It was the eve of the 2016 presidential election, and Natalie Jackson, senior polling editor for the *Huffington Post*, seemed supremely confident. She pegged Hillary Clinton as the near-certain winner. So did many pollsters, analysts, pundits, and journalists. But Jackson had ample data to support her confidence. Or so she thought.

Jackson coordinated the *Huffington Post*’s polls-based statistical model that was designed to forecast the election’s outcome. She did not equivocate in her final analysis. “The Huff Post presidential forecast model,” she wrote, “gives Democrat Hillary Clinton a 98.2 percent chance of winning the presidency. Republican Donald Trump has essentially no path to an Electoral College victory. Clinton’s win will be substantial, but not overwhelming. The model projects that she’ll garner 323 electoral votes to Trump’s 215.” Clinton, she added, “should fairly easily hold onto Michigan, Wisconsin and Pennsylvania” and stood better than an 80 percent chance of carrying Florida and North Carolina.1

A 98.2 percent probability of winning seemed unassailable. And yet, Jackson’s polls-based forecast wasn’t even the most adamant. Samuel Wang, a neuroscientist who forecasts election outcomes using his Princeton Election Consortium model, set Clinton’s win probability slightly higher, at 99 percent. Both models were derived from pre-election polls, including surveys conducted in key states.

Both were dramatically wrong. Trump did find a “path to an Electoral College victory,” an unlikely path in which he carried Michigan, Wisconsin, North Carolina.1

---

1. Of Pollsters, Journalists, and Presidential Elections

---

Introduction
Pennsylvania, and Florida, as well as North Carolina. Jackson returned to her forecast two days after the election and wrote, “It gutted me to realize I had been wrong.” No one, she added, “wants to feel like this—to be so utterly and publicly mistaken. People on Twitter have been calling for me to be fired. But what happened is done.” Within six months, she had left the Huffington Post.

“The problem,” Jackson wrote in her after-election assessment, “was that I placed way too much faith in polls. I assumed they would be right. . . . I kept looking at the consistency of the polls. They wavered in the exact margins, sure, but always showed Clinton winning in the key states that she needed to win. I saw no reason to question that the polls would be accurate overall. So I defended and stood by the numbers—as anyone who trusts their work does. That’s left me eating some crow.”

Jackson’s embarrassment was acute but hardly exceptional or without precedent in U.S. presidential elections. The 2016 election cycle abounded
with overoptimistic predictions that were buoyed by pre-election polls and proved embarrassingly wrong. They included a bold assertion by the chief political correspondent for Slate, Jamelle Bouie, whom the Poynter Institute for journalism studies proclaimed a “breakout star” of the campaign.4 “There is no horse race here,” Bouie wrote in late summer 2016. “Clinton is far enough ahead, at a late enough stage in the election, that what we have is a horse running by itself, unperturbed but for the faint possibility of a comet hitting the track. Place your bets accordingly.”5

Over the past eighty-five years and more, polls and poll-based forecasts have misfired in many ways in U.S. presidential elections, leaving pollsters, journalists, and pundits baffled or humiliated, and often without immediate explanation as to what went wrong. Polls in presidential elections do not always go wrong. Or dramatically wrong. But they have been wrong often enough to invite skepticism and wariness. Indeed, it is a rare election that does not produce polling controversies of some sort. As Mack C. Shelley and Hwarng-Du Hwang wrote more than thirty years ago, “The accuracy of presidential election polls has been argued in every presidential election year since polls first gained wide recognition” in the 1930s.6 Their observation is relevant still.

Just as no two presidential elections are quite alike, no two polling failures are precisely analogous. Not all polling failures are akin to the shock result of 2016. Their distinctiveness notwithstanding, polling failures tend to produce broadly similar effects—surprise, anger, bewilderment, and frustration at their failing to provide the American public with accurate clues about the most consequential of all U.S. elections.

Such reactions are hardly surprising, given that polls drive, color, and help fix news media narratives of presidential elections in the United States. They set expectations. They are central to how journalists, and Americans at large,
understand the dynamics of presidential campaigns. Polls are critical to shaping conventional wisdom about the competitiveness of those races. And they have long been recognized as such. Years ago, in a series of articles about polling, the *New York Post* recalled that in "the fall of 1948, the press of the nation, acting on the advisory of the political pollsters, in effect recorded an event that never took place—the election of Thomas E. Dewey as President of the United States." Or as George H. Gallup, opinion polling’s tireless evangelist, said about that polling failure, “We gave birth to a monster in 1948, the year when all of us pollsters elected Tom Dewey,” who lost to President Harry Truman.

This is not to say that polling failures are typically on the order of the epic fiasco of 1948 or the shock of 2016. In fact, some political scientists have argued that the record of election polling has been admirable in the United States and elsewhere. But polls have been in error often enough, and are beset by so many variables and potential contaminants (which pollsters don’t always discuss), that treating them cautiously is sensible. Election polls are not always accurate prophesies, and they certainly are not beyond challenge.

This book addresses in detail polling’s checkered record in U.S. presidential elections since 1936—a record that rarely has been considered in depth and never collectively. It is a history that is not especially well known. American journalists and, indeed, the American public are largely oblivious to the catalog of polling flubs and miscalls. They may be faintly familiar with the “Dewey defeats Truman” polling debacle of 1948. They may have a nebulous sense that election polls were in error in 2016. And in 2020. But little else. This unfamiliarity—along with the certitude that can attach to polls and poll-based forecasts—surely contributed to the immense surprise that greeted Trump’s split-decision victory over Clinton. He handily won the Electoral College; she clearly won the popular vote, and yet lost the election.

This book does not dwell on trivial shortcomings. The polling failures considered here—whether spectacular or somewhat more modest—were all surprising and controversial when they occurred. They were much commented on at the time. The book also addresses, and offers a fresh assessment about, the intriguing, intricate, and sometimes exasperating interplay between pollsters and journalists, a complex relationship that over the years has given rise both to sustained collaboration and to expressions of bitter disdain. The virus of poll-hatred is not especially potent in the news media nowadays, but for decades many prominent journalists reveled in their contempt for opinion polling and poll-takers. And yet, paradoxically, there has never been a time when polling was not
of some fascination to journalists—a fascination that persists, even if journalists are not entirely at ease with the intricacies or dynamics of survey research.

This book revisits prominent cases of polling failure while illuminating and presenting fresh insight into some of the characters, colorful and otherwise, who have shaped election polling and how it has been covered. Among these figures are George Gallup, the prickly founding father of public opinion research, who did much to promote and solidify expectations about the accuracy of election polling; Elmo Roper, a one-time retail jeweler and contemporary of Gallup, who by 1948 thought that election polling had little left to prove; and Warren J. Mitofsky, a brash yet brilliant innovator whose admonition to pollsters rings true across generations: “There’s a lot of room for humility in polling. Every time you get cocky, you lose.” Once-prominent journalists also enter the narrative; they include David Lawrence, Haynes Johnson, and Jimmy Breslin—all of whom were practitioners of “shoe-leather” journalism, an intensive and revered kind of out-of-the-newsroom reporting that frequently has been considered a response or alternative to indulgence in election polls.

The book does not consider at great length the well-documented troubles afflicting contemporary survey research—notably, the sharp and sustained declines in would-be respondents’ willingness to answer polls conducted by telephone, which once was the industry’s “gold standard” methodology. Since 1999 at least, pollsters have been experimenting with and incorporating internet-based approaches, generally with mixed results. Nor is the book steeped in the jargon and the opaque methodological arcana that pollsters and polling experts are keen to invoke, and for which they occasionally have been rebuked. Gallup, for example, once was taken to task for promoting impenetrable terms like “quindimensional analysis.”

This study is not one of methodologies, of the mechanics and different ways of conducting election polls. As such, only occasional attention will be devoted to topics such as “mode effects” or “nonresponse bias” or “non-probability internet panels.” Or “multilevel regression and post-stratification”—a statistical technique that pollsters know as “MRP” and refer to as “Mister P.” This is not to say such matters are trifling. They are not. But they are not central to this narrative study of the interplay of pollsters, journalists, and polling failure in modern U.S. presidential elections.

Polling failures and controversies arise from no singular template. Pollsters have forecast tight elections when landslides have occurred. They have pointed to the wrong winner in closer elections. The work of venerable pollsters has
been singularly and memorably in error. Exit polling has thrown Election Day into confusion. Failed state-level polls have upended widely anticipated national outcomes. Poll-based data aggregators have miscalled elections. Narratives like these make for compelling accounts of expectations dashed and hubris exposed; they are centerpieces of the chapters of this book that address polling failure in the elections of:

1936: It was the dawn of modern election polling, and it was a year of “polls gone mad,” as one newspaper described it. Polls seemed everywhere as the election drew near in 1936, and their conflicting findings fortified a sense that the race between President Franklin D. Roosevelt and Alf Landon was to be the closest in twenty years. The venerable Literary Digest magazine, whose mass mail-in survey never had been wrong, forecast a comfortable victory for Landon. Roosevelt swept to victory in one of the most lopsided elections in presidential history. Gallup, Roper, and Archibald Crossley debuted their quasi-scientific election polls that year, and all of them outperformed the Digest.

1948: The “Dewey defeats Truman” election humiliated Gallup, Roper, and Crossley, whose respective polls failed dramatically. Roper called the election for Dewey in early September, pledging to publish no further poll results unless a political catastrophe intervened. In late October, Roper quietly took another survey, but did not publicize the results because they showed Dewey still far ahead. President Harry S. Truman, who denounced the “sleeping polls” during the campaign, won by 4.5 percentage points. Journalists flayed themselves afterward for having leaned so heavily on the findings of the polls—and for having effectively delegated their legwork to pollsters.

1952: Mindful of their excruciating failure four years earlier, pollsters turned exceedingly cautious in interpreting the first Dwight Eisenhower–Adlai Stevenson election. At the end, they said the race was close, but Eisenhower was slightly ahead. They also said Stevenson, who seemed to be surging as the campaign closed, could win. He didn’t. Eisenhower swept to the presidency in a thirty-nine-state landslide that no pollster saw coming. As an editorial writer in Massachusetts observed, the pollsters in 1952 were “unable to tell a tidal wave from a photo finish.”

1980: Large news organizations had conspicuously entered pre-election surveying by 1980, and polls were more numerous than ever. The incumbent, Jimmy Carter, seemed locked in a tight race with Republican Ronald Reagan and the election may have turned on a late in the campaign surge of support for Reagan. Pollsters couldn’t agree. Reagan won handily—by near-landslide proportions that no poll had anticipated. To some pollsters,
the miscall evoked a lesson they presumably had learned in 1948: do not stop polling too soon.

2000: The closest presidential race in American history was decided by the U.S. Supreme Court, thirty-seven days after the election, in favor of George W. Bush. The delayed result revolved around the ambiguous outcome in Florida, a dispute that capped a trifecta of polling errors in 2000: the pre-election polls mostly signaled a tight race, but mostly pointed to the wrong popular vote winner; the Gallup organization’s daily tracking polls swung wildly, even implausibly, during the campaign; and exit polls in Florida erroneously indicated Al Gore was headed to victory in that pivotal state.

2004: Exit polls figured prominently in the 2004 election, erroneously indicating that Democrat John Kerry was destined to win the presidency. Exit-poll results seemed so certain that they inspired a top adviser to Kerry to refer to him as “Mr. President.” They also prompted George Bush to mope around the White House for a while, thinking he had been turned out of office after a single term, much as his father had. And exit-poll results touched off fierce but never-substantiated charges that the election had been stolen in Ohio, the state pivotal to the outcome.

2012: In a faint echo of the 1936 election, the most venerable election pollster—the Gallup Organization—miscalled the outcome, and was chastised and outperformed by a sports-loving data journalist named Nate Silver. He became Gallup’s bête noire. Silver’s poll-based statistical model correctly forecast the outcomes in all fifty states in 2012; Gallup picked the wrong winner. It estimated that Mitt Romney held a one-point advantage over the incumbent, Barack Obama. Romney lost by nearly four points. Gallup thereafter left election polling.

2016: Trump upset Clinton in an outcome that wasn’t supposed to happen. The statistical forecast models of Jackson, Wang, Silver, and others veered off target, largely because they relied on state polls that misfired in key places such as Michigan, Pennsylvania, and Wisconsin. Pollsters and journalists learned anew that election results can be correlated—that if a candidate does poorly in one state, she may do poorly in other states having similar demographic characteristics.

2020: Pre-election polls collectively turned in their worst performance in forty years by markedly underestimating support for Trump, who nonetheless lost reelection to Joe Biden. Some national surveys signaled that Biden would win by double-digit margins, but the race was much closer. Biden won the popular vote margin by 4.5 percentage points. But the well-placed shift of forty-three thousand votes to Trump in Arizona, Georgia, and Wisconsin would have produced a tie in the Electoral College.
To be sure, the universe of polling failure in U.S. presidential elections is not confined to those nine episodes. Although the book’s principal focus is on prominent polling missteps and errors, less flagrant miscalls, such as those of 1968, 1976, and 1996, won’t be ignored. They will be addressed, though not at chapter length: they will be incorporated into the narrative as relevant.16

This book argues that polling’s uneven, messy, and controversial past merits being addressed collectively and in its sweep. It proceeds in the recognition that polling failure often is correlated with journalistic failure. The correlation was pronounced in 2016, when poll-based prediction models and key state polls anticipated victory by Hillary Clinton. The correlation was likewise strong in 1948, 1952, and 1980.

So why write a narrative and interpretative study of prominent polling failure? Why focus on the interplay of journalists, pollsters, and failed polls in presidential elections? The motives are several. Presidential elections are the most consequential of campaigns, the most closely followed and most anticipated of all American elections. They represent the country’s only nationwide political race. No other campaigns are as intensely covered and analyzed by the news media. And polls are never more conspicuous—or commented on—than in presidential campaigns. Polling failures are most magnified in such races, and the stakes for pollsters and poll-based statistical analysts are seldom higher.

Polling, moreover, occupies an intriguing niche in American popular culture. Opinion polls exert an undeniable allure, even if their workings seem mysterious. Gimmick polls dreamed up in presidential elections—in the “Pullet Poll” of 1948, for example, farmers buying chicken feed preferred sacks stamped “a vote for the Democratic candidate” over the Republican option17—are hardly sophisticated or meant to be taken seriously. But they are expressions of popular interest in election polls and what they presume to foretell. That presumption can also invite sarcasm and satire—and occasionally the graphic wit of the cartoonist’s pen.

Examining the interplay of journalists, pollsters, and failed polls serves to highlight the priorities they share of accuracy and timeliness. Polls have long been integral to American journalism’s election coverage in part because they lend an impression of assurance amid the confusion and rhetorical conflict of political campaigns. As Irving Crespi, a former senior official for the Gallup Organization, once noted, “Interest in pre-election polls has always been based on the expectation that they can provide accurate advance indications of election outcomes.”18
Addressing the interplay also highlights recurring or unresolved phenomena in election polling. Opinion polling is a dynamic field given to experimentation, especially so in recent years as once-controversial methodologies like internet-based polling have entered mainstream practice. Still, the field is troubled by infirmities that seem chronic and can affect polling accuracy. “Pollster cockiness” is one of the infirmities.

It is a recurrent ailment that predates the first known use of “pollster”; the term appeared in Time magazine in 1939.cockiness certainly afflicted the Literary Digest, which declared in touting the launch of its ill-fated mail-in poll in 1936: "Who will win—Roosevelt or Landon? Will the country repudi-
ate the New Deal or give its leader a new, four-year mandate? To-day, nobody knows. But the Digest is seeking the answer—in the same way that has enabled it, time after time, to tell the country exactly what was going to happen when the voters went to the polls.”

The polling debacle of 1948 was another high moment in pollster cockiness. Elmo Roper was conspicuously adamant in predicting Dewey’s victory, delivering repeated assurances of the accuracy of his forecast in speeches, in newspaper columns, on radio programs. “In so far as the Presidential contest goes,” he told an audience at the Waldorf Astoria Hotel twelve days before the election, “the outcome is settled and has been settled for at least six or eight weeks. Mr. Dewey will be the next president, barring any error of catastrophic dimensions which he might make between now and election day. . . . Mr. Dewey is in, and we have found the campaigning to be so much Sound and Fury signifying little in the way of change of sentiment on the part of the voters.”

Campaigning, Roper figured, didn’t alter election trajectories much at all. After the election, Roper conceded that pollsters “had gotten pretty smug, and I was one of the smuggest of the lot.” He chided the public, though, for having “put too much dependence in the accuracy of the polls.”

The shock of 1948 left pollsters gun shy for years. In an unpublished memoir written in the late 1960s, Arch Crossley declared that the “polls were given a black eye in 1948 from which they have never fully recovered.” But the metaphoric black eye certainly didn’t prevent pollster cockiness from reemerging from time to time after 1948. In the afternoon on Election Day in 2004, exit-poll data pointing to John Kerry’s victory circulated among journalists and analysts. John Zogby, a pollster who was hailed for his accurate estimates in elections in 1996 and 2000, issued an updated prediction late in the day, as votes were still being cast. Kerry, he said, would win handily, carrying at least 311 electoral votes. Kerry won 251 electoral votes, and lost the election.

Zogby was nonchalant about his unconventional forecast. “I don’t know that anyone was hospitalized over my prediction,” he told the New York Times. “If there are any orphans that are out there, from the bottom of my heart, I apologize. We’ll try to start up a fund.”

The 2016 election also gave rise to misplaced cockiness. Samuel Wang, a poll-based forecaster at Princeton University, vowed to eat a bug on live television should Donald Trump win more than 240 electoral votes. Afterward, Wang conceded to having been “excessively certain” in forecasting Clinton’s victory and went on CNN to fulfill his pledge. With a hint of reluctance
stealing across his face, Wang dipped a spoon into a can of gourmet crickets, dug out a sample, and redeemed his promise.26

Late in the 2016 campaign, Ryan Grim, then the Washington bureau chief for the Huffington Post, declared that Clinton’s prospects of winning the election were unshakable. His assertion followed a widely noticed, late-campaign dustup on Twitter with Nate Silver, the data journalist who founded the FiveThirty-Eight predictions and analysis site. Grim accused Silver of interpreting polling data so that Trump’s prospects appeared stronger than they deserved to be.27 “If you want to put your faith in the numbers,” Grim wrote, alluding to the Huffington Post forecast model that showed Clinton almost certain of victory, “you can relax. She’s got this.” 28

Another recurring feature of election polling is vulnerability to surprise, which can be tied to developments late in presidential campaigns—the gaffes, dramatic gestures, and eleventh-hour disclosures that can confound poll estimates and expectations. These episodes materialize unpredictably, but with some frequency. Political history, as Politico has noted, “is littered with the charred remains of . . . late-in-the-election bombshells that scramble political calculus just as the stakes are at their highest.” 29

Some analysts, for example, have argued that Clinton lost the 2016 election when FBI director James Comey announced in late October that the FBI had reopened its inquiry into the private email server she had used while secretary of state.30 Barack Obama may have ensured his reelection in 2012 in a high-profile response to the devastation of Superstorm Sandy, which battered New Jersey days before the election. George W. Bush may have lost the popular vote to Al Gore in 2000 when a television reporter in Maine disclosed shortly before the election that Bush had been arrested for drunk driving in 1976— an episode Bush had never disclosed. President Jimmy Carter probably ensured defeat in 1980 by interrupting his campaign days before the election to return to the White House and focus on the prospective release of American diplomatic personnel taken hostage in Iran a year earlier. Carter’s move was futile, and had the effect of redirecting public attention to his administration’s most conspicuous foreign policy failure.

Dwight Eisenhower’s dramatic pledge late in the 1952 campaign to “go to Korea” in seeking an end to the war there may have altered the trajectory of what was mistakenly thought to be a close race. Thomas Dewey let the mask slip one afternoon in October 1948 when he berated as a “lunatic” the engineer of his “Victory Special” campaign train after it backed toward a crowd surging
to see him. The outburst did not cost Dewey the election, but the flash of temper offered insights into a buttoned-up candidate who ran a tightly controlled campaign that emphasized few specifics and abided no surprises.31

In a way, polling failure in presidential elections is not especially surprising. Indeed, it is almost extraordinary that election polls do not flop more often than they do, given the many and intangible ways that error can creep into surveys. And these variables may be difficult or impossible to measure or quantify.32

Unintended bias in wording the questions is an example. “Nothing matters if you’ve got a bad question,” Stephanie Marken, a survey methodologist for the Gallup Organization, once noted.33 “You can ask a question in such a way as to get any answer you want,” Elmo Roper said in an oral history interview more than fifty years ago.34 More subtly, there is some evidence that including, or excluding, a candidate’s job title can shape an election poll’s results.35

The order in which candidate-preference questions are posed, the day of the week interviews are conducted, and the gender of the interviewer can inject distortion into poll results. The way the survey is conducted—whether by telephone, internet, or mail, or in person—also can give rise to differing results. These are “mode effects,” as pollsters call them. Survey respondents can be a source of distortion, too. Pollsters have long recognized that respondents sometimes give answers that may be socially desirable or acceptable—while keeping their true feelings or unpopular opinions to themselves. This phenomenon was suspected as a factor in Trump’s unexpected victory in 2016. But pollsters and public opinion researchers report having found little evidence that a “shy Trump” effect was extensive, let alone decisive.36

Opinion polls can never flawlessly reflect the views of the entire population. It’s a statistical fact of life that some amount of error resides in every poll taken of some portion of a target group. This is true even when rigorous and reliable polling techniques are applied, such as taking pains to ensure that everyone in the target population has theoretically an equal chance of being interviewed, a key element in what is called probability sampling. The inevitable distortion in sample surveys is called the margin of error (or, more precisely, the margin of sampling error). A description of the margin of error often can be found in the small type accompanying newspaper reporting about election polls. Usually this caveat is described in phrases such as “This poll’s margin of error is plus or minus three or four percentage points.” Such a notice is an expression of the precision and reliability of the poll result.