Pew Research Center 💥

FOR RELEASE April 19, 2023

# How Public Polling Has Changed in the 21st Century

61% of national pollsters in the U.S. used methods in 2022 that differed from those of 2016

BY Courtney Kennedy, Dana Popky and Scott Keeter

FOR MEDIA OR OTHER INQUIRIES:

Courtney Kennedy, Vice President, Methods & Innovation Scott Keeter, Senior Survey Advisor Nida Asheer, Communications Manager Talia Price, Communications Associate 202.419.4372

www.pewresearch.org

**RECOMMENDED CITATION** 

Pew Research Center, April 2023, "How Public Polling Has Changed in the 21st Century"

# **About Pew Research Center**

Pew Research Center is a nonpartisan, nonadvocacy fact tank that informs the public about the issues, attitudes and trends shaping the world. It does not take policy positions. The Center conducts public opinion polling, demographic research, computational social science research and other data-driven research. It studies politics and policy; news habits and media; the internet and technology; religion; race and ethnicity; international affairs; social, demographic and economic trends; science; research methodology and data science; and immigration and migration. Pew Research Center is a subsidiary of The Pew Charitable Trusts, its primary funder.

© Pew Research Center 2023

# How we did this

This study looks at how national public opinion polling in the United States changed from 2000 to 2022. It focuses on two aspects: the *sample source(s)* – that is, where the survey takers came from – and the *mode(s)*, or how they were interviewed. The study tracks both the growth in the number of national pollsters and changes to how they conduct their public polls.

The unit of analysis is the polling organization. The <u>dataset</u> accompanying this report lists the 78 organizations included. The <u>Documentation URLs</u> file provides the webpages used to code the methods. The <u>Codebook</u> details what each code means.

Center researchers compiled the poll information in several stages and using a variety of sources. The initial coding of sample source and mode of interview was based on information available from sources such as pollster websites, news articles, press releases and the <u>Roper iPoll</u> data archive. After compiling the data, Center staff attempted to contact each organization and ask them to confirm whether the information for that organization was accurate. Most organizations that responded confirmed that the information gathered was accurate. When organizations provided corrections or additions, staff updated

the study database with that information. Each of these steps is described in greater detail in the <u>Methodology</u>.

### Number of active national public pollsters in the study dataset

Year	National public
rear	pollsters active
2000	29
2002	28
2004	33
2006	30
2008	42
2010	45
2012	49
2014	43
2016	54
2018	57
2020	73
2022	69

Source: Pew Research Center analysis of external data. See Methodology for details. "How Public Polling Has Changed in the 21st Century"

**PEW RESEARCH CENTER** 

# Terminology

**Address-based sampling (ABS).** ABS refers to recruiting survey takers by selecting a random sample of residential addresses from a master list, such as that maintained by the U.S. Postal Service.

**Interactive voice response (IVR).** IVR refers to a poll that entails automatic dialing of telephone numbers and a recorded voice asking a series of survey questions. They are sometimes referred to as "robo-polls."

**Mode of interview.** This refers to the format in which respondents are presented with and respond to survey questions. Surveys can be administered by an interviewer, or they can be self-administered. The most common interviewer-administered mode in public polling is live telephone. The most common self-administered modes include online, text message and paper.

**Method.** This study uses the term "method" broadly, referring to the source of the respondents (the "sampling frame") and how they were interviewed (the "mode"). This study describes a change in either of those as a change in method.

**Multiple methods.** Sometimes pollsters use <u>multiple sample sources</u> or <u>multiple interview</u> <u>modes</u> *within* a poll. Other times pollsters use multiple sample sources or multiple interview modes within the same year but in separate polls. The study describes any of the above as a pollster using multiple methods that year.

**Survey panel.** This is a group of people who have agreed to take surveys on an ongoing basis. The survey panels documented in this study each have thousands (and in some cases tens of thousands) of members.

**Probability-based panel.** This refers to a national survey panel recruited using random sampling from a database that includes most people in the population. Today, most such panels in the U.S. recruit by drawing random samples of residential addresses or telephone numbers. Typically, data collection with these panels is done online. However, some of these panels interview a small fraction of respondents (usually about 5% or fewer) using an offline mode such as live telephone. These panels are "probability-based" because the chance that each address or phone number was selected is known. However, the chance that each selected person will join the panel or take surveys after joining is not known.

#### **4** PEW RESEARCH CENTER

**Online opt-in polls.** These polls are recruited using a variety of methods that are sometimes referred to as "convenience sampling." Respondents are not selected randomly from the population but are recruited from a variety of online sources such as ads on social media or search engines, websites offering rewards in exchange for survey participation, or self-enrollment in an opt-in panel. Some opt-in samples are sourced from a panel (or multiple panels), while others rely on intercept techniques where respondents are invited to take a one-off survey.

**Registration-based sampling.** Some election polls sample from an official list of registered voters. All states are required to maintain a computerized and up-to-date list of voters and to make these lists publicly available for non-commercial purposes such as voter outreach and research.

**Sampling frame.** A sampling frame is a list of the population of interest. For a survey of the public, it is typically telephone numbers or residential addresses and ideally includes all members of the population (though in practice there are often gaps and omissions). The survey sample is selected from this list.

**Sponsor.** In this report, a survey sponsor is an organization that publicly releases results from a poll conducted on its behalf. Survey sponsors typically conceive of the study and either provide or obtain funding for it. Sponsors and vendors sometimes share the practical tasks involved in conducting a survey, such as when a sponsor drafts the questionnaire, and the vendor creates the sample and collects the data. This report uses "sponsor" and "pollster" interchangeably.

**Text message polling.** Text messaging is used by some organizations to contact a sample of cellphone numbers for the purpose of either directing respondents to an online survey or asking them a set of survey questions using a series of text messages and responses.

**Vendor.** In this report, the survey vendor is the organization that collects the survey data. The full set of tasks necessary for a survey are often shared between the sponsor and vendor, with the exact mix being determined by the specific expertise of the two parties and other factors. Sometimes vendors are also sponsors, either alone or in partnership with other sponsors.

# How Public Polling Has Changed in the 21st Century

# 61% of national pollsters in the U.S. used methods in 2022 that differed from those of 2016

The 2016 and 2020 presidential elections left many Americans wondering whether polling was broken and what, if anything, pollsters might do about it. A new Pew Research Center study finds that most national pollsters have changed their approach since 2016, and in some cases dramatically. Most (61%) of the pollsters who conducted and publicly released national surveys in both 2016 and 2022 used methods in 2022 that differed from what they used in 2016. The study also finds the use of multiple methods increasing. Last year 17% of national pollsters used at least

#### Polling has entered a period of unprecedented diversity in methods Number of national public pollsters in the U.S. using method(s) By 2022 the number of active pollsters more than doubled 7 Other and methods diversified 3 Online opt-in, probability-based panel and live phone Online opt-in and live phone About 29 national public 28 Online opt-in only pollsters were active in 2000, and nearly all used live phone IVR alone or with other methods ABS (USPS) with multiple modes alone or with other methods Probability-based panel and live phone 7 Probability-based panel only Live phone (RBS) only Live phone (RDD) only 2000 2004 2008 2012 2016 2020 '22

Note: Figures represent the number of active national public pollsters in each year and the method(s) that they used. IVR refers to interactive voice response, also known as robo-polling. ABS refers to address-based sampling. RBS refers to voter registration-based sampling. RDD refers to random-digit-dial sampling.

Source: Pew Research Center analysis of external data. See Methodology for details.

"How Public Polling Has Changed in the 21st Century"

#### **PEW RESEARCH CENTER**

three different methods to sample or interview people (sometimes in the same survey), up from 2% in 2016.

This study captures *what* changes were made and approximately *when*. While it does not capture *why* the changes were made, public commentary by pollsters suggests a mix of factors – with some adjusting their methods in response to the profession's recent election-related errors and others reacting to <u>separate industry trends</u>. The cost and feasibility of various methods are likely to have influenced decisions.

This study represents a new effort to measure the nature and degree of change in how national public polls are conducted. Rather than leaning on anecdotal accounts, the study tracked the methods used by 78 organizations that sponsor national polls and publicly release the results. The organizations analyzed represent or collaborated with nearly all the country's best-known national pollsters. In this study, "national poll" refers to a survey reporting on the views of U.S. adults, registered voters or likely voters. It is not restricted to election vote choice (or "horserace") polling, as <u>the public opinion field is much broader</u>. The analysis stretches back to 2000, making it possible to distinguish between trends emerging before 2016 (e.g., migration to online methods) and those emerging more recently (e.g., reaching respondents by text message). Study details are provided in the <u>Methodology</u>. Other key findings from the study include:

**Pollsters made more design changes after 2020 than 2016.** In the wake of the <u>2016</u> presidential election, it was unclear if the polling errors were an anomaly or the start of a longerlasting problem. <u>2020 provided an answer</u>, as most polls understated GOP support a second time. The study found that after 2020, more than a third of pollsters (37%) changed how they sample people, how they interview them, or both. This compares with about a quarter (26%) who made changes after 2016. As noted above, though, these changes did not necessarily occur because of concerns about election-related errors.

#### The number of national pollsters relying exclusively on live phone is declining

**rapidly.** Telephone polling with live interviewers dominated the industry in the early 2000s, even as pollsters scrambled to adapt to the rapid growth of cellphone-only households. Since 2012, however, its use has fallen amid <u>declining response rates</u> and increasing costs. Today live phone is not completely dead, but pollsters who use it tend to use other methods as well. Last year 10% of the pollsters examined in the study used live phone as their *only* method of national public polling, but 32% used live phone *alone or in combination* with other methods. In some cases, the other methods were used alongside live phone in a single poll, and in other cases the pollster did one poll using live phone and other polls with a different method.

**Several key trends, such as growth of online polling, were well underway prior to 2016.** While the <u>2016</u> and <u>2020</u> elections were consequential events for polling, the study illustrates how some of the methodological churn in recent years reflects longer-term trends. For example, the growth of online methods was well underway before 2016. Similarly, some live phone pollsters had already started to sample from registered voter files (instead of RDD, random-digit dialing) prior to 2016.

Use of probability-based panels has become more prevalent. A growing number of pollsters have turned to sampling from a list of residential addresses from the U.S. Postal Service database to draw a random sample of Americans, a method known as address-based sampling (ABS). There are two main types of surveys that do this: one-off or standalone polls and polls using survey panels recruited using ABS or telephone (known as probability-based panels). Both are experiencing growth. The number of national pollsters using probabilitybased panels alone or in combination with other methods tripled from 2016 to 2022 (from seven to 23). The number of national pollsters conducting one-off ABS surveys alone or in combination with other methods during that time rose as well (from one in 2016 to seven in 2022).

#### Polling on probability-based panels is becoming more common

Number of public pollsters in the U.S. using probabilitybased panels alone or in combination with other methods



Note: A probability-based panel is a group of people who agree to take surveys on an ongoing basis and who are recruited via random sampling of home addresses and/or phone numbers, depending on the panel.

Source: Pew Research Center analysis of external data. See Methodology for details.

"How Public Polling Has Changed in the 21st Century"

#### **PEW RESEARCH CENTER**

The growth of online opt-in among national pollsters appears to have paused after 2020. The number of national pollsters using convenience samples of people online ("opt-in sampling") – whether alone or in combination with other methods – more than quadrupled between 2012 and 2020 (from 10 to 47). In 2022, however, this number held flat, suggesting that the era of explosive growth could be ending.

Whether changes to sample sources and modes translate into greater accuracy in presidential elections remains to be

**seen.** The fact that pollsters are expanding into new and different methods is not a guarantee that the underrepresentation of GOP support occurring in 2016 and 2020 preelection polls has been fixed. Polling accuracy <u>improved in</u> <u>2022</u>, but this represents only one nonpresidential election.

# Growth of online opt-in methods in national public polls paused between 2020 and 2022

Number of national public pollsters in the U.S. using online opt-in samples alone or in combination with other methods



#### Notable study limitations

A study of this nature requires difficult

decisions about what exactly will be measured and what will not. This study focuses on two key poll features: the *sample source(s)* – that is, where the respondents came from – and the *mode(s)*, or how they were interviewed. While important, these elements are not exhaustive of the decisions required in designing a poll. The study did not attempt to track other details, such as weighting, where public documentation is often missing. Because the study only measured two out of all possible poll features, estimates from this study likely represent a *lower bound* of the total amount of change in the polling industry.

Another limitation worth highlighting is the fact that state-level polls are not included. Unfortunately, attempting to find, document and code polling from all 50 states and the District of Columbia would have exceeded the time and staff resources available. A related consideration is that disclosure of methods information tends to be spottier for pollsters who exclusively work at the state level, though there are some exceptions. It is not clear whether analysis at the level of detail presented in this report would be possible for state-only pollsters.

While not necessarily a limitation, the decision to use the *polling organization* rather than individual *polls* as the unit of analysis has implications for the findings. The proliferation of *organizations* using online methods implies but does not prove that online *polls* grew as well. However, research conducted by the American Association for Public Opinion Research (AAPOR) following the 2016 and 2020 elections reveals an explosion in the <u>share of all polling done using</u> <u>online methods</u>. AAPOR estimated that 56% of national polls conducted shortly before the 2016 election used online methods; the comparable share for 2020 was 84%. More details on the strengths and weaknesses of the study are presented in the <u>Methodology</u>.

#### Changes in methods are driven by many considerations, including costs and quality

In an attempt to verify the accuracy of the categorization of polling methodologies, researchers attempted to contact all organizations represented in the database. Several pollsters contacted for this study noted that use of a particular method was not necessarily an endorsement of methodological quality or superiority. Instead, design decisions often reflect a multitude of factors. Survey cost – especially the increasing cost of live phone polling – came up repeatedly. Timing can also be a factor, as a design like address-based sampling can take weeks or even months to field. As noted above, this study does not attempt to address *why* each organization polled the way they did. It aims only to describe major changes observable within the polling industry. Nor does it evaluate the quality of different methods, as a multitude of other studies address that question.

# Changes to polling after 2020 differed from those after 2016

The study found a different kind of change within the polling industry after 2020 versus 2016. After 2020, changes were both more common and more complex. More than a third (37%) of pollsters releasing national public polls in both 2020 and 2022 changed their methods during that

interval. By contrast, the share changing their methods between 2016 and 2018 was 26%.

The nature of the changes also differed. About half of the changes observed from 2016 to 2018 reflected pollsters going online – either by adding online interviewing as one of their methods or fully replacing live phone interviewing. By contrast, the changes observed from 2020 to 2022 were more of a mix. During that period, some added an approach like text messaging (e.g., Change Research, Data for Progress), probability-based panels (Politico, USA Today) or multiple new methods (CNN, Wall Street Journal). About a quarter of the change observed from 2020 to 2022 reflected pollsters who had already moved online dropping live

# More than a third of national public pollsters changed how they poll after 2020

% of pollsters who changed how they sample or interview people in national public polls in the U.S. during the two-year interval



Note: In this study change refers to using a different sample source or a different mode of interviewing. Figures for each interval are based on the set of pollsters that released at least one national public poll in both the starting year and ending year of the interval. Source: Pew Research Center analysis of external data. See Methodology for details. "How Public Polling Has Changed in the 21st Century"

**PEW RESEARCH CENTER** 

phone as one of their tools (e.g., CBS News, Pew Research Center).

A look at change over the entire recent period – from 2016 to 2022 – finds that more than half of national public pollsters (61%) used methods in 2022 that differed from those they used in 2016. As noted above, if features like weighting protocols were included in the analysis, that rate would be even higher.

A longer view of modern public polling (going back to 2000) shows that methodological churn began in earnest around 2012 to 2014. That was a period when about a third of national pollsters changed their methods. Change during that period was marked by pollsters starting to migrate away from live telephone surveys and toward online surveys.

## Pollsters increasingly use multiple methods - sometimes three or more

Pollsters are not just using *different* methods, many are now using *multiple* methods, the study found. Here again there is a discernable difference in how polls changed after 2016 and how they

changed after 2020. After 2016, the share of pollsters using multiple methods remained virtually unchanged (30% in both 2016 and 2018). After 2020, however, the share climbed to 39%. Notably, the share of pollsters using three or more different methodologies in their national public polls tripled from 5% in 2020 to 17% in 2022.

In this analysis, "multiple methods" refers to use of multiple sample sources (e.g., registered voter files and random-digit dial) or multiple interview modes (e.g., online, mail, live telephone). In some cases, several methods were used in a single poll. In other cases the pollster did one poll using one method and another poll using another method.

As an example, in 2014 Pew Research Center switched from

# Growing share of national pollsters are using multiple methods

% of national public pollsters in the U.S. using this many methods in polls they released each year

0	1 method	2 methods	3+ methods
2022	61%	22%	17%
2020	62	33	5
2018	70		26 4
2016	70		28 2
2014	81		19
2012	84		16
2010	87		13
2008	93		7
2006	93		7
2004	97		3
2002	93		7
2000	97		3

Note: A pollster is coded as using more than one method if they used more than one type of sample source (e.g., registered voter file, random-digit dial) or more than one interview mode (e.g., online, live phone).

Source: Pew Research Center analysis of external data. See Methodology for details. "How Public Polling Has Changed in the 21st Century"

PEW RESEARCH CENTER

exclusively using live phone with random-digit-dial sample to also <u>using a probability-based panel</u>. In 2020 the Center added an additional method, one-off address-based sample surveys offering

online or mail response. By 2022, the Center dropped live phone polling. Pollsters that used at least three different methods in 2022 include CNN, Gallup, NPR, Politico and USA Today.

### Text messaging and address-recruited panels see growth after 2020

An overarching theme in the study is the growth of new methods. Analysis earlier in this report aimed to describe trends for the most prominent methods. In the past, pollsters often used just one method (e.g., live phone with random-digit dial). That has changed. Today pollsters tend to use new methods (such as text) as one of several ways that they reach people. To track the trajectory of these newer methods, it helps to consider the number of pollsters using the method by itself *or* in combination with other methods.

A prime example is text message polling. An extremely small share of pollsters conduct national public polls exclusively by text. A larger share use text alongside another method, such as online opt-in.



#### Texting gains some traction in national polling in 2022

Number of national public pollsters in the U.S. using method alone or in combination with other methods

Note: Pollsters vary in how they use text messaging. In some cases respondents receive a text with a web link for an online survey. In other cases, respondents answer the questions via text.

Source: Pew Research Center analysis of external data. See Methodology for details.

"How Public Polling Has Changed in the 21st Century"

#### PEW RESEARCH CENTER

How texting is used varies. In some cases respondents receive a text with a web link for an online survey. In other cases, respondents answer the questions via text. Among the pollsters in this study, just one used texting in a national public survey in 2020. In 2022 that number rose to nine, representing 13% of the active national pollsters tracked that year. These figures reflect the number of pollsters using texting alone or in combination with other methods like live phone.

Analysis looking at methods used either alone or in combination with other approaches also suggests a change in the trajectory of online opt-in polling. While online opt-in usage grew tremendously between 2006 and 2020, that growth appears to have slowed if not stopped in 2022 for national polling.

By contrast, the share of national pollsters turning to probability-based panels continues to grow. In 2022 a third (33%) of national pollsters used probability-based panels either alone or in combination with other methods. This is up from roughly 10% during most of the 2010s.

Live phone was once the dominant method of polling but has been in decline since 2016. As of 2022, about a third of national pollsters used live phone alone or in combination (32%), while a much smaller share relied on it as their only method (10%).

The study also tracked the adoption of a specific kind of opt-in sample – members of an online opt-in panel who are matched to a record in a national registered voter file. This study first observed that approach in 2018. In 2018, 2020 and 2022, about 3% to 5% of national public pollsters used online opt-in samples matched to registered voter files, the study found.

# **Acknowledgments**

This report was made possible by The Pew Charitable Trusts. Pew Research Center is a subsidiary of The Pew Charitable Trusts, its primary funder.

This report is a collaborative effort based on the input and analysis of the following individuals:

Research team	Communications and editorial
Courtney Kennedy, Vice President, Methods and Innovation Dana Popky, Associate Panel Manager Scott Keeter, Senior Survey Advisor	Rachel Drian, Associate Director, Communications Nida Asheer, Communications Manager Talia Price, Communications Associate David Kent, Senior Copy Editor
Methodology	
Arnold Lau, Research Methodologist	Graphic design and web publishing
	Bill Webster, Senior Information Graphics Designer Sara Atske, Associate Digital Producer

We are grateful to the many colleagues across the field who responded to our request to confirm or correct information we had gathered about their national public polling. Several individuals went above and beyond, offering insights about why various changes occurred. We appreciate everyone who took the time to respond. Within Pew Research Center, we thank Carroll Doherty and Claudia Deane for providing useful feedback on a draft of the report.

We are also grateful to organizations that maintain databases of public polls, which facilitated some of the data collection for this study. These include the Roper iPoll, the Inter-university Consortium for Political and Social Research, the Louis Harris Data Center operated by the University of North Carolina Odom Institute, and FiveThirtyEight.com.

# Methodology

This study was designed to examine changes in national public opinion polling in the United States from 2000 to 2022. The study focuses on two key features: the sample source(s) (where the respondents came from) and the mode(s) (how they were interviewed).

The study's unit of analysis is the organization sponsoring the polling. In total, 78 such organizations are included. Pew Research Center staff coded the sample sources and interview modes used in national public polls for each organization during each even-numbered year from 2000 to 2022. Odd-numbered years were excluded on purely practical grounds. Hundreds if not thousands of national public polls are released each year, and processing them for this report required substantial labor. Focusing on even-numbered years cut the manual labor roughly in half. As shown in the report, the even-numbered-year approach was able to successfully track major changes in the industry.

The initial coding of sample source and mode of interview was based on information available from a variety of sources including pollster websites, news articles, and the <u>Roper iPoll</u> opinion poll data archive. After this data was compiled, Center staff attempted to contact each organization and ask them to confirm whether the information for that organization was accurate. Most organizations that responded confirmed that the information gathered was accurate. When organizations provided corrections or additions, the study database was updated with that information. Each of these steps is described in greater detail below.

# Inclusion criteria

The study aimed to examine change over time in national public polls. To be included, an organization needed to sponsor at least one national public poll in two or more of the years studied (i.e., the even-numbered years from 2000 to 2022). Organizations that sponsored a national public poll in only one of these years are not included. This criterion helped to reduce the influence of organizations that were not consistently involved in polling.

The national polls examined in this study are those based on the general public (e.g., U.S. adults ages 18 and older), registered voters or likely voters. Polls that exclusively surveyed a special population (e.g., teachers) were not included, as they often require unique designs. Additionally, polls described as experimental in their methods or research goals are not included.

# Who is a 'pollster' in this study

One important choice in designing this type of project is deciding whether to study the organization *sponsoring a poll* or the organization *collecting the data* (known as the "vendor"). For example, many Pew Research Center polls are fielded by Ipsos. Pew Research Center is the sponsor, and Ipsos is the vendor. This study focuses on sponsors and refers to them as the "pollster." There are several reasons why sponsors are the focus rather than vendors, including:

- The sponsor is typically the organization commissioning the poll and attaching their institutional name to its public reporting. Sponsors decide whether to make poll results public or keep them private. Generally speaking, sponsors are the entity most responsible for any public reporting from the poll.
- Sponsors dictate, in broad terms, the budget available for a poll, and the vendor works within that constraint to find a design that fits within that budget. While the vendor will often decide the exact final price tag, they usually are reacting to information from the sponsor as to whether the budget available is, say, \$10,000 or closer to \$100,000. In other words, whether a poll uses a very expensive method or very inexpensive method is generally dictated by the sponsor.

One complication is that sometimes the sponsor and vendor are one and the same. Continuing with the Ipsos example, in addition to the polls they conduct on behalf of clients, Ipsos also conducts national polling and releases the results by themselves. Accordingly, Ipsos is among the 78 pollsters in the analysis based on polling that they (co-)sponsor. Other vendors only conduct work on behalf of clients and never put themselves in the role of sponsor. This explains why a few major companies that conduct polling do not necessarily appear in the study dataset. Their work is represented in the rows associated with their clients.

Admittedly, focusing on the sponsor rather than the vendor feels more appropriate in some cases than others. For example, some sponsors are very engaged in decisions about how their polls are conducted, down to the smallest detail. Referring to such sponsors as the "pollster" feels accurate. Other sponsors are much more hands-off, letting the vendor make most or all decisions about how the poll is conducted. In these cases, it is tempting to think of the vendor as the pollster.

Ultimately, to execute a study like this, researchers must rely on information in the public domain. Nuanced records of who made which decisions are simply not available, nor is it possible to gather such information about polls fielded as early as 2000. Of the options available, focusing on the sponsor was the best fit for this study, but we acknowledge that in some cases the sponsors were likely not deeply engaged with decisions about sampling frames and interview mode.

Another complication is that some polls have multiple sponsors. For example, ABC News and The Washington Post have a <u>long-standing partnership</u> for live phone polling. In addition, they sponsor polls either solo or with other partners. In this study, each pollster is credited with using the methods employed for national public polls that they either sponsored or co-sponsored. For example, this study's records for both ABC News and The Washington Post reflect the use of live phone with random-digit-dial sample for the years in which they jointly sponsored such polls.

A few well-known, recurring national surveys were considered but ultimately excluded from this analysis because they are too different from a typical public opinion poll. <u>The General Social</u> <u>Survey</u> and <u>American National Election Survey</u> both measure public opinion, but their budgets and timelines are an order of magnitude different from a typical public opinion poll. On that basis, we decided it was unhelpful to include them. Similarly, the <u>National Election Pool (NEP) poll</u> and <u>VoteCast</u> are not included because they are designed specifically to cover election results not just nationally but at the state level. Their methodological and budget considerations are quite different from those of ordinary opinion polls. These studies are very valuable to opinion research in the U.S., but they are not comparable to the polling studied in this project.

# Which polling organizations were included

The analysis is based on 78 organizations. Each organization sponsored and publicly released national poll results in at least two of the years studied (i.e., the even-numbered years from 2000 to 2022). There is no authoritative list of such organizations, and experts might disagree on whether certain edge cases should be included. Several gray areas required resolution.

#### Inclusion based on content

In the broadest sense, public opinion exists for many topics – from politics and the economy to pop culture and brand preferences. That said, the national dialogue around "polls" and "polling" is generally understood to be focused more narrowly on public affairs, politics and elections. For instance, a market research survey on public preferences in ice cream flavors is probably not what comes to mind when someone is asked about public opinion polling. At the other extreme, surveys measuring support for candidates running for public office (known as "the horserace") perhaps represent the prototypical conception of a poll. Many polls fall in between – measuring attitudes about important national issues but not measuring the horserace.

Each of the 78 organizations included in this study has a track record of measuring public attitudes about public affairs, politics and elections. Not all organizations included here specifically measure

the horserace, but all of them have asked the public about factors influencing how they might vote in an election.<sup>1</sup>

#### Academic research versus public opinion polling

The decision to consider a sponsoring organization as the pollster raised two practical questions when considering academic organizations. Colleges and universities may have multiple individuals or entities independently conducting polls within them, and so one question is whether these separate polling efforts should be considered as different pollsters or not. The second question is whether (or which) national polls conducted by faculty primarily for academic purposes should be included.

Although some academic institutions have more than one branded entity conducting public polls on politics and policy, publicly released surveys typically carried the institutional name. Moreover, there was no practical way to ascertain the degree of independence of the entities within a university. As a result, the decision was made to code all polling that met the study's content criteria for inclusion as polling by that university.

However, this study does not contain all the surveys conducted by every college and university. Indeed, it would be nearly impossible to locate all such polling, since most of it is made public only through academic papers or conference presentations. We acknowledge that as a limitation of the study. This study does not purport to represent every national survey whose results can be found somewhere in the public domain. The ones that are included tend to have a news media partnership increasing their visibility; maintain a public website updated regularly with the latest polling results; and/or are archived in the Roper iPoll public opinion poll repository. It is worth underscoring that the goal of this study is to describe the nature and degree of changes in national public polling from 2000 to 2022. Research designed primarily for an academic audience (e.g., peer-reviewed journals) is not the type of polling people have in mind when they question whether polling still works after the 2016 and 2020 elections.

Among the news media organizations included and contacted in this study, only one raised the concern that multiple units within the organization might be conducting polls. Most of their polling was coordinated through their political unit, but a few polls over the years were not. Mapping out the decision structures within each organization was outside the scope of this project. In the interest of applying the same standard to each organization, this study includes the methods for any public polls we found during the years studied. In some cases, this yields a track record

<sup>&</sup>lt;sup>1</sup> Two organizations in the initial list were judged to focus on topics too far afield of politics, elections and policy and were excluded. <u>The</u> <u>Hollywood Reporter</u> sponsors national public polls focused on the entertainment industry. The <u>University of Michigan Surveys of Consumers</u> measures attitudes and perceptions about prices and economic conditions.

that does not reflect centralized decision-making. In all cases, though, this approach reflects what the public sees – either a poll was or was not sponsored by "Organization X."

# **Data collection**

#### Creating a list of national public pollsters

There is no authoritative list of organizations sponsoring and releasing results from national public opinion polls, so researchers constructed one using a variety of sources. First, researchers compiled the names of organizations releasing national estimates for U.S. presidential approval or horserace estimates for U.S. presidential elections going back to 2000. For this task, researchers used polling databases and summaries from <u>The Polling Report</u>, <u>FiveThirtyEight.com</u>, <u>RealClearPolitics.com</u> and <u>Wikipedia</u>. Researchers then expanded the list with the names of prominent national polling organizations (e.g., <u>Kaiser Family Foundation</u>, the American Enterprise Institute's <u>Survey Center on American Life</u>, <u>PRRI</u>) that do not necessarily appear in those sources. Finally, additional sponsors were identified through polling partnerships. For example, if researchers saw that Pollster A co-sponsored a few polls with Pollster B, then the researchers investigated all the polling associated with Pollster B. If Pollster B qualified for inclusion, they were then added to the study.

#### Determining which methods each pollster used in national public polling for each year

For each pollster in the study, researchers set out to document which sampling frames and interview modes the pollster used for national public polls in each even-numbered year from 2000 to 2022. Unfortunately, this kind of information cannot be found in any one location. Indeed, one of the main motivations for this study was that existing databases are insufficient for understanding key distinctions in modern polling. Existing resources might indicate whether a poll was done by "phone" or "online," but there is often no information about the sample source. Consequently, Center researchers scoured the internet for more detailed documentation. Researchers executed this work in several steps.

1. Internet search for pollster and year. Researchers conducted a Google search for "[POLLSTER NAME] poll [YEAR]" starting with 2022 and working backwards, doing even-numbered years only. For each year, they limited the search time frame to 01/01/[YEAR] to 12/31/[YEAR]. They then investigated each of the hits on the first page of results. The poll in question could be disregarded if it was not sponsored by the pollster of interest and/or if the poll was fielded in the previous year (an odd-numbered year). Next, researchers conducted a Google search for "[POLLSTER NAME] survey [YEAR]" in the same manner. This was done because some organizations use the term "survey" instead of

"poll." This searching often yielded poll reports, press releases, methodology statements and other useful documentation.

2. Search the pollster's website for documentation of polls and methodology. Some pollsters had a public webpage where poll results or documentation were posted. For some pollster websites this was productive, but for others it was not. In some cases, a poll was listed with a broken link. In those cases, researchers added the additional information from the pollster website to the Google search methods listed above.

*3. Search the Roper iPoll Archive.* Researchers entered the pollster and year in the search fields and looked for documentation of polls and methodology.

*4. Search the FiveThirtyEight.com pollster database*. This resource was helpful to find instances of polls and methodologies that had not been identified through the prior step.

*5. Additional internet searches for missing information.* In some cases, additional, more specific internet searches were needed to look for missing information. Researchers conducted Google searches for "[POLLSTER NAME] poll [YEAR] [MODE]" to confirm the year that use of a particular method began or ended. For example, "CBS News poll 2010 online" was used to check that the pollster did not do online polling in 2010.

Two researchers performed the steps above independently for each pollster. The team also created a codebook assigning a number to each combination of methods observed from a pollster in a given year. For example, code 1 denotes that the pollster used only live phone with random-digit-dial sample for at least one national public poll that year, while code 25 denotes that the pollster did at least one poll using online opt-in and at least one poll using a probability-based panel. The team conducted a reliability analysis on the two independently gathered sets of data. The Cohen's kappa was 0.7, which is typically considered an acceptable level of agreement in social science content coding. A senior researcher then resolved any conflicts and produced a single dataset reflecting the best information available from the searching phase.

To record the data, researchers created a spreadsheet in which each column was a year (2000, 2002, ... 2022) and each row was an instance of a pollster using a particular method during that year. Researchers archived the URL documenting each instance of a pollster using a given method in a particular year. If the pollster did multiple polls using the same method, only one instance was archived per year.

#### Pollster outreach to verify the information gathered

As a final quality check, a senior researcher emailed each pollster to verify the information. These emails explained the study goals and presented the methods information observed from 2000 to 2022 specifically for the pollster. The email asked for any additions or edits. Most organizations (47 of the 78) responded. Pollsters were very generous with their time. Among those responding, most (81%) confirmed that the information recorded looked accurate. Others offered corrections, which were then applied to the study dataset. In some cases, a pollster correction could be corroborated by a URL. In a small number of instances, staff could not find a corroborating URL, but the information provided by the pollster was taken as authoritative. Such instances are recorded in the documentation dataset as "pollster email" instead of a URL.

## Assumptions made when information was incomplete

Sometimes the best available documentation of a poll did not clearly describe the sample source and/or interview mode. However, it was often the case that circumstantial evidence supported a reasonable educated guess. The team applied several guidelines in such situations. Each of these guidelines proved well-founded based on the input received during pollster outreach.

- If mode was not specified, but the questionnaire contained interviewer instructions, such as to read or not read certain response options, the poll was coded as *live phone*.
- If mode was specified as live phone, the sample source was not specified, but the poll was described as a sample "of Americans," the poll was coded as *live phone with random-digit-dial sample*.
- If the poll was described as having been conducted "online" but there was no other information, the poll was coded as *online opt-in*. (Our experience was that any time a poll used a probability-based panel, the pollster always disclosed the name of the panel.)
- If the poll documentation reported a "credibility interval" or "modeled margin of error" but did not disclose the sample source, the poll was coded as *online opt-in*.
- If the methodology was not disclosed but the pollster had a clear track record of consistently using a certain methodology, the poll was coded consistent with the pollster's known methods.
- Documentation of live phone polls before roughly 2012 often did not disclose the sample source, such as whether it was random-digit dial or registered voter records. The information that is available suggests that most national public polls conducted from 2000 to 2012 were probably using random-digit-dial sample. If a poll was live phone but the sample source was not specified, the poll was coded as *live phone with random-digit-dial sample* (unless the pollster had a record of using registration-based sampling).

# Strengths and weaknesses of the research design

Like any study, this one has its strengths and limitations. The strengths include:

#### 22 PEW RESEARCH CENTER

- Offering more insight than common industry characterizations. Existing databases that code
  polling methods tend to use terms like "online" or "telephone" that gloss over major
  distinctions, such as whether an online sample was recruited using convenience approaches or
  random sampling from a high-coverage frame. This study distinguishes between online
  surveys fielded with opt-in sample and those fielded on probability-based panels. This study
  also distinguishes between live phone surveys using registered voter records or random-digitdial procedures.
- *Documenting growth in the number of methods individual pollsters use.* This study captures not just changes in methods but the increasing use of multiple methods within and across polls by the same pollster.
- Offering a timeline long enough to see how trends unfolded. While other reports have
  provided a snapshot of the methods used in a particular year or election, they generally do not
  shed light on the trajectory of those methods. This study distinguishes between changes
  emerging in recent years from those with longer arcs.
- *Studying new approaches*. This study coded methods information at a level granular enough to see the emergence of new techniques such as using text as a supplemental mode or matching an online opt-in sample to registered voter records.

The study limitations include:

- *Only national polling was considered.* Unfortunately, attempting to find, document and code polling from all 50 states and the District of Columbia would have exceeded the time and staff resources available.
- *Not all methods details were considered*. This study focuses on two key poll features: where the respondents came from (the sample source or sources) and how they were interviewed (the mode or modes). While important, they are not exhaustive of the important decisions in designing a poll. The study did not attempt to track other details, such as weighting. Because the study only measured two out of all possible poll features, estimates from this study likely represent a *lower bound* of the total amount of change in the polling industry.
- *Odd-numbered years were excluded.* The study reflects data for even-numbered years only, which was a purely practical decision. Hundreds if not thousands of national public polls are released each year, and processing them for this report required substantial labor. Focusing on even-numbered years cut the manual labor roughly in half.

- *Given the size and fluidity of the industry, the dataset may be missing some pollsters.* While the research team went to great lengths to include all the pollsters that were eligible under the study criteria, it is very possible that some were missed. The online public record is only so complete, and the further back in time one goes the more broken links one encounters. Moreover, technology has destroyed the barriers to public polling that once existed, and so the number of potential pollsters feels almost without limit. This study includes many of the most prominent polling organizations, but probably misses some of the less prominent pollsters with smaller digital footprints.
- The study does not measure the volume of polls conducted with each sample source and mode of interview. The low cost of online opt-in polling has led to a dramatic increase in the number of such polls during the period of study. In practical terms, this means that the growth in the number of *organizations* conducting online opt-in polls could understate the growth in the share of all *polls* conducted with this methodology.

## Supplemental materials

The <u>dataset</u> accompanying this report lists the 78 organizations included. The <u>Documentation</u> <u>URLs</u> file provides the webpages used to code the methods. The <u>Codebook</u> details what each code means.