American Association for Public Opinion Research

Testing Polls in Official Election Booths
Author(s): Harry H. Field and Gordon M. Connelly
Source: *The Public Opinion Quarterly,* Vol. 6, No. 4 (Winter, 1942), pp. 610–616
Published by: Oxford University Press on behalf of the American Association for Public
Opinion Research
Stable URL: https://www.jstor.org/stable/2745450
Accessed: 22-01-2024 00:09 +00:00

TESTING POLLS IN OFFICIAL ELECTION BOOTHS

By HARRY H. FIELD and GORDON M. CONNELLY, National
   Opinion Research Center

ALTHOUGH sampling surveys since the 1936 presidential election have succeeded in predicting the outcome of several hundred local, state and national elections with a range of error seldom exceeding six per cent, critics have maintained that the results of surveys on social and economic questions do not express the true opinions people would reveal if allowed actually to vote on them.

To test the justification of this criticism and to throw light on the reliability of opinion surveys on issues, the National Opinion Research Center attempted to locate a city which could be transformed into an open-air laboratory. The city of Boulder, Colorado, was selected, largely because the University of Colorado expressed a willingness to participate in the experiment and other authorities were ready to cooperate.

This report discloses the opinions secured from a sample of Boulder's electorate on three public issues, and comparable opinions secured in the actual polling stations on election day on the three identical questions, which were presented to all voters. Also, for purposes of comparing results of a sampling forecast with actual election returns, the representative respondents were asked their pre-election preferences in the senatorial and gubernatorial contests. Results of both the survey and the election are for the city of Boulder only. Similarity of the findings provides a measure of the accuracy of the sampling technique whether applied to candidates or issues.

WHY OFFICIAL ELECTION BOOTHS USED

To determine whether a stratified sampling survey on issues represents not only the honest opinions of those interviewed but also of those not interviewed, the Center had to find some system for ascertaining the secret opinions of all voters (or a preponderant number of them). Since the critic, especially he who thinks the average citizen will not reveal his honest opinion to an interviewer, would hardly accept a 100 per cent house-to-house canvass by interviewers as a test of the reliability of sampling results, the Center was obliged to devise another plan for "total enumeration." The logical choice was the private polling booth, where Americans, without fear of authoritarian reprisal or public censure or individual coercion, traditionally express their own secret opinions.

Be it understood, the Center does not necessarily subscribe to the notion that elections *per se* are the final authorities on the reliability of surveys. As a matter of fact, in attempting to predict an election, a research organization faces a world of problems not directly related to opinion at all. Many opinion analysts contend that scientific surveys, more than elections themselves, are the most reliable expression of public opinion. They claim that surveys are not subject to political-machine manipulations in getting out a disproportionately large Republican or Democratic vote, to "short-pencil artists," to dead persons' voting, to gerrymandering, to the effect

of weather keeping the elderly and infirm from the polls, nor to the poll tax. Scientific surveys, experts maintain, represent all types of persons in the degree to which they exist in the entire population, while elections seldom do. It is common for less than half of the registered voters to go to the polls, not to mention the unregistered. The Center realizes, however, that to test the reliability of surveys, it must use the popular election as the yardstick, for no other measure would have general acceptance. In so doing, N.O.R.C. runs the risk of having its work discredited in the public eye without its accuracy actually being disproved.

While N.O.R.C. generally uses a social cross section, representing all types of persons according to their actual numbers in the total adult population, in this case, where it desired to show results comparable to those of an election, the Center found it necessary to sample only those who were eligible, and by their own statements intended voters. How well it anticipated male and female voters under and over forty years of age may be seen in the following table, where a comparison of the sample and vote is made. Probably never before in American history had representatives of an unofficial organization been permitted to operate in any way whatsoever within an official polling station. Before the Center could enter such stations to ask voters to record their opinions on a special ballot, many steps were taken.

ENLISTING OFFICIAL COOPERATION

After enlisting the cooperation of the regents, president and faculty of the University of Colorado at Boulder, the Center received permission to enter the polls from the following officials: The Secretary of State of Colorado, the Attorney General of Colorado, the county commissioners of Boulder County, the county clerk of Boulder County, the election judges of Boulder and the Republican and Democratic county committees. The Boulder Chamber of Commerce and the daily newspaper, respectively, endorsed the experiment and promised sufficient publicity to acquaint voters with the reason for the Center's appearance at the polls.

FACTUAL COMPARISONS

| | | | | | Vote on N.O.R.C. Ballot at Polls | |
	1940 Census	Pre-election Survey	Unofficial 1942 Vote	Total	Voters Interviewed In Survey	Voters Not Interviewed Previously
Men						
21-40 Years	19.7%	15.1%	10.7%	11.8%	12.4%	11.6%
Over 40 Years	24.8	34.1	31.3	30.3	31.6	29.8
Women						
21-40 Years	21.5	20.1	16.4	17.8	18.7	17.4
Over 40 Years	34.0	30.7	41.6	40.1	37.3	41.2
In cases						
100% equals	9,077	1,221	3,926	3,300	969	2,331

Questions as diverse as possible seemed desirable for this test. Finally chosen were those dealing with a national sales tax, old age pensions and a world union. Inasmuch as the extent of voter cooperation in filling out opinion ballots had never previously been determined, the Center limited its questions to three, rather than provoke the doubtfully-cooperative citizen. Also, as a part of the sampling survey, it seemed desirable to forecast two political contests involving personalities. Twenty-nine interviewers, twenty-five of whom were University students, were especially trained to interview a sample of voters. Interviewing was done in the customary manner. All questions were asked exactly as worded, and no explanations or interpretations were offered. Opinions were gathered by personal calls.

A "pilot study" of 196 cases was made October 23 and 24, after which a number of improvements was made in the questionnaire. Consequently, none of the figures from this first sample were included in the results recorded the evening before the election, or are printed herein. The sample upon which all predictions were made included 1,224 cases gathered on the following dates: October 26, 406; October 28, 397; October 30-31, 421.

PROCEDURE AT THE POLLS

In each of the sixteen official precinct polling stations on November 3, an authorized representative of the Center was present during all voting hours from 7 a.m. to 7 p.m. to hand to every voter one ballot to be marked secretly and placed in a sealed ballot box. In order not to interfere in any way with the election, voters were not given their opinion ballots until they had completed the official voting. N.O.R.C. representatives in each station were permitted to display prominently a large poster informing voters of their part in the experiment. Regardless of whether or not they were interviewed in the sampling survey, all voters were requested to mark a ballot and place it in a special sealed ballot box. The representatives recorded by sex and age all those who for any reason whatsoever declined to "vote." After the polls had closed, the sixteen sealed ballot boxes were delivered to a special University of Colorado committee, which supervised the tabulation.

RESULTS OF EXPERIMENT

The significance of the Boulder experiment is that scientific methods of opinion measurement can predict within a small degree of error how people would vote on issues, as well as how they would vote on candidates. While one test is far from conclusive evidence of the reliability of sampling surveys, the fact cannot be overlooked that in this one instance, at least, a survey proved capable of ascertaining public opinion on three separate issues with a reasonable degree of accuracy, particularly on two of the three.

The following table includes the predictions based on the sample, the actual vote, the error within which the results should presumably have fallen provided the cross section was an accurate miniature of the Boulder voting population with opinions and provided there were no other biasing factors, and the actual error in predicting the opinion of those who voted:

Topic	Answer Given	Prediction	Actual Result	Margin Of Error On Prediction[1]	Actual Error
National Sales Tax	Affirmative	42.4%	35.2%	3.7%	7.2%
Old Age Pensions	Federal	60.5%	61.6%	3.7%	1.1%
Union of Nations	Good Idea	81.8%	78.6%	3.0%	3.2%
Carr vs. Johnson	Carr	53.3%	53.0%	4.2%	0.3%
Vivian vs. Bedford	Vivian	61.6%	61.4%	4.2%	0.2%

Assuming that those who cast the official vote in Boulder on election day were actually representative of all registered persons, it may be argued that even this vote was nothing more than a sample itself, being subject to a margin of error of 1.4 per cent on candidates and 1.8 per cent on issues, insofar as it represented all registered voters. That is, if all those registered had voted, the result might have varied as much as 1.4 per cent on candidates and 1.8 per cent on issues from the actual vote, merely on the basis of a mathematical sampling error. (Boulderites who voted for Carr and Johnson totaled approximately 5,000 out of about 8,900 registered voters in the city, and those who participated in the experiment at the polls totaled a little over 3,700.)

Specific results showed 61.6 per cent of those voting at the polls on election day in favor of federal rather than state control of old age pensions, whereas 60.5 per cent of those interviewed in their homes favored federal control. In the election 78.6 per cent supported American participation in a union of nations after the war, compared to 81.8 per cent in the sample. The secret ballot boxes yielded a 64.8 per cent opposition to a national sales tax of 2 per cent "on everything that people buy," as against a 57.6 per cent negative vote given to interviewers. The Center's prediction of the outcome of the United States Senatorial election between Governor Ralph L. Carr, Republican, and Senator Edwin C. Johnson, Democratic incumbent, was virtually perfect, 53.3 to 53.0 per cent in the city of Boulder. Even closer was the forecast that John C. Vivian, Republican, would receive 61.6 per cent of the gubernatorial vote in Boulder. Actually he got 61.4 per cent.[2]

VALUE OF EXPERIMENT

Never before have issues used in surveys been put to a "total enumeration" before there has been extensive campaigning and public discussion about them. Surveys covering issues already coming to a vote have been conducted with amazing accuracy, but in such cases public thinking had become crys-

[1] For the benefit of the statistician, the margin of error is based on three Standard Deviations from the obtained sample percentage. This formula was used:

$$S.E. = \sqrt{\frac{PQ}{\frac{(n(N-1))}{(N-n)}}}$$

[2] Final returns differ slightly from the figures released in the Center's preliminary reports.

tallized through propaganda and education.

To determine the reliability of the usual survey covering topics not coming to an immediate vote, however, the most logical device appears to be a public vote on corresponding questions without the voter's foreknowledge of the decision he must reach. In Boulder, for example, neither respondent nor voter knew beforehand what questions he would be answering. This situation exactly parallels that in the usual interview.

This statement is not meant to infer that surveys on impending plebiscites concerning matters of opinion do not test the reliability of survey results on issues, but merely that such tests are not exactly comparable to surveys on issues which are not to be referred to the voters.[3] It is in this sphere that sampling surveys can perform their most valuable service to democracy, measuring opinions which would not otherwise be measured.

"Public opinion polls do not exist for the purpose of predicting elections," Dr. George Gallup, Director of the American Institute of Public Opinion, wrote recently, "but these tests do provide a means of evaluating the accuracy and reliability of methods. The true function of polls is to measure and report the current trends of public opinion on the many vital issues which confront the nation."

Should further experiment prove the reliability of opinion surveys as decisively as they have proved the accuracy of election predictions,[4] statesmen and academicians will be able to determine whether an endorsement of a candidate is also an endorsement of his policies. Politicians no longer will be able to claim that an overwhelming vote for a certain candidate is a mandate for the victor to continue this or that of his policies. Also, the Center hopes that accredited surveys will perfect the processes of democratic government by rendering citizens articulate at all times on all issues, rather than just every two or four years generally on personalities only.

DISCREPANCIES CONSIDERED

Two important exceptions to the almost perfect predictions of the sample occurred. First, the percentage of Boulderites approving a national sales tax registered a 7.2 per cent decline from the sample to the election. This discrepancy, which exceeds the margin of error of the sample must be attributed to more than chance. Conceivably it may be due to any one or combination of three things; prestige, position on the forms, and/or some inherent element of the question for which the sample was not properly representative.

Did respondents tell interviewers they thought a national sales tax would be "all right" for fear that opposing the tax would brand them unpatriotic,

[3] One such plebiscite prediction was made by the Canadian Institute of Public Opinion early in 1942, when the vote on conscription of Canadians for overseas duty was forecast with an error under 4 per cent.

[4] In seven years the Gallup Poll has predicted 114 elections with an average error between 3 and 4 per cent, and the *Fortune* Survey in 1940 called the outcome of the Roosevelt-Willkie popular vote with an error of only 0.2 per cent.

especially in this war period?[5] If so, such a patriotic prestige factor might not be present when a secret ballot is marked. Or, were answers less reliable because this was the first opinion question on both the questionnaire and ballot, soliciting an answer before the respondent or voter was "warmed up" to thought and expression?

Second, the percentage of undecided persons showed a decline on every question at the polls, dropping 6.1 per cent on the sales tax, 6.0 per cent on the pensions and 6.7 per cent on the world union. Chance cannot explain all the discrepancy. Either or both of two factors seems responsible in this case. Either the secret ballot drew opinions from people who were reluctant to express them to interviewers, or most people who had no opinions were not sufficiently interested to vote, although they had said they would.

Supporting the first possibility is an experiment conducted by the American Institute of Public Opinion in Lisbon Township, Maine, during the 1940 presidential campaign, in which a secret ballot technique used by interviewers reduced from 16 to 2 per cent the number of undecided persons interviewed in the usual questioning manner. Supporting the second possibility is the fact that, while 80 per cent of all adults willing to grant interviews[6] claimed they were registered and were planning to vote, only 55 per cent of Boulder's

adults actually voted. Because there was virtually no difference in the percentages of no answers between voters who had and had not been interviewed, the possibility of the survey acting as an educating factor does not seem likely.

PRINCIPAL SHORTCOMINGS

Another weakness was the small size of the city of Boulder,[7] which necessitated a sample which was too great a part of the total enumeration to be as convincing a test of the sampling technique as would have been possible in a larger city. For statistical reliability, nevertheless, the laws of probable error make it necessary to secure nearly as many cases in a small city as in a large one. In Boulder the Center interviewed such a proportionately large sample that 28 per cent of the 3,713 voters who marked opinion ballots had already been interviewed in their homes. Because some voters of this 28 per cent were undoubtedly interviewed in the 300-odd test cases secured in the "pilot study" and in training interviewers, it is likely that only about 23 per cent were included in the final sample of 1,224 cases. Such an assumption is based upon the probability that the interviewed voters were evenly distributed over all the sample surveys. Where a larger electorate could be sampled, as accurate a cross section could have been found in less than 1 per cent of the total voters. Foreseeing this

[5] To avoid this prestige element, the Center might have asked: "Do you think the federal government should put a national sales tax of 2 per cent on everything that people buy, or do you think there are better ways for the government to raise money?"

[6] Although 1,526 adults were willing to be interviewed, only the 1,224 who said they intended to vote were asked all questions and

were included in the reported sample. Of those who said they would vote, however, only 69 per cent went to the polling stations, according to those who indicated on their secret ballots that they had been interviewed previously.

[7] According to the 1940 census, Boulder has a total population of 12,958, of whom 9,077 are 21 years of age or older.

situation, N.O.R.C. sought the consent of an election commission in a larger city. When this effort failed, the directorate preferred accepting Boulder's proffered cooperation to postponing the experiment indefinitely.

Added to the relatively oversize sample was the fact that the November 3 vote provided much less than a total enumeration. In the first place, only 69 per cent of the city's 1940 voters returned for this off-year election. In the second place, only 81 per cent of the local[8] vote participated in the experiment, although 95 per cent were approached by N.O.R.C. representatives.

[8] Excluded from the base were 441 absentee voters, whom the sample could not pretend to represent.

While considerably short of a total response, the cooperation of 81 per cent was considered rather successful by the Center, which looked upon this phase as more or less of an experiment within an experiment. Of those actually handed ballots at the polls, 86 per cent marked them, compared to 90 per cent of the eligible voters who answered the questions for interviewers.

Perhaps more through oversight of voters than non-cooperation, a fairly high percentage of omissions was found in the factual data on the back of the opinion ballots, probably because these ballots had to be turned over to be completed. Ten per cent failed to indicate sex, 9 per cent omitted the age classification and 17 per cent skipped evaluation or rental of their homes.