In the previous handout, we discussed potential biases that may arise because of who responded to a survey. Survey results should be interpreted not only in the context of who responded, but also with regard to what questions were asked. Factors such as question type, question wording, and the order in which questions were presented can influence the results. Careful consideration should be given to these issues when designing questionnaires. In this handout, we’ll discuss strategies for creating good survey questionnaires that will accurately elicit the desired information from the respondents.

Creating a good survey requires keeping the study objectives firmly in mind. Some have suggested that researchers should think about how they would report the results before creating the final draft of a survey. If a question is not essential for meeting the goals of the study and it doesn’t make the cut for the final report, then it probably should not be included in the survey. Likewise, thinking about the final report beforehand might help identify whether additional questions are needed to address the study objectives.

Furthermore, in order for the survey results to be useful, the survey must demonstrate both **reliability** and **validity**.

|  |
| --- |
| Definitions |
| **Reliability** – This is the extent to which repeatedly measuring the same thing produces the same result. |
| **Validity** – This is the extent to which a survey question measures what it is supposed to measure. |

Both of these concepts will be discussed in more detail in the future. For now, recognize that our goal is to write good questions that will maximize the validity and reliability of the survey. Also, through careful design of the survey questions, we want to minimize any nonsampling errors that may occur.

**POSSIBLE SOURCES OF RESPONSE BIAS IN SURVEYS**

In their book *Mind on Statistics*, Utts and Heckard (2007) identify the following seven sources of response bias that may occur during a survey study.

* Deliberate bias in questions
* Unintentional bias in questions
* Desire of respondents to please
* Asking the uninformed
* Unnecessary complexity
* Ordering of questions
* Confidentiality and anonymity concerns

Each of these is discussed in more detail below.

**Deliberate Bias in Questions**

If the objective of a survey study is to support a certain cause, the creators of the survey may deliberately try to influence the results by using questions that are purposefully worded in a biased manner. For example, consider the following.

| **Example: Half of America Wants Obama Impeached?** |
| --- |
| In May of 2013, Wenzel Strategies in association with WorldNetDaily conducted a poll to investigate American adults’ opinions regarding whether Obama should be impeached. Though both claim to be independent organizations, many people feel that these agencies have a biased political agenda. Here is the headline that accompanied their press release describing the survey results.    Here is one question from their survey:    *Source:* [*http://www.wnd.com/2013/05/half-of-america-wants-obama-impeached*](http://www.wnd.com/2013/05/half-of-america-wants-obama-impeached)  Given the wording of this question, how much faith do you have in the results of this survey study? |

In general, avoid writing loaded or leading questions. You should always make sure your questions are worded neutrally.  
  
Deliberate bias is not always introduced as blatantly as it was in the previous example. Some survey creators with an agenda simply take advantage of the proven fact that human beings, in general, like to be agreeable and are inclined to answer “yes” when a question leads them in that direction. For example, consider the following question posed in an online survey.

| **Example: Have Obama and Democrats Bounced Back?** |
| --- |
| The following question was posed on msnbc.com in December of 2013.    *Source:* [*http://www.msnbc.com/msnbc/poll-do-you-agree-obama-and-the-democrats-have-bounced-back*](http://www.msnbc.com/msnbc/poll-do-you-agree-obama-and-the-democrats-have-bounced-back) |

Studies have shown that a question like the one above will receive many more yes responses than it should simply because a “yes” response seems to agree with the interviewer’s notion of a correct response. To avoid this, questions should be asked in a balanced form. For example, consider the next set of survey questions periodically administered by Gallup to U.S. adults.

| **Example: Gallup Question - Which Party is Better Able to Protect U.S. from Terrorism?** |
| --- |
| *Source:* [*http://www.gallup.com/poll/175727/republicans-expand-edge-better-party-against-terrorism.aspx*](http://www.gallup.com/poll/175727/republicans-expand-edge-better-party-against-terrorism.aspx) |

| **Example: Gallup Question – American Opinions on Smoking Bans** |
| --- |
| *Source:* [*http://www.gallup.com/poll/174203/americans-favor-ban-smoking-public-not-total-ban.aspx*](http://www.gallup.com/poll/174203/americans-favor-ban-smoking-public-not-total-ban.aspx) |

Once again, note that both of the above Gallup questions are examples of good wording because they don’t indicate to the respondent which response is preferable.

Exercises

1. Temporarily forget the fact that the poll results from the msnbc.com example are not generalizable to all American adults in the first place (why not?). Instead, suppose that this question was to be included on a legitimate survey. Change the wording of the question to remove the bias.
2. Though this is not ethical and would not be advised in a legitimate survey study, for illustrative purposes only, reword the Gallup questions described above to try to bias the results one way or the other.

**Unintentional Bias in Questions**

One way that this can occur is when questions are worded in such a way that the meaning is misinterpreted by a large proportion of the respondents. This can often happen inadvertently when words have multiple meanings or when words are interpreted differently than intended.

| **Example: The Current Population Survey (CPS) and Layoffs** |
| --- |
| The CPS changed their survey in 1994 because of issues identified with question wording. Some questions on the survey referred to being laid off, and researchers found that their definition of layoff was different from that of the respondents. This introduced bias in the results. The following is a quote from an article describing these changes.  *Source:* [*http://www.bls.gov/cps/revisions1994.pdf*](http://www.bls.gov/cps/revisions1994.pdf) |

To protect against such bias, it is good practice to clearly define terms in the survey that might otherwise be ambiguous. In the book *Elementary Survey Sampling* byScheaffer et al. (1990), the authors provide the following examples of poorly worded questions and possible improvements to those questions.

|  |  |
| --- | --- |
| **Poorly Worded Question** | **Improved Version** |
| “How much water do you drink?” | “Here is an eight-ounce glass. (Hold one up.) How many eight-ounce glasses of water do you drink each day?” |
| “How many children are in your family?” | “How many persons under the age of 21 live in your household and receive more than one-half of their financial support from you?” |

Note that all components of the improved versions of the questions have been well-defined.

Unintentional bias can also arise because of the choice of wording. Even small differences in wording can change the responses people provide. For example, consider two different versions of a question asked on the General Social Survey.

| **Example: The General Social Survey and Opinions on Welfare Spending** |
| --- |
| One question on the General Social Survey investigates opinions regarding welfare spending. Two versions of what some might argue are the same question are asked:  Version 1: Are we spending too much, too little, or about the right amount on *welfare*?  Version 2: Are we spending too much, too little, or about the right amount on *assistance to   the poor*?  The following table summarizes the percent who responded with “too little” for each version of the question across various years.   |  |  |  | | --- | --- | --- | |  | **Assistance to the Poor** | **Welfare** | | 1984 | 64.1 | 25.2 | | 1985 | 65.2 | 19.8 | | 1986 | 62.8 | 23.1 |   *Source: Rasinski, Kenneth A. 1989. “The Effect of Question Wording on Public Support for Government Spending.” Public Opinion Quarterly, 53(3).* |

What if only one of the above versions of the question had been asked on the survey? What is the real opinion of American adults, and would this survey with only one version of the question have captured this opinion?  
  
The General Social Survey now asks both versions of the question (half of their sample receives Version 1 and the other half Version 2). The aforementioned author, Kenneth Rasinksi, also describes differences in how subjects responded to questions regarding opinions of amount spent on “dealing with drug addiction” versus “drug rehabilitation.” How do you think the responses differed? Other studies have cited differences in survey questions that use the word “forbid” versus “not allow.” Again, how do you think the responses differed in those studies?  
  
In general, it is important to note that certain words may elicit a stronger response from the survey subjects. If possible, try not to use words that might provoke strong reactions from the respondents.

**Desire of Respondents to Please**

This is also referred to as **social desirability bias**. In general, respondents tend to respond to questions with answers they think will be viewed favorably by others. This involves both understating behaviors/opinions that are viewed negatively and overstating behaviors/opinions that are viewed positively. Careful wording of questions can help to reduce this bias.

| **Example: Polls Regarding Household Budgets** |
| --- |
| A few years ago, a Pew Social Trends poll asked respondents, “Do you have a budget to guide your household expenses or don’t you rely on a formal budget?” About 48% of those surveyed said they kept a budget.  In 2013, Gallup conducted their own poll on this issue, but their question was better worded to reduce social desirability bias. They asked if the respondent or their spouse or partner “prepare a detailed written or computerized household budget each month that tracks your income and expenditures.” This time, 32% of those surveyed answered “yes.”  *Sources:* [*http://www.theblaze.com/stories/2013/07/16/so-what-percentage-of-americans-actually-keep-a-budget-gallup-results-may-surprise-you/*](http://www.theblaze.com/stories/2013/07/16/so-what-percentage-of-americans-actually-keep-a-budget-gallup-results-may-surprise-you/)  *and* [*http://www.gallup.com/poll/162872/one-three-americans-prepare-detailed-household-budget.aspx*](http://www.gallup.com/poll/162872/one-three-americans-prepare-detailed-household-budget.aspx) |
| **Example: Gallup Poll Regarding Voting in Last Election** |
| People often claim to have voted in a past election, regardless of whether or not they really did. To reduce social desirability bias when soliciting this information from a respondent, Gallup worded their question after the last election as follows: “In the election in November 2012, did things come up which kept you from voting, or did you happen to vote?”  *Source:* [*http://www.gallup.com/poll/111268/how-gallups-likely-voter-models-work.aspx*](http://www.gallup.com/poll/111268/how-gallups-likely-voter-models-work.aspx) |

**Asking the Uninformed**

Sometimes, respondents claim to know something about an issue because they are embarrassed to admit that they don’t. For example, Utts and Heckard (2007) describe surprising results of a study carried out to investigate this type of bias. “When the American Jewish Committee studied Americans’ attitudes toward various ethnic groups, almost 30% of the respondents had an opinion about the fictional Wisians, rating them in social standing above a half dozen other real groups including Mexicans, Vietnamese, and African blacks.”

To combat this type of bias, be sure to keep your survey questions clear of technical jargon, and clearly define any terms that need to be defined. Also, be sure that, when necessary, survey questions are worded so that the respondent is given a chance to admit that they don’t have the knowledge required to answer the question. For example, Gallup is usually careful to include “don’t know” or “not sure” as potential responses for most of their questions.

**Unnecessary Complexity**

Questions should be worded as simply and clearly as possible. Using negative wording like “Do you agree or disagree that students should not be required to take a comprehensive exam to graduate?” are difficult to process and are likely to introduce measurement error. Avoid the use of negatives, if possible.   
  
Also, don’t write questions with multiple parts (called “double-barreled” questions) that ask respondents for a single answer. For example, consider the following question: “Do you support Obamacare, since it would ensure that college students stay on their parents’ health plans until age 26?” What if you support the plan, but not for that reason? What if you support that reason but disagree with the remainder of the plan? How should you answer this question? Another example is as follows: “Do you sometimes find that you have arguments with your family members and co-workers?” What if you argue with only one family member? How would you answer this question?  
  
In summary, to avoid introducing bias in your results, use short, simple words that are familiar. If technical jargon must be used, it is best to define the term specifically in the survey. Finally, avoid using double-barreled questions.

**Ordering of Questions**

When writing a survey, it is recommended that you start with items that are easy to understand and to answer (so it is wise to avoid highly controversial or sensitive topics early on in the questionairre). This helps ensure that the respondent stays comfortable, interested, and motivated to provide you with high-quality data.  
  
If a survey is fairly long, it makes sense to group questions by topic. Be aware, however, that the first question in a sequence might influence their answers to later questions.

| **Example: Gallup Questions Regarding Gay Marriage and Civil Unions** |
| --- |
| Consider the following questions used in a Gallup opinion poll:   * “Would you favor or oppose a law that would allow homosexual couples to legally form civil unions, giving them some of the legal rights of married couples?” * “Do you think marriages between homosexuals should or should not be recognized by the law as valid, with the same rights as traditional marriages?”   In one study, Gallup rotated the order of these questions. Half of the sample heard the questions in one order and half in reverse order. The results are as follows:        *Source:* [*http://www.gallup.com/poll/11662/revisiting-gay-marriage-vs-civil-unions.aspx*](http://www.gallup.com/poll/11662/revisiting-gay-marriage-vs-civil-unions.aspx) |

As mentioned in the previous example, to combat the effect of question ordering, it is good practice to rotate question order across respondents within blocks of related questions.

**Confidentiality and Anonymity Concerns**

Behaviors and opinions related to sensitive topics such as sex, illegal drug use, etc., are difficult to measure because people consider these to be private matters and are hesitant to answer honestly (which biases the results). One way to reduce this type of bias is to conduct an anonymous survey (i.e., the researcher does not know the identity of the respondents). If this is not possible, researchers can promise confidentiality (i.e., the researcher promises not to release identifying information about respondents). A variety of techniques have been developed to help ensure confidentiality, but surveys on such issues are still hard to conduct accurately.  
  
The next example describes a study that investigated the effects of anonymity.

| **Example: National Survey of Adolescent Males** |
| --- |
| Abstract:  *Source: Turner et. al. (1998), Science, Vol. 280, pp. 867-873. “Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. 1995 National Survey of Adolescent Males.”* |

**CHOOSING TYPES OF QUESTIONS**

When designing a questionnaire, one not only has to take careful steps to avoid bias but also has to make a series of decisions about each question’s format. For example, will it be open-ended or closed? Should rating scales be used? If so, how many points should be on the scale and how should the scale be labeled? As stated earlier, these decisions should be made so as to maximize the validity and reliability of the survey. The remainder of this handout will discuss some of these issues.

**Open-Ended vs. Closed Questions**

First, the researcher must decide whether a question should be open-ended or closed. An open-ended question is one in which respondents are allowed to provide their own answer without any prompting; a closed question is one in which respondents are given a list of alternatives from which to choose their answer. Usually, the closed form offers a choice of “other” in which the respondent is allowed to fill in the blank.

Problems with closed questions: Note that responses to closed questions may be influenced by the options provided and results could differ drastically from what would be obtained via an open-ended question on the same issue. This is illustrated in the next example.

| **Example: Differences in Responses to an Open-ended vs. a Closed Question** |
| --- |
| In one study, half of a random sample was asked to respond to an open-ended question and the other half to a closed question on the same issue. The results are shown below.    *Source: Schuman and Scott (1987), Science, Vol. 236, pp. 957-959. “Problems in the use of survey questions to measure public opinion.”*  Questions: Are you surprised by these differences? What if only the closed form of the question had been used? |

Problems with open-ended questions: Someone is tasked with coding the responses! To analyze, the responses must be grouped together in a relatively small number of categories. Also, compiling the results often requires the researcher to interpret the meaning of the responses, which can introduce some researcher bias. In short, open-ended questions are harder to analyze and summarize, which has led many researchers to use closed questions.

Which *should* you use: open-ended or closed?   
Closed questions will provide more uniformity of the responses and will be much easier to summarize. As the previous example illustrates, however, respondents tend to confine their answers to the choices offered, even if the researcher includes an “other” option. Closed questions can be used effectively only if the options provided are comprehensive, and this can be assured only if an open-ended version is administered in a pretest with a reasonably large sample. Some researchers are of the opinion that they have to deal with coding at this point, anyways, so they simply opt to include the open-ended question in the final version of the survey. Moreover, some studies have shown that open-ended questions have higher reliability and validity than closed questions, which is another reason they might be preferable. In any case, if a researcher decides to go with closed questions, they should first be presented as open-ended questions to a test sample before the real survey (this is called a “pilot survey”) and the most common responses should be included in the list of choices for the closed question.

**Rating vs. Ranking Questions**

Suppose a researcher decides to use a closed question to investigate peoples’ preferences of late night TV hosts. They could use the following question: “Do you prefer Jimmy Fallon or Jimmy Kimmel?” Respondents would then choose one. When respondents answer this question, they are essentially ranking the choices. An alternative approach is to ask them to rate their attitudes toward these two hosts separately, and the researcher could then infer which is preferred.

Asking the single ranking question seems to be the most straightforward approach, but note that ranking a large number of objects would be much more difficult (e.g., what if you were asked to rank your preferences of 10 different vegetables?). Also, ranking might force respondents to choose between items toward which they feel the same, which can be problematic. In such cases, respondents might try to give items toward which they feel the same the same rank value. Or, they may just leave some items unranked, which complicates things for the researcher.

Asking the respondents to rate choices, on the other hand, can reveal not only which item a respondent prefers but also how different his or her opinions of the items are. Ratings don’t come without trouble, however. Studies have shown that when subjects are asked to rate a large number of items on a single scale, many people rate multiple items identically (this is called *nondifferentiation*). They choose a reasonable point on the scale at which to rate most items and just select that point over and over again. As a result, reliability and validity of ranking data are often superior to those of rating data.

**Rating Scales**

If a researcher chooses to use a rating scale, one choice they must make is with regard to the number of points on the scale. Studies have shown that for “bipolar” scales running from positive to negative with neutral in the middle (e.g., scales running from “very satisfied” to “very dissatisfied”), using seven points maximizes reliability and validity. For “unipolar” scales (e.g., scales running from “no importance” to “high importance”), using five points is optimal.  
  
Furthermore, many studies suggest that measurement error is reduced when all points on the scale are labeled with words. These words should divide the scale equally. For example, one would not want to use the scale “very good, good, and poor.”

**Order of the Responses**

Another choice that researchers have to make with closed questions is with regard to the order of the responses. Just as the ordering of *questions* can influence results, so can the ordering of *responses*. So, it is best to rotate the order of the response choices across respondents whenever possible.

**REFERENCES\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

Scheaffer, Mendenhall, and Ott. (1990). *Elementary Survey Sampling, 4th Edition.* PWS-KENT Publishing Company, Boston, MA.

Survey Fundamentals – A guide to designing and implementing surveys. Office of quality improvement, University of Wisconsin-Madison. Retrieved from: http://oqi.wisc.edu/resourcelibrary/uploads/resources/Survey\_Guide.pdf

Utts and Heckard. (2007). *Mind on Statistics, 3rd edition*. Thomson Brooks/Cole, Belmont, CA.

Visser, P. S., Krosnick, J. A., & Lavrakas, P. J. (2000). Survey research. In H. T. Reis &

C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 223-252). Cambridge, UK: Cambridge University Press.