

Chatbot-driven Voting Aid Applications Increase Knowledge about Party Positions but Do Not Change Party Evaluations

Yamil R. Velez¹, Donald P. Green¹, and Semra Sevi²

¹Columbia University

²University of Toronto

June 12, 2025

Abstract

Voting Aid Applications (VAAs) are interactive tools that communicate information about elections, yet their effectiveness in enhancing political knowledge and participation remains understudied. Moreover, traditional VAAs may disproportionately attract politically engaged users with already well-formed ideological views, limiting their potential to inform a broader and less engaged electorate. This paper introduces a novel “VAA Bot” that employs large language models (LLMs) and retrieval-augmented generation (RAG) to deliver balanced, personalized information drawn from official party platforms and public documents. We evaluate the VAA Bot’s impact across three experimental studies aimed at young politically unaffiliated adults. The findings provide evidence that the VAA Bot improves knowledge of party stances on issues of great importance to each user. However, when we focus on respondents whose primary issue position aligns closely with one of the parties, we observe weaker effects on downstream outcomes such as vote preferences and party evaluations. These findings contribute to ongoing debates about the role of political information in shaping behavior and underscore both the promise and the limitations of LLM-based tools for civic learning.

Y.V. built the VAA bot. S.S. applied for and obtained grants. Both S.S. and Y.V. programmed the surveys on Qualtrics, contributed to the analysis, and prepared the original draft. Y.V., S.S., and D.G. contributed to the design of the study, IRB and ethics approval, data collection, and writing and editing the paper. To whom correspondence should be addressed. Email: semra.sevi@utoronto.ca

We thank Alexander Coppock, Andy Hall, Vincent Hopkins, Allen Jeeyong Kim, Bryn McCarthy, Can Mekik, Niki Monazzam, B. Pablo Montagnes, Ethan Porter, Daniel Rubenson, Connor Gardner Sharkey, Participants of 2024 APSA Annual Meeting, and 2025 CPSA Annual Meeting for helpful comments. Author names are listed in alphabetical order.

1 Introduction

Since *The American Voter* (Campbell et al., 1960) first conceptualized party identification as a stable psychological attachment to a political group, scholars have treated it as one of the most important predictors of political behavior. Across countries and time periods, this identity has shaped how individuals vote, whether they turn out, how they interpret political issues, and how they evaluate candidates and fellow citizens (Brader, Tucker and Duell, 2013; Green, Palmquist and Schickler, 2002; Huddy, Mason and Aarøe, 2015).

Yet a striking number of people, especially young people, do not identify with any political party. According to the Comparative Study of Electoral Systems, nearly half (45%) of Americans express no party preference, a figure that rises to 54% among those under 25 (CSES, 2016). This pattern is not unique to the U.S.: younger citizens and recent immigrants across democracies are significantly less likely to report a party attachment (Dalton and Wattenberg, 2002). These groups are far from marginal. In recent U.S. elections, individuals under 25 have accounted for roughly 10% of the electorate, with millions of new young voters entering the electorate each election cycle.

Despite their political importance, we know relatively little about how young people come to form party identities. Decades of political socialization research have shown that family and peer influences are critical in shaping how young people learn about politics and parties (Hyman, 1959; Jennings and Niemi, 1968). Yet recent work finds that the intergenerational transmission of party identification is often incomplete and ephemeral (Niemi and Jennings, 1991; Jennings, Stoker and Bowers, 2009; Hatemi and Ojeda, 2021). Many young adults enter the electorate with limited political interest, low turnout propensities, and little sense of party attachment (Campbell et al., 1960; Green, Palmquist and Schickler, 2002; Blais and Rubenson, 2013; Grumbach and Hill, 2022).

Why do so many young citizens remain politically unaffiliated? And what conditions help them develop a meaningful party identification? While much attention has focused on pre-adult socialization experiences, this paper explores a complementary mechanism: issue proximity. Building on classic models of issue voting (Downs, 1957; Stokes, 1963) and an extensive empirical literature (Franklin and Jackson, 1983; Jackson, 1975; Kollman and Jackson, 2021; Goren and Chapp, 2017), we examine whether young people can be drawn toward parties that more closely match their issue positions, especially when those positions are clarified through direct engagement. Given that many voters are uncertain about where parties stand (Alvarez, 1998), increasing the clarity and accessibility of party positions may foster a sense

of alignment that leads to stronger partisan identities.

This study introduces and evaluates an innovative approach: a conversational, AI-powered Voting Aid Application (VAA) chatbot designed to help politically disengaged users, particularly young independents, identify which parties best align with their policy preferences on issues of greatest importance to them. Unlike traditional VAAs that rely on static online questionnaires, our chatbot engages users in personalized, interactive conversations based on official party platforms, state voter guides, and speeches at party conventions.

We conducted three experiments in the United States to assess whether this AI-assisted VAA improves young independents' understanding of party positions and encourages partisan identification. We find that the VAA bot significantly increases subjects' accuracy when locating parties on key issues, especially on their self-identified "core concerns." However, this boost in knowledge does not appear to produce stronger partisan attachments or changes in vote intentions. These findings contribute to ongoing debates about the role of political knowledge in party alignment and the promise and limits of AI tools for enhancing democratic engagement.

Age, Political Engagement, and Partisan Identity

Age remains one of the most consistent predictors of political behavior. Across democracies, younger voters are significantly less likely to participate in elections than older cohorts, a gap that persists across time and countries (Jennings and Niemi, 1978; Smets, 2012; Angelucci, Carrieri and Improta, 2024; Sevi, 2021). In the United States, this disparity is especially stark, raising concerns about the representativeness and inclusiveness of democratic participation.

Age-related gaps extend well beyond turnout. Young voters are also less likely to identify with a political party and more prone to switching their partisan allegiances over time (Niemi and Jennings, 1991; Green, Palmquist and Schickler, 2002). These patterns reflect both life-cycle effects, where individuals gradually form stronger political identities with age, and generational effects shaped by distinct historical and political environments (Tilley and Evans, 2014).

Another dimension of this age gap concerns political knowledge. On average, younger individuals score lower on measures of political information than older voters (Stockemer and Rocher, 2017; Lupia, 2016). The size of these gaps depends on the type of knowledge being assessed. For example, the differences are smaller when measuring political knowledge during an election, as politics becomes more

salient (Jennings, 1996; Furnham and Gunter, 1987; Rosenqvist, 2020; Soule, 2001; Lau and Redlawsk, 2008; Stiers and Hooghe, 2021).

These differences in engagement, partisanship, and political knowledge raise important questions about how younger and less politically involved individuals access and process political information, and what tools might help bridge these divides.

Voting Aid Applications: A Tool for Bridging Gaps?

In response to these disparities, researchers and policymakers have developed Voting Aid Applications, digital tools that match users' policy preferences with the positions of political parties or candidates. Rather than persuading users to support specific parties or candidates, VAAs help individuals identify the party most aligned with their views. These tools are now widely used in democracies such as the Netherlands, Germany, Switzerland, the U.K., Canada, and the United States. In some countries, their adoption is exceptionally high: in the Netherlands, for example, the “StemWijzer” was used by nearly 40% of the electorate in 2006 (Kleinnijenhuis et al., 2007), and Germany’s “Wahl-O-Mat” reached over 15 million users during the 2017 federal election (Manavopoulos et al., 2018, p. 107).

The effectiveness of VAAs, however, remains a subject of debate. A recent meta-analysis reviewed 22 studies (17 observational and five experimental) using VAAs and found strong effects on self-reported turnout and moderate effects on vote choice, though only small improvements in issue knowledge (Munzert and Ramirez-Ruiz, 2021). Notably, these findings vary significantly by research design: observational studies show large, positive impacts, while experimental studies often find modest or null effects (Munzert and Ramirez-Ruiz, 2021). This pattern suggests that more politically engaged individuals are more likely to use VAAs, limiting the tool’s reach among the disengaged.

The key challenge is thus not whether VAAs work in general, but whether they can reach and meaningfully influence those with lower levels of political engagement, especially younger voters. If VAAs primarily serve the already-engaged, their democratic potential may be limited (Alvarez et al., 2014). New approaches might be needed to make these tools more accessible, personalized, and appealing to a broader range of users.

From Static Surveys to Smart Conversations

Recent advances in natural language processing have paved the way for Conversational Agent Voting Aid Applications (CAVAAs), chatbot-based tools that provide a more interactive and user-friendly experience (Kamoen, McCartan and Liebrecht, 2021). Unlike traditional VAAs, which rely on static questionnaires, CAVAAs engage users in dynamic conversations, adapting to their knowledge level and interests. Early versions of these tools used deterministic rule-based systems, but new developments in large language models (LLMs) allow for far more sophisticated interactions.

LLMs are trained on massive text datasets and are capable of producing human-like responses tailored to individual users. They can explain political issues, clarify party stances, and guide users through complex topics, all using accessible language. However, LLMs face a well-documented problem: they can generate false or misleading information, or “hallucinations,” that may misinform users, particularly in politically sensitive contexts.

To address this concern, our study uses a retrieval-augmented generation (RAG) framework that grounds chatbot responses in verified political content.¹ We constructed a corpus that includes 73 state party platforms (39 Democratic, 34 Republican), 19 state voter guides from 2022, and 12 national party speeches or town halls from 2020 and 2023. Each document was divided into 512-token chunks and embedded using OpenAI’s text-embedding-ada-002 model, which maps each chunk into a 1,536-dimensional vector space. When users interact with the chatbot (OpenAI’s GPT-4 in Studies 1 and 2; GPT-4o in Study 3), their queries are mapped onto the same embedding space, retrieving relevant documents and ensuring that responses are grounded in official party positions (For a schematic representation of the workflow, see SI Appendix, Fig. 4). It is worth noting that even without RAG, advanced LLMs such as those that we use in our studies exhibit high rates of factual accuracy, with professionally assessed accuracy rates of 99.2% observed when debunking conspiracy theories, many of which are political in nature Costello, Pennycook and Rand (2024).

To enhance engagement, the chatbot interface includes gamified features, such as unlockable dialogues that encourage users to identify which party is closest to them based on the conversation (see Figure 1). By combining official sources, open-ended dialogue, and engaging user experiences, the tool is designed to meet younger and less engaged voters where they are.

Using this tool, we assess whether advanced, personalized VAAs can improve political knowledge

¹While RAG-enabled VAAs are growing in popularity (Dieing, 2024; Vassos et al., 2024; Gittmann et al., 2024; Gemenis, 2024), existing studies focus more on implementation, factuality, and ethical considerations than on estimating their effects.

and foster stronger party attachments among a demographic often overlooked in traditional political outreach. More broadly, it contributes to the growing conversation about how AI can support informed and inclusive democratic participation. While recent applications of large language models in politics have stressed more troubling capabilities such as their ability to manipulate voters (Velez and Liu, 2024; Costello, Pennycook and Rand, 2024; Bai et al., 2025; Crabtree et al., 2025; Chen et al., 2025), our study points to the strengths of LLMs as information retrieval assistants (see also Ash, Galletta and Opocher (2025)). Despite concerns over their political slant, these tools can provide balanced information with proper prompting, aligning with recent work emphasizing the effectiveness of prompting on reducing perceived political bias (Westwood, Grinner and Hall, 2025).

Data and Methods

To evaluate the impact of a dynamic, conversational Voting Aid Application (VAA) on political knowledge and attitudes, we conducted three studies. Each study focused on politically unaffiliated participants under 35, those who did not identify as either Democrat or Republican.²

- Study 1 includes 383 participants, aged 18–34, recruited through CloudResearch Connect. It was conducted between November and December 12, 2023.
- Study 2 involved 505 participants, aged 18–34, recruited via Prolific. It ran from January 16 to 24, 2024.
- Study 3 included 2,000 participants, aged 18–34, recruited through YouGov. This study took place from October 18 to November 4, 2024.

We opted for three studies for several reasons: First, Study 1 had a relatively small sample (383 participants), so we added two larger studies, especially Study 3, to improve statistical power. Second, finding ample numbers of politically unaligned participants proved to be challenging in both Study 1 and Study 2; we reached the maximum feasible sample sizes through those platforms. Finally, we partnered with three reputable survey firms, CloudResearch, Prolific, and YouGov, to ensure high data quality across studies.

All three studies were pre-registered prior to data collection (see Supplementary Materials for pre-registration details). In each study, participants first completed a pre-treatment survey assessing their

²Participants were screened for political affiliation using data provided by each of the sample providers.

Figure 1: Example of the chatbot interface. User messages are shown in blue, while bot responses are in gray. After four exchanges, users are alerted to a “new feature” and prompted to ask which party aligns closest with their views.

Where do the Republicans and Democrats stand on immigration?

The Republican Party generally favors stricter immigration controls and policies that prioritize national security, often advocating for stronger border security and enforcement of existing immigration laws. The Democratic Party tends to support more comprehensive immigration reform, including a pathway to citizenship for undocumented immigrants and protection for Dreamers. They also focus on humane treatment of immigrants and refugees. Let me know if you need more specifics on any particular aspect of immigration policy!

Where do the parties stand on environmental policy?

The Democratic party generally emphasizes strong environmental protections and actions to combat climate change, often supporting regulations and international agreements. In contrast, the Republican party typically prioritizes economic growth and energy independence, and may advocate for fewer environmental regulations, emphasizing market-based solutions and the responsible development of natural resources. If you're looking for more specifics, feel free to ask!

What about carbon capture?

The Democratic Party generally supports the development and implementation of carbon capture technologies as part of a broader strategy to address climate change. The Republican Party may also recognize the potential of carbon capture but often emphasizes energy independence and economic impacts, possibly favoring less regulation on the energy sector. If you're weighing the importance of carbon capture in your decision, consider how each party integrates it into their environmental and energy policies.

Which party comes closest to my views?

To help you determine which party aligns with your views, could you tell me some of your top priorities or issues you care about? The Democratic and Republican parties often have different stances on topics such as healthcare, the economy, and social issues.

I care about reducing climate change.

The Democratic Party typically emphasizes strong action on climate change, advocating for policies to reduce greenhouse gas emissions and invest in renewable energy. In contrast, the Republican Party often prioritizes economic growth and energy independence, and while some members support climate action, the party as a whole tends to be more cautious about regulations that could affect businesses. Let me know if you need information on specific climate policies or other issues!

Which party comes closest to my views? Ask

You've unlocked a new feature! Ask me which party is closest to you.

Knowledge Level

political knowledge, issue preferences, party placements, and their self-identified “most important issue.” They were then randomly assigned to one of two experimental conditions.

- Control: A placebo interaction with a “trivia bot” posing as Benjamin Franklin, which asked users to guess its identity.
- Treatment: An interactive VAA bot that nudged participants to discuss differences between Republicans and Democrats on their “core issue.”³

Following the interaction, all participants completed a post-treatment survey about party positions on their issue and other polarized topics (e.g., gun rights, abortion), along with questions measuring party attachments and attitudes. The participants were then thanked for their participation and debriefed.

Measuring participant engagement

A key concern in encouraging chatbot conversations is distinguishing meaningful from perfunctory user dialogue. To assess participant engagement, we classified conversations using a locally hosted large language model (Mistral’s Mistral Small 3.1 24B). Most interactions were constructive and involved genuine information-seeking behaviors, with 70-86% of participants classified as “engaged.” Notably, engagement levels were generally higher among those with party-aligned issues, suggesting that the bot was most relevant to participants whose concerns could be connected to differences in party platforms (see SI Appendix, section on Conversation Details).

Outcome Measures and Estimation Strategy

To assess the effects of the VAA bot, we estimate two sets of models. When estimating issue knowledge, we simply compute the average treatment effect (ATE), with and without covariates. The covariate-adjusted ATE is estimated using a linear model with HC2 robust standard errors:

$$Y_i = \beta_0 + \beta_1 \text{VAA Bot}_i + \mathbf{X}'_i \boldsymbol{\beta} + \epsilon_i, \quad (1)$$

³We also assessed the performance of an “Ideology Bot” that mirrored the nature of the “VAA Bot,” but encouraged participants to consider ideological differences between the parties. Exposure to this arm was crossed with the “VAA Bot” arm in a 2x2 factorial design, such that participants could be assigned to two treatment bots, two placebo bots, or one of each. Effect estimates are comparable across the different experimental conditions. Therefore, for this analysis, we pool these treatment groups into a single experimental condition. Key findings are unaffected by analyzing the treatment arms separately.

where Y_i is a measure of issue knowledge and X is a set of covariates including age, education, ideology, political interest, gender, race (non-white = 1), pre-treatment measures of political knowledge, and a scale capturing party placements for polarized issues such as abortion.

We rely on two measures of issue knowledge: core issue accuracy and general issue accuracy. Core issue accuracy captures whether the participant can correctly identify which party is closest to them on their core issue. This is coded as a binary item, where 1 reflects a correct answer. We expect β_1 (i.e., the treatment effect) to be positive in this model, reflecting accuracy increases due to the focused interaction with the bot. To examine if there are second-order effects on a more general set of policies, we also compute an accuracy score for policies such as immigration, abortion, and gun rights. This score is the share of correct party placements. While we anticipate a positive treatment effect (β_1) on general issue accuracy as well, we expect it to be smaller in magnitude relative to the effect on core issue accuracy. The VAA Bot’s focus on the individual’s core issue may lead to less pronounced learning about other policy areas. However, it is plausible that the enhanced understanding of party positions gained through the interaction could generalize to some extent, resulting in a modest improvement in accuracy on the broader set of issues.

To examine how exposure to the VAA Bot affects evaluations of politically proximate parties, we estimate Conditional Average Treatment Effects (CATEs) within subgroups defined by the proximity of their core issue to either the Democratic or Republican Party. We deviate from our pre-analysis plan, which specified estimating Average Treatment Effects (ATEs), because on reflection we realized that the impact of proximity information will vary depending on individuals’ pre-existing issue alignments. If an individual’s “core issue” aligns more closely with the Democratic (Republican) party, we expect the treatment to promote Democratic (Republican) Party identification. Given these directional expectations, we subset on participants who report an issue that is either aligned with the Democratic or Republican Party when carrying out our analyses.

Table 1: Distribution of Human-Coded Issue Stance Categories, by Study

Issue Alignment	CloudResearch	Prolific	YouGov
Democratic-aligned Issue	69.2%	64.5%	42.4%
Republican-aligned Issue	11.5%	19.9%	13.9%
Both parties	9.2%	7.3%	15.7%
Neither Party	8.0%	7.1%	13.7%
No Issue Provided	2.0%	1.2%	14.4%

We hand-classify respondents’ positions on their “core issue” as being more proximate to either

the Democratic or Republican Party. Across our three samples, between 42.4% and 69.2% reported a Democratic-aligned issue (e.g., gun restrictions, abortion rights), while 11.5% to 19.9% cited a Republican-aligned issue (e.g., reducing immigration). The remaining participants expressed positions that were either supported by both parties (e.g., growing the economy; 7.3% to 15.7%), reflected issues without clear party stances (e.g., artificial intelligence; 8.0% to 13.7%), or did not express an issue stance at all (2.0% to 14.4%). In our analyses, the latter three are collapsed into a “no clear party” subgroup, whose results are presented in the Supplementary Materials.

We estimate CATEs for each subgroup, presenting regression estimates with and without covariates. We rely on two measures of downstream political outcomes: partisan affect and self-reported vote choice.⁴ These outcomes are moderately correlated, with Pearson correlation coefficients of 0.45, 0.47, and 0.35 across the three samples, respectively.

For party affect, we construct an additive index comprised of 101-point feeling thermometers toward popular Republican and Democratic figures (Donald Trump, Joe Biden, Barack Obama, Bernie Sanders) and general affect toward each party (Cronbach’s $\alpha = .67$).⁵ The additive scale is rescaled to range from a minimum of zero to a maximum of 1 ($\bar{x} = .54$; $s = .23$), with the lowest score indicating the most pro-Republican sentiment and the maximum score indicating the most pro-Democratic sentiment.

For vote choice, we use a binary item measuring intention to vote for the Democratic Party in the 2024 general election (1 = Democratic Party; 0 = Republican Party).

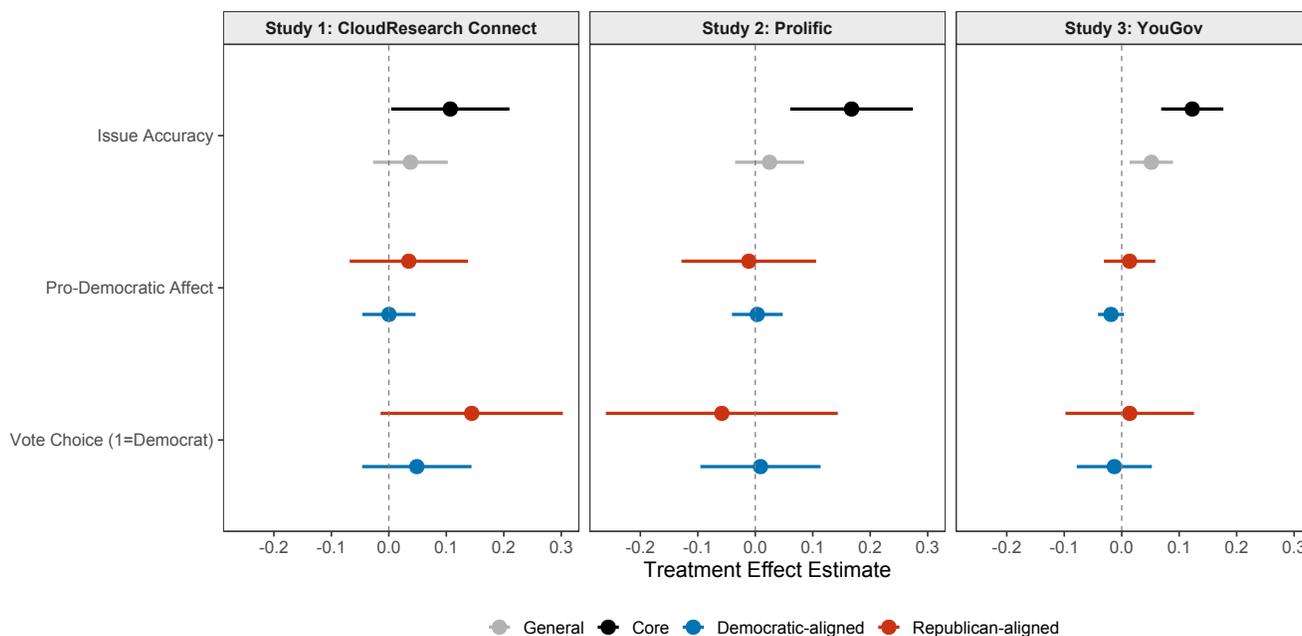
In the case of downstream outcomes such as party affect and vote choice, we expect β_1 to be positive for those reporting a Democrat-aligned issue, reflecting more positive Democratic sentiment and an increased desire to vote for Democrats. In contrast, we anticipate a negative β_1 for those with a Republican-aligned core issue (e.g., gun rights), indicating more positive Republican sentiment and a higher likelihood of voting for Republican candidates. We do not have clear expectations about β_1 among those who disclose a valence issue or one where party positions are undifferentiated, and their results are relegated to the Supplementary Materials.

⁴We also examine effects on the traditional 7-point party identification scale in the CloudResearch and Prolific samples, but find no significant effect.

⁵The latter is measured using seven-point favorability scales, ranging from “least favorable” to “most favorable.”

Results

Figure 2: Main Results



1.1 Study 1: CloudResearch Connect

We assess whether the VAA Bot improves participants’ understanding of party positions on their self-identified core issue. As shown in the left facet of Figure 2, being assigned to the VAA Bot condition significantly increases the accuracy of party placements on these core issues (see Tables 1, 2, and 3 in Supplementary Materials for detailed estimates). Controlling for covariates, the estimated average treatment effect (ATE) is a 10.7 percentage point (pp) increase in accuracy (SE = 5.2pp). To place these estimates in context, we administered a three-item political knowledge battery in which participants were asked to identify the offices held by Rishi Sunak, Kamala Harris, and Mike Johnson. The gap in accuracy between participants with the lowest and highest scores on this scale was 8.6 percentage points.

In contrast, the intervention has only small and statistically imprecise effects on participants’ understanding of other politically salient issues, such as gun rights, immigration, and abortion (labeled as “General” in the figure). These effects are minimal (a 3.8pp increase) with a standard error of similar size (SE = 3.3pp). This suggests the VAA is effective at improving understanding of parties’ stances on participants’ “most important issue” but has limited effects on knowledge about other policy areas.

We also examine how the VAA affects partisan sentiment. Figure 2 (middle row) presents our subgroup estimates. For participants whose core issue aligns with the Democratic Party, the effect on partisan

sentiment is minimal ($\beta_1 = .02\text{pp}$; $\text{SE} = 2.3\text{pp}$). Among those aligned with a Republican-leaning issue, the shift is modest ($\beta_1 = 3.5\text{pp}$; $\text{SE} = 5\text{pp}$). Overall, the VAA has limited impact on affective evaluations of the parties. We also consider effects on intended vote choice (coded as 1 = Democrat, 0 = Republican). The VAA leads to a modest shift in favor of Democrats among participants aligned with Democratic issues: 4.9pp in the adjusted model ($\text{SE} = 4.8\text{pp}$). Among Republican-aligned participants, the shift toward Democrats is larger but imprecisely estimated, 14.4pp ($\text{SE} = 8.0\text{pp}$) in the adjusted model.

Overall, the left facet of Figure 2 highlights that the VAA Bot effectively improves participants' understanding of party positions on the issue they care about most. However, its influence on broader issue knowledge, partisan sentiment, and vote intention is generally limited and statistically imprecise, particularly when analyzed by subgroup.

1.2 Study 2: Prolific

We conducted a second study to obtain a more ideologically-balanced sample. As our CloudResearch Connect sample skewed liberal, we relied on the sample provider Prolific to recruit young moderates and conservatives. These efforts were largely successful, as we were able to nearly double the proportion of self-identified conservatives in our sample from 9.9% in the CloudResearch Connect study to 19.0% in the Prolific study.

As shown in the middle facet of Figure 2, we replicate our key findings from Study 1. Focusing on the first row in the middle facet, we find that the effect of a VAA Bot on “core issue” accuracy is 16.76pp ($\text{SE} = 5.4\text{pp}$), but weaker for general issue accuracy ($\beta_1 = 2.5\text{pp}$; $\text{SE} = 3.1\text{pp}$). Next, we recover limited effects on partisan affect. Among participants with Democratic-aligned issues, the effect of a VAA Bot on pro-Democratic sentiment is .36pp ($\text{SE} = 2.2\text{pp}$), whereas among those with Republican-aligned issues, the effect is -1.1pp ($\text{SE} = 5.9\text{pp}$). Turning to vote choice, we find directionally consistent but imprecisely estimated effects, with an effect of 0.9pp ($\text{SE} = 5.3\text{pp}$) among those with a Democratic-aligned issue and -5.8pp ($\text{SE} = 10.2\text{pp}$) for those with a Republican-aligned issue.

1.3 Study 3: YouGov

Building on our promising findings regarding issue accuracy, the imprecise estimates for partisan leanings, and the challenge of reaching additional young independents on CloudResearch Connect and Prolific, we conducted a third study using the sample provider YouGov. YouGov is widely regarded as a leading high-quality survey vendor (Rivers, 2016) and, for this study, was able to provide a large sample of

young independents ($N = 2,000$). In contrast to Prolific and CloudResearch, participants in our YouGov sample were less likely to select *any* ideological category, with 26.8% choosing that they had not “thought much about” their ideology (compared to 5.2% and 5.5% in the CloudResearch and Prolific samples, respectively).

Turning to the first row of the Figure 2 and the right-most facet, we once again observe an effect of the VAA Bot on “core issue accuracy,” with a 12.3pp increase ($SE = 2.8pp$) in accuracy over the placebo. We also detect a small and statistically significant increase in “general issue accuracy” (5.2pp; $SE = 1.9pp$). Next, turning to the middle row of the right-most facet, we once again observe minimal effects on pro-Democratic sentiment, with Democratic-aligned participants shifting by -1.8pp ($SE = 1.1pp$) and Republican-aligned participants moving by 1.3pp ($SE = 2.3pp$) when exposed to the VAA Bot. Similarly, effects on vote choice are also small in magnitude. Those with Democratic-aligned issues are 1.3pp ($SE = 3.3pp$) less likely to report intending to vote for Democrats, whereas those with Republican-aligned issues are 1.4pp ($SE = 5.7pp$) more likely to do so.

1.4 Overall Effects

We use random-effects meta-analysis with restricted maximum likelihood (REML) to summarize effects across the three studies. On average, the VAA bot improves participants’ accuracy in locating the parties on issues of personal importance by 12.7pp (95% CI: 8.4 to 17.1), with no statistically discernible evidence of between-study heterogeneity ($I^2 = .02\%$; $Q(2) = 0.73$, $p = 0.69$). General issue accuracy also improves, albeit by a smaller amount, with a meta-analytic estimate of 4.3pp (95% CI: 1.4 to 7.1). In contrast, the pooled effect on pro-Democratic affect among Democratic-aligned participants is -1.2pp (95% CI: -3.0 to 0.7; $I^2 = 0\%$; $Q(2) = 1.08$, $p = 0.58$), and among Republican-aligned participants is 1.4pp (95% CI: -2.4 to 5.3; $I^2 = 0\%$; $Q(2) = 0.35$, $p = 0.84$), with both effects statistically indistinguishable from zero. Similarly, the effect on Democratic vote choice among Democratic-aligned participants is 0.7pp (95% CI: -4.0 to 5.5; $I^2 = 0\%$; $Q(2) = 1.10$, $p = 0.58$). Among Republican-aligned participants, the estimated effect is 4.0 percentage points (95% CI: -6.0 to 14.0; $I^2 = 27.4\%$; $Q(2) = 2.93$, $p = 0.23$). In sum, our findings suggest a robust effect of VAA Bots on improving knowledge of where the parties stand on personally relevant issues, with smaller effects on general issue accuracy and negligible effects on partisan sentiment and vote choice.

Discussion

The findings of this paper speak to a wide array of practical and theoretical concerns. The rapid development of LLMs allows for scalable, personalized, and interactive instructional aids, while at the same time raising the specter of algorithmic bias that distorts the information that is communicated. Here, we show the feasibility of creating an even-handed LLM voting aid application that is trained directly on the parties' own public-facing documents. Although this tailored retrieval system does not rule out the possibility of bias, our extensive pretesting of the VAA Bot suggests that this bespoke LLM accurately characterizes the parties' public stances on the myriad of topics. When put to the task of guiding large samples of respondents who were asked about issues important to them, the VAA Bot acquitted itself well, providing cogent and succinct characterizations of party stances on topics that divide the parties. When respondents named core concerns that were vague or ideologically indistinct, the VAA Bot was appropriately cautious about describing party differences. Moreover, the VAA Bot was sufficiently interactive and engaging that roughly three-quarters of our respondents asked it one or more follow-up questions. Levels of engagement with LLMs such as ours may improve in the future as user interfaces allow for audio-visual animation.

LLMs offer a flexible and extensible form of voting aid application, but what are the effects of such VAAs on those who interact with them? To what extent do VAAs make users more knowledgeable, in this case about party platforms? In particular, to what extent do they transmit political information to users such as young adults who lack party attachments? A dearth of randomized experiments has left such questions unresolved.

Our study provides a fresh perspective on VAAs' potential and limitations. We find clear evidence that a randomly assigned conversation with a VAA Bot does increase subjects' understanding of where the parties stand on the issue that the respondent regards as most important to them. The estimated average treatment effect of 12.7 percentage points is larger than the apparent knowledge gap in our sample between those with a bachelors degree or higher and those lacking a high school diploma (11.5 percentage points).

At the same time, our experiments reveal the limited effects that interaction with the VAA Bot had on "general" knowledge of party stances on topics outside the user's core concern. Public opinion scholars since Converse (Converse, 2006) have noted that only a small portion of the public have a clear sense of the ideological underpinnings of the parties' platforms. Lacking a broad understanding of how political

ideas fit together, many subjects who learned where the parties stand on an issue of personal importance evidently had difficulty inferring the parties' positions on other issues.

What about the downstream effects of learning the parties' stances on core issues? Theories of issue proximity-driven party evaluations imply that a substantial exogenous change in subjects' perceived closeness to one of the parties on a topic of personal importance would be consequential. Subjects who learn that the Republican (Democratic) Party is closer to them on this core issue should express greater support for the Republican (Democratic) Party and its leaders. Yet, in all three of our experiments, subjects in the assigned treatment group do not change their ratings of the parties or partisan figures in the predicted direction. Average treatment effects instead hover near zero, and the same is true for vote preferences. Although intuition might suggest that voters would hew toward the party that is revealed to be closer to them on their key issue, the average effect appears to be negligible. To our knowledge, this experiment offers one of the clearest experimental assessments of the issue proximity hypothesis insofar as it does not involve descriptions of hypothetical candidates or persuasive attempts to call attention to issues to paint one of the parties in a favorable light. Here, the intervention is simply information about the parties' stances on an issue of importance to the subject.

That said, we acknowledge that we have focused on one particular segment of the electorate, young adults who have yet to form a party attachment. Given their weak priors, we expected this group to be especially responsive to information about party positions. However, it may be that subjects who are recruited to a study in which they are randomly exposed to information about the parties are less motivated to assimilate this information than the typical user of a voting aid application. A fruitful line for future research is to measure (or encourage) intrinsic motivation at baseline in order to assess whether it interacts with exposure to information about the parties' positions.

References

- Alvarez, R Michael. 1998. *Information and elections*. University of Michigan Press.
- Alvarez, R Michael, Ines Levin, Alexander H Trechsel and Kristjan Vassil. 2014. “Voting advice applications: How useful and for whom?” *Journal of Information Technology & Politics* 11(1):82–101.
- Angelucci, Davide, Luca Carrieri and Marco Improta. 2024. “‘No Participation Without Representation’: The Impact of Descriptive and Substantive Representation on the Age-Related Turnout Gap.” *Political Studies* p. 00323217241229316.
- Ash, Elliott, Sergio Galletta and Giacomo Opocher. 2025. “BallotBot: Can AI Strengthen Democracy?” *Available at SSRN* .
- Bai, Hui, Jan G Voelkel, Shane Muldowney, johannes C Eichstaedt and Robb Willer. 2025. “AI-Generated Messages Can Be Used to Persuade Humans on Policy Issues.”
URL: osf.io/stakv6
- Blais, André and Daniel Rubenson. 2013. “The source of turnout decline: New values or new contexts?” *Comparative Political Studies* 46(1):95–117.
- Brader, Ted, Joshua A Tucker and Dominik Duell. 2013. “Which parties can lead opinion? Experimental evidence on partisan cue taking in multiparty democracies.” *Comparative Political Studies* 46(11):1485–1517.
- Campbell, Angus, Philip E. Converse, Warren E. Miller and Donald E. Stokes. 1960. *The American Voter*. University of Chicago Press.
- Chen, Zhongren, Joshua Kalla, Quan Le, Shinpei Nakamura-Sakai, Jasjeet Sekhon and Ruixiao Wang. 2025. “A Framework to Assess the Persuasion Risks Large Language Model Chatbots Pose to Democratic Societies.”
URL: <https://arxiv.org/abs/2505.00036>
- Converse, Philip E. 2006. “The nature of belief systems in mass publics (1964).” *Critical review* 18(1-3):1–74.
- Costello, Thomas H, Gordon Pennycook and David G Rand. 2024. “Durably reducing conspiracy beliefs through dialogues with AI.” *Science* 385(6714):eadq1814.

- Crabtree, Charles, John Holbein, Mitchell Bosley and Semra Sevi. 2025. “Can AI Help Reduce Prejudice? Evaluating the Effectiveness of AI-Powered Personalized Persuasion on Support for Transgender Rights.” *Evaluating the Effectiveness of AI-Powered Personalized Persuasion on Support for Transgender Rights* (April 24, 2025) .
- CSES. 2016. “The Comparative Study of Electoral Systems; CSES MODULE 5.” www.cses.org.
- Dalton, Russell J and Martin P Wattenberg. 2002. *Parties without partisans: Political change in advanced industrial democracies*. Oxford University Press, USA.
- Dieing, Thilo I. 2024. AI-Driven Dialogue: Leveraging Generative AI in Conversational Agent Voting Advice Applications (CAVAAs). In *International Symposium on Chatbots and Human-Centered AI*. Springer pp. 161–180.
- Downs, Anthony. 1957. “An economic theory of political action in a democracy.” *Journal of political economy* 65(2):135–150.
- Franklin, Charles H and John E Jackson. 1983. “The dynamics of party identification.” *American political science review* 77(4):957–973.
- Furnham, Adrian and Barrie Gunter. 1987. “Young people’s political knowledge.” *Educational studies* 13(1):91–104.
- Gemenis, Kostas. 2024. “Artificial intelligence and voting advice applications.” *Frontiers in Political Science* 6:1286893.
- Gittmann, Yannick, Ing Martin Schiele, Stefan Ilchmann, Ante Gojsalić, Dominik Jurinčić and Phylis Klempt. 2024. “Voting Advice Applications: Implementation of RAG-supported LLMs.” *Authorea Preprints* .
- Goren, Paul and Christopher Chapp. 2017. “Moral power: How public opinion on culture war issues shapes partisan predispositions and religious orientations.” *American Political Science Review* 111(1):110–128.
- Green, DP, B Palmquist and E Schickler. 2002. “Partisan Hearts and Minds New Haven, CT: Yale Univ.”.
- Grumbach, Jacob M and Charlotte Hill. 2022. “Rock the registration: Same day registration increases turnout of young voters.” *The Journal of Politics* 84(1):405–417.

- Hatemi, Peter K and Christopher Ojeda. 2021. "The role of child perception and motivation in political socialization." *British Journal of Political Science* 51(3):1097–1118.
- Huddy, Leonie, Lilliana Mason and Lene Aarøe. 2015. "Expressive partisanship: Campaign involvement, political emotion, and partisan identity." *American Political Science Review* 109(1):1–17.
- Hyman, Herbert. 1959. "Political socialization."
- Jackson, John E. 1975. "Issues, party choices, and presidential votes." *American Journal of Political Science* pp. 161–185.
- Jennings, M Kent. 1996. "Political knowledge over time and across generations." *Public Opinion Quarterly* 60(2):228–252.
- Jennings, M Kent, Laura Stoker and Jake Bowers. 2009. "Politics across generations: Family transmission reexamined." *The Journal of Politics* 71(3):782–799.
- Jennings, M Kent and Richard G Niemi. 1968. "The transmission of political values from parent to child." *American political science review* 62(1):169–184.
- Jennings, M Kent and Richard G Niemi. 1978. "The persistence of political orientations: An over-time analysis of two generations." *British Journal of Political Science* 8(3):333–363.
- Kamoen, Naomi, Tessa McCartan and Christine Liebrecht. 2021. Conversational agent voting advice applications: a comparison between a structured, semi-structured, and non-structured chatbot design for communicating with voters about political issues. In *International Workshop on Chatbot Research and Design*. Springer pp. 160–175.
- Kleinnijenhuis, Jan, Otto Scholten, Wouter H van Atteveldt, Anita MJ van Hoof, APM Krouwel, Dirk Oegema, Jan A de Ridder, PC Ruigrok and Janet Takens. 2007. "Nederland vijfstromenland: De rol van de media en stemwijzers bij de verkiezingen van 2006."
- Kollman, Ken and John E Jackson. 2021. *Dynamic partisanship: How and why voter loyalties change*. University of Chicago Press.
- Lau, Richard R and David P Redlawsk. 2008. "Older but wiser? Effects of age on political cognition." *The Journal of Politics* 70(1):168–185.

- Lupia, Arthur. 2016. *Uninformed: Why people know so little about politics and what we can do about it*. Oxford University Press.
- Manavopoulos, Vasilis, Vasiliki Triga, Stefan Marschall and Lucas Constantin Wurthmann. 2018. “The impact of VAAs on (non-voting) aspects of political participation: Insights from panel data collected during the 2017 German federal elections campaign.” *Statistics, Politics and Policy* 9(2):105–134.
- Munzert, Simon and Sebastian Ramirez-Ruiz. 2021. “Meta-analysis of the effects of voting advice applications.” *Political Communication* 38(6):691–706.
- Niemi, Richard G and M Kent Jennings. 1991. “Issues and inheritance in the formation of party identification.” *American Journal of Political Science* pp. 970–988.
- Rivers, Doug. 2016. “Pew Research: YouGov consistently outperforms competitors on accuracy.” *YouGov*, May 13.
- Rosenqvist, Olof. 2020. “Rising to the occasion? Youth political knowledge and the voting age.” *British Journal of Political Science* 50(2):781–792.
- Sevi, Semra. 2021. “Do young voters vote for young leaders?” *Electoral Studies* 69:102200.
- Smets, Kaat. 2012. “A widening generational divide? The age gap in voter turnout through time and space.” *Journal of Elections, Public Opinion & Parties* 22(4):407–430.
- Soule, Suzanne. 2001. “Will They Engage? Political Knowledge, Participation and Attitudes of Generations X and Y.”.
- Stiers, Dieter and Marc Hooghe. 2021. “A picture is worth a thousand words. A visual test of political knowledge: Does it reduce the effect of age and gender?” *International Journal of Public Opinion Research* 33(1):137–146.
- Stockemer, Daniel and Francois Rocher. 2017. “Age, political knowledge and electoral turnout: a case study of Canada.” *Commonwealth & Comparative Politics* 55(1):41–62.
- Stokes, Donald E. 1963. “Spatial models of party competition.” *American political science review* 57(2):368–377.
- Tilley, James and Geoffrey Evans. 2014. “Ageing and generational effects on vote choice: Combining cross-sectional and panel data to estimate APC effects.” *Electoral studies* 33:19–27.

Vassos, Stavros, Stratos Goudelis, Dimi Balaouras, Giannis Vitalis, Vasilis Nakos, Glykeria Pigka, Loukia Tsagkli, Menia Hatzikou, Zachos Tsionas, Alexandros Chasanis et al. 2024. Now I know! Empowering Voters with RAG-enabled LLMs to Eliminate Political Uncertainty. In *Proceedings of the 13th Hellenic Conference on Artificial Intelligence*. pp. 1–7.

Velez, Yamil Ricardo and Patrick Liu. 2024. “Confronting core issues: A critical assessment of attitude polarization using tailored experiments.” *American Political Science Review* pp. 1–18.

Westwood, Sean J., Justin Grinner and Andrew B. Hall. 2025. Measuring Perceived Slant in Large Language Models Through User Evaluations. Working Paper 4262 Stanford Graduate School of Business.

Supplementary Materials

“Chatbot-driven Voting Aid Applications Increase Knowledge about Party Positions but Do Not Change Party Evaluations”

Pre-registration



What are the attitudinal effects of chatting with a bot about politics? (#150713)

Author(s)

This pre-registration is currently anonymous to enable blind peer-review.
It has 3 authors.

Pre-registered on: 2023/11/12 - 08:56 AM (PT)

1) Have any data been collected for this study already?

No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?

What is the effect of a conversation with an artificial intelligence expert on respondents' knowledge of where parties stand on important current issues and their own reported party identification?

3) Describe the key dependent variable(s) specifying how they will be measured.

-- Views on several policy positions (measured pre and post conversation). Our key outcome measure is whether participants can correctly identify which party is aligned with them on a deeply-held issue (binary). Additionally, we will measure whether participants can accurately match the parties to a variety of issue positions, including supporting gay marriage, supporting gun rights, and opposing abortion (share of correct answers).

-- Respondents' party identification (measured pre and post conversation with the artificial intelligence expert). Five measures will be used to gauge shifts in partisanship: a favorability measure asking participants to report their favorability toward the two parties; a measure of how closely the participant identifies with the party; a set of seven items capturing policy preferences; a prospective vote choice measure about the 2024 election; and a set of feeling thermometers capturing attitudes toward prominent members of the Democratic and Republican Party.

4) How many and which conditions will participants be assigned to?

Two conditions: talk to an artificial intelligence expert about where the parties stand on important current issues (treatment) or play a trivia game with an artificial intelligence expert (control). Both conditions will talk to an artificial intelligence expert.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

We will estimate linear models with HC2 robust standard errors. Our models will include a treatment indicator and set of pre-treatment covariates (age, gender, ideology, education, race, political interest, political knowledge, and pre-treatment scores on a party-policy matching task). We expect a positive average treatment effect (ATE), such that assignment to the political AI will (1) increase knowledge of where the parties stand, especially on an issue that is important to the respondent and (2) increase the likelihood of identifying with one of the two parties as well as the strength of identification.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

We will exclude any cases where participants do not click our link on Qualtrics (and therefore are never assigned a condition).

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.

Our goal is to recruit 400 individuals who have a conversation with an artificial intelligence expert.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

We will examine whether the treatment effect varies substantially with age; we expect stronger effects among younger respondents and, more generally, those with little information about or interest in politics.

We will measure the number of exchanges with the artificial intelligence expert, as well as the time taken to complete the task as intermediate outcomes akin to manipulation checks.

Version of AS Predicted Questions: 2.00

Available at <https://aspredicted.org/qw67-23cr.pdf>

Figure 1: Pre-registration for Study 1

What are the attitudinal effects of chatting with a bot about politics? Study 2 (#158149)

Author(s)

This pre-registration is currently anonymous to enable blind peer-review.
It has 3 authors.

Pre-registered on: 01/16/2024 06:40 AM (PT)

1) Have any data been collected for this study already?

No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?

What is the effect of a conversation with an artificial intelligence expert on respondents' knowledge of where parties stand on important current issues and their own reported party identification?

3) Describe the key dependent variable(s) specifying how they will be measured.

-- Views on several policy positions (measured pre and post conversation). Our key outcome measure is whether participants can correctly identify which party is aligned with them on a deeply-held issue (binary). Additionally, we will measure whether participants can accurately match the parties to a variety of issue positions, including supporting gay marriage, supporting gun rights, and opposing abortion (share of correct answers).

-- Respondents' party identification (measured pre and post conversation with the artificial intelligence expert). Five measures will be used to gauge shifts in partisanship: a favorability measure asking participants to report their favorability toward the two parties; a measure of how closely the participant identifies with the party; a set of seven items capturing policy preferences; a prospective vote choice measure about the 2024 election; and a set of feeling thermometers capturing attitudes toward prominent members of the Democratic and Republican Party.

4) How many and which conditions will participants be assigned to?

Factorial design: First, participants will be randomized to either (1) talk to an artificial intelligence (AI) expert about the meaning of liberalism and conservatism (ideology bot condition) or (2) play a trivia game. Then, participants will be randomized to either (1) talk to an artificial intelligence expert about where the parties stand on important current issues (VAA bot condition) or (2) play another trivia game. All subjects will interact with a chatbot in some way, whether in the treatment or placebo condition.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

We will estimate linear models with HC2 robust standard errors. We will adjust for a set of pre-treatment covariates (age, gender, ideological self-identification, scores on a policy attitudes scale, education, race, political interest, political knowledge, and scores on a party-policy matching task). We will separately estimate main effects for the ideology bot and VAA bot conditions. We will also estimate an interaction model with treatment-by-treatment interactions, adjusting for the covariates above.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

We will exclude any cases where participants do not click our link on Qualtrics (and therefore are never assigned a condition).

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.

Our goal is to recruit 500 individuals.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

We plan to assess CATEs across levels of political interest (median split), political knowledge (median split), scores on the party-policy matching task (median split), ideological self-identification (liberal, moderate, conservative), and scores on the policy attitudes scale (median split).

Figure 2: Pre-registration for Study 2

What are the attitudinal effects of chatting with a bot about politics? (#194783)

Author(s)

This pre-registration is currently anonymous to enable blind peer-review.
It has 2 authors.

Pre-registered on: 10/18/2024 01:35 PM (PT)

1) Have any data been collected for this study already?

No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?

What is the effect of a conversation with an artificial intelligence expert on respondents' knowledge of where parties stand on important current issues and their own reported party identification?

3) Describe the key dependent variable(s) specifying how they will be measured.

-- Views on several policy positions (measured pre and post conversation). Our key outcome measure is whether participants can correctly identify which party is aligned with them on a deeply-held issue (binary). Additionally, we will measure whether participants can accurately match the parties to a variety of issue positions, including supporting gay marriage, supporting gun rights, and opposing abortion (share of correct answers).

-- Respondents' party identification (measured pre and post conversation with the artificial intelligence expert). Five measures will be used to gauge shifts in partisanship: a favorability measure asking participants to report their favorability toward the two parties; a measure of how closely the participant identifies with the party; a set of seven items capturing policy preferences; a prospective vote choice measure about the 2024 election; and a set of feeling thermometers capturing attitudes toward prominent members of the Democratic and Republican Party.

4) How many and which conditions will participants be assigned to?

Two-arm design: Participants will be randomized into one of two groups: (1) interact with an artificial intelligence expert (VAA bot) to learn about where political parties stand on important issues; (2) play a trivia game against a chatbot.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

When estimating effects on "issue accuracy," or the capacity for participants to correctly identify where the parties stand on core issues and more distal issues, we will estimate linear models with HC2 robust standard errors. We will adjust for a set of pre-treatment covariates (age, gender, ideological self-identification, scores on a policy attitudes scale, education, race, political interest, political knowledge, and scores on a party-policy matching task).

When estimating effects on party evaluations and identification, we will divide participants into three groups based on each party's proximity to the voter on their core issue: those with Republican-aligned core issues, those with Democrat-aligned core issues, and those who report a valence issue or one where the parties do not diverge (Neither). We will estimate the same models as above within each of these subgroups. Additionally, we will estimate an alternative specification where party-relevant outcome variables are recoded so that higher values indicate greater support for the party more closely aligned with the participant's core issue. In these analyses, we will subset only on participants in the Republican-aligned and Democrat-aligned subgroups, since these are the only groups where a match between the party and issue stance can be established.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

We will exclude any cases where participants do not click our link on YouGov (and therefore are never assigned a condition).

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.

Our goal is to recruit 2,000 unaligned individuals.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

We plan to assess CATEs across levels of political interest (median split), political knowledge (median split), scores on the party-policy matching task (median split), ideological self-identification (liberal, moderate, conservative), and scores on the policy attitudes scale (median split).

Figure 3: Pre-registration for Study 3

Main Model Estimates

Study 1: Cloud Research

Table 1: Issue Accuracy Analysis (Study 1)

	Core Issue Accuracy		General Issue Accuracy	
	No Covariates	With Covariates	No Covariates	With Covariates
Treatment	0.110* (0.051)	0.107* (0.052)	0.029 (0.033)	0.038 (0.033)
Num.Obs.	311	311	311	311
R2	0.015	0.066	0.002	0.100
Control Mean	0.656	0.656	0.818	0.818

Note: This table presents the results of the Issue Match Analysis in Study 1. The models estimate the effect of treatment on both core and other issue matches, with and without covariate adjustment.
Significance levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: Conditional Average Treatment Effects (CATEs) on Pro-Democratic Party Affect by Core Issue Subgroup (Study 1)

	Democrat-aligned		Republican-aligned	
	No Covariates	With Covariates	No Covariates	With Covariates
Treatment	0.008 (0.024)	0.000 (0.024)	0.047 (0.054)	0.035 (0.051)
Num.Obs.	265	265	46	46
R2	0.000	0.125	0.017	0.353
R2 Adj.	-0.003	0.080	-0.006	0.143
Control Mean	0.608	0.608	0.242	0.242

Note: This table reports the Conditional Average Treatment Effects (CATEs) for different subgroups (Democrat-aligned, Republican-aligned, and No clear party) on an additive index measuring party affect. This scale ranges from pro-Republican sentiment to pro-Democratic sentiment. Models are estimated both with and without covariates with HC2 clustered standard errors.
Significance levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Conditional Average Treatment Effects (CATEs) on Vote Choice by Core Issue Subgroup (Study 1)

	Democrat-aligned		Republican-aligned	
	No Covariates	With Covariates	No Covariates	With Covariates
Treatment	0.062 (0.051)	0.049 (0.048)	0.190* (0.080)	0.144+ (0.080)
Num.Obs.	265	265	46	46
R2	0.006	0.170	0.113	0.399
R2 Adj.	0.002	0.127	0.093	0.204
Control Mean	0.750	0.750	0.000	0.000

Note: This table reports the Conditional Average Treatment Effects (CATEs) for different subgroups (Democrat-aligned, Republican-aligned, and No clear party) on vote choice (1 = Democratic Party; 0 = Republican Party). Models are estimated both with and without covariates.
Significance levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Study 2: Prolific

Table 4: Issue Match Analysis (Study 2)

	Core Issue Accuracy		General Issue Accuracy	
	No Covariates	With Covariates	No Covariates	With Covariates
Treatment	0.169** (0.055)	0.168** (0.054)	0.021 (0.034)	0.025 (0.031)
Num.Obs.	426	426	426	426
R2	0.024	0.070	0.001	0.164
Control Mean	0.518	0.518	0.832	0.832

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

Table 5: Conditional Average Treatment Effects (CATEs) on Pro-Democratic Party Affect by Core Issue Subgroup (Study 2)

	Democrat-aligned		Republican-aligned	
	No Covariates	With Covariates	No Covariates	With Covariates
Treatment	0.009 (0.023)	0.004 (0.022)	-0.023 (0.068)	-0.011 (0.059)
Num.Obs.	326	326	100	100
R2	0.000	0.130	0.002	0.264
R2 Adj.	-0.003	0.096	-0.008	0.162
Control Mean	0.574	0.574	0.364	0.364

Note: This table reports the Conditional Average Treatment Effects (CATEs) for different subgroups (Democrat-aligned, Republican-aligned, and No clear party) on an additive index measuring party affect. This scale ranges from pro-Republican sentiment to pro-Democratic sentiment. Models are estimated both with and without covariates.

Significance levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Conditional Average Treatment Effects (CATEs) on Vote Choice by Core Issue Subgroup (Study 2)

	Democrat-aligned		Republican-aligned	
	No Covariates	With Covariates	No Covariates	With Covariates
Treatment	0.013 (0.057)	0.009 (0.053)	-0.086 (0.103)	-0.058 (0.102)
Num.Obs.	324	324	99	99
R2	0.000	0.197	0.009	0.227
R2 Adj.	-0.003	0.166	-0.001	0.119
Control Mean	0.703	0.703	0.211	0.211

Note: This table reports the Conditional Average Treatment Effects (CATEs) for different subgroups (Democrat-aligned, Republican-aligned, and No clear party) on vote choice (1 = Democratic Party; 0 = Republican Party). Models are estimated both with and without covariates.

Significance levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Study 3: YouGov

Table 7: Issue Accuracy Analysis (Study 3)

	Core Issue Accuracy		General Issue Accuracy	
	No Covariates	With Covariates	No Covariates	With Covariates
Treatment	0.140*** (0.029)	0.123*** (0.028)	0.069** (0.022)	0.051** (0.019)
Num.Obs.	1124	1124	1124	1124
R2	0.020	0.145	0.008	0.294
Control Mean	0.517	0.517	0.540	0.540

Note: This table presents the results of the Issue Match Analysis in Study 3. The models estimate the effect of treatment on both core and other issue matches, with and without covariate adjustment.
Significance levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Conditional Average Treatment Effects (CATEs) on Pro-Democratic Party Affect by Core Issue Subgroup (Study 3)

	Democrat-aligned		Republican-aligned	
	No Covariates	With Covariates	No Covariates	With Covariates
Treatment	-0.016 (0.012)	-0.018 (0.011)	0.012 (0.024)	0.014 (0.023)
Num.Obs.	847	847	277	277
R2	0.002	0.065	0.001	0.105
R2 Adj.	0.001	0.050	-0.003	0.061
Control Mean	0.502	0.502	0.454	0.454

Note: This table reports the Conditional Average Treatment Effects (CATEs) for different subgroups (Democrat-aligned, Republican-aligned, and No clear party) on an additive index measuring party affect. This scale ranges from pro-Republican sentiment to pro-Democratic sentiment. Models are estimated both with and without covariates with HC2 clustered standard errors.
Significance levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Conditional Average Treatment Effects (CATEs) on Vote Choice by Core Issue Subgroup (Study 3)

	Democrat-aligned		Republican-aligned	
	No Covariates	With Covariates	No Covariates	With Covariates
Treatment	-0.007 (0.034)	-0.015 (0.033)	0.017 (0.057)	0.015 (0.056)
Num.Obs.	844	844	274	274
R2	0.000	0.054	0.000	0.085
R2 Adj.	-0.001	0.039	-0.003	0.039
Control Mean	0.616	0.616	0.313	0.313

Note: This table reports the Conditional Average Treatment Effects (CATEs) for different subgroups (Democrat-aligned, Republican-aligned, and No clear party) on vote choice (1 = Democratic Party; 0 = Republican Party). Models are estimated both with and without covariates.
Significance levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Full Model Estimates

Study 1: Cloud Research

Table 10: Issue Accuracy Analysis (Study 1)

	Core Issue Accuracy		General Issue Accuracy	
	No Covariates	With Covariates	No Covariates	With Covariates
Treatment	0.126** (0.048)	0.129** (0.048)	0.016 (0.030)	0.029 (0.029)
Num.Obs.	381	381	381	381
R2	0.018	0.069	0.001	0.116
Control Mean	0.594	0.594	0.826	0.826

Note: This table presents the results of the Issue Match Analysis in Study 1. The models estimate the effect of treatment on both core and other issue matches, with and without covariate adjustment.
Significance levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: Conditional Average Treatment Effects (CATEs) on Pro-Democratic Party Affect by Core Issue Subgroup (Study 1)

	Democrat-aligned		Republican-aligned		No clear party	
	No Covariates	With Covariates	No Covariates	With Covariates	No Covariates	With Covariates
Treatment	0.008 (0.024)	0.000 (0.024)	0.047 (0.054)	0.035 (0.051)	0.083+ (0.046)	0.107* (0.044)
Num.Obs.	265	265	46	46	70	70
R2	0.000	0.125	0.017	0.353	0.046	0.360
R2 Adj.	-0.003	0.080	-0.006	0.143	0.032	0.225
Control Mean	0.608	0.608	0.242	0.242	0.411	0.411

Note: This table reports the Conditional Average Treatment Effects (CATEs) for different subgroups (Democrat-aligned, Republican-aligned, and No clear party) on an additive index measuring party affect. This scale ranges from pro-Republican sentiment to pro-Democratic sentiment. Models are estimated both with and without covariates with HC2 clustered standard errors.
Significance levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12: Conditional Average Treatment Effects (CATEs) on Vote Choice by Core Issue Subgroup (Study 1)

	Democrat-aligned		Republican-aligned		No clear party	
	No Covariates	With Covariates	No Covariates	With Covariates	No Covariates	With Covariates
Treatment	0.062 (0.051)	0.049 (0.048)	0.190* (0.080)	0.144+ (0.080)	0.086 (0.121)	0.170 (0.116)
Num.Obs.	265	265	46	46	70	70
R2	0.006	0.170	0.113	0.399	0.007	0.314
R2 Adj.	0.002	0.127	0.093	0.204	-0.007	0.170
Control Mean	0.750	0.750	0.000	0.000	0.457	0.457

Note: This table reports the Conditional Average Treatment Effects (CATEs) for different subgroups (Democrat-aligned, Republican-aligned, and No clear party) on vote choice (1 = Democratic Party; 0 = Republican Party). Models are estimated both with and without covariates.
Significance levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Study 2: Prolific

Table 13: Issue Match Analysis (Study 2)

	Core Issue Accuracy		General Issue Accuracy	
	No Covariates	With Covariates	No Covariates	With Covariates
Ideology Bot	0.155* (0.060)	0.165** (0.060)	-0.005 (0.039)	-0.005 (0.035)
VAA Bot	0.141* (0.064)	0.133* (0.063)	0.057 (0.039)	0.045 (0.035)
Ideology Bot x VAA Bot	-0.120 (0.087)	-0.110 (0.086)	0.000 (0.053)	0.017 (0.048)
Num.Obs.	505	505	505	505
R2	0.021	0.072	0.009	0.195
Control Mean	0.477	0.477	0.807	0.807

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

Table 14: Conditional Average Treatment Effects (CATEs) on Pro-Democratic Party Affect by Core Issue Subgroup (Study 2)

	Democrat-aligned		Republican-aligned		No clear party	
	No Covariates	With Covariates	No Covariates	With Covariates	No Covariates	With Covariates
Ideology Bot	0.015 (0.030)	0.009 (0.029)	-0.066 (0.073)	-0.027 (0.063)	0.049 (0.066)	0.054 (0.066)
VAA Bot	-0.001 (0.030)	-0.014 (0.029)	0.037 (0.073)	0.025 (0.065)	0.041 (0.064)	0.011 (0.072)
Ideology Bot x VAA Bot	-0.003 (0.043)	0.017 (0.042)	0.000 (0.090)	-0.024 (0.081)	0.039 (0.089)	0.060 (0.095)
Num.Obs.	326	326	100	100	79	79
R2	0.001	0.132	0.035	0.273	0.060	0.275
R2 Adj.	-0.008	0.093	0.005	0.153	0.023	0.117
Control Mean	0.574	0.574	0.364	0.364	0.402	0.402

Note: This table reports the Conditional Average Treatment Effects (CATEs) for different subgroups (Democrat-aligned, Republican-aligned, and No clear party) on an additive index measuring party affect. This scale ranges from pro-Republican sentiment to pro-Democratic sentiment. Models are estimated both with and without covariates.

Significance levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 15: Conditional Average Treatment Effects (CATEs) on Vote Choice by Core Issue Subgroup (Study 2)

	Democrat-aligned		Republican-aligned		No clear party	
	No Covariates	With Covariates	No Covariates	With Covariates	No Covariates	With Covariates
Ideology Bot	-0.001 (0.070)	-0.001 (0.065)	-0.111 (0.111)	-0.046 (0.107)	-0.026 (0.160)	0.059 (0.142)
VAA Bot	0.094 (0.069)	0.068 (0.063)	-0.011 (0.126)	-0.036 (0.125)	-0.187 (0.162)	-0.260 (0.171)
Ideology Bot x VAA Bot	-0.134 (0.100)	-0.099 (0.091)	-0.009 (0.149)	-0.013 (0.152)	0.238 (0.229)	0.204 (0.222)
Num.Obs.	324	324	99	99	78	78
R2	0.011	0.203	0.027	0.231	0.024	0.301
R2 Adj.	0.001	0.166	-0.004	0.103	-0.015	0.145
Control Mean	0.703	0.703	0.211	0.211	0.500	0.500

Note: This table reports the Conditional Average Treatment Effects (CATEs) for different subgroups (Democrat-aligned, Republican-aligned, and No clear party) on vote choice (1 = Democratic Party; 0 = Republican Party). Models are estimated both with and without covariates.

Significance levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Study 3: YouGov

Table 16: Issue Accuracy Analysis (Study 3)

	Core Issue Accuracy		General Issue Accuracy	
	No Covariates	With Covariates	No Covariates	With Covariates
Treatment	0.069** (0.022)	0.062** (0.022)	0.069*** (0.018)	0.051*** (0.014)
Num.Obs.	2000	2000	2000	2000
R2	0.005	0.052	0.008	0.360
Control Mean	0.373	0.373	0.476	0.476

Note: This table presents the results of the Issue Match Analysis in Study 3. The models estimate the effect of treatment on both core and other issue matches, with and without covariate adjustment.
Significance levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 17: Conditional Average Treatment Effects (CATEs) on Pro-Democratic Party Affect by Core Issue Subgroup (Study 3)

	Democrat-aligned		Republican-aligned		No clear party	
	No Covariates	With Covariates	No Covariates	With Covariates	No Covariates	With Covariates
Treatment	-0.016 (0.012)	-0.018 (0.011)	0.012 (0.024)	0.014 (0.023)	0.017 (0.014)	0.023+ (0.013)
Num.Obs.	847	847	277	277	876	876
R2	0.002	0.065	0.001	0.105	0.002	0.053
R2 Adj.	0.001	0.050	-0.003	0.061	0.001	0.038
Control Mean	0.502	0.502	0.454	0.454	0.430	0.430

Note: This table reports the Conditional Average Treatment Effects (CATEs) for different subgroups (Democrat-aligned, Republican-aligned, and No clear party) on an additive index measuring party affect. This scale ranges from pro-Republican sentiment to pro-Democratic sentiment. Models are estimated both with and without covariates with HC2 clustered standard errors.
Significance levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 18: Conditional Average Treatment Effects (CATEs) on Vote Choice by Core Issue Subgroup (Study 3)

	Democrat-aligned		Republican-aligned		No clear party	
	No Covariates	With Covariates	No Covariates	With Covariates	No Covariates	With Covariates
Treatment	-0.007 (0.034)	-0.015 (0.033)	0.017 (0.057)	0.015 (0.056)	0.120*** (0.034)	0.113*** (0.034)
Num.Obs.	844	844	274	274	862	862
R2	0.000	0.054	0.000	0.085	0.014	0.055
R2 Adj.	-0.001	0.039	-0.003	0.039	0.013	0.041
Control Mean	0.616	0.616	0.313	0.313	0.467	0.467

Note: This table reports the Conditional Average Treatment Effects (CATEs) for different subgroups (Democrat-aligned, Republican-aligned, and No clear party) on vote choice (1 = Democratic Party; 0 = Republican Party). Models are estimated both with and without covariates.
Significance levels: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Large Language Model (LLM) Details

This section details the parameters and the prompt structure used for interacting with the LLM. The front-end of the chat was programmed using HTML and JavaScript, whereas PHP was used for the back-end (e.g., saving conversations, interacting with the API). Each interaction with the chatbot was saved on an external SQL server. Each time a participant queried the chatbot, an API call including a system prompt describing the preferred behavior of the bot, seven nearest neighbor text chunks semantically related to the user query, and the chat history was made. The LLM was explicitly instructed to provide neutral information without attempting to persuade the user.

API Configuration and Parameters

The system interacts with the LLM via an OpenAI API endpoint using specific model and generation parameters:

- **LLM Model:** openai/gpt-4-1106-preview (Studies 1-2) openai/chatgpt-4o-latest (Study 3)
- **Temperature:** 0.7
- **Value of K (Context Retrieval):** 7. This refers to the number of nearest neighbor document chunks retrieved from an external vector database storing embeddings and metadata (Pinecone) to be used as context for the LLM.

LLM Prompt Structure

```
[
  {
    "role": "system",
    "content": "Your name is Polly. You are an AI chatbot that helps voters make their \
minds up about the American political parties. The context are party platforms for the \
two major parties. Please be helpful. Be as accurate as possible. Do not include any \
persuasive arguments for the other side. Your goal is not to convince people of one \
side or the other. Your goal is to help people make up their minds. If the discussion \
about an issue gets stale, ask follow up questions guiding the user to the next issue. \
Return the information in an accessible reading level (high school). Always end in \
complete sentences. Use details from conversation history to seem more personable. \
This chat is for people on the go. Be a cool person who is affable and nice. Do not \
seem overly chat-botty. If anyone says anything harmful, please tell them to seek \
professional help. Respond in 2-3 sentences. Pay attention to the query. Reject \
irrelevant context. Always begin by drawing a clear contrast between the two parties \
(e.g., The Republican and Democratic parties [specific difference, avoid mealymouthed \
responses]). Important: If the question is generic and does not mention states, do not \
return context referring a US state. Do not always repeat the same thing. Vary the \
language a bit every time. Pretend you are a neutral expert having a natural \
conversation. If the user asks which party is closest to their views, use the chat \
history to determine which party aligns most closely. Do your best to guess which party \
they would most likely support. Vary the language a bit every time. Pretend you are a \
neutral expert having a natural conversation. Please keep the conversation going. \
Take it step by step. Context: [RETRIEVED_CHUNKS]"
  },
  {
    "role": "user",
    "content": "Conversation History: [USER_CONVERSATION_HISTORY_STRING]\
User Query: [USER_QUERY_STRING]"
  },
]
```

```
{
  "role": "assistant",
  "content": "Assistant:"
}
```

Dynamic Content Placeholders:

- [RETRIEVED_CHUNKS_JSON_STRING]: This placeholder is replaced by a JSON string representation of the $K = 7$ document chunks retrieved from Pinecone. Each chunk includes metadata such as document ID, text snippet, political party, year, relevance score, and state.
- [USER_CONVERSATION_HISTORY_STRING]: This placeholder is populated with the past dialogue between the user and the assistant, formatted with "User:" and "Assistant:" prefixes.
- [USER_QUERY_STRING]: This placeholder contains the current sanitized query input by the user.

Retrieval Augmented Generation (RAG) Architecture

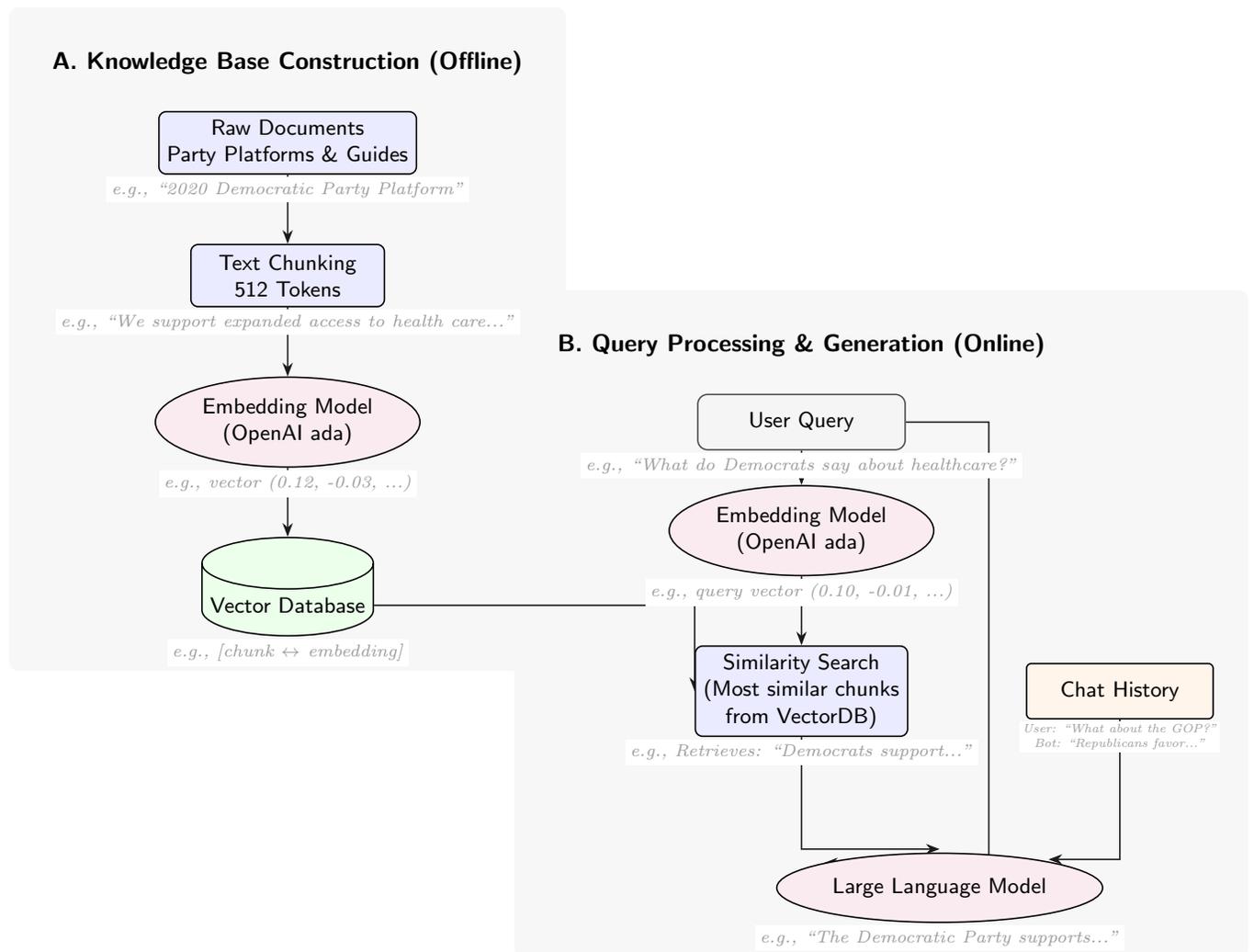


Figure 4: System architecture of the Retrieval Augmented Generation (RAG) model for the political chatbot. The architecture consists of two main phases: **(A) Knowledge Base Construction (Offline)** where raw political documents (party platforms and voter guides) are processed through text chunking (512 tokens), converted to 1536-dimensional vector embeddings using OpenAI’s text-embedding-ada-002 model, and stored in an external vector database (Pinecone); and **(B) Query Processing and Generation (Online)** where user queries are embedded into the same 1536-dimensional semantic space, matched against document chunks through cosine similarity search to retrieve the $K=7$ most relevant chunks, and combined with conversation history to provide contextual information for the Large Language Model’s response generation. White-background boxes at each step show examples, demonstrating how political information is retrieved and presented to users in a conversational format.

Sample Composition

Table 19: Demographic Summary of Survey Respondents

Characteristic	Category	CloudResearch Connect	Prolific	YouGov
Sample Size	Total Respondents (N)	383	505	2000
Age Group	18-24 years	90 (23.5%)	132 (26.1%)	765 (38.2%)
Age Group	25-35 years	289 (75.5%)	373 (73.9%)	1235 (61.8%)
Age Group	Prefer not to say/Other	4 (1.0%)	0 (0.0%)	0 (0.0%)
Gender	Female	168 (43.9%)	197 (39.0%)	969 (48.4%)
Gender	Male	197 (51.4%)	288 (57.0%)	1031 (51.5%)
Gender	Other/Not Listed	16 (4.2%)	20 (4.0%)	0 (0.0%)
Gender	Unknown	2 (0.5%)	0 (0.0%)	0 (0.0%)
Race/Ethnicity	White, non-Hispanic	204 (53.3%)	275 (54.5%)	907 (45.4%)
Race/Ethnicity	Black/African-American	64 (16.7%)	60 (11.9%)	582 (29.1%)
Race/Ethnicity	Hispanic/Latino	45 (11.7%)	69 (13.7%)	258 (12.9%)
Race/Ethnicity	Other	68 (17.8%)	101 (20.0%)	253 (12.7%)
Race/Ethnicity	Unknown	2 (0.5%)	0 (0.0%)	0 (0.0%)
Education	Less than High School	6 (1.6%)	9 (1.8%)	129 (6.5%)
Education	High School/GED	64 (16.7%)	82 (16.2%)	730 (36.5%)
Education	Some College	141 (36.8%)	179 (35.4%)	674 (33.7%)
Education	4-year College Degree	146 (38.1%)	194 (38.4%)	370 (18.5%)
Education	Graduate School	24 (6.3%)	41 (8.1%)	97 (4.9%)
Education	Unknown	2 (0.5%)	0 (0.0%)	0 (0.0%)
Political Ideology	Liberal	125 (32.6%)	145 (28.7%)	285 (14.2%)
Political Ideology	Moderate	198 (51.7%)	236 (46.7%)	982 (49.1%)
Political Ideology	Conservative	38 (9.9%)	96 (19.0%)	197 (9.8%)
Political Ideology	DK/Haven't thought	20 (5.2%)	28 (5.5%)	536 (26.8%)
Political Ideology	Unknown	2 (0.5%)	0 (0.0%)	0 (0.0%)

Issue Categories

We classify the open-ended issue text by first producing summaries of the issue topics using a 4-bit quantized locally hosted large language model (MistralAI’s Mistral Small 3.1 24B) and then applying a keyword-based method using phrases associated with popular issues to consolidate the LLM-generated topics.¹ This produces 78 discrete issue topics, ranging from salient issues such as abortion to less salient issues such as consumer protection and gaming. When asked about their “most important issue,” participants across the three studies mentioned a range of economic and social issues. Despite different percentages of issue mentions, the most salient issues were comparable across the YouGov, Prolific, and CloudResearch Connect samples, with abortion, healthcare, gun rights, the economy, and immigration appearing among the most commonly mentioned issue areas. The prevalence of the “no issue” category was significantly higher in the YouGov sample (10.5%) versus 0.20% and 0.25% in the Prolific and CloudResearch Connect samples, respectively. The relative absence of stated issue priorities in the YouGov sample also dovetails with lower rates of ideological self-identification, with 26.8% refraining from selecting an ideological category (relative to 5.2% 5.5% in the CloudResearch Connect and Prolific samples, respectively).

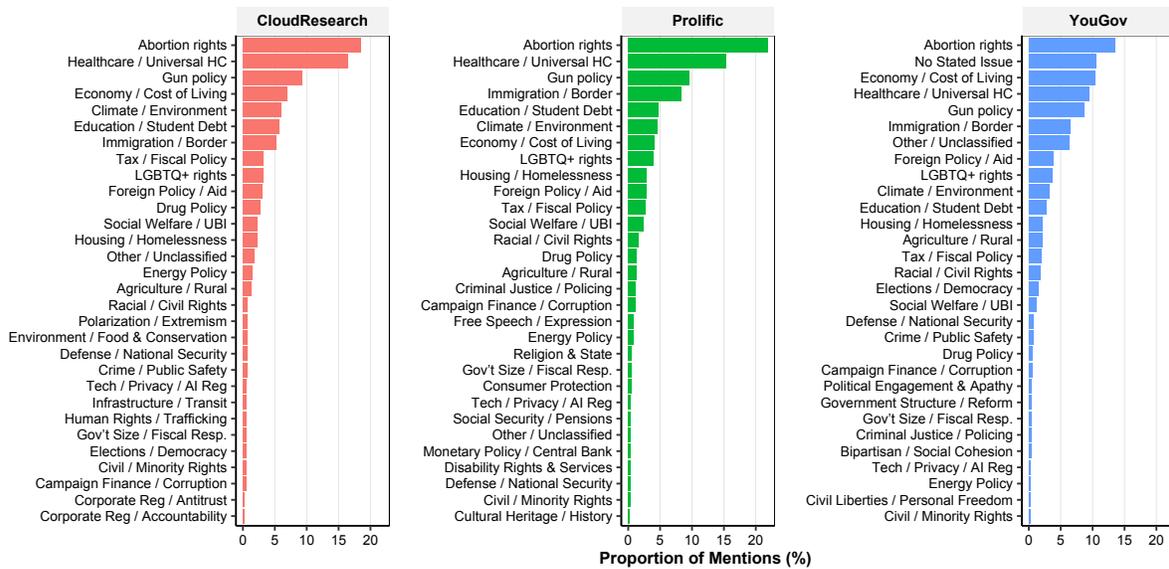


Figure 5: Prevalence of issue topics across the CloudResearch Connect, Prolific, and YouGov samples

¹Our use of a locally hosted LLM is in line with emerging best practices for replicating *classification* tasks.

Conversation Details

Experimental Chatbot Conditions

VAA Bot After indicating their most important political issue, participants entered a chatbot interface with pre-populated chat bubbles tailored to their stated concern. For example, a participant who mentioned abortion would be prompted to ask, “Where do Republicans and Democrats stand on key issues?” After each conversational turn, a “Knowledge Level” progress bar advanced incrementally beneath the text box. Upon reaching the fourth turn, participants were notified: “You’ve unlocked a new feature! Ask me which party is closest to you.” In Studies 2 and 3, an auto-complete feature also nudged users to type, “Which party comes closest to my views?” At the seventh turn, participants were instructed to continue to the next page.

The screenshot displays a chatbot interface with the following elements:

- Three pre-populated chat bubbles at the top: "Where do the Republicans and Democrats stand on Nuclear Policy?", "Do the parties disagree on Nuclear Policy?", and "What are the core values of the two parties?".
- A text input field containing the question: "I care about investing in alternative sources of energy".
- A response bubble from the bot: "The Republican and Democratic parties have very different views on alternative energy. Democrats strongly support transitioning to clean, renewable sources like wind and solar. They want major investments in green infrastructure and aim to move the U.S. toward 100% renewable electricity. Republicans, on the other hand, tend to prioritize energy independence using a mix of sources, including fossil fuels. They generally oppose government mandates that push renewables over traditional energy. Since you're into alternative energy, you're probably going to connect more with the Democratic platform on this issue. Want to dig into how each party views climate change or energy jobs next?".
- A second text input field containing: "I'm worried about an overreliance on oil".
- A second response bubble: "The Republican and Democratic parties take very different approaches to America's reliance on oil. Republicans generally support expanding domestic oil production, including drilling and pipeline development, as part of an 'all-of-the-above' energy strategy. Democrats, on the other hand, focus on reducing oil dependence by promoting clean energy, phasing out coal power, and investing in energy conservation and renewables. Since you're concerned about overreliance on oil, you might want to look closer at each party's energy and climate goals. Want to dive into how they differ on renewable energy or climate change?".
- A third text input field containing: "Provide a richer description of where the parties stand on energy policy".
- A third response bubble: "The Republican and Democratic parties take very different approaches to energy policy. Republicans support an 'all-of-the-above' strategy, which includes expanding oil and gas exploration, building new refineries, and cutting federal environmental regulations that they see as burdensome. In contrast, Democrats focus on transitioning away from fossil fuels, aiming for 90% renewable energy by 2050, with strong support for clean energy investment, conservation, and environmental justice. Are you more interested in how these policies might affect jobs, energy prices, or the climate?".
- A fourth text input field containing: "What about expanding healthcare?".
- A fourth response bubble: "The Republican and Democratic parties have very different views on expanding healthcare. Democrats generally support expanding government-funded healthcare programs like the Affordable Care Act and even moving toward a single-payer system in the long term. Republicans, on the other hand, tend to favor reducing the federal role in healthcare, promoting private insurance, and giving more control to states and individuals. Do you think the government should be more involved in healthcare, or do you prefer a more market-based approach?".
- A text input field at the bottom containing: "Which party comes closest to my views?".
- A green "Ask" button next to the input field.
- A yellow notification bar: "You've unlocked a new feature! Ask me which party is closest to you."
- A green progress bar labeled "Knowledge Level" at the bottom.

Figure 6: Simulated interaction with the VAA Bot. Pre-populated bubbles prompt a participant who mentioned nuclear power to begin the conversation by asking about party differences on a personally meaningful issue. The knowledge level progress bar advances after every turn. A highlighted dialog below the text box indicates when a new feature is unlocked.

Ideology Bot (Prolific Only) The progression of the Ideology Bot mirrored the VAA Bot, but participants were instead prompted to select questions about liberal and conservative positions (e.g., “Which issue positions are traditionally considered liberal or conservative?”). As with the VAA Bot, the fourth conversational turn unlocked a new feature: “You’ve unlocked a new feature! Ask me which party is more liberal or conservative than the other.” After seven turns, participants advanced to the next question.

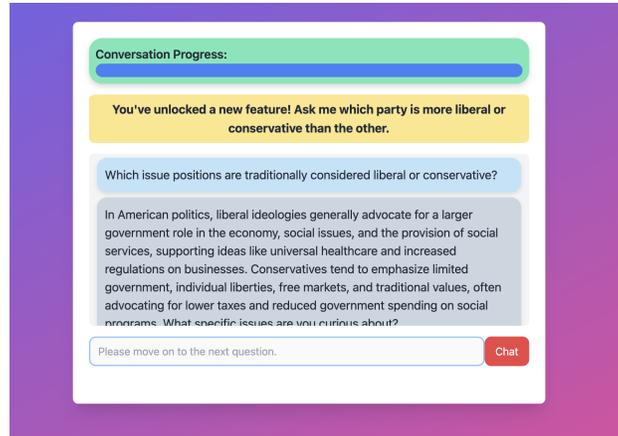


Figure 7: Simulated interaction with the Ideology Bot. Pre-populated bubbles prompt discussion of liberal and conservative positions. The knowledge level progress bar advances after each turn. A highlighted dialog below the text box indicates the new feature.

Trivia Bot (Placebo) Participants assigned to the placebo condition encountered a similar interface, but the task was to identify a historical figure — either Benjamin Franklin or, in a second Prolific arm, Paul Revere. Pre-set bubbles included questions such as “Where was this person born?” After five turns, the identity of the character was revealed.

The image shows a simulated interaction with a Trivia Bot (placebo) interface. It consists of a vertical sequence of bubbles. The first bubble is light blue and contains the text "Tell me one random fact about this person." Below it is a light grey bubble with the text "I once called the city of brotherly love my home, where I engaged in the pressing matters of the day." The second bubble is light blue and contains the text "What was your occupation?" Below it is a light grey bubble with the text "In my time, I was known to have a hand in crafting words that would be read widely, shaping public opinion as I arranged type." The third bubble is light blue and contains the text "Where were you born?" Below it is a light grey bubble with the text "I first drew breath in a place known for its baked beans and a certain infamous tea party." At the bottom is a white input field with the placeholder text "Ask me for clues." and a green arrow button to its right.

Figure 8: Simulated interaction with the Trivia Bot (placebo). Pre-populated bubbles prompt participants to ask questions about the character’s identity.

Classifying Interactions

We calculate the number of conversational turns (i.e., a single exchange between a human and a bot) across the three studies. Engagement with the bot ranged from an average of 3.82 turns in the CloudResearch Connect study to 5.77 turns in the Prolific study, with YouGov between the two (4.2). In all of the studies, a majority of participants reached the phase of the experiment where they were encouraged to ask which party was most proximate to them (51% in the YouGov study; 53% in the CloudResearch study; 77% in the Prolific study).

Engagement generally varied based on whether the participant possessed a party-aligned issue, with longer conversations observed among the Democratic-aligned and Republican-aligned participants when compared to the No clear party subgroup (with the exception of those in the YouGov sample). Whereas 55% to 57% of participants in the CloudResearch study with an aligned issue reached the dialogue encouraging them ask which party was closest to them, only 39% of the unaligned participants did so. Similarly, between 77% and 84% of those in the Prolific sample who possessed a party-aligned issue reached the party proximity phase, whereas only 74% of those in the unaligned subgroup did so. Finally, in the YouGov sample, between 49% and 51% of participants reached the party proximity phase of the experiment if they possessed an aligned issue, whereas 52% of those in the unaligned condition did so. In the CloudResearch (Prolific) study, the average number of conversational turns ranged from 3.44 (5.30) to 3.95 (5.96), with unaligned participants engaging in shorter conversations. In the YouGov study, average conversational turns ranged from 4.12 to 4.27, with slightly shorter conversations when issues were Republican-aligned versus Democratic-aligned and unaligned.

Table 20: Descriptive statistics for conversational length (turns) by platform.

Source	Conversation Length (Turns)					
	≥ 4 Turns	Mean	SD	25th %	Median	75th %
CloudResearch	0.52	3.82	2.42	2	4	6
Prolific	0.81	5.77	2.12	5	7	7
YouGov	0.51	4.20	2.57	2	4	7

Table 21: Descriptive statistics for conversational length (turns) by platform.

Source	Classification	Conversation Length (Turns)					
		≥ 4 Turns	Mean	SD	25th %	Median	75th %
CloudResearch	Democrat-aligned Issue	0.55	3.90	2.45	2	4	6.00
CloudResearch	No clear party	0.39	3.44	2.25	2	3	5.25
CloudResearch	Republican-aligned Issue	0.57	3.95	2.62	2	4	7.00
Prolific	Democrat-aligned Issue	0.84	5.96	1.99	6	7	7.00
Prolific	No clear party	0.74	5.30	2.38	3	7	7.00
Prolific	Republican-aligned Issue	0.77	5.54	2.25	5	7	7.00
YouGov	Democrat-aligned Issue	0.51	4.27	2.48	2	4	7.00
YouGov	No clear party	0.52	4.17	2.57	2	4	7.00
YouGov	Republican-aligned Issue	0.49	4.12	2.90	1	3	7.00

Though these findings suggest a modest level of engagement with the VAA Bot, a more qualitative assessment can distinguish between genuine engagement and more perfunctory interactions. We used a 4-bit quantized version of a local LLM (MistralAI’s Mistral Small 3.1 24B) to classify all of the VAA Bot conversations across the three studies as part of a non-pre-registered exploratory analysis. The LLM was instructed to assign each conversation to one of four categories:

- **Engaged:** Users who ask thoughtful questions, show interest, and follow up.
- **Perfunctory/Disengaged:** Users with terse, minimal responses showing little interest.

- **Inattentive:** Users providing gibberish, nonsensical, or completely irrelevant responses.
- **Adversarial/Skeptical:** Users who challenge claims, express disagreement, or question neutrality of bot.

These categories capture a range of interactions, from more genuine information gathering conversations to shorter interactions suggesting a desire to quickly move to the next section of the survey. These are instructive as they provide a glimpse into how such tools might be received among the general public. Across the three studies, we find that the modal category is “Engaged” (69.5% in YouGov; 81.0% in CloudResearch Connect; 85.7% in Prolific) suggesting genuinely informative interactions with the VAA Bot, followed by “Perfunctory/Disengaged” (8.0% in Prolific; 14.9% in CloudResearch Connect; 20.3% in YouGov). The remaining categories of “Adversarial/Skeptical” and “Inattentive” comprise 4.0% in the CloudResearch Connect sample, 6.2% in the Prolific sample, and 10.2% in the YouGov sample. Dividing the sample into issue alignment subgroups, we generally observe more engaged responding among participants with party-aligned issues. In the CloudResearch Connect study, 83.7% of those with a Democratic-aligned issue could be classified as engaged versus 66.7% of those with a Republican-aligned issue and 78.8% of those with an unaligned issue. Moving to the Prolific study, 86.3% and 86.7% of those with Democratic and Republican-aligned issues could be characterized as engaged relative to 81.8% of those without an aligned issue. Finally, in the YouGov sample, 77.8% and 73.4% of participants with Democratic and Republican-aligned issues could be classified as engaged, relative to 60.3% without an aligned issue.

This suggests that the vast majority of participants approached the VAA Bot interactions constructively, rather than attempting to game or test the system. The variation across platforms, with Prolific showing the highest engagement rates and YouGov showing more perfunctory interactions, may reflect differences in participant pools, with Prolific’s sample being comprised of “professional survey takers” who are more accustomed to thoughtful survey participation. Moreover, the slight differences in engagement across issue-based subgroups could be symptomatic of lower levels of political engagement among those who prioritize issues with no clearly discernible partisan stance. As their preferences may not be properly reflected in the political system, engagement with tools that are predicated on established partisan differences may be seen as less personally relevant.

Table 22: Distribution of conversation classifications across study platforms

Classification	CloudResearch Connect	Prolific	YouGov
Engaged	141 (81.0%)	426 (85.7%)	684 (69.5%)
Perfunctory/Disengaged	26 (14.9%)	40 (8.0%)	200 (20.3%)
Adversarial/Skeptical	5 (2.9%)	27 (5.4%)	58 (5.9%)
Inattentive	2 (1.1%)	4 (0.8%)	42 (4.3%)

Table 23: Conversation classification by issue subgroup

Classification	Democrat-aligned Issue	No clear party	Republican-aligned Issue
CloudResearch			
Engaged	103 (83.7%)	26 (78.8%)	12 (66.7%)
Perfunctory/Disengaged	17 (13.8%)	4 (12.1%)	5 (27.8%)
Adversarial/Skeptical	2 (1.6%)	2 (6.1%)	1 (5.6%)
Inattentive	1 (0.8%)	1 (3%)	0 (0%)
Prolific			
Engaged	272 (86.3%)	63 (81.8%)	91 (86.7%)
Perfunctory/Disengaged	25 (7.9%)	7 (9.1%)	8 (7.6%)
Adversarial/Skeptical	16 (5.1%)	5 (6.5%)	6 (5.7%)

Classification	Democrat-aligned Issue	No clear party	Republican-aligned Issue
Inattentive	2 (0.6%)	2 (2.6%)	0 (0%)
YouGov			
Engaged	332 (77.8%)	261 (60.3%)	91 (73.4%)
Perfunctory/Disengaged	62 (14.5%)	123 (28.4%)	15 (12.1%)
Adversarial/Skeptical	17 (4%)	27 (6.2%)	14 (11.3%)
Inattentive	16 (3.7%)	22 (5.1%)	4 (3.2%)

We provide a set of example conversations that exemplify these different engagement patterns.

Engaged

- User: What is the difference between a Democrat and a Republican

Bot: The Democratic and Republican parties in the U.S. have different core beliefs. Democrats generally advocate for a government that plays a more active role in providing social and economic support, emphasizing equality, and environmental protection. Republicans, on the other hand, focus on individual rights, free enterprise, and limited government intervention, stressing personal responsibility and traditional values. Both parties have diverse opinions within their ranks, so it's always good to look at specific issues you're interested in. What issues are you curious about?
- User: What are the core values of the two parties?

Bot: The core values of the Democratic and Republican parties differ in several key areas. Democrats typically emphasize social equality, environmental protection, and a more significant role for the government in ensuring social welfare and economic support. On the other hand, Republicans prioritize individual freedoms, a free-market economy, and a limited government role, with a focus on personal responsibility and traditional values. These foundational beliefs shape their policies and approaches to governance. Is there a particular issue you want to explore further?
- User: What is the biggest difference between Democrat and Republican

Bot: The biggest difference between Democrats and Republicans often revolves around the role of government in society. Democrats tend to support a larger government role in providing social services and regulating economic activities to promote equality and protect the environment. Republicans, conversely, advocate for smaller government, emphasizing personal responsibility, individual liberties, and free-market principles. These differing views influence their policy-making and priorities. Are there any specific areas or issues you would like more information on?
- User: What are the core values of the two parties?

Bot: The core values of the Democratic and Republican parties have some distinct differences. Democrats generally emphasize social equality, environmental protection, and a significant government role in ensuring social welfare and economic support. Republicans, on the other hand, prioritize individual freedoms, a free-market economy, and a limited government role, with a focus on personal responsibility and traditional values. These differing values shape their policy-making and approach to governance. If you want to dive into a specific issue, let me know!
- User: What his the difference between a Democrat and a Republican

Bot: The Democratic and Republican parties differ primarily in their views on the role of government. Democrats generally support a larger government role in providing social services, promoting equality, and protecting the environment. In contrast, Republicans emphasize a smaller government, focusing on personal responsibility, individual liberties, and free-market principles. These differences shape their policies and priorities. Let me know if there's a specific issue you're curious about!
- User: What are some of the main topics in the election this yeT

Bot: In the current election, some of the main topics being discussed include the economy, healthcare, climate change, and immigration. Democrats often focus on expanding healthcare access, addressing

climate change through environmental regulations, and reforming immigration policies to be more inclusive. Republicans typically emphasize economic growth through tax cuts and deregulation, securing borders and reforming immigration laws, and maintaining a strong national defense. Each party has its approach to these issues, reflecting their core values. What specific topic would you like to explore further?

- User: What are the core values of the two parties?

Bot: The Republican and Democratic parties have fundamentally different core values. Democrats emphasize collective responsibility, social programs, and government's role in addressing inequality. Republicans focus on individual liberty, free markets, limited government intervention, and traditional values. Would you like to explore how these values shape specific policies like healthcare, education, or economic approaches?

- User: What is democratic view on abortion

Bot: The Democratic and Republican parties have opposing views on abortion rights. Democrats generally support reproductive rights and access to abortion services, viewing it as a fundamental healthcare right and personal choice. Republicans typically oppose abortion and support restrictions, viewing it as a moral issue that should be limited. Would you like to know more about how these positions translate into specific policies?

- User: The Republican view on abortion

Bot: The Republican and Democratic parties have opposing stances on abortion. Republicans generally oppose abortion, advocating for restrictions or bans to protect what they view as the right to life of the unborn, with exceptions varying by candidate. Democrats support abortion access as a fundamental healthcare right. Would you like to learn more about specific policy proposals from either party on this issue?

Perfunctory/Disengaged

- User: What are the core values of the two parties?

Bot: The core values of the Republican Party typically emphasize limited government, individual liberties, free-market capitalism, and a strong national defense. They often focus on traditional social values and personal responsibility. On the other hand, the Democratic Party generally values social equality, environmental protection, healthcare access, and a more active role of government in economic regulation and social welfare. They often advocate for diversity and inclusion, as well as expanded civil rights. Is there a specific issue you'd like to explore further?

- User: Got it

Bot: Great! If you have any specific issues or topics you'd like to dive into, feel free to ask. Whether it's healthcare, education, the economy, or something else, I'm here to help you understand the parties' positions.

- User: Like how

Bot: It sounds like you're interested in how each party approaches specific issues. For example, on healthcare, the Democratic Party generally supports expanding access through government programs like the Affordable Care Act, while the Republican Party often favors market-based solutions and reducing government involvement. If there's a particular issue you're curious about, just let me know!

- User: What's the difference

Bot: The Republican and Democratic parties have different approaches to many key issues. For instance, in terms of economic policy, Republicans typically advocate for lower taxes and less regulation to stimulate business growth, while Democrats often support higher taxes on the wealthy to fund social programs and greater regulation to protect consumers and the environment. If you're interested in another specific topic, feel free to ask!"

Adversarial/Skeptical

- User: What are the core values of the two parties?
Bot: The Democratic Party generally emphasizes social equality, environmental protection, and a stronger social safety net, while the Republican Party tends to prioritize limited government, free market principles, and national defense. Both parties hold a range of views within them, but these are some of their core values that guide their platforms and policies.
- User: Ha! The democrats and republicans are barely different from each other. Name one social equality issue democrats have actually voted to solve
Bot: Democrats often focus on expanding healthcare access, and one significant piece of legislation they passed to address social equality in healthcare is the Affordable Care Act (ACA). The ACA aimed to reduce the number of uninsured Americans and make healthcare more affordable.
- User: ACA is barely healthcare at all, let alone something that makes it more affordable. Why is it that so many democrats support sending military aid to Israel?
Bot: The Democratic and Republican parties diverge on their approach to foreign aid, including to Israel. Many Democrats have supported aid to Israel as part of a longstanding alliance and commitment to Israel's security, while also advocating for a two-state solution and diplomacy in the region. Republicans typically emphasize strong support for Israel's military defense capabilities as a key ally in the Middle East.
- User: Doesn't sound like democrats care too much about social equality at all if any of them could possibly consider funding war a good thing.
Bot: The Democratic Party's stance on social equality encompasses various domestic policies focused on reducing inequality and promoting justice. However, international relations, including military aid, can be complex and involve strategic alliances, security considerations, and diplomatic efforts that some Democrats believe contribute to broader global stability, which they may view as indirectly supporting social equality by promoting peace and democracy internationally.
- User: Ah yes, the number one contributor to peace: war. that was sarcasm.
Bot: The Republican and Democratic parties both engage in foreign policy decisions that can be contentious, such as military aid. While some Democrats see military aid as a part of international obligations and strategic alliances, others within the party call for more restraint and focus on diplomacy and development. The debate within the party reflects a range of views on how best to achieve peace and security.
- User: which party is closest to me
Bot: The Republican Party generally emphasizes limited government, free enterprise, and strong national defense. The Democratic Party tends to focus on social equality, environmental protection, and a stronger social safety net. Reflect on which set of values aligns more closely with your own to determine the party that is closest to you.
- User: Neither, they are both truly evil oligarchs with no care for citizens. On a political scale, they'd both be right wing.
Bot: The Republican and Democratic parties both have complex platforms that can be perceived differently by various individuals. On the political spectrum, the Republican Party is typically seen as more conservative, advocating for limited government and free markets, while the Democratic Party is often viewed as more liberal, advocating for social programs and regulation for social equity. If you feel that neither aligns with your views, you might explore third parties or independent candidates that could better represent your perspectives.

Inattentive

- User: Where do the Republicans and Democrats stand on Justice System?
Bot: The Republican Party generally advocates for strict law enforcement measures, emphasizing the need for tough sentencing to deter crime, while the Democratic Party tends to focus on criminal justice reform, including reducing mandatory minimum sentences and addressing systemic racial disparities.

- User: Where do the Republicans and Democrats stand on Justice System?
Bot: The Republicans lean towards stringent law enforcement and tough sentencing to combat crime. In contrast, the Democrats push for reforms aimed at reducing sentences, particularly mandatory minimums, and correcting systemic racial imbalances within the justice system.
- User: Where do the Republicans and Democrats stand on Justice System?
Bot: The Republican Party traditionally supports strong law enforcement and strict sentencing as a deterrent to crime, while the Democratic Party focuses more on reforming the criminal justice system to address issues like racial inequality and reducing mandatory minimum sentences.