

Civic Education and Public Confidence in U.S. Elections: A Field Experiment in Contra Costa County*

Susan D. Hyde[†] Jennie Barker[‡] Oren Samet[§]

October 17, 2026

Abstract

Public confidence in American election administration has fallen sharply, and a substantial share of the electorate doubts that votes are counted accurately. We propose a randomized field experiment evaluating a civic education program built from condensed videos of the Contra Costa County Clerk-Recorder-Elections Department's Certified Election Observer program. The program and the videos are intended to increase transparency and therefore increase public confidence in elections by showing residents how their elections are actually administered. To study the effects of this program, residents recruited by the Elections Office will attend a three-hour session at a community venue, watch the video series as a group, and discuss it with a facilitator. Given the rolling recruitment throughout different locations in the county and the current absence of an acceptable placebo, every participant will eventually receive the program. Identification comes from randomizing, within each session, which participants complete the survey instrument before the program begins. Participants measured at the start are untreated at the moment of measurement, participants measured at the end are treated with the program, and random assignment of survey timing allows us to estimate the effect of the program. We will also administer the post-test survey to those who were pretested, which allows us in the context of this pilot to measure how much administering a pretest distorts the estimated effect. The 2026 study is designed as a pilot that will measure the enrollment, attendance and effect-size parameters needed to power a larger roll-out of this program and study. If time allows the research team will separately field a study with the same design, but in which the videos are watched individually.

*Pre-analysis plan.

[†]Robson Professor, Travers Department of Political Science, University of California, Berkeley. Email: susanhyde@berkeley.edu

[‡]Johns Hopkins University. Email: jbarke21@jh.edu

[§]Rice University. Email: os36@rice.edu

1 INTRODUCTION

Public distrust in the electoral process in the United States has reached potentially unprecedented levels. Despite the absence of significant and consequential electoral irregularities, a substantial portion of the electorate harbors doubts about the integrity of elections, driven in part by elite claims of malfeasance and the spread of misinformation.¹ These trends pose a series of challenges to election administration—a fact recognized by local election officials throughout the country. They also pose a serious challenge to democratic governance. The willingness of citizens to accept election outcomes, even when voters’ preferred candidates lose, is a critical element of self-enforcing democracy (Przeworski 2005; Fearon 2011; Weingast 2004). The “loser’s consent” is also theorized as a critical component of peaceful elections.

In prior work, we have examined whether nonpartisan election observation can bolster public confidence in U.S. elections. In a series of survey experiments we found that information about the presence of nonpartisan domestic election observers increased respondents’ trust in hypothetical election results (Barker, Samet, and Hyde 2024). That work was conducted in a survey setting, using vignettes and measuring respondents’ stated attitudes. But what about real-world exposure to election administrators and information about how the process actually works? Does structured factual information about how elections are administered and how votes are counted change what residents believe, and what they are willing to do, in the context of a real election?

Considerable effort and money now go into restoring public confidence in elections, and little of it has been systematically evaluated.² Practitioners are often working without evidence about what they can do to increase trust in processes that merit it, and election offices throughout the United States have expressed interest in what they can do to make their process more transparent and better understood among the voting public. This study proposes an evaluation of a condensed version of a real-world program implemented by election officials in Contra Costa County, California, aimed at bolstering trust in the local election process. The Certified Election Observer (CEO) program is typically run over many weeks, with participants spending a total of 12-14 hours in the program, mostly during weekdays. This is a level of commitment that most citizens cannot afford to make, so the Contra Costa County Elections Department produced videos drawn from their more extensive CEO program. This research aims to study the pilot of the more condensed version of the CEO program. In multiple sites throughout the county, the extended pilot program will show the training videos in a three-hour group-based setting held in community venues. This condensed program also marks an inexpensive way to scale up the existing CEO program, which is constrained by space within the Elections Division building in Contra Costa County. Establishing whether it works is valuable for the county and for other entities interested in increasing trust in elections more generally.

1. See, for example, Kousser, Thad, Lauren Prather, Jack Stewart, Seth Hill, Laura Uribe, Mindy Romero, Jennifer Merolla, Jennifer Gaudette, Cheryl Boudreau, Mackenzie Lockhart, and Theodoros Ntounias. 2026. “Fewer than Half of Americans Expect Free and Fair Midterms.” Center for Transparent and Trusted Elections, September. https://elections.ucsd.edu/_files/surveys/Trust-in-Elections-Sept-2026.pdf

2. Though see Brown et al. (2024), Gaudette et al. (2025), Suttman-Lea and Merivaki (2023), and Kousser and Prather (2024)

2 THE CEO PROGRAM

Contra Costa County (CCC) has, since 2022, run a Certified Election Observer (CEO) Program offering residents the opportunity to learn about and examine the steps the county takes to ensure the transparency and integrity of the electoral process. It reflects what CCC Elections Division would like citizens to know before attempting to observe elections. Despite the name, the program is not itself nonpartisan election observation; it is closer to a civic education program about local election administration aimed at making the elections process very transparent. It has been popular and has attracted participants from across the political spectrum. The version of the program held at the Elections Department is live, not video-based, and is time intensive: it covers 12-14 hours across repeated weekday visits, including direct observation of processes such as machine testing and verification and attending a poll worker training. Because it is demanding of time and because space inside the Elections Department is constrained, in practice its participants have been mainly limited to individuals who take significant time off during the week, and has had approximately 10-20 participants per election cycle.

To extend that reach, the county produced a video series with Contra Costa Television (CCTV) covering voter registration and verification, voting equipment testing and certification, mail ballot receipt and processing, vote counting procedures, and the role of nonpartisan election observers. The series also shows Contra Costa residents participating in the CEO program during the 2022 election cycle. Together the videos run approximately ninety minutes. The series was produced at public expense and is not publicly distributed. The county controls access to the videos, so there is not a concern that anyone has already watched them.

The program we evaluate is a three-hour session in which a group of residents watches that video series together and discusses it with a facilitator. The shorter program is designed to be distributed throughout the county. Rather than requiring residents to travel repeatedly to the Elections Department, it can be delivered in one sitting at a community venue anywhere in the county. Because participants often have many questions, pairing the videos with group discussion is the best condensed approximation of the current program. The longer program and the shorter videos of how elections are run at the county level could potentially serve as models for other jurisdictions.

2.1 Legal Context

California conducts elections under a transparent framework that provides for a relatively higher degree of voter access to observe the electoral process. Under the state’s universal vote-by-mail system, all registered voters receive a mail ballot, and counties may begin processing vote-by-mail ballot return envelopes as early as 29 days before the election (Cal. Elec. Code §15101). Processing includes signature verification, opening envelopes, duplicating damaged ballots, and preparing ballots for machine tabulation, though no vote count may be accessed or released until 8 p.m. on election day (§15101(b)(2)). Critically, all vote-by-mail ballot processing is required by law to be open to public observation, both before and after the election (§15104). Election officials must provide at least 48 hours advance notice of processing dates, times, and locations. Observers are entitled to sufficiently close

access to observe return envelopes and signatures (§15104). Signature verification follows a presumption of validity, and rejection of a voter signature requires independent review by multiple officials (§3019).

Though all voters in California can vote-by-mail, voters can also choose to vote in-person, either before election day at early voting centers (in counties that participate in the Voter’s Choice Act, of which Contra Costa is one) or in polling sites on election day. Observers are also entitled to observe proceedings at polling sites and early voting centers.

Thus, the election process in California is designed to be observable. However, most voters have never seen it up close, many do not know the legal provision for citizen observation exists, and would not necessarily know what to look for or how elections are supposed to be run.

2.2 Partnership

This study is conducted in partnership with the Contra Costa County Clerk-Recorder-Elections Department, which produced the video materials and runs the underlying program. The county leads recruitment of participants and delivery of the program. The research team designs the measurement and conducts the analysis.

3 RESEARCH QUESTIONS

The study seeks to answer a series of questions about the effectiveness of the condensed video-based version of the county’s CEO program:

1. Does the video-based program increase residents’ factual knowledge about how elections are conducted?
2. Does the video-based program increase how well residents feel they understand the electoral process, and their comfort explaining it to someone else?
3. Does the video-based program increase trust in the electoral process?
4. Does the video-based program increase willingness to sign up as a nonpartisan election observer?
5. Does the video-based program make participants more likely to pass information about the program on to others?
6. Are these effects moderated by partisan affiliation or prior political engagement (if measurable)?
7. How much does administering a baseline survey itself alter the measured effect?

4 HYPOTHESES

Hypothesis 1 *Relative to participants measured before the program, the program should increase participants' factual knowledge about election administration processes, including voter registration, ballot processing, voting equipment testing and vote counting.*

Hypothesis 2 *Relative to participants measured before the program, the program should increase how well participants feel they understand the electoral process in Contra Costa County, and their comfort explaining it to someone else.*

Hypothesis 3 *Relative to participants measured before the program, the program should increase trust in the electoral process.*

Hypothesis 4 *Relative to participants measured before the program, the program should make participants more likely to sign up as a nonpartisan election observer.*

Hypothesis 5 *The program should make participants more likely to send information about the program to someone else, relative to participants asked before the program.*

Hypothesis 6 *Effects vary by party registration, which is recorded in the voter file and can be matched to the confidential responses using a merge key. Based on our prior findings we expect effects among registrants of both parties, and larger effects in the party whose members have reported less confidence in election administration in recent years.*

Hypothesis 7 *If there is pretest sensitization, participants who completed a pretest are more likely to report different post-session outcomes than participants who only completed the post-test.*

5 RESEARCH DESIGN

5.1 Overview

We will conduct a randomized field experiment in Contra Costa County, California, in the weeks before the November 2026 election. The 2026 study is a pilot. Its primary purpose is to better understand session attendance and completion, outcome variance in this population, within-session attrition, and the magnitude of pretest sensitization.

5.2 Setting and Recruitment

Sessions are held at community venues distributed across the county. The Elections Office will lead the recruitment process which will be open to anyone interested in learning more about elections in CCC. The county will publicize the sessions through its own channels and will route interested individuals to a public registration page where any county resident can sign up for the location they prefer. There is no eligibility screen and no invitation list. The research team will not select participants. Assignment is blocked on pairs of consecutive arrivals within each session. After the sessions are held we will match participants to the voter register to obtain party registration and the number of the last four elections in which each participant voted. These are pre-treatment characteristics whatever point we observe them, so they are legitimate covariates, but the match will not succeed for every participant and we will not have them at the moment of randomization. We therefore use them in the heterogeneity analysis and report the match rate, and we do not rely on them in the primary specification.

The sample is defined by individuals who show up for the training program. The treated and untreated comparisons are randomized within each session, and the study is intended to apply to the universe of willing participants rather than the entire voting public in the county. Host venues are chosen to span the county’s political and demographic geography, and estimates will account for host-site fixed effects.

Recruitment will be rolling. Sessions are scheduled and filled sequentially over the weeks before the election, and randomization happens within each session as it is held. Adding a session later therefore does not disturb earlier assignments or require re-randomization, which is convenient for a rolling recruitment process aimed at bringing sessions closer to where voters live. Sessions will run at a mix of evening and weekend times at various locations, unlike the existing weekday program, which takes place at the Elections Office headquarters in Martinez, California.

Sessions can arise in two ways. County-recruited sessions are held at venues the county and the research team select and are filled through open registration. In recruitment, it will also be clear that groups can request a session. Group-organized sessions are those in which an existing group, such as a congregation, community organization, service club, union, student group, or senior center, asks the county to deliver a session and provides both the participants and (ideally) the room. Recruitment route will be recorded as a session characteristic. We report effects by route as a pre-specified subgroup analysis (though we are not guaranteed to have any sessions requested by existing groups).

Under this recruitment design attrition and non-compliance should be minimized. We will record: recruitment volume by channel of recruitment; registrations by channel, from a required item at registration asking how the registrant heard about the session; and the registration-to-attendance rate. Group-organized sessions will be excluded from the channel-level rates, since their participants arrive through no county outreach, but their attendance rate will be recorded.

All sessions will be held before the November 2026 election. Recruitment has no stopping rule based on enrollment or on interim results. Sessions are scheduled with the county as venues and staff allow, and the election date fixes the end of enrollment. We report the number of sessions held and realized attendance at each. No interim look at outcomes will

determine whether further sessions are added.

5.3 Intervention

Each session runs approximately three hours and consists of the full CEO video series, viewed as a group, interspersed with facilitated discussion and questions. We planned sessions of between 15 and 30 participants, with 20 as the target. Registration for the four scheduled sessions has run below that, between 3 and 13. Each session is supplied with twenty-four assignment cards, which caps attendance at twenty-four. We report realized attendance for each session. Facilitation will be provided by county staff or by a facilitator following a common protocol. The trainings are designed to be nonpartisan and focused on factual information about the electoral process. The video content will be identical across sessions, providing a consistent directed treatment; questions and group discussion create some naturalistic variation.

Every participant receives the program. We randomize the timing of a survey, not access to the intervention, so the 2026 pilot has no untreated control group. Sessions have a fixed capacity, so registrations beyond it will be waitlisted and offered a later session.

For each session we will record the affiliation of the facilitator, session composition, and attendance. Facilitator affiliation will be recorded because it is plausibly part of the treatment: an official county presenter and a community organization carry different credibility with different participants. That variation will be common to everyone within a session and therefore cannot bias the within-session contrast, but it is informative in its own right.

5.4 Random Assignment

On arrival, before any program content is delivered, each attendee will be randomly assigned to one of two arms:

1. **Arm A** completes the survey instrument at the start of the session and again at its end.
2. **Arm B** completes the instrument at the end of the session only.

We assign each individual to either Arm A or Arm B with pre-printed cards. Each session is supplied with twenty-four cards in twelve pairs, one arm A card and one arm B card in each pair, held in label order. Staff take the next pair, and give one to each of the next two people to arrive without knowledge of which is Card A or Card B, affixed to the back of their name badge (which the participants write their name on themselves). Each card carries a participant identifier and a QR code that opens the survey with the identifier and the arm already set. Staff record which card went to which participant on a roster held by the research team.

In the pilot there are three quantities of interest:

1. **The program's effect**, which will be estimated by comparing Arm A pre-tests with Arm B post-tests.

2. **The pretest sensitization effect** will be estimated by comparing Arm A’s post-tests with Arm B’s post-tests. This is an estimate of whether completing the pre-test survey influenced post-test responses.
3. **Within-participant change**, for Arm A, which is the naive pre-post estimate. The comparison of the program’s effect (1) to the within-participant change (3) will also be helpful in future designs.

The most obvious alternative design is a placebo-control in which participants are randomly assigned between the groups. Randomly selected respondents would instead participate in a second arm attending an equivalent session on non-election civic content, which would separate the effect of the election-specific material from the effect of attending a civic session at all. We have not yet identified suitable placebo content and have concerns about challenges in recruitment. For reasons related to power in the smaller pilot with recruitment concerns, we elected an individual level randomization on arrival within the group.

Randomizing survey timing means the relevant comparison is within a single session, so session fixed effects can be estimated, and we can add sessions as feasible if requested by community groups. Arm A’s post-test will not be used in the primary specifications, which will compare Arm A’s pre-tests and Arm B’s post-tests. Arm A’s post-tests will be used to identify pretest sensitization, and supply a within-participant (non-experimental) naive pre-post estimate. The instrument is short enough to complete twice, about ten minutes each time.

Pretest Sensitization. The concern motivating this structure is that answering questions about ballot verification and trust in election officials immediately before a program on those topics may itself alter what participants take from the program or report afterward (Solomon 1949). Evidence on whether assessment reactivity of this kind meaningfully biases behavioral studies is mixed (McCambridge et al. 2011). Because pretest administration will be randomized across arms, sensitization does not threaten the internal validity of the primary estimate. However, it is relevant to external validity, and to the interpretation of studies in this area that pretest their subjects.

Primary Test and Identifying Assumption Using Arm A’s start-of-session responses as the untreated counterfactual for Arm B’s end-of-session responses assumes that nothing other than the program moved the outcome across the intervening interval. Because that interval is a few hours within a single sitting, we judge this to be a weak assumption and a tradeoff worth making in the context of a pilot.

5.5 Data Collection

All measurement occurs at the event. Respondents take the survey on their own devices by scanning the QR code on their card, and we provide a small number of loaner devices. Some evidence suggests that respondents are more comfortable responding on their own devices.

The instrument is delivered as two linked surveys in Qualtrics. The first asks whether the respondent is arriving for the session or finishing it, then collects a name and email address

so that a respondent’s two measurements can be linked. In the analysis, any identifying information will be kept separate from analysis data, per the procedure defined in the IRB approval. Arm B respondents who are arriving register for the event and are thanked and sent no further questions. Everyone else continues to the second survey, which carries the consent question and the full set of outcome items and takes about ten minutes.

Consent is requested once. Arm A consents at the start of the session and Arm B at the end, and Arm A is not asked again at the end.

The arm comes from the card. Whether a response is a pre-test or a post-test comes from the participant’s answer to the first question, which asks whether they are arriving or finishing. Every response carries a submission time, and sessions have a fixed start and end, so we check the reported phase against the time of submission. If the two disagree we use the submission time, and we report how often that occurs. A participant who does not scan at the end leaves no post-test. In Arm A that participant still contributes a pre-test to the primary contrast, and in Arm B that participant contributes nothing to it.

If the identifier on the card does not pass through to the survey, the respondent can type the identifier printed beside the QR code, and staff can open the link from the roster. Ideally the surveys are administered by a researcher who is not also the facilitator.

6 OUTCOME MEASURES

6.1 Primary Outcomes

1. **Factual knowledge about election administration.** An index based on factual questions about how elections are conducted in Contra Costa County and California, for example how mail ballots are verified and what role nonpartisan observers play, constructed as the proportion of correct answers. The instrument includes three such questions.
2. **Self-assessed understanding of the electoral process.** Two items on a 0–10 scale: how well the respondent feels they understand the electoral process, asked separately for Contra Costa County, California, and the United States; and how comfortable the respondent would be explaining to another person how elections work in Contra Costa County. We treat the Contra Costa items as primary.
3. **Trust in the electoral process.** A composite of the following, each on a 0–10 scale:
 - Confidence in the accuracy and integrity of election results in Contra Costa County
 - Confidence in the election system and processes in Contra Costa County
 - Confidence that the respondent’s own ballot is properly counted and documented

The first two items are also asked about California and about the United States. We report those separately as a check on whether any movement is specific to local administration.

4. **Sign-up as a nonpartisan election observer.** The instrument records two things. Respondents are asked whether they would like to receive more information about serving as an election observer, which we treat as a survey response. Those who say yes are offered a link to a sign-up page on the training website where they enter an email address, and the survey records whether that link was opened and when. The sign-up page is outside the survey instrument, and the consent materials state that the email address is collected there.
5. **Referring someone else to the program.** The instrument asks whether the participant would be interested in sharing a link to the training with family and friends. Those who say yes are shown a share page offering three routes: copying the link, opening a pre-filled email, or invoking the device’s native share sheet. Sharing is participant-initiated. The study does not send messages on a participant’s behalf and does not collect the contact details of anyone they share with.

For each participant we record whether the share control was engaged, how many times it was engaged, which routes were used, the time of first use, and whether a native share was reported complete. These are the referral measures used in analysis. They record what the participant did with the link, and not whether anyone received it.

The shared link carries a code identifying the participant’s arm and measurement phase, not the participant. Landings on the registration page are logged with that code, so downstream traffic and any resulting registrations are attributable to a randomization group, not to an individual referrer. We therefore treat group-level registration counts as a secondary, descriptive quantity and do not link referral to registration at the individual level. The registration page’s “how did you hear about this” item includes a personal-referral option, which gives a self-reported second route to identifying referred registrants when the coded link is not used. This is closer to a behavioral outcome, like the observer sign-up.

6.2 Covariates

- Party registration
- Voting history
- Demographic characteristics: age, gender, education, race and ethnicity
- Session characteristics: facilitator affiliation, venue, session size, and recruitment route (county-recruited or group-organized)

6.3 Exploratory Measures

The instrument includes one further item, asked after the observer and referral questions, on whether the respondent would be interested in being paired with someone who holds different political beliefs when observing. We report it as exploratory.

7 ANALYSIS

Our analysis follows the recommendations of Gerber and Green (2012) for the analysis of field experiments and adopts the covariate adjustment procedure proposed by Lin (2013).

7.1 Estimand

Our estimand is the average treatment effect of the program among attendees surveyed before the session compared to those only surveyed after the session. Because we randomize among individuals who have already shown up and who are willing to participate in the study, compliance should be close to 100%.

The study population (to which the estimate generalizes) is residents who volunteer for a three-hour civic education session and who attend. It is not intended to be the full population of registered voters, but rather those who respond to a normal recruitment process run by the county. Because recruitment is open and enrollment is self-selected, this population is likely to be more politically engaged than the county electorate at large, though volunteers could be attracted to the sessions because they have doubts about the electoral process or because they are already very trusting. Characterizing this population and understanding successful recruitment strategies is one of the goals of the pilot. We will also estimate the pretest sensitization effect, which is identified directly because pretest administration is randomized.

7.2 Primary specification

For each primary outcome we estimate the treatment effect by comparing Arm B’s post-tests to Arm A’s pre-tests, using the Lin (2013) estimator with pre-treatment covariates mean-centered and interacted with the treatment indicator, and HC2 robust standard errors (Lin 2013; Aronow and Samii 2017):

$$Y_i = \beta_0 + \beta_1 Z_i + \tilde{\mathbf{X}}_i' \boldsymbol{\gamma} + Z_i \tilde{\mathbf{X}}_i' \boldsymbol{\delta} + \epsilon_i \quad (1)$$

where Z_i indicates the treated measurement, Arm B’s post-test, and the untreated measurement is Arm A’s pre-test, and $\tilde{\mathbf{X}}_i$ is the vector of mean-centered pre-treatment covariates. We include session fixed effects throughout.

Assignment is blocked on pairs of consecutive arrivals within each session. We pre-register the specification above and report a version absorbing pair fixed effects alongside it.

Pre-registered covariates are those available for all participants regardless of arm from the voter register or from the survey: party registration, vote propensity from the voter file, age, education, and session block indicators.

Arm A’s post-tests are excluded from this estimate. Pooling both arms’ post-tests would raise the observation count but contaminate the treated mean, since Arm A’s post-tests are primed by construction, which makes them informative about sensitization and unsuitable for the primary contrast. We pre-register the unpooled estimator and report the pooled version only as a diagnostic for future research.

As a robustness check we report the unadjusted difference in means with randomization-inference p -values.

7.3 Sensitization

The sensitization effect is estimated by comparing Arm A’s post-tests to Arm B’s post-tests under the same specification. We additionally report the naive pre-post estimate computed from Arm A alone, which is in addition to the experimental estimate, to quantify how far a conventional one-group pre-post design would have departed from the randomized result. This comparison within the pilot will be useful for designing the larger study.

7.4 Attendance and dose

Attendance and the portion of the session each participant was present for will be recorded. Those who leave early will still be asked to complete the post-test.

7.5 Heterogeneous effects

We estimate heterogeneous effects by including interaction terms and by estimating separate Lin regressions within subgroups defined by pre-specified moderators:

$$Y_i = \beta_0 + \beta_1 Z_i + \beta_2 M_i + \beta_3 (Z_i \times M_i) + \tilde{\mathbf{X}}_i' \boldsymbol{\gamma} + Z_i \tilde{\mathbf{X}}_i' \boldsymbol{\delta} + \epsilon_i \quad (2)$$

in order of priority: party registration, vote propensity, education, and session characteristics. Following Gerber and Green (2012, ch. 9), subgroup analyses are inherently exploratory and will be flagged as such. We do not adjust p -values for heterogeneity tests but distinguish confirmatory from exploratory analyses clearly.

7.6 Binary outcomes, multiple comparisons, attrition

For binary outcomes we use the same Lin specification, which provides consistent estimates of the average treatment effect for binary outcomes in randomized experiments, and report logistic regression as a robustness check. With three primary attitudinal outcomes and two primary behavioral outcomes we apply a Benjamini-Hochberg correction controlling the false discovery rate at 5%, reporting both unadjusted and adjusted p -values, along with an inverse-covariance-weighted summary index across the attitudinal outcomes as a single pre-registered composite.

Attrition within the session is expected to be minimal, since both measurements occur before participants leave, but it will be reported by arm.

Loss from the primary contrast is one-sided. Arm A contributes its pre-test, which is collected before any program content, so a participant who leaves without scanning again still contributes to the primary comparison. Arm B contributes only if a post-test exists, which makes Arm B the only arm that can be lost from it. We report the number assigned to each arm and the number contributing to each contrast, and we bound the estimate following Lee (2009). The sensitization comparison uses Arm A’s post-tests, so it can lose observations from both arms, and we report its attrition separately.

7.7 Experimenter demand

The principal threat to the attitudinal outcomes is experimenter demand, which is asymmetric by construction. Participants completing the instrument at the end of a session have just been presented with material supporting confidence in the process, while those completing it at the start have not. The threat is more significant for trust and confidence and less of a concern for factual knowledge, which is not readily produced by social desirability.

We mitigate it in two ways: wherever possible the survey instrument administration is separated from program delivery, with the facilitator blind to which arm each participant is in. All measurement is at the event, and because of this, we do not have a longer-term follow-up in this design, so we interpret the trust and confidence effects with caution.

8 POWER

Because 2026 is a pilot, we do not treat it as powered for confirmatory tests. Its purpose is to estimate the parameters for a larger study including the enrollment funnel, attendance and session completion, outcome variance in this population, within-session attrition, and the magnitude of pretest sensitization.

Recruitment is open and continuing, so we give power across the range the four scheduled sessions can produce. Each session is supplied with twenty-four assignment cards, so the four sessions seat at most 96. Attendance will be lower than registration, and the attendance rate is one of the quantities the pilot will estimate.

The primary contrast compares arm A’s pre-tests with arm B’s post-tests. It is a comparison between participants, so the effective sample is half of attendance. Table 1 reports power for a two-sided test at the 0.05 level across attendance from 24 to capacity.

Attendees	Per arm	$d = 0.15$	$d = 0.25$	$d = 0.30$	MDE at 80%
24	12	0.06	0.09	0.11	1.20
34	17	0.07	0.11	0.14	0.99
48	24	0.08	0.14	0.17	0.83
60	30	0.09	0.16	0.21	0.74
72	36	0.10	0.18	0.24	0.67
96	48	0.11	0.23	0.31	0.58

Table 1: Power for the primary contrast, by realized attendance. Effects in standard deviations.

Our earlier survey-experimental work leads us to anticipate effects between 0.15 and 0.30 standard deviations on the attitudinal outcomes. The pilot cannot distinguish effects of that size from zero at any attendance these sessions can reach. If all four sessions fill, the smallest effect detectable with 80 percent power is about 0.58 standard deviations. It is not impossible that an in person three hour session will have larger effects than a survey experimental prime, which is also why the pilot is helpful.

The within-participant pre-post comparison in arm A is paired. With a pre-post correlation of 0.7, power against an effect of 0.25 standard deviations is about 0.18 at 12 pairs and about 0.33 at 24 pairs.

Reaching 80 percent power against an effect of 0.25 standard deviations would require about 253 participants per arm, or roughly 500 attendees. At 0.30 standard deviations it would require about 176 per arm. We will report estimates with confidence intervals, and we will not interpret a null result from these sessions as evidence that the program has no effect.

Attrition is not expected to be an issue in this design because we randomize after participants have arrived.

9 ETHICAL CONSIDERATIONS

- Participation is entirely voluntary, and participants may decline any question or withdraw at any time without affecting their participation in the county program.
- We randomize the timing of a survey, not access to the intervention. Every participant who attends receives the full program unless they choose to leave.
- The intervention consists of factual, nonpartisan civic education content produced by a county government agency. It does not advocate for any party, candidate, or political position.
- Recruitment is conducted by the county’s election administrator through its own channels and host venues.
- Where a sign-up item transmits participant information to the county or to an outside organization, that disclosure is stated in the consent materials and is kept separate from survey responses.
- Survey data are de-identified for analysis, and identifying information is held separately from the analysis data. Identifiers are retained only to link a participant’s two measurements and to deliver the sign-up items participants opt into. They are destroyed once both are complete.
- The study is conducted under review by the UC Berkeley Committee for Protection of Human Subjects.

10 REGISTRATION

This pre-analysis plan will be registered with OSF prior to randomization.

11 POTENTIAL ADDITIONS

- **An individual-viewing comparison.** One possible extension is to have individuals watch the videos at home rather than in a group with a facilitator. This isolates how much of the program’s effect comes from the content and how much from collective viewing and discussion, and it is the version that could scale at near-zero marginal cost (though at this point we are not interested in randomizing between modalities). It is

a separate question, would be based on its own sample rather than cross-randomized, and would be a separate study. Nevertheless comparing treatment effects from an individual vs. group based study would be potentially informative.

- **Following up on observer sign-ups.** Participants who ask for information about serving as an election observer state an intention during the session. Whether they act on it is a separate question. A follow-up after election day could establish how many of them went on to observe, and could give them a way to report what they saw back to the county. The county keeps its own record of who observes, which would answer the first part without relying on self-report. Observable activities in Contra Costa run into late November, so the window extends well past election day.

REFERENCES

- Aronow, Peter M., and Cyrus Samii. 2017. “Estimating Average Causal Effects Under General Interference, with Application to a Social Network Experiment.” *The Annals of Applied Statistics* 11 (4): 1912–1947.
- Barker, Jennie, Oren Samet, and Susan D. Hyde. 2024. *Can Citizen Election Observation Increase Public Confidence in U.S. Elections?* SSRN Scholarly Paper. Rochester, NY, November.
- Brown, Mitchell, Kathleen Hale, Soren Jordan, and Ryan D. Williamson. 2024. “Restoring Trust in US Elections through Effective Election Administrator Messaging.” *Public Opinion Quarterly* 88, no. SI (August): 632–655.
- Fearon, James D. 2011. “Self-Enforcing Democracy.” *The Quarterly Journal of Economics* 126 (4): 1661–1708.
- Gaudette, Jennifer, Seth J. Hill, Thad Kousser, Mackenzie Lockhart, and Mindy Romero. 2025. “Can Official Messaging on Trust in Elections Break Through Partisan Polarization?” *British Journal of Political Science* 55, no. e16 (January): 1–19.
- Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. W. W. Norton.
- Kousser, Thad, and Lauren Prather. 2024. *Can election facility tours increase trust in elections?* Research Summary. Cambridge: MIT Election Data and Science Lab.
- Lee, David S. 2009. “Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects.” *The Review of Economic Studies* 76 (3): 1071–1102.
- Lin, Winston. 2013. “Agnostic Notes on Regression Adjustments to Experimental Data: Reexamining Freedman’s Critique.” *The Annals of Applied Statistics* 7 (1): 295–318.
- McCambridge, Jim, Kanan Butor-Bhavsar, John Witton, and Diana Elbourne. 2011. “Can Research Assessments Themselves Cause Bias in Behaviour Change Trials? A Systematic Review of Evidence from Solomon 4-Group Studies.” *PLoS ONE* 6, no. 10 (October): e25223.

- Przeworski, Adam. 2005. "Democracy as an Equilibrium." *Public Choice* 123 (3): 253–273.
- Solomon, Richard L. 1949. "An extension of control group design." *Psychological Bulletin* (US) 46 (2): 137–150.
- Suttmann-Lea, Mara, and Thessalia Merivaki. 2023. "The Impact of Voter Education on Voter Confidence: Evidence from the 2020 U.S. Presidential Election." *Election Law Journal: Rules, Politics, and Policy* 22, no. 2 (June): 145–165.
- Weingast, Barry R. 2004. "Constructing self-enforcing democracy in Spain." In *From Anarchy to Democracy*, edited by Joe Oppenheimer and Irwin Morris. Stanford: Stanford University Press.