaction.

Primary Hypotheses

Hypothesis 1: Outgroup Belief Accuracy

We hypothesized that participants would report more accurate beliefs about the outgroup's environmental attitudes after interacting with the AI chatbot. Accuracy was operationalized as the absolute difference between participants' estimates and the outgroup's actual mean green attitudes. The time coefficient in the mixed-effects model indicates whether accuracy improved (error decreased) from pre to post, while the interaction term tests whether this change differed between Democrat and Republican learners.

Hypothesis 2: Outgroup Warmth

We hypothesized that participants would feel warmer toward the political outgroup after the interaction. The time coefficient indicates whether warmth increased from pre to post, while the interaction term tests whether Democrats and Republicans showed different patterns of change.

Exploratory Findings

Secondary analyses examined confidence in outgroup judgments and perceptions of the bot's informativeness and empathy. Confidence ratings provide insight into participants' metacognitive awareness of their knowledge about the outgroup. Bot perception ratings help illuminate potential mechanisms—if participants found the bot informative and empathic, this may mediate changes in accuracy and warmth.

Implications for Marketing Practice

If the hypothesized effects are supported, this research suggests that AI-mediated interactions can serve as a scalable tool for reducing political misperceptions in consumer research. Marketers often struggle to understand ideologically diverse consumer segments, relying on stereotypes or costly focus groups. AI chatbots could provide an accessible method for gaining insight into outgroup consumer preferences and values.

This approach may be particularly valuable for marketing environmentally responsible products, where political divides can lead to misperceptions about consumer attitudes. Understanding that political outgroups may be more environmentally conscious than stereotypes suggest could inform more effective cross-partisan messaging strategies.

Limitations

Several limitations should be noted. The pilot sample consists primarily of test and preview responses, limiting statistical power and generalizability. The environmental attitude domain may not generalize to other consumer preferences. Belief accuracy depends on having adequate sample sizes within each political group to compute reliable actual outgroup means. The

Actual Environmental Attitudes by Political Party
Self-reported green attitudes (1-5 scale)

Political Party	Mean Green Attitude	SD	N
Democrat	4.31	0.69	115
Republican	3.95	0.94	96

study examines only immediate post-interaction effects; durability of any changes remains unknown.

Additionally, participants interacted with an AI chatbot rather than actual outgroup members. While this approach offers scalability advantages, it raises questions about whether insights gained from AI interactions transfer to understanding real outgroup members. The chatbot was prompted to represent the outgroup but may not fully capture the diversity of views within each political group.

Next Steps

The full study will recruit N = 500 participants from a representative sample. Key priorities include examining whether bot informativeness and empathy mediate changes in warmth and accuracy, analyzing participant-generated marketing slogans for energy-saving appliances, and testing for heterogeneous effects by political extremity and prior outgroup attitudes. Exploratory analyses will examine engagement (words typed during chat) as a predictor of outcomes.

Future research should examine the durability of effects over time, test generalizability to other product categories and consumer attitudes beyond environmental responsibility, and compare AI-mediated learning to alternative interventions such as reading testimonials from outgroup members or interacting with human outgroup representatives.

Appendix

Descriptive Statistics

Data Availability

Raw data: data/raw/qualtrics.parquet

Cleaned data: output/cleaned_data.parquet

Outgroup Belief Accuracy
Absolute error (lower = more accurate)

Learner Party	Time	Mean Error	SD	N
D→R	pre	1.825	0.924	115
D→R	post	1.083	0.884	115
R→D	pre	0.788	0.724	96
R→D	post	0.523	0.396	96

Outgroup Warmth (Feeling Thermometer)
By learner party and time point (0-100 scale)

Learner Party	Time	Mean Warmth	SD	N
D→R	pre	23.50	23.34	115
D→R	post	29.48	24.79	115
R→D	pre	31.71	23.28	96
R→D	post	37.08	25.83	96

Political Extremism by Learner Party
Distance from political center (0-50 scale)

Learner Party	Mean	SD	Min	Max
D→R	39.80	13.56	2.00	50.00
R→D	35.33	15.78	0.00	50.00

Bot Perception Ratings
By learner party (1-5 scale)

Learner Party	Measure	Mean Rating	SD	N
D→R	Empathy	3.32	1.10	115
D→R	Informativeness	3.70	1.16	115
R→D	Empathy	3.80	0.99	96
R→D	Informativeness	4.16	0.87	96

<div align="center">

Confidence in Outgroup Judgments

By learner party and time point (1-5 scale)

</div>

Learner Party	Time	Mean Confidence	SD	N
D→R	pre	2.98	1.06	115
D→R	post	3.05	1.03	115
R→D	pre	3.33	1.00	96
R→D	post	3.61	1.05	96

Code Availability

Data cleaning script: src/clean_data.R

Pre-Registration

This study was pre-registered prior to data collection. See docs/pre-registration.md for complete details including hypotheses, measures, sample size justification, and planned analyses.

Session Information

```
R version 4.5.0 (2025-04-11)
Platform: aarch64-apple-darwin20
Running under: macOS 26.1

Matrix products: default
BLAS:   /Library/Frameworks/R.framework/Versions/4.5-arm64/Resources/lib/libRblas.0.dylib
LAPACK: /Library/Frameworks/R.framework/Versions/4.5-arm64/Resources/lib/libRlapack.dylib;

locale:
[1] C.UTF-8/C.UTF-8/C.UTF-8/C/C.UTF-8/C.UTF-8

time zone: America/New_York
tzcode source: internal

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base

other attached packages:
 [1] patchwork_1.3.2    ggpubr_0.6.0      gt_1.0.0           modelsummary_2.3.0 lmerTest_3.
 [7] Matrix_1.7-3       arrow_19.0.1.1    lubridate_1.9.4    forcats_1.0.0      stringr_1.5
```

```
[13] purrr_1.1.0          readr_2.1.5        tidyr_1.3.1        tibble_3.3.0        ggplot2_3.5

loaded via a namespace (and not attached):
 [1] tidyselect_1.2.1   farver_2.1.2        fastmap_1.2.0        bayestestR_0.16.1   egg_0.4
 [6] digest_0.6.37       timechange_0.3.0   estimability_1.5.1   lifecycle_1.0.4     magritt
[11] compiler_4.5.0      rlang_1.1.6        tools_4.5.0          yaml_2.3.10         data.tal
[16] ggsignif_0.6.4      knitr_1.50         labeling_0.4.3       bit_4.6.0           xml2_1.3
[21] RColorBrewer_1.1-3  abind_1.4-8        tinytable_0.9.0      withr_3.0.2         numDeriv
[26] grid_4.5.0          datawizard_1.2.0   fansi_1.0.6          xtable_1.8-4        emmeans_
[31] scales_1.4.0        MASS_7.3-65        insight_1.4.0        cli_3.6.5           mvtnorm_
[36] rmarkdown_2.29      reformulas_0.4.1   generics_0.1.4       performance_0.15.0  tzdb_0.5
[41] parameters_0.28.0   minqa_1.2.8        splines_4.5.0        assertthat_0.2.1    effects_
[46] vctrs_0.6.5         boot_1.3-31        carData_3.0-5        jsonlite_2.0.0      car_3.1-
[51] litedown_0.7        hms_1.1.3          rstatix_0.7.2        bit64_4.6.0-1       Formula_
[56] glue_1.8.0          nloptr_2.2.1       stringi_1.8.7        gtable_0.3.6        tables_0
[61] lmtest_0.9-40       pillar_1.11.0      htmltools_0.5.8.1    R6_2.6.1            Rdpack_2
[66] evaluate_1.0.4      lattice_0.22-6     rbibutils_2.3        backports_1.5.0     broom_1
[71] Rcpp_1.0.14         coda_0.19-4.1      gridExtra_2.3        nlme_3.1-168        checkmat
[76] mgcv_1.9-1          xfun_0.52          zoo_1.8-14           pkgconfig_2.0.3
```