

Pre-analysis plan

How do colors convey political information and affect individual attitudes?

Damon C. Roberts 
University of Colorado Boulder
damon.roberts-1@colorado.edu

ABSTRACT Colors are important to politics as a form of political information. Building upon existing theories of political information processing, attitude formation, and affect in neuroscience, I present a snap-judgement model of political information processing. In this model, colors provide automatic information about a politically-relevant object that may shape subsequent processing of more complex information. The model has important implications for how we consider the role that visual information has on political information processing and attitude formation. The model additionally provides clarity on motivations behind party branding and the ways in which information may activate partisan biases pre-consciously. I test this model using a survey experiment tracking participants' mouse movements to view different parts of yard signs that vary the use of partisan and non-partisan colors. I additionally leverage congressional redistricting to examine changes in color use for political yard signs as a strategic choice for attracting voters by communicating partisanship.

Table 1: Summary of hypotheses

Hypotheses	Expectation
H_1	People notice color
H_2	Red is associated with Republicans and Blue is associated with Democrats
H_3	Candidates using Red are more supported by Republicans; candidates using Blue are more supported by Democrats
H_4	Republicans spend less time evaluating candidates using Red; Democrats spend less time evaluating candidates using Blue
H_5	Campaigns recognize the importance of color to voters' evaluations of candidates and respond strategically



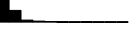

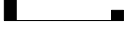


From my review of the literature in political science, neuroscience, and communication, I expect the following: that the average potential voter pays attention to the colors used in campaign branding (H_1); that these colors that they notice shape perceptions about the person and ideological symbol represented in the branding – meaning that they express different levels of preference for receiving more information that is similar to what they saw and their levels of preference for supporting such a campaign (H_2); that the consistency of information explains more positive perceptions – simple visual information with more complex “traditional” information (H_3); this positive and consistent information is processed quicker than negative and inconsistent, negative and consistent, and positive and inconsistent information (H_4); and finally that campaigns make strategic choices about their branding to attract voters (H_5) in line with their primary objective of reelection (Fenno 1973; Mayhew 1974). A summary of these hypotheses are included in Table 1.

Pre-test

I conducted a pre-test in November 2019 with a sample of over 400 undergraduate students at a large university in the northwestern region of the United States. I recruited students enrolled in a political science course and were offered extra credit for their participation in the study. The study asked participants to participate in 5 survey experiments administered by those affiliated with the university’s college-level unit. These other survey experiments focus on capturing local policy issues around urban design and criminal justice and probing participants about political participation in local and national-level elections. Subjects participated in my survey experiment after one that examined their levels of political participation in local, state, and national elections.

Using the `modelsummary` (Arel-Bundock 2022b) package implemented in the R language (R Core Team 2022), I create Table 2 which presents the descriptive characteristics of the sample. The sample is primarily White with over 80% self-reporting that they are White(coded as: 0 = non-White, 1 = White). The sample also skews slightly female on sex, with about 60% reporting that they are female

Table 2: Descriptive statistics of pre-test measures

	Unique (#)	Missing (%)	Mean	SD	Min	Median	Max	
Female	3	3	0.6	0.5	0.0	1.0	1.0	
White	3	3	0.8	0.4	0.0	1.0	1.0	
age	30	2	21.7	5.6	18.0	20.0	68.0	
PartyId	8	15	0.1	2.2	-3.0	0.0	3.0	
BlueTreatment	2	0	0.3	0.5	0.0	0.0	1.0	
RedTreatment	2	0	0.3	0.5	0.0	0.0	1.0	
Vote	4	9	1.9	0.4	1.0	2.0	3.0	

Data source: Pre-test experiment.

"Unique" column includes missing values.

(coded as 0 = Male, 1 = Female). The sample also, unsurprisingly, skews young, with the average respondent reporting an age of about 22 years old. The average respondent also appears to be an independent but leans Republican (coded as -3 = strong Democrat, -2 = Democrat, -1 = leans Democrat, 0 = Independent, 1 = leans Republican, 2 = Republican, 3 = strong Republican).

I randomly assign participants to three conditions. The conditions prompt subjects to "Imagine that [they] are driving along a road and see this yard sign" with the same message "Vote for Riley." The conditions vary on the color of the background for the image.¹ In the control condition, the background was white. Then I had a red yard sign and a blue yard sign condition. I asked participants questions operationalized as outcomes of interest on a separate screen.

To provide a preliminary test of H_1 and H_2 , I ask participants to report whether the candidate was a "Republican, Democrat, or Independent." I test whether participants presumed that the candidate was of a particular partisan persuasion based only on the color choice of the yard sign alone. I created two indicator variables of the treatment the subjects received: whether or not they had the blue or red yard sign treatment. The placebo condition (the white yard sign treatment) is treated as the baseline condition when including both treatment indicator variables in the model.

To examine whether subjects presumed a partisan affiliation of the fictional candidate, I use the `brms` (Bürkner 2017) implementation of `STAN` to fit an ordinal logistic regression model using the cumulative PDF with a inverse logit link function. My model specification is reflected in Equation 1. I use posterior prediction checks after fitting the model to examine whether the predicted values from the posterior draws resemble my observed values. I calculate the average marginal effects across my posterior draws with High Density Intervals (as op-

1. The Appendix contains the images used for the treatments and the particular wording for the dependent variables.

posed to Equal Tail Intervals) and plot the distributions of these. In Figure 1, the bars reflect the 90% and 95% credible intervals whereas the dot reflects the median posterior draw.

$$\begin{aligned}\alpha &\sim \text{Student} - T(3, 0, 2.5) \\ \beta_i &\sim \text{Normal}(0, 1) \\ \hat{Party}_i &= \text{Cumulative}(\text{logit}^{-1}(\alpha_1 + \alpha_2 + \beta_1 \times \text{RedTreatment} + \beta_2 \times \text{BlueTreatment}))\end{aligned}\tag{1}$$

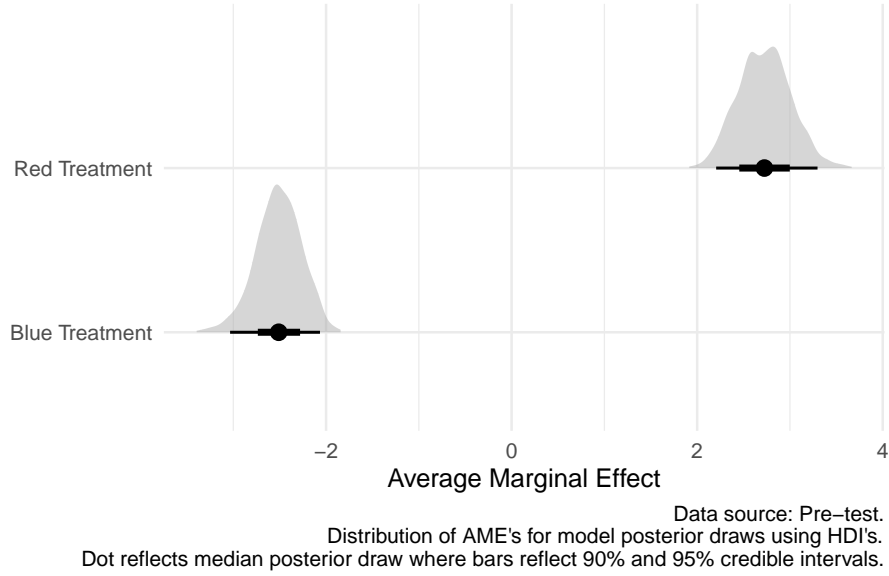


Figure 1: Effect of treatments on party perceptions

The results suggest that for those in the red yard sign treatment condition perceive the candidate to be a Republican. Whereas, those assigned to the blue yard sign treatment condition perceived the owner of the yard sign to be a Democrat. This is without *any* information about the political candidate. Both of these treatments have a quite large effect on conveying the partisanship of the candidate. This finding fits with the expectation that individuals have internalized the use of color in party branding and are independently capable and willing to make that connection. This preliminary evidence suggests that color likely conveys important information to the public.

Individuals associate Republicans with the color Red and Democrats with the color Blue. Do these associations influence their attitudes? In addition to asking participants about whether they perceived the candidate was Republican or Democratic, prior to that question, I asked participants whether they would be willing to vote for the candidate and whether they would want to avoid engaging

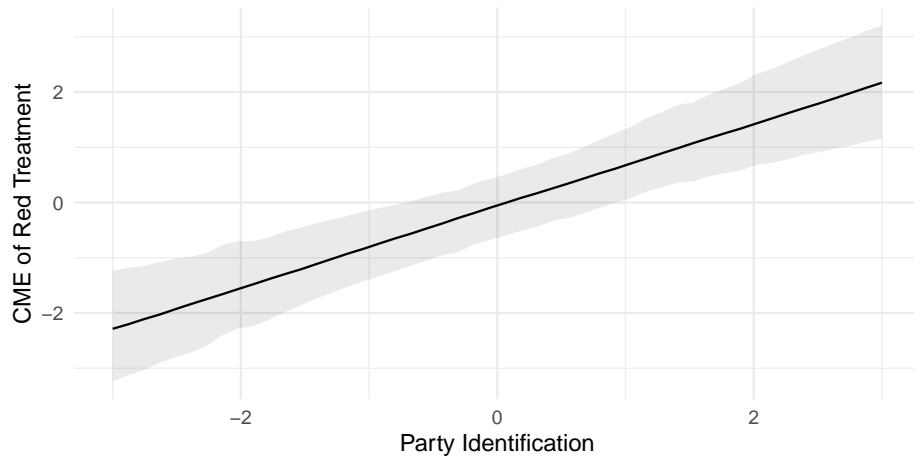
with the candidate’s campaign materials in the future. I used these two questions to construct a measure of willingness to engage with the fictional candidate. Those who reported that they do not want to vote for the candidate *and* want to avoid the candidate in the future I coded with a value of 1. Those who reported a mix of responses were coded as 2. Finally, those that reported that they were willing to vote for the candidate *and* wanted to receive more information about the candidate were coded as 3. Again, the *only* information they have about the candidate is conveyed on the yard sign. I fit an ordinal logistic regression model again using the **brms** (Bürkner 2017) implementation of **STAN**. My model specification is conveyed in Equation 2.

$$\begin{aligned}\alpha &\sim Student - T(3, 0, 2.5) \\ \beta_i &\sim Normal(0, 1) \\ \hat{Party}_i &= Cumulative(logit^{-1}(\alpha_1 + \alpha_2 + \beta_1 \times RedTreatment + \beta_2 \times BlueTreatment \\ &\quad + \beta_3 \times PartyIdentification(3 - item) \\ &\quad + \beta_4 \times RedTreatment \times PartyIdentification(3 - item) \\ &\quad + \beta_5 \times BlueTreatment \times PartyIdentification(3 - item))) \\ &\quad (2)\end{aligned}$$

I expect that the effect of the treatment on willingness to engage more with the candidate, either in terms of learning more (Lodge and Taber 2013) and voting (Broockman and Kalla 2022) is dependent on whether that candidate is a co-partisan. I also examined this model with a variety of priors and settled on priors I believe to reflect maximum entropy principles (McElreath 2020) and those that produced posterior predictive checks with predicted values from the model resembling those observed in my data. I use the **marginaleffects** package (Arel-Bundock 2022a) to calculate and plot the conditional marginal effects.

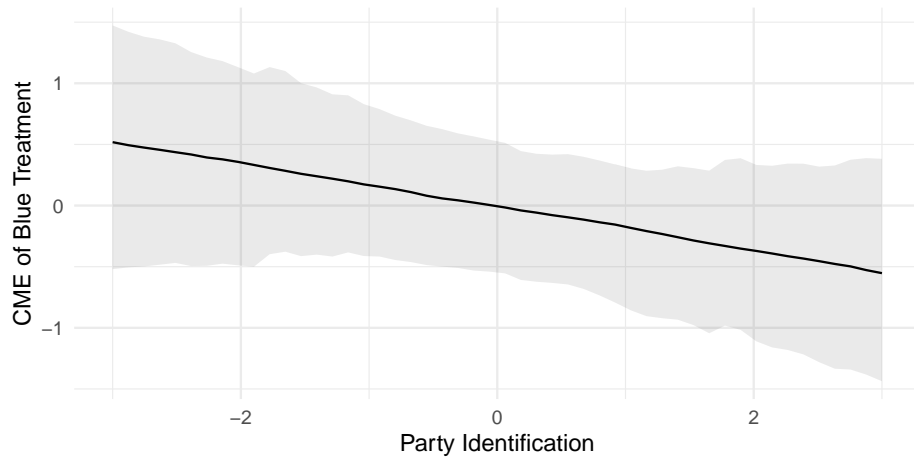
Figure 2 presents results suggesting that among Republicans receiving the red treatment, they are more likely to indicate a positive valence toward the candidate than Democrats, who are more likely to report a negative evaluation of the candidate. While Democrats receiving the blue treatment are more likely to report a positive valence toward the candidate relative to Republicans, the effect is plausibly zero. This may be an artifact of asymmetric political polarization. Some scholars suggest that Republicans are much more group-oriented than Democrats (see Lupton, Smallpage, and Enders 2020); these results may fit with such a narrative. Republicans are reactive to those they may presume to be co-partisan in a way that Democrats do not appear to be as reactive in a similar magnitude.

The evidence from this pre-test is limited. The experimental design is not testing pre-conscious evaluations. The treatments are explicitly political and rely upon a convenience sample of those enrolled in political science classes. The ability of individuals to associate partisanship with the treatments is likely overstated. The sample is also quite unrepresentative, so the inference is indeed threatened. As I discussed with the third model, I also have omitted variable



Data source: Pre-test.
Conditional Marginal Effects of Red Treatment upon support for candidate.
Uncertainty reflected by predicted posterior draws at the 95% level with HDI's.

(a) Republicans would vote for candidate with Red yard sign



Data source: Pre-test.
Conditional Marginal Effects of Blue Treatment upon support for candidate.
Uncertainty reflected by predicted posterior draws at the 95% level with HDI's.

(b) No difference among partisans in support for candidate with Blue yard sign

Figure 2: Effect of yard sign color on candidate evaluation

bias as the result of my outcome is a better measure of approach behaviors for those who are extroverted.

Though there are problems with the design here, it does serve some purpose for this project. It demonstrates that despite its problems, color can betray information at some level when thinking of politics. It also provides a functional first test of a possible research design to examine what this particular design can and cannot tell me about my proposed mechanism.

Study 1

The purpose of [Study 1](#) is to examine the average treatment effect ² of color in a common form of campaign branding – yard signs – on the attitudes of the observer. Evidence suggests that yard signs *do* matter in shaping political attitudes (Makse, Minkoff, and Sokhey 2019), vote intentions (Makse, Minkoff, and Sokhey 2019), and even electoral outcomes (Green et al. 2016). As they are simple, cheap, intrusive, and common forms of campaign branding, they provide a conservative test of the effects of color in campaign branding. Using yard signs in this study, I test the first four hypotheses I derive from the snap-judgment model. First, [Study 1](#) tests the claim that individuals do indeed notice the color of electoral yard signs. Next, it tests the claim that these colors that individuals detect influence the viewer’s evaluations of the yard sign and the candidate represented on it. [Study 1](#) then tests the claim that the effects on perception are moderated by how consistent the color is with the more complex information displayed on the yard sign. Furthermore, finally, [Study 1](#) examines whether positive information that contains consistency is processed quicker than negative or inconsistent information. That is, do partisans quickly detect and encode information from a clearly co-partisan political candidate?

Design

I recruit participants from Prolific.³ After providing informed consent to participate in the study, Prolific redirects subjects to Pavlovia⁴ and are provided with a demographics and political attitudes questionnaire. As I am concerned about priming effects introduced by these questions, while I am simultaneously concerned about bias introduced by post-treatment control (Montgomery, Nyhan, and Torres 2018), I randomly select half of the participants to receive the questionnaire post-treatment and the other half receive it pre-treatment. This questionnaire includes common questions about the participant’s ascriptive and descriptive characteristics and about the participant’s political ideology, partisan identification, interest in politics, patriotism, and political knowledge.

2. $\frac{1}{n} \sum i(Y_i(d=1) - Y_i(d=0))$ (see Lundberg, Johnson, and Stewart 2021)

3. I pay subjects a rate of \$12.00 per hour. On top of the price per participant, Prolific charges a 30% servicing fee and an additional fee to guarantee a nationally representative sample.

4. Pavlovia allows for researchers to host and run open source experiments for about \$0.20 per participant (to cover their server costs). I use it primarily to integrate the JavaScript components from the jsPsych package (Leeuw 2015) for my experimental design.

I include those questions in the questionnaire due to expectations that they may act as confounds in my hypotheses. As political knowledge intertwines with the strength to which an individual identifies with a political party (Delli Carpini and Keeter 1996), I expect that political knowledge is an important confounder in my tests of H_2 , H_3 , and H_4 . Therefore, I include a standard battery (the American National Election Study’s battery) for assessing political knowledge.⁵ As political knowledge is shaped by levels of interest in politics as well (Delli Carpini and Keeter 1996), I include the American National Election Study’s question to assess levels of self-reported interest in politics. I include a battery to measure patriotism. For this measure, participants can respond to values between 1 (not proud at all) and 4 (very proud) to the following questions: “When I hear the American national anthem, it makes me feel” and “when I say the American pledge of allegiance, it makes me feel” (Huddy and Khatib 2007; Pérez, Deichert, and Engelhardt 2019). I additionally include some questions collecting information on participants’ ascriptive and descriptive characteristics such as age, education, gender identity, and racial identity, as a number of these are correlates with partisan identification (see Campbell et al. 1960; Mason 2018).

Additionally, I include a question about the respondent’s sex assigned at birth and about whether they have received a diagnosis of any color blindness. As some individuals may possess undiagnosed colorblindness, asking about their sex assists in covariate balance. I additionally include four open-ended questions asking participants to describe (1) their “first memory of a political event,” (2) “what occurred in the most recent political event” they encountered, (3) how they “keep up-to-date with current political events,” and (4) “what occurred in the most recent non-political event” they encountered. The use of multiple open-ended questions helps provide an attention check and identify duplicated responses for those spoofing IP addresses with a VPN (see Kennedy et al. 2021). Given estimates using these exclusion criteria, I would expect that upwards of 40% of my original sample will fail these attention checks (Kennedy et al. 2021); thus, the large sample size.

I then present participants with an instruction screen informing them of the task for the experiment. In the first trial of the experiment, participants I randomly presented participants with two of three possible yard signs, one at a time. These yard signs are simple with the text “Vote for Riley” and a solid background color of either “Republican Red,” “Democratic Blue,” or White.⁶ There is an added component to this, however.

Rather than use eye-tracking devices and software, I instead use Mouseview.js (see Anwyl-Irvine, Armstrong, and Dalmaijer 2022), which either blocks out or blurs a large portion of the participant’s screen and encourages them to move their mouse to view different parts of the screen in isolation. As the participants move their cursor around the screen, it tracks the coordinates of the cursor along with the “dwell” time of the cursor in that particular coordinate. One primary

5. A table including the wording of this, and all, measures are included in the Supplementary Information.

6. See the supplementary information to view all of the stimuli used in [Study 1](#).

benefit of Mouseview.js is that it allows researchers to field their experiments outside of a lab-based setting – while providing results that robustly correlate with the results from a design employing eye-tracking hardware (Anwyl-Irvine, Armstrong, and Dalmaijer 2022). This allows researchers to rely less on convenience samples, which are common with eye-tracking studies. For my design, I am particularly concerned about reliance on a convenience sample due to variations in participants’ ability to detect and process color in the U.S. population relative to a student sample. This means that Mouseview.js is a handy tool for my study. A published pilot version of the experiment used in [Study 1](#) on [Pavlov](#).

When viewing each yard sign in the first trial, there is a blur over a substantial portion of the screen. At any given point in time, participants can view only 8% of the image without an obstruction, which simulates the observation that we typically foveate on about 8% of our available visual field at any given time (Wedel and Pieters 2008).⁷ Participants move their cursor to explore the yard sign. I allot 5000 ms to perform the exploration until the image goes away to encourage a consistent and short duration to explore the image.⁸ After exploring each image, I asked participants what colors were on the yard sign and whether they felt that the candidate represented on the yard sign was a Democrat, a Republican, or Neither. After viewing both images, I ask subjects to indicate their preference among the two images. To ensure that participants have a standardized initial placement of their cursor, I display a blank page before viewing the yard sign that requires participants to click a “Next” button. Immediately after clicking “Next,” participants are shown the yard sign. The goal is to ensure that variation in where participants explore the image is not dependent on a non-standard starting point for their cursor. I additionally utilize a gaussian blur for the overlay of the image rather than a solid overlay obstructing the participant’s view. This gaussian blur allows participants to see a blurred visual field beyond the cursor. This allows participants to see enough to take purposeful action to explore blurred parts of the image that attract them (Anwyl-Irvine, Armstrong, and Dalmaijer 2022). The use of the gaussian blur requires that participants use a web browser other than Safari because of a known issue (Anwyl-Irvine, Armstrong, and Dalmaijer 2022). This will require participants to either switch browsers or to not participate in the study if they are using Safari at the time they are recruited by Prolific to participate in the study. As this requirement is enforced *before* joining the study, this should not have an effect on the number of excluded participants from my original sample nor on a nationally representative original sample.

There are three more trials that are much like the first trial. What is different between the two other trials is that I vary the amount of color that is on the yard

7. I include a screenshot providing an example of what the participants are able to see with the blur included in the supplementary materials.

8. In marketing research, some studies give participants about 6000 milliseconds in eye-tracking studies to examine a brand and to formulate an intention to purchase a product or not (Wedel and Pieters 2008). With Mouseview.js, a study examining the tool’s correlation with optical responses to viewing disgust-and-pleasure-evoking images uses 10000 milliseconds; but is intended to be an extended amount of time (Anwyl-Irvine, Armstrong, and Dalmaijer 2022).

signs (trials 2 and 3). I provide more textual information that deviates from the association of Republicans with red and Democrats with blue (trial 4). Examples of all of these yard signs are included in the supplementary materials.

Do individuals notice color in political branding?

To address this first question, I use two measures of a participant’s attention toward the colors on the yard sign. I collected the first measure through a question posed to the participant after viewing each yard sign, “what color was the yard sign?”. The second measure is more implicit than the first: it accounts for the time someone’s mouse hovered over the non-text elements of the yard sign relative to how long their mouse hovered over the text. The self-reported measure allows us to examine the conscious detection of color, while the more implicit measure allows us to examine where individuals’ attention goes: toward the color or the text.

This is primarily a descriptive exercise, and I intend to compare the differences in measures between the “non-partisan” (white) yard signs and the partisan (red and blue) yard signs. I will examine the differences in measures between the partisan yard signs among self-identified partisan respondents. I additionally include additional models that interact partisanship with age and political knowledge. As the consistent usage by the parties to use the colors red and blue did not occur until the 2000 presidential election, those that were in their early adult years during that time experienced partisan politics where colors were not a strong cue. I additionally expect that those who are more politically knowledgeable should rely on these partisan cues more as they tend to be stronger partisans (Delli Carpini and Keeter 1996).

Do colors shape perceptions of political objects?

To address the following question of whether the colors affect perceptions of the candidate and the yard sign, I ask participants to report whether they perceived the candidate to be a partisan – either Republican, Democrat, or as non-partisan immediately after viewing each image. Everything on the yard signs remains constant except for the color. As representations of ideology are associated with more than just political views but things like space (Mills et al. 2016) and color (Maestre and Medero 2022), differences between respondents on the perceived political affiliations of the candidate should be more than stochastic differences. However, differences occur based on the associations between red and blue with partisanship and the lack of political information that the color white conveys.

I examine differences in respondents’ reported perceptions of the candidate’s partisan affiliation. As respondents are making the choice between two yard signs, my estimand of interest is of a pairwise comparison of the two yard signs they see in each trial. While these are most common in ranking athletic teams, there is recent academic interest in these estimands as they allow researchers to examine the “popularity” of a particular treatment when experimental participants are asked to choose between two options (see Hopkins and Noel 2022). As color is the only information that changes between yard signs in a given trial, pairwise

comparisons are a more appropriate estimand than those from conjoint analyses which often produce treatments that simultaneously vary a number of dimensions of a profile (Bansak et al. 2021). There are a handful of options about how to calculate a pairwise comparison, but I elect to estimate a Bradley-Terry score. Bradley-Terry scores are designed for determining “winners” that account for a variety of predictors (Schnakenberg and Penn 2013). Unlike Elo Scores, a popular alternative, Bradley-Terry scores do not require every pairwise matchup to determine a winner.

Do these perceptions require consistency between information types?

Another hypothesis derived from the snap-judgment model suggests that inconsistency in the yard sign’s visual information will lead to more mixed perceptions of the candidate’s political stances. The design of the latter trials in the study presents yard signs with mixtures of non-partisan and out-partisan colors. For example, primarily red and presumably Republican yard signs, but has some blue or white in them.

I take the same analytical strategy to address this question as I did before. I estimate the Bradley-Terry scores among respondents. I should expect that the trials that use less “consistent” visual information demonstrate more ambivalence among respondents in their reported perceptions of the politician’s positions. Specifically, I expect that participants will perceive the yard signs that have both red and blue on them as more moderate and that the higher proportion of the color red or blue among the two colors will lead respondents, on average, to be more likely to believe that the candidate leans more Republican or Democratic.

Do partisans process co-partisan branding faster?

The other analytical tasks do not examine expectations derived from the motivated reasoning portion of the model. That is, do people process in-group information faster than out-group information? There is evidence of these tendencies in political circumstances (Lodge and Taber 2013).

To examine whether motivated reasoning is also active with the processing of politically-relevant color, I examine the difference between the amount of time between the start of viewing a stimulus and clicking “Next” to stop viewing the stimulus among those who were viewing a presumed co-partisan yard sign relative to those viewing a presumed out-partisan yard sign. As motivated reasoning tends to be more prevalent among strong partisans and those highly knowledgeable and interested in politics (Lodge and Taber 2013), I control for responses to the pre-treatment knowledge, interest, and partisan strength batteries.

Checking the robustness of my analytical strategy with simulations

The [Pre-test](#) provides a lot of useful information that I might use to construct my expectations for what I’ll discover in [Study 1](#). Rather than evaluating my design

choices based on expected long-run tendencies, I explicitly use the results from the [Pre-test](#) and adjust that information with assumptions about differences between the [Pre-test](#) and [Study 1](#)'s design to construct expectations that directly feed into my analytical choices. That is, rather than evaluating my design in the Frequentist framework, I rely on results from the [Pre-test](#) to help with constructing my priors that I will use in a Bayesian model specification. Bayesian models often face criticism for the seemingly subjectivity introduced into a model specification by including one's priors into the calculation of one's estimates. While there has been a lot of ink spilt debating the validity of this claim, a Bayesian model specification seems particularly useful in the context of pre-registering a study. The [Pre-test](#) provides results which give me a benchmark by which I might refine my expectations about the effect of color on people's perceptions about politics; it gives me a concrete β coefficient. I can then use this information to construct my priors for my model. As I admit to a number of problems with my [Pre-test](#), my priors will reflect my uncertainty about whether those particular estimates should hold in [Study 1](#) and allows me to define an interval of effect sizes as opposed to a point estimate. These priors are updated by my likelihood function that is constructed by the data I collect in [Study 1](#). What this implies is that the more data I have, even if my priors are incorrect, I will converge upon the effect in the population (Johnson, Ott, and Dogucu 2021). However, since my priors require uncertainty, the weight of my priors on my eventual estimate are weaker the more uncertainty I specify. The lessons this provides for a pre-registered report suggests that I can use the results from my [Pre-test](#) to inform choices for [Study 1](#), but that, though I use that information to inform my choices for the pre-registered report, I explicitly define how I expect that the results for [Study 1](#) are different than those in [Pre-test](#) and that the data for [Study 1](#) directly update my expectations set by [Pre-test](#).

As my estimation strategy does not assume long-run convergence to the population, many assume that my Bayesian models are agnostic to the sample size and often eschew power analyses. However, as estimates from a Bayesian model are weighted by the observed data and my priors, I want to examine the robustness of my model specification using simulated data that vary on sample size. While I do expect that my priors will shift slightly depending on my pre-registered analyses here and upon exploratory data analysis once I collect my data, using simulated data to examine my analytical pipeline offers the opportunity to demonstrate my model specification's ability to converge upon parameters defined by a hypothetical data generating process.

Using the `fabricatr` (Blair et al. 2022) package, I simulate a population with an N of 1,00,000. The specified data generating processes for each of the variables are included in [Listing 1](#) and [Listing 2](#). I then generate 500 random samples for each sample size of 200, 400, 600, and 800 participants. This sample size is not the size of the total sample I intend to recruit but is the total sample that I have data on after excluding those who fail attention checks and those who provide duplicate responses. Some estimate that insincere responses account for about 40% of a researcher's original sample (Kennedy et al. 2021). Therefore the total

sample that I recruit from Prolific should be about 40% larger than the total sample size used in these simulations.

To examine my analytic strategy’s ability to converge upon the population’s data generating process for respondents’ perceptions of the fictional candidate’s partisanship, I specify a ordinal logistic regression where I set relatively constrained priors upon the beta coefficients, as depicted in Equation 3.

$$\begin{aligned}
\alpha &\sim Student - T(3, 0, 2.5) \\
\beta_i &\sim Normal(0, 1) \\
\hat{Party}_i &= Cumulative(logit^{-1}(\alpha_1 + \alpha_2 \\
&+ \beta_1 \times RedTreatment + \beta_2 \times BlueTreatment \\
&+ \beta_3 \times age + \beta_4 \times Redtreatment \times age + \beta_5 \times Bluetreatment \times age \\
&+ \beta_7 \times Attention + \beta_7 \times Knowledge))
\end{aligned} \tag{3}$$

These priors suggest that I believe 68% of my beta coefficients, given the data, should be between -1 and 1; with the median of the estimates at 0.

After fitting each model, I construct 95% high density interval credible intervals from the model’s posterior draws. I then record a value of 1 if the credible interval does not contain zero (true positive) and a value of 0 if the credible interval does (false negative) for each parameter. Once I have run each of my models for the specified sample size, I calculate the average of true positive and false negatives. This gives me a percentage of the time that I would come to the correct conclusion that there is a non-zero effect on that parameter. Table 3 documents these calculations for each sample size.

Table 3: True positive rate

Sample size	True positive rate
n = 200	<ul style="list-style-type: none"> • Red treatment: 0.348 • Blue treatment: 0.274 • Age: 0.732 • Red treatment \times Age: 0.694 • Blue treatment \times Age: 0.002 • Attention: 1.000 • Knowledge: 1.000

Sample size	True positive rate
n = 400	<ul style="list-style-type: none"> • Red treatment: 0.712 • Blue treatment: 0.718 • Age: 0.960 • Red treatment \times Age: 0.946 • Blue treatment \times Age: 0.064 • Attention: 1.000
n = 600	<ul style="list-style-type: none"> • Knowledge: 1.000 • Red treatment: 0.890 • Blue treatment: 0.892 • Age: 0.992 • Red treatment \times Age: 0.980 • Blue treatment \times Age: 0.320 • Attention: 1.000
n = 800	<ul style="list-style-type: none"> • Knowledge: 1.000 • Red treatment: 0.962 • Blue treatment: 0.956 • Age: 0.994 • Red treatment \times Age: 0.992 • Blue treatment \times Age: 0.620 • Attention: 1.000 • Knowledge: 1.000

The results of my simulation suggest that I should aim for a sample of 600 respondents once I have excluded those that do not meet the inclusion criteria (i.e., fail attention checks, duplicated IP addresses and provide insincere responses to the open-ended prompts). Given estimates of poor responses for about 40% of some samples (Kennedy et al. 2021), I should aim to recruit a total of 1500 participants for Study 1.

Study 2

[Study 1](#) presents evidence in support of the snap-judgment model in a relatively artificial way. The yard signs presented as stimuli are elementary in information content and design. Furthermore, viewing yard signs in real-world scenarios does not occur through a computer screen with an overlay allowing individuals to observe only a fraction of the screen simultaneously.

One way to examine the real-world effects of color on decisions about campaign branding is to go directly to the experts. If the effects of [Study 1](#) hold up in natural conditions, we should expect that campaigns make strategic decisions to nurture this behavior. Descriptive evidence suggests that female candidates in the United States make different branding decisions based on the constituency and whether their opponent is a man or a woman (Williams, Horsting, and Ramirez 2022). If the color of a yard sign does indeed matter and its electoral effects depend on the perceptions of potential voters, we should expect that campaigns are less likely to place yard signs with a large amount of Red or Blue in more moderate districts. However, in districts where there is a strong partisan base, we should expect that the candidate of the popular party will use the party's color more, while the disadvantaged candidate will use their party's color less as a way to either play up or down their loyalty to their party.

Are colors on yard signs a strategic choice by campaigns?

In this first part of [Study 2](#), I do a number of informal interviews with managers of national, state, and local campaigns. In these interviews, I simply ask the campaign managers: “what goes into your decision for the colors to put on your yard sign?” The goal of this open-ended question is to hear from the practitioners whether they perceive the yard sign color as a strategic decision for the campaign. I then follow-up to ask “do you think that the color choices you make for the yard signs matter to voters?” The goal of this question is to examine whether practitioners intuitively perceive their strategic choices to have efficacy in the election. I then ask them “if colors matter and you choose to use them strategically, do you use certain colors to attract certain groups of voters that may not be reliably in your base of support?” The goal of this question is to get at the argument that the colors that campaigns use are meant as a way to communicate the partisan and ideological alignment of the candidate with the voters. Finally, I get directly to the point “what information about the candidate are you trying to convey to potential voters when you choose the colors on your yard signs?”

While it is useful to hear from practitioners whether they strategically use colors for their yard signs and whether they believe that their strategic choices matter, there are many open questions. Firstly, the selection of whom to interview is dependent on convenience and the accessibility of the campaign manager. In other words, some campaign managers may be quite hard to contact due to sparse information about them or they may be hard to contact as a result of their schedules. Second, it is doubtful that campaign managers will openly admit that the choices they make are not strategic or that they think their strategic choices have little effects on the campaign's outcome.

To address these remaining questions, I employ a second part of [Study 2](#) which relies upon observational data and leverages congressional redistricting following the 2020 National Census to examine whether the partisan context of a district is associated with changes to color choices for campaigns in that same district. As this is likely a very endogenous process – where voters are influenced by the colors used and the voters influence the colors that are used – the primary goal of this second part of the study is to demonstrate that the parties are *responding* to their intuition that suggests that colors matter to the campaign; it is not to make any directional claim that partisan composition of a district *determines* the colors on a yard sign.

Do branding choices reflect electoral context?

To examine whether the use of colors on yard signs varies systematically depending on the electoral context, using the `selenium` ([Project 2022](#)) and `time` ([Foundation, n.d.](#)) `python` ([Foundation, n.d.](#)) packages, I collect images from 2018, 2020, and 2022 Congressional elections for the House of Representatives across the United States. These yard signs are pulled together on one website by the Center for American Politics and Design⁹. With these data, I detect the percentage of the “Republican Red” and “Democratic Blue” on the yard signs. I then use redistricting data from the [MIT election lab and Redistricting Data Hub](#) to determine districts that did and did not have changes in the 5-year rolling average partisan vote share but that are also nationally representative. As there are a number of external factors that may effect choice in yard sign color besides perceptions of how it may attract voters, this threatens my ability to make claims about any estimation of an Average Treatment Effect¹⁰. Instead these data are used to perform a synthetic control design to give me causal leverage over the role that electoral context has on the proportion of partisan colors (i.e., red and blue) on yard signs relative to non-partisan colors by estimating a Local Average Treatment Effect¹¹ on districts that were redistricted but vary in whether the redistricting leads to a partisan gerrymandered district or not. This analysis aims to bear evidence on the hypothesis that campaigns respond to the preferences of partisan voters and adjust their branding as a result. In this case, the branding is the color of the yard sign. First, I’ll explain how I collected data on the proportion of partisan and non-partisan colors on these yard signs from the [Center for American Politics and Design](#).

I collected the GOP logo used on their official Twitter account during the 2022 midterm election cycle to provide an example of how color detection works. With the `opencv` package ([Bradski 2000](#)), I load this image and convert it to a three-dimensional `numpy` ([Harris et al. 2020](#)) array that contains information about the GBR (reversed RGB) values for the pixels in that image. I resized the images to be a standardized 224 \times 224 pixels. I train the computer to detect a

9. See: <https://www.politicsanddesign.com/>

10. $\frac{1}{n} \sum i(Y_i(d=1) - Y_i(d=0))$ (see Lundberg, Johnson, and Stewart 2021)

11. $E[(Y_i^1 - Y_i^0)|d_i^1 - d_i^0 = 1]$ [see Cunningham (2021)]

range of GBR values that encompass the official “Republican Red”¹². For the broader exercise, I do it for the color white¹³ and “Democratic blue”¹⁴. Once this range of values is specified, the computer detects the pixels that do not contain values within this pre-specified range and converts those values to represent the color black. Figure 3 presents this process using the `opencv` (Bradski 2000) and `matplotlib` (Hunter 2007) packages.

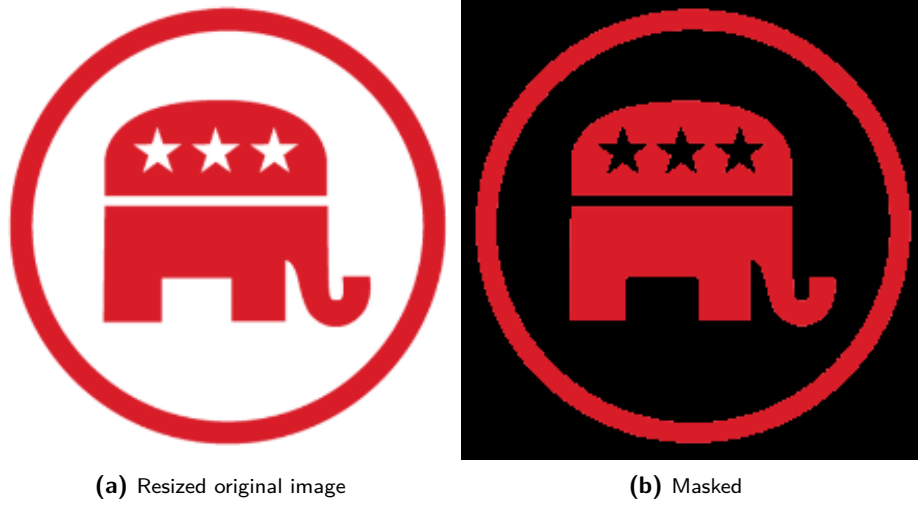


Figure 3: Detecting colors in the GOP logo

I then extract the values in the array that are non-black and calculate the percentage of non-black pixels (as depicted in Equation 4).

$$\text{Color}\% = \frac{\text{Non-black}}{\text{Transformed}} \times \frac{\text{Original}_{\text{Height}} + \text{Original}_{\text{Width}}}{2\text{Transformed}_{\text{Height}} + 2\text{Transformed}_{\text{Width}}} \quad (4)$$

For the example in Figure 3, about 32.26 of the image is red.

With these data, I estimate a mixed effects model with intercept random effects at the district level and intercept fixed effects at the election year level to examine the district and within election year differences between choices to employ more “partisan” colors on a yard sign based on historical data on the electoral performance of partisan and non-partisan candidates. That is, I examine whether there are differences between how campaigns respond to redistricting that make the electorate for that constituency more partisan or less partisan while controlling for a number of election-year and campaign specific factors. Support for H_5 would suggest that campaigns in districts that are more partisan

12. lower values: (93, 9, 12), higher values: (236, 69, 75)

13. upper and lower values: (255, 255, 255)

14. lower values: (0, 18, 26), higher values: (102, 212, 255)

will shift to use more traditionally-partisan colors while campaigns in districts that do not shift to have a more historically partisan constituency will use fewer of those traditionally-partisan colors.

Discussion

Not only do yard signs matter for electoral strategy, outcomes, and voter attitudes and perceptions, but I suspect that the yard signs' color partially explains these phenomena.

While yard signs are a simple form of party branding, richer theorizing about how diverse types of potentially politically-relevant information allows us to fill some of these gaps. In the case of yard signs, while some may be skeptical of whether yard signs and campaigns, more generally, matter for electoral outcomes, an application of the snap-judgment model on yard signs suggests that they indeed have potential and provide an explanation as to how it occurs.

Outside of color on yard signs, I suspect that color provides valuable information that is processed efficiently by the public. Due to this, voters can connect colors like Red and Blue to make assumptions about a political candidate's partisan affiliation. With this information, voters can then evaluate the candidate. This is a much simpler task than hearing the candidate's positions on immigration, for example, and then making a connection to their partisan affiliation.

References

- Allaire, J.J., Charles Teague, Carlos Scheidegger, and Yihui Xie. 2022. *Quarto*. <https://doi.org/10.5281/zenodo.5960048>. <https://github.com/quarto-dev/quarto-cli>.
- Anwyl-Irvine, Alexander L., Thomas Armstrong, and Edwin S. Dalmaijer. 2022. “MouseView.js: Reliable and valid attention tracking in web-based experiments using a cursor-directed aperture.” *Behavior Research Methods* 54:1663–1687. <https://doi.org/10.3758/s13428-021-01703-5>.
- Arel-Bundock, Vincent. 2022a. *marginaleffects: Marginal Effects, Marginal Means, Predictions, and Contrasts*. R package version 0.8.1. <https://CRAN.R-project.org/package=marginaleffects>.
- . 2022b. “modelsummary: Data and Model Summaries in R.” *Journal of Statistical Software* 103 (1): 1–23. <https://doi.org/10.18637/jss.v103.i01>.
- Bansak, Kirk, Jens Hainmueller, Daniel J. Hopkins, and Teppei Yamamoto. 2021. “Conjoint Survey Experiments.” In *Advances in Experimental Political Science*, 19–41. New York: Cambridge University Press. <https://doi.org/10.1017/9781108777919.004>.
- Blair, Graeme, Jasper Cooper, Alexander Coppock, Macartan Humphreys, Aaron Rudkin, and Neal Fultz. 2022. *fabricatr: Imagine Your Data Before You Collect It*. R package version 1.0.0. <https://CRAN.R-project.org/package=fabricatr>.
- Bradski, G. 2000. “The OpenCV Library.” *Dr. Dobb’s Journal of Software Tools*.
- Broockman, David E., and Joshua L. Kalla. 2022. “When and Why are Campaigns’ Persuasive Effects Small? Evidence from the 2020 U.S. Presidential Election?” *American Journal of Political Science*, <https://doi.org/10.1111/ajps.12724>.
- Bürkner, Paul-Christian. 2017. *brms: An R Package for Bayesian Multilevel Models Using Stan*. 1. <https://doi.org/10.18637/jss.v080.i01>.
- Campbell, Angus, Phillip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. Chicago: John Wiley / Sons, Inc.
- Cunningham, Scott. 2021. *Causal Inference: The Mixtape*. Yale University Press. <https://doi.org/10.2307/j.ctv1c29t27>.
- Delli Carpini, Michael X, and Scott Keeter. 1996. *What Americans Know About Politics and Why It Matters*. New Haven, CT: Yale University Press.
- Fenno, Richard F. 1973. *Congressmen in Committees*. Little Brown.
- Foundation, Python Software. n.d. *Python Language Reference, version 3.11*. <http://www.python.org>.
- . n.d. *time: Time access and conversions*. <https://docs.python.org/3/library/time.html>.
- Green, Donald P., Jonathan S. Krasno, Alexander Coppock, Benjamin D. Farrer, Brandon Lenoir, and Joshua N. Zingher. 2016. “The effects of lawn signs on vote outcomes: Results from four randomized field experiments.” *Electoral Studies* 41:143–150. <https://doi.org/10.1016/j.electstud.2015.12.002>.
- Harris, Charles R., K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, et al. 2020. “Array programming with NumPy.” *Nature* 585 (7825): 357–362. <https://doi.org/10.1038/s41586-020-2649-2>. <https://doi.org/10.1038/s41586-020-2649-2>.
- Hopkins, Daniel J., and Hans Noel. 2022. “Trump and the Shifting Meaning of “Conservative”: Using Activists’ Pairwise Comparisons to Measure Politicians’ Perceived Ideologies.” *American Journal of Political Science* 116 (3): 1133–1140. <https://doi.org/10.1017/S0003055421001416>.
- Huddy, Leonie, and Nadia Khatib. 2007. “American Patriotism, National Identity and Political Involvement.” *American Journal of Political Science* 51 (1): 63–77. <https://doi.org/10.1111/j.1540-5907.2007.00237.x>.
- Hunter, J. D. 2007. “Matplotlib: A 2D graphics environment.” *Computing in Science & Engineering* 9 (3): 90–95. <https://doi.org/10.1109/MCSE.2007.55>.
- Johnson, Alicia A., Miles Q. Ott, and Mine Dogucu. 2021. *Bayes Rules! An Introduction to Applied Bayesian Modeling*. CRC Press. <https://www.bayesrulesbook.com/>.
- Kennedy, Courtney, Nicholas Hatley, Arnold Lau, Andrew Mercer, Scott Keeter, Joshua Ferno, and Dorene Asare-Marfo. 2021. “Strategies for Detecting Insincere Respondents in Online Polling.” *Public Opinion Quarterly* 85 (4): 1050–1075. <https://doi.org/10.1094/poq/nfab057>.

- Leeuw, Josh de. 2015. "jsPsych: A JavaScript library for creating behavioral experiments in a web browser." *Behavior Research Methods* 47 (1): 1–12. <https://doi.org/10.3758/s13428-014-0458-y>.
- Lodge, Milton, and Charles S. Taber. 2013. *The Rationalizing Voter*. New York: Cambridge University Press.
- Lundberg, Ian, Rebecca Johnson, and Brandon M. Stewart. 2021. "What Is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory." *American Sociological Review* 86 (3): 532–565. <https://doi.org/10.1177/00031224211004187>.
- Lupton, Robert N., Steven M. Smallpage, and Adam M. Enders. 2020. "Values and Political Predispositions in the Age of Polarization: Examining the Relationship between Partisanship and Ideology in the United States, 1988-2012." *British Journal of Political Science* 50 (1): 241–260. <https://doi.org/10.1017/S0007123417000370>.
- Maestre, Roberto Losada, and Rubén Sánchez Medero. 2022. "Color War. Does Color Influence the Perception of Political Messages?" *Psychological Reports*, <https://doi.org/10.1177/00332941221114418>.
- Makse, Todd, Scott L. Minkoff, and Anand E. Sokhey. 2019. *Politics on Display: yard signs and the politicization of social spaces*. New York: Oxford University Press.
- Mason, Lilliana. 2018. *Uncivil Agreement*. Chicago, IL: University of Chicago Press.
- Mayhew, David R. 1974. *Congress: The Electoral Connection*. Second Edi. New Haven, CT: Yale University Press.
- McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Second edition. Boca Raton, FL: Chapman & Hall/CRC.
- Mills, Mark, Frank J. Gonzalez, Karl Giuseffi, Benjamin Sievert, Kevin B. Smith, John R. Hibbing, and Michael D. Dodd. 2016. "Political conservatism predicts asymmetries in emotional scene memory." *Behavioural Brain Research* 306:84–90. <https://doi.org/10.1016/j.bbr.2016.03.025>.
- Montgomery, Jacob M., Brendan Nyhan, and Michelle Torres. 2018. "How Conditioning on Posttreatment Variables Can Ruin Your Experiment and What to Do about It." *American Journal of Political Science* 63 (3): 760–775. <https://doi.org/10.1111/ajps.12357>.
- Pérez, Efrén O., Maggie Deichert, and Andrew M. Engelhardt. 2019. "E Pluribus Unum? How Ethnic and National Identity Motivate Individual Reactions to a Political Ideal." *The Journal of Politics* 81 (4): 1420–1433. <https://doi.org/10.1086/704596>.
- Project, The Selenium. 2022. *Selenium WebDriver*. <https://www.selenium.dev/documentation/webdriver/>.
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Schnakenberg, Keith E., and Elizabeth Maggie Penn. 2013. "Scoring from Contests." *Political Analysis*, <https://doi.org/10.1093/pan/mpt018>.
- Wedel, Michel, and Rik Pieters. 2008. "A Review of Eye-Tracking Research in Marketing." In *Review of Marketing Research*, edited by Naresh K. Malhotra. Armonk, NY: M.E. Sharpe, Inc.
- Williams, Alexandra M., Trudy Horsting, and Mark D. Ramirez. 2022. "What's in a Campaign Logo? Exploring Differences in Candidate Self-Presentation through Campaign Logos." *Journal of Political Marketing*, <https://doi.org/10.1080/15377857.2022.2040691>.

Listing 1 Code to generate simulated population data

```
dgp <- fabricate(  
  N = 100000, # N in the population  
  E = rnorm(N), # epsilon term  
  age = round( # define age variable  
    runif(N, 18, 85)  
  ),  
  gender = draw_binary( # define binary gender identity variable  
    N,  
    prob = 0.5  
  ),  
  white = draw_binary( # define white indicator variable  
    N,  
    prob = 0.6  
  ),  
  PartyId = draw_ordered( # define party identity variable as 3-item ordinal  
    x = rnorm(  
      N,  
      mean = 0.4 * age - 0.6 * gender + 0.7 * white + E  
    ),  
    breaks = c(  
      -Inf, 20.14, 23.01, Inf  
    )  
  ),  
  Attention = draw_ordered( # define Attention variable as 5-item ordinal  
    x = rnorm(  
      N,  
      mean = 0.5 * age - 0.3 * gender + 0.1 * white + E  
    ),  
    breaks = c(  
      -Inf, 16.5, 28.26, 36.54, 43.82, Inf  
    )  
  ),  
  Knowledge = rnorm( # define Knowledge variable  
    N,  
    mean = 0.6 * age - 0.5 * gender + 0.2 * white + 0.8 * Attention + E  
  )/100,  
)
```

...

Listing 2 Code to generate simulated population data (continued)

```
...
  RedTreatment = draw_binary( # Simulate treatment assignment with prob 1/3
    N,
    prob = 1/3
  ),
  BlueTreatment = draw_binary( # Simulate treatment assignment with prob 1/3
    N,
    prob = 1/3
  ),
  PartyGuess = draw_ordered( # Define PartyGuess outcome variable as 3-item ordinal
    x = rnorm(
      N,
      mean = (
        2 * RedTreatment + -2 * BlueTreatment - 0.01 * age
        - 0.1 * RedTreatment * age + 0.1 * BlueTreatment * age
        + 0.1 * Attention + 0.1 * Knowledge + E
      )
    ),
    breaks = c(
      -Inf, -0.5, 0, Inf
    )
  ),
  PartyGuessTrialTwo = draw_ordered( # Define PartyGuess outcome variable as 3-item cat
    x = rnorm(
      N,
      mean = (
        2 * RedTreatment + -2 * BlueTreatment - 0.01 * age
        - 0.1 * RedTreatment * age + 0.1 * BlueTreatment * age
        + 0.1 * Attention + 0.1 * Knowledge + E
      )
    ),
    breaks = c(
      -Inf, -0.5, 0, Inf
    )
  ),
  Vote = draw_ordered( # Define Vote outcome variable as indicator variable
    x = rnorm(
      N,
      mean = (
        1 * RedTreatment + - 1 * BlueTreatment - 0.01 * age
        + 0.1 * Attention + 0.1 * Knowledge + E
      )
    ),
    breaks = c(
      -Inf, 0, Inf
    )
  )
)
```