

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. 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.

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.

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.

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.

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.

One other challenge in developing questionnaires is what is called “social desirability bias.” People have a natural tendency to want to be accepted and

liked, and this may lead people to provide inaccurate answers to questions that deal with sensitive subjects. Research has shown that respondents understate alcohol and drug use, tax evasion and racial bias; they also may overstate church attendance, charitable contributions and the likelihood that they will vote in an election. Researchers attempt to account for this potential bias in crafting questions about these topics. For instance, when Pew Research Center surveys ask about past voting behavior, it is important to note that circumstances may have prevented the respondent from voting: “In the 2012 presidential election between Barack Obama and Mitt Romney, did things come up that kept you from voting, or did you happen to vote?” The choice of response options can also make it easier for people to be honest; for example, a question about church attendance might include three of six response options that indicate infrequent attendance. Research has also shown that social desirability bias can be greater when an interviewer is present (e.g., telephone and face-to-face surveys) than when respondents complete the survey themselves (e.g., paper and web surveys).

Lastly, because slight modifications in question wording can affect responses, identical question wording should be used when the intention is to compare results to those from earlier surveys (see [measuring change over time](#) for more information). Similarly, because question wording and responses can vary based on the mode used to survey respondents, researchers should carefully evaluate the likely effects on trend measurements if a different survey mode will be used to assess change in opinion over time (see [collecting survey data](#) for more information).

Question order

Once the survey questions are developed, particular attention should be paid to how they are ordered in the questionnaire. The placement of a question can have a greater impact on the result than the particular choice of words used in the question.

When determining the order of questions within the questionnaire, surveyors must be attentive to how questions early in a questionnaire may have unintended effects on how respondents answer subsequent questions. Researchers have demonstrated that the order in which questions are asked can influence how people respond; earlier questions – in particular those directly preceding other questions – can provide context for the questions that follow (these effects are called “order effects”).

One kind of order effect can be seen in responses to open-ended questions. Pew Research surveys generally ask open-ended questions about national problems, opinions about leaders and similar topics near the beginning of the questionnaire. If closed-ended questions that relate to the topic are placed before the open-ended question, respondents are much more likely to mention concepts or considerations raised in those earlier questions when responding to the open-ended question.

For closed-ended opinion questions, there are two main types of order effects: contrast effects, where the order results in greater differences in responses, and assimilation effects, where responses are more similar as a result of their order.

An example of a contrast effect can be seen in a Pew Research Center poll conducted in October 2003 that found that people were more likely to favor allowing gays and lesbians to enter into legal agreements that give them the same rights as married couples when this question was asked after one about whether they favored or opposed allowing gays and lesbians to marry (45% favored legal agreements when asked after the marriage question, but 37% favored legal agreements without the immediate preceding context of a question about gay marriage). Responses to the question about gay marriage, meanwhile, were not significantly affected by its placement before or after the legal agreements question.

Another experiment embedded in a December 2008 Pew Research poll also resulted in a contrast effect. When people were asked “All in all, are you satisfied or dissatisfied with the way things are going in this country today?” immediately after having been asked “Do you approve or disapprove of the way George W. Bush is handling his job as president?”; 88% said they were dissatisfied, compared with only 78% without the context of the prior question. Responses to presidential approval remained relatively unchanged whether national satisfaction was asked before or after it. A similar finding occurred in December 2004 when both satisfaction and presidential approval were much higher (57% were dissatisfied when Bush approval was asked first vs. 51% when general satisfaction was asked first).

Several studies also have shown that asking a more specific question before a more general question (e.g., asking about happiness with one’s marriage before asking about one’s overall happiness) can result in a contrast effect. Although some exceptions have been found, people tend to avoid redundancy by excluding the more specific question from the general rating.

Assimilation effects occur when responses to two questions are more consistent or closer together because of their placement in the questionnaire. We found an example of an assimilation effect in a Pew Research poll conducted in November 2008 when we asked whether Republican leaders should work with Obama or stand up to him on important issues and whether Democratic leaders should work with Republican leaders or stand up to them on important issues. People are more likely to say that Republican leaders should work with Obama when the question was preceded by the one asking what Democratic leaders should do in working with Republican leaders (81% vs. 66%). However, when people were first asked about Republican leaders working with Obama, fewer said that Democratic leaders should work with Republican leaders (71% vs. 82%).

The order questions are asked is of particular importance when tracking trends over time. As a result, care should be taken to ensure that the context

is similar each time a question is asked. Modifying the context of the question could call into question any observed changes over time (see [measuring change over time](#) for more information).

A questionnaire, like a conversation, should be grouped by topic and unfold in a logical order. It is often helpful to begin the survey with simple questions that respondents will find interesting and engaging to help establish rapport and motivate them to continue to participate in the survey. Throughout the survey, an effort should be made to keep the survey interesting and not overburden respondents with several difficult questions right after one another. Demographic questions such as income, education or age should not be asked near the beginning of a survey unless they are needed to determine eligibility for the survey or for routing respondents through particular sections of the questionnaire. Even then, it is best to precede such items with more interesting and engaging questions.