Seim
AP US Government

U.S. Survey Research

Questionnaire design

Perhaps the most important part of the survey process is the creation of questions that accurately measure the opinions, experiences and behaviors of the public. Accurate random sampling and high response rates will be wasted if the information gathered is built on a shaky foundation of ambiguous or biased questions. Creating good measures involves both writing good questions and organizing them to form the questionnaire.

Questionnaire design is a multistage process that requires attention to many details at once. Designing the questionnaire is complicated because surveys can ask about topics in varying degrees of detail, questions can be asked in different ways, and questions asked earlier in a survey may influence how people respond to later questions. Researchers also are often interested in measuring change over time and therefore must be attentive to how opinions or behaviors have been measured in prior surveys.

Surveyors may conduct pilot tests or focus groups in the early stages of questionnaire development in order to better understand how people think about an issue or comprehend a question. Pretesting a survey is an essential step in the questionnaire design process to evaluate how people respond to the overall questionnaire and specific questions.

For many years, surveyors approached questionnaire design as an art, but substantial research over the past thirty years has demonstrated that there is a lot of science involved in crafting a good survey questionnaire. Here, we discuss the pitfalls and best practices of designing questionnaires.

Question development

There are several steps involved in developing a survey questionnaire. The first is identifying what topics will be covered in the survey. For Pew Research Center surveys, this involves thinking about what is happening in our nation and the world and what will be relevant to the public, policymakers and the media. We also track opinion on a variety of issues over time so we often ensure that we update these trends on a regular basis so we can understand whether people's opinions are changing.

At Pew Research Center, questionnaire development is a collaborative and iterative process where staff meet to discuss drafts of the questionnaire several times over the course of its development. After the questionnaire is drafted and reviewed, we pretest every questionnaire and make final changes before fielding the survey.

Measuring change over time

Many surveyors want to track changes over time in people's attitudes, opinions and behaviors. To measure change, questions are asked at two or more points in time. A cross-sectional design, the most common one used in public opinion research, surveys different people in the same population at multiple points in time. A panel or longitudinal design, frequently used in other types of social research, surveys the same people over time. Pew Research Center launched its own random sample panel survey in 2014; for more, see the section on the American Trends Panel.

Many of the questions in Pew Research surveys have been asked in prior polls. Asking the same questions at different points in time allows us to report on changes in the overall views of the general public (or a subset of the public, such as registered voters, men or African Americans).

When measuring change over time, it is important to use the same question wording and to be sensitive to where the question is asked in the questionnaire to maintain a similar context as when the question was asked previously (see question wording and question order for further information). All of our survey reports include a topline questionnaire that provides the exact question wording and sequencing, along with results from the current poll and previous polls in which the question was asked.

Open- and closed-ended questions

One of the most significant decisions that can affect how people answer questions is whether the question is posed as an open-ended question, where respondents provide a response in their own words, or a closed-ended question, where they are asked to choose from a list of answer choices.

Fewer People Mention Economy in Open-Ended Version

What one issue mattered most to you in deciding how you voted for president?

	Open-ended[1]	Closed-ended[2]
The economy	**35%**	**58%**
The war in Iraq	5	10
Health care	4	8
Terrorism	6	8
Energy policy	*	6
Other	43	8
Candidate mentions	9	-
Moral values/social issues	7	-
Taxes/dist. of income	7	-
Other issues	5	-
Other political mentions	3	-
Change	3	-
Other	9	-
Don't know	7	2
	100	100

Data from Pew Research November 2008 Post-election survey

[1] Unprompted first responce to open-ended question.

[2] First choice from five options read to respondents.

For example, in a poll conducted after the presidential election in 2008, people responded very differently to two versions of this question: "What one issue mattered most to you in deciding how you voted for president?" One was closed-ended and the other open-ended. In the closed-ended version, respondents were provided five options (and could volunteer an option not on the list).

When explicitly offered the economy as a response, more than half of respondents (58%) chose this answer; only 35% of those who responded to the open-ended version volunteered the economy. Moreover, among those asked the closed-ended version, fewer than one-in-ten (8%) provided a response other than the five they were read; by contrast fully 43% of those asked the open-ended version provided a response not listed in the closed-ended version of the question. All of the other issues were chosen at least slightly more often when explicitly offered in the closed-ended version than in the open-ended version. (Also see "High Marks for the Campaign, a High Bar for Obama" for more information.)

Researchers will sometimes conduct a pilot study using open-ended questions to discover which answers are most common. They will then develop closed-ended questions that include the most common responses as answer choices. In this way, the questions may better reflect what the public is thinking or how they view a particular issue.

When asking closed-ended questions, the choice of options provided, how each option is described, the number of response options offered and the order in which options are read can all influence how people respond. One example of the impact of how categories are defined can be found in a Pew Research poll conducted in January 2002: When half of the sample was asked whether it was "more important for President Bush to focus on domestic policy or foreign policy," 52% chose domestic policy while only 34% said foreign policy. When the category "foreign policy" was narrowed to a specific aspect – "the war on terrorism" – far more people chose it; only 33% chose domestic policy while 52% chose the war on terrorism.

In most circumstances, the number of answer choices should be kept to a relatively small number – just four or perhaps five at most – especially in telephone surveys. Psychological research indicates that people have a hard time keeping more than this number of choices in mind at one time. When the question is asking about an objective fact,

such as the religious affiliation of the respondent, more categories can be used. For example, Pew Research Center's standard religion question includes 12 different categories, beginning with the most common affiliations (Protestant and Catholic). Most respondents have no trouble with this question because they can just wait until they hear their religious tradition read to respond.

What is your present religion, if any? Are you Protestant, Roman Catholic, Mormon, Orthodox such as Greek or Russian Orthodox, Jewish, Muslim, Buddhist, Hindu, atheist, agnostic, something else, or nothing in particular?

In addition to the number and choice of response options offered, the order of answer categories can influence how people respond to closed-ended questions. Research suggests that in telephone surveys respondents more frequently choose items heard later in a list (a "recency effect").

Because of concerns about the effects of category order on responses to closed-ended questions, many sets of response options in Pew Research Center's surveys are programmed to be randomized (when questions have two or more response options) to ensure that the options are not asked in the same order for each respondent. For instance, in the example discussed above about what issue mattered most in people's vote, the order of the five issues in the closed-ended version of the question was randomized so that no one issue appeared early or late in the list for all respondents. Randomization of response items does not eliminate order effects, but it does ensure that this type of bias is spread randomly.

Questions with ordinal response categories – those with an underlying order (e.g., excellent, good, only fair, poor OR very favorable, mostly favorable, mostly unfavorable, very unfavorable) – are generally not randomized because the order of the categories conveys important information to help respondents answer the question. Generally, these types of scales should be presented in order so respondents can easily place their responses along the continuum, but the order can be reversed for some respondents. For example, in one of the Pew Research Center's questions about abortion, half of the sample is asked whether abortion should be "legal in all cases, legal in most cases, illegal in most cases, illegal in all cases" while the other half of the sample is asked the same question with the response categories read in reverse order, starting with "illegal in all cases." Again, reversing the order does not eliminate the recency effect but distributes it randomly across the population.

Question wording

The choice of words and phrases in a question is critical in expressing the meaning and intent of the question to the respondent and ensuring that all respondents interpret the question the same way. Even small wording differences can substantially affect the answers people provide.

An example of a wording difference that had a significant impact on responses comes from a January 2003 Pew Research Center survey. When people were asked whether they would "favor or oppose taking military action in Iraq to end Saddam Hussein's rule," 68% said they favored military action while 25% said they opposed military action. However, when asked whether they would "favor or oppose taking military action in Iraq to end Saddam Hussein's rule *even if it meant that U.S. forces might suffer thousands of casualties,*" responses were dramatically different; only 43% said they favored military action, while 48% said they opposed it. The introduction of U.S. casualties altered the context of the question and influenced whether people favored or opposed military action in Iraq.

There has been a substantial amount of research to gauge the impact of different ways of asking questions and how to minimize differences in the way respondents interpret what is being asked. The issues related to question wording are more numerous than can be treated adequately in this short space. Here are a few of the important things to consider in crafting survey questions:

First, it is important to ask questions that are clear and specific and that each respondent will be able to answer. If a question is open-ended, it should be evident to respondents that they can answer in their own words and what type of response they should provide (an issue or problem, a month, number of days, etc.). Closed-ended questions should include all reasonable responses (i.e., the list of options is exhaustive) and the response categories should not overlap (i.e., response options should be mutually exclusive).

It is also important to ask only one question at a time. Questions that ask respondents to evaluate more than one concept (known as double-barreled questions) – such as "How much confidence do you have in President Obama to handle domestic and foreign policy?" – are difficult for respondents to answer and often lead to responses that are difficult to interpret. In this example, it would be more effective to ask two separate questions, one about domestic policy and another about foreign policy.

In general, questions that use simple and concrete language are more easily understood by respondents. It is especially important to consider the education level of the survey population when thinking about how easy it will be for respondents to interpret and answer a question. Double negatives (e.g., do you favor or oppose *not* allowing gays and lesbians to legally marry) or unfamiliar abbreviations or jargon (e.g., ANWR instead of Arctic National Wildlife Refuge) can result in respondent confusion and should be avoided.

Similarly, it is important to consider whether certain words may be viewed as biased or potentially offensive to some respondents, as well as the emotional reaction that some words may provoke. For example, in a 2005 Pew Research survey, 51% of respondents said they favored "making it legal for doctors to give terminally ill patients the means to end their lives," but only 44% said they favored "making it legal for doctors to assist terminally ill patients in committing suicide." Although both versions of the question are asking about the same thing, the reaction of respondents was different. In another example, respondents have reacted differently to questions using the word "welfare" as opposed to the more generic "assistance to the poor." Several experiments have shown that there is much greater public support for expanding "assistance to the poor" than for expanding "welfare."

One of the most common formats used in survey questions is the "agree-disagree" format. In this type of question, respondents are asked whether they agree or disagree with a particular statement. Research has shown that, compared with the better educated and better informed, less educated and less informed respondents have a greater tendency to agree with such statements. This is sometimes called an "acquiescence bias" (since some kinds of respondents are more likely to acquiesce to the assertion than are others). A better practice is to offer respondents a choice between alternative statements. A Pew Research Center experiment with one of its routinely asked values questions illustrates the difference that question format can make. Not only does the forced choice format yield a very different result overall from the agree-disagree format, but the pattern of answers among better- and lesser-educated respondents also tends to be very different.

Source: Pew Research Center