Restricted

# REPORT OF THE SOCIAL SCIENCE RESEARCH COUNCIL COMMITTEE* ON ANALYSIS OF PRE-ELECTION POLLS AND FORECASTS

## I. SUMMARY OF CONCLUSIONS AND RECOMMENDATIONS OF THE COMMITTEE

The Committee has made a study of available data on the 1948 election forecasts and has come to the following conclusions:

1. The pollsters over-reached the capabilities of the public opinion poll as a predicting device in attempting to pick, without qualification, the winner of the 1948 presidential election. They had been led by false assumptions into believing their methods were much more accurate than in fact they are. The election was close. Dewey could have won by carrying Ohio, California, and Illinois which he lost by less than 1 percent of the vote. In such a close election no polls, no advance information of any kind, could have predicted a Truman or Dewey victory with confidence. The failure of the polls was due to neglecting the possibility of a close election and the necessity of measuring preferences very accurately just before the election to determine whether a flat forecast could be made with confidence.

2. The pollsters could have foreseen the possibility of a close contest had they looked more carefully at their data and past errors. They acted in good faith but showed poor judgment in failing to apply in 1948 what they knew about their past errors, and failing to ascertain late campaign shifts.

3. The over-all operation of making election predictions from pre-election polls is a complex one involving eight major steps at each of which error may enter. It is very difficult to unscramble the total error and allocate components of it to these various steps. The evidence indicates that there were two major causes of errors (a) errors of sampling and interviewing and (b) errors of forecasting, involving failure to assess the future behavior of undecided voters and to detect shifts of voting intention near the end of the campaign.

---

*James Phinney Baxter, 3rd, President of Williams College

Philip M. Hauser, Associate Dean, Division of Social Sciences, University of Chicago

Carl I. Hovland, Chairman of the Department of Psychology, Yale University

V. O. Key, Professor of Political Science, The Johns Hopkins University

Isador Lubin, Chairman of the Committee on Statistical Standards of the American Statistical Association

Frank Stanton, President, Columbia Broadcasting System

Frederick F. Stephan, Professor of Social Statistics, Princeton University

Samuel A. Stouffer, Professor of Sociology, Harvard University

Samuel S. Wilks, CHAIRMAN, Professor of Mathematical Statistics, Princeton University

4. These sources of error were not new. While Gallup and Crossley were more successful in picking the winner and his electoral vote in 1940 and 1944 than in 1948, their average errors state by state were at least half as great in the two preceding elections as in 1948. Hence, it is possible that their errors in 1948 were due to much the same causes as those that produced the earlier forecast errors, but if so, these causes operated more strongly. Roper's wide discrepancy in 1948 cannot be explained so readily in terms of factors present in his earlier close estimates of the national vote. It appears to be due to the upsetting of the balance previously maintained among various factors in his polling operations.

5. To improve the accuracy of pre-election poll predictions satisfactorily it is necessary to reduce the error at every step in the over-all polling process. The error at some of the steps, notably sampling and interviewing, could be reduced by using methods now available. But reduction at other steps depends on further basic research in psychological and political behavior.

6. The manner in which the pre-election polls were analyzed, presented and published for public consumption contributed materially to the widespread misinterpretation of the results of the poll and to the great public reaction to their failure to pick the winner. This led to a poor understanding of the lack of accuracy of the polls and of the nature of the errors residing in the polls, with the result that the public placed too much confidence in polls before the 1948 election and too much distrust in them afterwards.

7. The public should draw no inferences from pre-election forecasts that would disparage the accuracy or usefulness of properly conducted sampling surveys in fields in which the response does not involve expression of opinion or intention to act. There are more appropriate methods to check the accuracy of such surveys.

Election forecasts from polls are hazardous. This fact might tempt some polling organizations in the future to limit their political studies to attitudes which cannot be checked up at elections. Thus, they could avoid the spotlight of a public audit of their work. The Committee would view such a step with concern, for all possible tests of opinion studies are needed. Election returns are not a direct or even a good test of the adequacy of polling on issues, like the Marshall plan, civil rights, etc. But elections are useful for testing the adequacy of polling methods for estimating the percentage of the vote going to each candidate from various groups in the population. No better test is now known. Even in the case of these estimates, it will not be possible to make election returns a satisfactory test unless pre-election studies and post-election studies are more rigorously designed so that different types of errors can be partitioned out. The responsibility rests squarely on polling agencies, in cooperation with social scientists, to work out such designs.

Apart from their use in election forecasting, opinion polls conducted in relation to elections have great value in giving social science and the public a better understanding of the American democratic process. Such polls do two things:

(a) They reveal the voting preferences and voting behavior of different parts of the population, such as men and women, old and young, etc.

(b) They reveal the relation between opinion on important issues and the

choice of the candidate for whom the voter casts his ballot. In addition, they reveal the opinions and preferences of those people who do not turn out to vote.

On the basis of its study the Committee makes the following recommendations:

1. To improve the accuracy of polls, increased use should be made of the better techniques now available, particularly in sampling and interviewing. Since the reduction of any part of the error greatly increases the chances of a successful forecast, the Committee urges that polling organizations exert every effort to adopt more reliable techniques.

2. Increased attention should be paid to the development of research on each step of the polling operation to attempt to improve methods used in opinion research. This would include research on sampling methods, interviewer bias, concealment of opinions, selection and training of interviewers, etc. Experimental studies of such problems should be planned well in advance of elections. Many of these experiments can be incorporated into regular survey operations and as a part of more general cooperative studies of specific communities.

3. Research should be expanded on the basic sciences, particularly social psychology and political science, which underlie the analysis of voters' behavior. Even if perfect sampling of individuals is employed, we now know too little about voting intentions, factors affecting change in opinion, prestige effects, and similar topics to predict who will translate his opinion into actual voting.

4. In view of the increasing amount of emphasis being placed on public opinion polls, the Committee considers it very important for the public to be effectively informed about the limitations of poll results so it can interpret them intelligently. It urges that polling organizations, newspapers and magazines, and social scientists who work with poll results help provide the public with more information about polls, their interpretation and limitations.

5. There should be more effective cooperation between research workers interested in opinion measurement on common problems of methodology, underlying theory, and research design, including more studies of validity and reliability than have been made heretofore. Such cooperation would benefit all parties through the wider dissemination of their findings, and more efficient use of research resources and opportunities for experimentation.

6. More extensive training facilities and opportunities for practical experience under effective supervision should be provided for students who may be interested in research careers in this field, which includes political behavior, psychological research, and opinion measurement and statistical methods.

7. Analysis of the elections is greatly hampered by long delays in the reporting of official returns. The present inadequacies with respect to the collection of election statistics should be remedied by effective organization for rapid and accurate reporting of election statistics at the local, state and national level.

## II. THE TASK OF THE COMMITTEE

The failure of the public opinion polls to predict correctly the outcome of the 1948 presidential election has created wide confusion and misgivings about the reliability of the polls. Reactions of the public to the polls have ranged from charges of outright fraud to expressions of personal sympathy for the pollsters. Reactions of experts ranged from condemnation for carelessness, unintentional bias, errors of judgment, and use of outmoded techniques, to a determination to make use of this experience to enlarge our knowledge of political behavior and to improve survey methodology.

The Committee on Analysis of Pre-election Polls and Forecasts was set up by the Social Science Research Council to make an immediate examination of available information about the polls and the election in order to discover as far as possible the causes of the errors in the forecasts and to make recommendations for the improvement of polling methods. In accepting this assignment, the Committee recognized the urgency of completing the analysis and reporting their findings while the relevant information was still fresh and before various gross misconceptions became crystallized. The Committee, as a technical group, has not been able to undertake an exhaustive study of the problems involved. It is hoped that further studies of many of their aspects will be pushed vigorously by students of polling methodology and political behavior.

The Committee wishes to acknowledge its appreciation to the various polling organizations and especially to Messrs. Crossley, Gallup, and Roper for their wholehearted cooperation. They have responded to every request of the Committee for data and explanations.

The Committee also wishes to commend the pioneering work that these men have done in the field of public opinion polling, a broader and more important activity than pre-election forecasting. Public opinion polling is only in its infancy and if soundly developed and properly utilized it will be of great assistance in understanding political and social behavior. Sound development requires the cooperation of the polling organizations and social scientists.

The Committee wishes to acknowledge the assistance of a staff of five able young social scientists and statistical experts who worked for five weeks under considerable strain and inconvenience on the detailed analysis of many problems which confronted the Committee. The work of the staff will be published in a separate monograph.

## III. THE PROBLEM OF PRE-ELECTION FORECASTING

### A. Complexity of the Problem

Forecasting election returns from pre-election polls is an exceedingly complex and difficult task. It involves two main problems:

First, selecting a sample of persons to be interviewed and ascertaining their voting intentions and political preferences as of the time of interview. Who expects to vote and how does he expect to vote?

Second, forecasting who will actually vote on election day and how they will vote. This requires anticipation of possible shifts from expressed political preference and voting intention between the time of interview and election day. What voters are most likely to change their plans?

### B. Steps Involved in Pre-election Polling and Forecasting

Each polling organization has its own procedures, but all of them take the following major steps:

(1) Design of a plan and instructions for selecting a sample of respondents

(2) Design of a questionnaire and instructions on interviewing procedure

(3) Selection of the respondents in the field

(4) Interviewing the respondents

(5) Decision as to which respondents will actually vote

(6) Decision on what to do about voters who say they are undecided or will not tell how they will vote

(7) Processing the data, including adjustments and corrections for trends and other factors

(8) Interpretation and presentation of results, including the projection of a forecast.

In polling, none of these operations can be performed perfectly. Some error will be made at each step. The total error between a pre-election poll prediction and the actual vote is a combination of the errors entering into these various steps. Some of these errors may cancel out.

A consistently high degree of accuracy of forecasting will be possible only if every step in the process is highly controlled and carefully executed. The Committee doubts that such control is possible in the present state of knowledge.

C. Chances for Successful Prediction

It is one thing to predict the national popular vote for a presidential candidate and quite another to predict the winner of a presidential election. The public has not been interested in the size of the error with which a pre-election poll predicts percent of popular votes that a given candidate will receive. It demands that the pre-election poll predict the winner. This is a very severe test, heavily affected by chance in close elections, and is complicated by the electoral college system.

In predicting the winner, the forecast either succeeds in naming the man who is elected or it doesn't. In forecasting a candidate's percentage of the popular vote, the success of the prediction is a matter of degree. It is usually measured by taking the difference between the percentage forecast for him, and the percentage he actually received.

All three national polls picked the wrong man as winner. They underestimated Truman's percentage of the national vote by 4 to 12 percentage points. Roper made no state forecasts. Crossley and Gallup came within 2 points of the correct state vote in 9 and 3 states respectively and within 4 points in 21 and 17 states respectively. Crossley's greatest errors were -10.8 and +8.2 points, Gallup's -11.5 and +1.7. Gallup did not attempt to forecast in 4 southern states and Crossley in 1.

These state errors can be regarded as the result of two tendencies. The first is a general tendency of the forecasts to underestimate Truman's vote. This is reflected in the average of state errors which is -4 of -5 percentage points. The second is the tendency of the errors to vary from state to state around the average error, reflecting special conditions in each state and the effects of chance in the polling and forecasts.

This error variation around the average error swings from 0 to 8 points in both directions. By itself it is seldom great enough to upset the forecasts except in very close state contests. Even then it tends to balance out in the totaling of the state results. Clearly the average error, affecting all states alike, was the more serious source of the failure of the forecasts. The problem, therefore, is to seek out the factors that caused it.

The average errors in earlier election forecasts were about 2 points in 1944 and 1940 and 5 to 6 points in 1936.

These average errors of state forecasts all resulted in underestimating the Democratic vote. How long such a tendency could be expected to continue was uncertain. All the presidential polls prior to 1948 had Roosevelt's candidacy as a common factor and it was possible that the new line-up would change the size and direction of the average state error. Moreover, some of the forecasts in interviewing Congressional elections had overestimated the Democratic vote.

The success of methods of prediction is often tested by comparison with what some simple predicting device would show. As a rule-of-thumb test of election forecasts, an analysis was made of the accuracy that could be attained by merely predicting that the vote in each state will divide exactly as it did in the preceding election. It seems reasonable to expect that polling forecasts should work at least as well as this simple method, over a series of elections,

if they are to be of any use. Allowance can be made for the appearance of new parties in 1948 by dealing only with the Republican vote. In the period during which polls have been conducted (1936-1948) the accuracy of this simple method of prediction is almost as good as that attained by polling. In periods of rapid political change, however, like 1916-1920 and 1928-1932, prediction from past elections is very poor. It is to be expected that in such cases polling would be significantly superior to forecasts based on past voting.

Another test can be applied to the performance of the polls in predicting elections. On the basis of average error and error variation found in the polls from 1936 to 1948 one might ask the following question: What are the chances that a poll having such errors would correctly predict the winning presidential candidate in an election having a Democratic victory as close as the 1948 election? The answer to this question is about 1 chance in 4. In elections as close as this when the Republican candidate is the victor, the chances would be about 3 out of 4 that polls with these errors would predict the winner correctly.

From this it appears that the tendency to underestimate the Democratic vote seriously disturbs the chances of success when the Democratic candidate wins a close election. With such a tendency, efforts to reduce the other errors actually decrease the chance of successfully predicting Democratic winners and increase the chances of successfully predicting Republican winners.

Since 1844 there were eight elections so close that the winning candidate had a plurality over his leading opponent of 4 percentage points or less. Forecasting the results of these elections by the methods used in 1948 would have been extremely hazardous though predicting the percentage vote would not necessarily have been more difficult than in other elections.

## IV. THE FORECASTS COMPARED WITH ELECTION RESULTS

The latest available returns show a discrepancy of between 4 and 5 percentage points between Crossley's and Gallup's predictions and the actual election results for Truman and Dewey. Roper's error was larger. The predictions and returns for the national election are as follows:

| | Percentage of total presidential vote | | | | |
|---|---|---|---|---|---|
| | Dewey | Truman | Thurmond | Wallace | Total* |
| National vote | 45.1 | 49.5 | 2.4 | 2.4 | 99.4 |
| Crossley | 49.9 | 44.8 | 1.6 | 3.3 | 99.6 |
| Gallup | 49.5 | 44.5 | 2.0 | 4.0 | 100.0 |
| Roper | 52.2 | 37.1 | 5.2 | 4.3 | 98.8 |

*Exclusive of percentages for minor candidates. Gallup percentages calculated on total vote for four principal candidates.

Crossley and Gallup both reported their state-by-state predictions as well as over-all prediction of popular votes. The state-by-state predictions and actual vote will be given in the Committee's staff report. Roper released only national averages. Gallup predicted the winning candidate correctly for 30 states

(with 323 electoral votes) of 44 states predicted and Crossley for 32 states (with 353 electoral votes) of 47 states predicted ( in two states equal percentages of those polled favored each candidate).

A number of individual city and state polls were analyzed. In general their errors were similar to the state-by-state errors made by Crossley and Gallup. While predictions were made with great accuracy in several instances, there were so many state-by-state predictions made that some predictions will be close just a matter of chance. Indistinguishable from close predictions due to chance, there may be some that are close due to superior methods.

The following table, based on preliminary election returns, shows the state-by-state errors in predicting the vote for Truman:

| Forecast percentage minus election percentage | Number of state errors in Crossley predictions | Number of state errors in Gallup predictions |
|---|---|---|
| From -12.0% to -10.1% | 2 | 4 |
| " -10.0% to - 8.1% | 4 | 2 |
| " - 8.0% to - 6.1% | 5 | 9 |
| " - 6.0% to - 4.1% | 13 | 12 |
| " - 4.0% to - 2.1% | 12 | 14 |
| " - 2.0% to - .1% | 7 | 2 |
| " 0 % to - 1.9% | 2 | 1 |
| " 2.0% to 3.9% | 0 | 0 |
| " 4.0% to 5.9% | 1 | 0 |
| " 6.0% to 7.9% | 1 | 0 |
| Total number of states forecasted | 47 | 44 |

## V. PRINCIPAL SOURCES OF ERROR IN 1948 PREDICTIONS

The total error in the 1948 election predictions is compounded from errors, some of them compensating, enumerated in the eight steps listed in section III B. It is not possible to measure accurately the various components of the total error. But it is possible, on the basis of the analysis made by the Committee, to identify the major sources of error.

The evidence indicates that there were two major causes of errors: (a) errors of sampling and interviewing and (b) errors of forecasting, involving failure to assess the future behavior of undecided voters and to detect shifts near the end of the campaign.

First is the class of errors arising from the sampling and interviewing methods used. Two alternative methods of sampling have been employed in such studies. In the first the individuals to be interviewed are determined in advance by probability methods. In the second, the quota method, the interviewer is left free, within the restrictions of assigned quotas, to select respondents largely in accordance with his own judgment. This is the method

primarily used by the major polling organizations. It is impossible to separate the error introduced by the quotas set from that arising from the process of selection by the interviewers.

All of the major polling organizations interviewed more people with college education than the actual proportion in the adult population over 21 and too few people with grade school education only. This practice has been defended on the ground that non-voters are found disproportionally in the classes with little education. If the problem posed were to learn the political preferences on a given day of all individuals over 21 years of age, irrespective of whether or not they would vote, the sampling problem would be theoretically simple. The sample of respondents should be so drawn as to correspond to know characteristics of the total adult population.

But there are only two ways to eliminate potential non-voters. One way is not to draw them in the sample (as in the case of people known to be under 21). The other way is to eliminate them through screening questions after they are drawn in the sample - questions whose answers will show that the respondent is not likely to vote. The first method is safe only if we know that there is an extremely close association between some objective characteristic (like age under 21) and non-voting. There is not sufficient factual evidence to justify eliminating a fraction of a group (like part of those with grade school education) when the relationship between the characteristic and non-voting is far from perfect. This easily could have been one of the most dangerous traps into which the polls fell.

On the other hand, the second method, of eliminating non-voters through screening questions, also used by the pollsters, is still a primitive procedure in the absence of sound theory and in the absence of experience with which to establish the value of the screening questions. The chief trouble with the quota method is the fact that there is no good procedure for digging into the results to find out what kind of selection the interviewer actually makes--hence the flaws in the several stages of the procedure cannot be clearly separated for analysis. The advantages of the probability method of sampling, advocated by some statisticians, lie primarily in the fact that the errors introduced by selection of respondents by the interviewer can be reduced to a minimum. But with this method, as with the quota method, there is still the difficult problem of screening out the people who will not actually vote.

A second major source of errors in the 1948 election forecasts is to be found in the failure of the polls to detect shifts in voting intentions during the later stages of the campaign. Roper erred by assuming that voting intentions would not change during the campaign, as evidenced by his announcement of September 9. Crossley and Gallup made no attempt to detect the shift in voting intentions in the last two weeks of the campaign. Post-election polls report that about 1 voter in 7 said he made his decision on how to vote within the two weeks preceding the election, and that about 3 out of 4 of these voters said they voted for Truman. Even if one makes allowance for errors in such reports, one must conclude that failure to detect and measure changes of mind about voting during the closing days of the campaign accounts for a considerable part of the total error of the prediction.

The problems of forecasting in 1948 were complicated by the relatively large proportion of voters who were undecided when interviewed. These respondents had to be eliminated as not likely to vote or allocated among the candidates. The

number of apparently undecided voters found by the pre-election polls was approximately 15 percent, which is nearly twice as large as that of 1944. This group remained large even after those "leaning" towards one candidate or the other were removed.

In general, the "undecided" vote was allocated to the different candidates in accordance with the proportions found among the decided voters or ignored on the assumption that the undecided probably would not vote. The evidence available suggests that a great majority of the "undecided" voters who voted actually voted for Truman and that a better allocation of the "undecided" might have been made from other data available to the pollsters. However, the indications are that defective allocation of the "undecided" vote contributed less than 1.5 percentage points to the over-all prediction error. Much more serious may have been the errors due to difficulty of estimating who would vote and who would actually shift allegiance before election.

The error in predicting the actual vote from expressed intention to vote was undoubtedly an imporatnt, although not precisely measurable, part of the over-all error of the forecast. The prediction of human behavior from an expression of intent is, in the present state of knowledge, and particularly with the actual methods used, a hazardous venture. This is a central problem for research, which has been largely ignored in pre-election poll predictions. The gap between an expression of intent and actual behavior will continue to be a major source of error in future election forecasts as well as in other attempts to predict human behavior; and in the present state of knowledge, must be recognized as a baffling and unsolved problem imposing serious limitations on opinion poll predictions.

It is a responsibility of social science to contribute much more to knowledge of the contingencies which must be taken into account. Extensive anaylsis by political scientists and social psychologists of election behavior in 1948 and in other elections is essential to provide more knowledge of how voters act under different circumstances. The poverty of our exact knowledge of how and why people vote in general severely handicaps efforts to predict how they will vote in a particular election.

## VI. INTERPRETATION AND PRESENTATION

In interpreting the results of the pre-election polls and presenting them to the public, the pollsters went far beyond the bounds of sound reporting of the results of pre-election polls . They attempted the spectacular feat of predicting the winner without qualification. The presentation of the results gave the impression of certainty as to the outcome. The final releases carried very little indication of the limitations of polling and the tendency in past election forecasts to underestimate the Democratic vote. Statements of conditions under which different outcomes of the election might occur were dropped almost completely before the end of the campaign.

The polls also failed to provide the public with sufficient information about the methods of poll operation to permit assessment of the degree of confidence that could be placed in the predictions. The number of cases used, the type of sampling employed, the corrections introduced, and how returns from individuals who did not know for whom they would vote were tabulated, were not discussed adequately. It is recognized that there is pressure from newspaper editors and readers to omit qualifications and "technicalities," but pollsters

and social scientists have an important responsibility for educating readers of poll results to evaluate them and understand their limitations.

## VII. SPECIFIC EVIDENCE OF SOURCES OF ERROR

The full facts supporting the Committee's findings will be published in the report of the Committee's staff. Enough of the evidence on the errors that occurred at the various stages of the polling and forecasting operation will be given here to show the kinds of weaknesses that were revealed.

### A. The Field Operation

#### 1. The plan or method of respondent selection

Most of the election forecasts were based on surveys in which respondents were selected by "quota sampling" methods, described in Section V. These methods were known to have certain inherent weaknesses but were used again in 1948, in part because they were less costly than the more highly controlled methods that had been advocated by some sampling experts, and in part because tests of quota sampling methods had convinced the pollsters that they would work satisfactorily.

These tests had been made in a limited number of situations which did not contain all the hazards present in the 1948 campaign. Various methods devised by the pollsters to eliminate or offset errors inherent in their techniques enabled them to pick the winner in the fair weather sailing of 1936, 1940, and 1944 but failed to preserve them from shipwreck in the rough weather of 1948.

Past experience with quota sampling has demonstrated that it under-samples the groups having the very highest and the lowest incomes in the population. The quotas assigned for four major economic groups do not control this fully. These groups were originally based on a crude measurement of economic level that may have been badly distributed by recent economic changes. There is, however, no direct test of the error due to the use of these economic groups in the 1948 forecasts. One test of the sampling is provided by questions on education. Gallup's sample under-represented the lower educational groups but he corrected the proportions to match the estimated educational distribution of the voting population. No question on education was asked by Crossley's interviewers. Similar distortions due to the method of selecting respondents probably occurred in other polls.

There is also evidence that some of the samples of rural population were not fully representative due to inadequate control of the tendency to take people who come to town frequently or are easy to reach.

A test of the selection of localities in which interviewing was performed by Gallup and Crossley was made in 3 states by comparing the actual election results for these localities with those for the entire state. It showed that the contribution from this source to the total error in 1948 was of secondary importance. Thus, it is likely that the principal weakness of the quota control method occurred at the local level at which respondents are selected by interviewers.

The size of the samples used by Gallup and Crossley for state estimates varied between 200 and 3000. Samples this small will have errors, due to their limited size alone, from 0 to 5 percentage points or more in either direction, the larger errors occurring, of course, less frequently. The larger the sample, the smaller the errors due to size alone. For Gallup's national samples of 3250 respondents each, it was estimated that from this source

errors of 1 percentage point or less would occur in about two-thirds of the forecasts and between 1 and 2 percentage points in about one-fourth of the forecasts. Hence in the national forecast of the percentage of the vote going to a major candidate the error introduced by the small size of the sample would counteract other errors half the time and would account for no more than 2 percentage points of the total error in 95% of the remaining instances.

The total errors of the national forecasts, pointing as they did all in the same direction as previous election polls, were not due to the limited size of the samples. On the contrary, the errors introduced by the smallness of the sample may have offset some of the errors in the opposite direction that were due to other causes. Hence the part of the design that established the size of the samples cannot be held accountable for more than a minor share in the failure of the predictions.

Another defect of some of the sampling designs was the use of quotas for states based not on the population of voting age but on the actual vote in 1944. When this was followed by the elimination from the calculations of persons who said they did not expect to vote, it tended to overcorrect for relative turnout thereby throwing the quotas out of line with actual turnout. While the effect was small, this was nonetheless a defect in design. The combined effect of several such defects may sometimes be of serious proportions.

2. The questionnaire design

The questionnaires as designed, while generally good and reflecting years of experience in polling, appeared to be inadequate to cope effectively with the difficult psychological problems involved in getting correct measures of intention actually to vote, and of the voter's choice of his candidate. No direct evidence of the effect of these inadequacies was available, except that the analysis of the poll results was hampered by the lack of information needed for a more accurate determination of which respondents would actually vote and how the undecided voters would most probably make their choice among the candidates. No questions were asked to determine between which candidates the undecided were debating their choice, i.e., on which fence each was sitting. Such a question would have made their opinions on various issues more useful guides to an estimate of the direction in which they would jump. Very little was obtained on the susceptibility of respondents to campaign influences that might lead them to switch at the last minute or the likelihood that they would actually come out and vote. Crossley's "filter" or screening questions, designed to eliminate persons who are unlikely to vote, were an interesting exception. Furthermore, some of the questions were changed from one poll to the next, obscuring the indications of trend.

While polling organizations have made much progress in question wording, the need for further research and experimentation in the improvement of questionnaires is evident. Few opportunities were taken to introduce experimental comparisons into the questionnaire in order to discover the effect of variations in the questions. Some important exceptions are noted in the case of Roper's comparison of questions loaded toward Dewey with other questions loaded toward Truman and Gallup's experimentation with secret ballots. But these only point to other tests that might have been made to improve the forecasts and measure their dependability.

3. Actual selection of respondents

Given the sampling plans as drawn and the quotas assigned, the actual selection of respondents proceeded in reasonable conformity to the plan. Exceptions were noted in certain instances. In Elmira, New York (a survey not intended for the prediction of the election but for analyzing voting behavior) about one assigned individual in five was missed even after repeated calls. In Washington State similar proportions of both probability and quota samples were not completed prior to the election. In the case of the national polling organizations, however, the quotas as set were completed.

4. Interviewing procedure

It has been suggested that the interviewer may have influenced respondents to reply in the direction of the interviewer's own political preferences. Such influences undoubtedly do exist to a minor degree but it is impossible to measure them from the data that are available. The interviewers were more representative of the middle class and higher educated part of the population than of other groups. The possible conscious or unconscious influence of their social characteristics and political preferences should be studied in future polls. Considerable variation was found between individual interviewers who had similar assignments, and between resident interviewers and travelling crews.

We found many possibilities of error due to the interviewing but very little evidence on their magnitude and direction. While the polling organizations have done a great deal of work to improve their interviewing operations, they face many limitations in their selection, training and supervision of a widely scattered field force of part-time or full-time interviewers. It is not surprising, therefore, if the results are subject to an important degree of uncontrolled error due to the interviewing. Improvement of the questionnaire may be restricted by the limited abilities of interviewers to conduct more elaborate interviews.

Finally, our inability to evaluate the combined effect of field operations as potential sources of error is mainly due to the sad state of our basic knowledge of what really goes on in the interviewing process. Field operations proceed largely by rule of thumb, by the experience of people who have interviewed --there is no proven knowledge of what is correct procedure. Responsibility for lack of basic knowledge in this field must be laid at the doorsteps of pollsters and social scientists alike.

B. Analysis of the Responses

1. Finding out who will vote

The forecast of an election must rest on the voting intentions of persons interviewed who will actually vote since the preferences of others, at most, can influence the vote only indirectly. Most of the polls eliminated probable non-voters by their response to a simple question "Do you plan to vote?" or its equivalent. Roper only excluded those who volunteered that they did not expect to vote, while Crossley used a battery of questions to screen out those who were unlikely to vote. Some small part of the error may be due to this phase of the polls but, again, it cannot yet be measured.

2. The undecided voter

After the analysis of the responses has set aside these responses from persons unlikely to vote, it must determine as best it can how those persons will vote who report that they will vote but have not yet made up their minds. This group clearly includes some persons who do have a preference but choose not to tell it, and others who might have given a definite answer if the question had been asked in some other way. The percentage of undecided voters among persons planning to vote was about 15 percent in Gallup's poll. This was reduced to about 8 percent by separating out those who would tell which way they were leaning. The group of "undecided" voters in 1948 appears to have been much higher than the corresponding figures for 1944, but exact comparison is difficult, since the effect of the questions may have been different for the two elections.

Various guesses can be made about the reasons for the indecision of the undecided voters, but very little evidence about them is available. Post-election surveys suggest that more than half of the undecided group did vote favoring Truman roughly two to one. Careful calculations of the possible effects of a reasonable pre-election estimate of how the undecided might divide show that the effect on the total error would hardly be more than 1.5 percentage points and probably much less. Here again, the contribution to the total error from the handling of one phase of the polling procedure was undoubtedly small, but several such contributions in the same direction can add up to a large aggregate error.

3. Processing of data and adjustments

The Committee did not audit the procedures used by Crossley, Gallup and Roper in the processing of the questionnaire, namely, editing, coding and tabulation, but in its extensive use of the results it found errors from these operations to be negligible. The clerical operations in all three organizations were, on the whole, adequately supervised and controlled and performed with competence. Moreover, the Committee's own tabulation of records used by Roper and Gallup produced results consistent with their own tabulations. The Committee did not make any special tabulation of Crossley's data.

The polling organizations varied in their use of external data for adjusting their survey results. Gallup increased the accuracy of his predictions somewhat by making adjustments for apparent biases in the representation of the various levels of education in his sample and previous voting in the 1944 election. Roper and Crossley did not make such adjustments, although the procedure would have improved the accuracy of their forecasts.

Adjustments of the survey results to available external data is a desirable procedure when performed by proper methods and consistent with other features of the polling operation. For example, Gallup's adjustment of his sample for education had the effect of decreasing Dewey's expected vote by about 2.5 percentage points and increasing Truman's by 1.3 percentage points in terms of the total major party vote. Moreover, Gallup also analyzed his sample on the basis of the vote reported for 1944--that is, whether the person voted for Roosevelt, Dewey or Other (didn't vote, too young, didn't remember, etc.). Each of these groups was studied separately for 1948 voting intentions and then combined by using actual 1944 vote proportions corrected for the inflation in the 1948 response to the question on the 1944 vote. The correction for past voting behavior reduced the Dewey percentage by 1.2 percentage points and increased the Truman percentage by 0.8 points. In preparing final estimates, Gallup averaged

the results of the two methods of adjustment. The net effect, therefore, was to reduce Dewey's expected vote by about 1.8 points and to increase Truman's expected vote by 1.1 points.

Roper did not make any adjustments in his polls. In an attempt to offset the weakness inherent in quota sampling methods, certain features had been "built in" his sampling to balance certain consistent errors. He knew, for example, that his sample over-represented the South, giving it 26% of the quota although it cast only 17% of the popular vote. This was, in his mind, a hedge against that underestimation of the Democratic vote in Northern areas which had plagued forecasters in earlier canvasses. The effects balanced out well in previous elections, but in 1948 the correction was less helpful. It cut the Dewey vote 1.4 percentage points, increased the Thurmond vote 1.3 percentage points, and the Truman vote only 0.2 percentage points. The error in the regional quotas thus offset part, only part, of the other errors.

Roper's sample over-represented not only the South, but the higher educational groups in both 1944 and 1948. He "built in" an underestimate of the lower educational groups to offset the fact that they tend to turn out at the polls in smaller proportion than the more educated. Having made these assumptions, he considered it illogical to correct his tabulated figures for these "built in" biases. His system of compensating errors had made possible a highly accurate estimate in 1944. By his system, it seemed to him unwise to correct his tabulated figures for the educational bias for which he thought he had already made allowance, but it is of interest to note that if he had done so, his estimate for Truman's share of the vote would have been 44.6 percent, agreeing fairly well with Gallup's figure and thus coming much closer to, though well short of, the actual vote.

C. The Forecast

1. Making predictions for poll data

After a pre-election poll has been made and the results processed and calculations made to take account of such factors as probable non-voters, undecided voters and trends, there remains the problem of actually making an election prediction. In most instances, it was considered sufficient to base the prediction on the last survey preceding the election. This is true of most of the state and local polls which have been examined by the Committee, However, there were various deviations from this procedure in the case of the national polling organizations.

All three national polling organizations followed procedures in making their final predictions which tended to disregard or smooth out changes in voting intention during the later stages of the campaign. Roper did not make his prediction on the basis of his last survey (October 25-28), but on a re-percentaging of his September 9th release based on a survey made in August, leaving out the undecided vote.

Gallup's final national prediction was based on a combination of his last two national samples taken in the middle of October. Crossley's final national prediction was based on the consolidation of his state predictions, which in turn were based on using the last available interviews varying by economic strata

for the 48 individual states polled from August to October. In each case, therefore, the forecast did not reflect fully the preferences of votes immediately before the election.

Gallup used four types of operations in collecting data for his final state predictions--several different methods for the same states. To make final state predictions for some states, Gallup had to choose among varying results for the same state based on different methods and for varying periods in time. The decision to use a given set of results for prediction purposes for a given state, as in 1944, was largely in accordance with subjective judgment rather than objective statistical rules.

Crossley obtained individual state interviews at three times: in approximately the middle of August, September, and October. In August a full cross-section was taken in the 48 states. In September a full cross-section was taken in special states where the August results pointed toward a close race. In October, interviews were largely restricted to the lower economic levels on the assumption that voting intention in the higher economic levels would be stable. The final prediction was based on a combination of the latest interviews for each economic level, thus combining data collected from August to October in a number of states. This procedure was hazardous and had the effect of masking any trends in intentions to vote and consequent measurement of such trends.

In arriving at their predictions no systematic attempt was made by the three polling organizations to take into account political activities of labor unions and other groups in the various states and the possibilities of "coat-tail" influences of strong state candidates on the presidential vote in the various states. The Committee is aware of the great difficulties of attempting to take such factors into account, because very little is known about quantitative measurement of this type of political phenomena. Nevertheless, these political factors can affect the outcome of state elections considerably; as, for example, in Massachusetts and Iowa in the 1948 election. Their influence on elections may become particularly important in the last stages of the campaign, thus creating for the polls the hazard of shifting voters. Any effective treatment of these factors will depend on the further development of political science in these directions, but until they are taken into account adequately in the forecasts, greater caution is necessary in making flat predictions from poll data.

2. The possibilities of "last minute" shift to Truman

The possibility of "last minute" shifts of voters from Dewey (or one of the other candidates) to Truman has been widely discussed as an explanation of at least a part of the error made by the 1948 pre-election polls in predicting the winner. The main evidence on the operation of this factor comes from post-election polls. The reliability of this type of information is limited by possible errors of several kinds. The Committee has made careful analyses of national polls conducted by Gallup and Roper, a state poll conducted by the Washington Public Opinion Laboratory and a local poll in Elmira, N.Y., as well as other scattered sets of data.

The Committee and its staff have examined these post-election surveys very carefully. They all show a greater proportion of respondents who said after the election that they had voted for Truman than reported before the election that they intended to do so. The analysis is complicated by the

presence of the undecided voters before the election and some persons who refused to say how they voted after the election. Except in the resurvey of the same persons, there are sampling errors and other disturbances that obscure the comparison and the samples are too small to warrant confidence in the meaningfulness of small differences.

"Panel" studies, wherein the same individuals are re-interviewed during the campaigns, are especially illuminating since they show what each person said before the election and afterward. Thus they show the apparent shifts directly for each individual who voted differently than he said he would. In some instances, however, these apparent shifts may reflect errors of interviewing and response rather than actual voting behavior. All resurveys agree in showing that the previously undecided tended to say afterwards that they swung to Truman; also more voters said that they shifted from Dewey to Truman in the last two weeks than said they shifted in the opposite direction. Some supporting evidence is provided by answers on national pre-election polls to the question "When did you make up your mind how you would vote?" These results are consistent with the results of all panel studies in showing evidence of a substantial swing to Truman at the close of the campaign.

Looking at all the evidence now available, the Committee tentatively concludes that in the last two weeks there was a net shift to Truman of 2 or 3 percentage points probably varying quite a bit from state to state. The complexity of the problem of allowing for all the kinds of errors of interpretation is such that it is possible that the shift at the end of the campaign was even greater than these figures suggest. Indeed, there is a remote possibility that the shift could have been large enough to make Gallup and Crossley's last pre-election survey not too far off the mark, as of two weeks before the election. But this would simply have meant that other polling errors tended fortuitously to compensate better than now seems likely. It would provide no guarantee of future accuracy unless these other errors can be better controlled.

Some additional light will be thrown on the probable size of the shift when local election data have been analyzed, the effect of last minute campaign efforts have been studied, and the examination of supplementary opinion questions on some of the past-election surveys has been completed. While the shift may have been large enough to account for more of the error than seems likely just now, it cannot account for all the error in all of the polls. This is made clear by a comparison of the forecasts made by different organizations for the same state or for the entire nation. Differences thus revealed show that the errors must be ascribed not only to late campaign effects but also to the sampling, interviewing, and other parts of the polling methods used.

## VIII. PROBLEMS AND QUESTIONS BEYOND THE SCOPE OF THE COMMITTEE

The Committee has confined itself almost entirely to the technical aspects of pre-election polling and the problem of finding out something about the nature and size of the errors between the poll predictions and election results. The Committee does not consider that it can properly attempt to deal with many questions and problems which have been raised by the public and the press concerning the polls. These questions range from charges of fraud and sell-out to the Republicans or Democrats to the problem of whether pre-election polls hinder or help the democratic process.

As far as fraud and similar questions are concerned, the Committee found no evidence of it at any time and it considers such charges groundless.

Public discussion of the causes of Truman's victory places great emphasis on the part played by labor and the farm vote, which may have been responsible for almost all the last minute changes that occurred. Some direct evidence about the response of the farmers and efforts of labor organizations in many localities has been obtained. To analyze their effect on the state and national returns, however, one must have separate counts of the vote for the different farm and industrial counties. Local election returns are very difficult to obtain promptly and except for a few states could not be obtained and analyzed in time for this report. They are needed for a study of the turnout as well as shifts in the vote.

There are many applications of the polling or survey technique outside the field of pre-election polls and there are probably many persons who may expect the Committee to consider the implications of the failure of the 1948 pre-election polls to these other applications. The Committee does not consider that its function extends to a full discussion of these implications nor does it feel fully qualified to do so. Instead of attempting such a discussion, the Committee has tried to bring into clear focus the causes of errors of the pre-election polls and what research has to be done to further understand these causes and reduce the errors. It is hoped that such a treatment will be useful to those concerned with applications of polling and survey methods in other fields in seeing the implications of the failure of the pre-election polls on their own applications.

At this juncture, the Committee would like to point out that pre-election poll predictions constitute a severe although not always revealing test of polling or survey methodology. While it is hard to conceive of a better criterion than elections as a test of the polling of preferences for candidates, better tests are needed for polling on political issues and opinions. There are more steps in pre-election polling at which error can enter than in many other applications of polling and survey methodology. Good tests of the accuracy of polling methods are important in the development of polling and survey methodology. Without suitable tests there is danger that users of those methods will drift into an unjustifiably high degree of confidence in the accuracy of their current methods.

The Committee would be remiss not to point out, however, that in other kinds of surveys the available sampling survey techniques properly utilized have produced results with surprisingly little difference from those which were obtained from a complete canvass of the population. Competent sampling survey techniques used within their known limitations, and particularly where response errors are adequately controlled, can provide highly useful information at relatively low cost, with great speed and with small margins of error. The failure of pre-election forecasts successfully to pick a winner in no way constitutes a yardstick for measuring either the accuracy or usefulness of sampling surveys in fields in which the response does not involve the expression of opinion of intention to act.