4.7 Types of Closed-Ended Survey Questions

Researchers can choose from a variety of questions including scales, dichotomous questions, and multiple-choice questions. In this section, this book presents the most important types of questions:

4.7.1 Scales

Scales are ubiquitous in social sciences questionnaires. Nearly all opinion-based questions use some sort of scale. An example would be: "Rate your satisfaction with your country's police forces from 0 (not satisfied at all) to 10 (very satisfied)." Many questions tapping into personal attributes use scales, as well [e.g., "how would you rate your personal health?" (0, poor; 1, rather poor; 2, average; 3, rather good; 4, very good)]. Finally, some behavioral questions use scales, as well. For example, you might find the question: "How often do you generally watch the news on TV?" Response options could be not at all (coded 0), rather infrequently (coded 1), sometimes (coded 2), rather frequently (coded 3), and very frequently (coded 4). The most prominent scales are Likert and Guttman scales.

Likert Scale A Likert Scale is the most frequently used ordinal variable in questionnaires (maybe the most frequently used type of questions overall) (Kumar 1999: 129). Likert Scales use fixed choice response formats and are designed to measure attitudes or opinions (Bowling 1997; Burns and Grove 1997). Such a scale assumes that the strength/intensity of the experience is linear, i.e., on a continuum from strongly agree to strongly disagree, and makes the assumption that attitudes can be measured. Respondents may be offered a choice of five to seven or even nine pre-coded responses with the neutral point being neither agree nor disagree (Kumar 1999: 132).

Example 1: Going to War in Iraq Was a Mistake of the Bush Administration
The respondent has five choices: strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, and strongly disagree.

Example 2: How Often Do You Discuss Politics with Your Peers?
The respondent has five choices: very frequently, frequently, occasionally, rarely, and never.

In the academic literature, there is a rather intense debate about the usage of the neutral category and the do not know option. This applies to Likert Scales and other survey questions as well. Including a neutral category makes it easier for some people to choose/respond. It might be a safe option, in particular for individuals, who prefer not taking a choice. However, excluding the neutral option forces individuals to take a position, even if they are torn, or really stand in the middle for the precise question at hand (Sapsford 2006; Neuman and Robson 2014). A similar logic applies

to the do not know option (see Mondak and Davis 2001; Sturgis et al. 2008). The disadvantage of this option is that many respondents go with "no opinion" just because it is more comfortable. In most cases, respondents choose this option due to conflicting views and not due to a lack of knowledge. So, in a sense, they usually do have an opinion and lean at least slightly to one side or another. That is why the inclusion of a "no opinion" option can reduce the number of respondents who give answers to more complex subjects. On the other hand, if this option is omitted, respondents can feel pressed to pick a side although they really are indifferent, for example, because they know very little about the matter. Pushing respondents to give arbitrary answers can therefore affect the validity and reliability of a survey.

For example, somebody might not be interested in economics and besides getting her monthly paycheck has no idea how the economy functions in her country. Such a person can only take a random guess for this question, if the "do not know" option is not included in the questionnaire. For sure, she could leave out the question, but this might feel challenging. Hence, the respondent might feel compelled to give an answer even if it is just a random guess.

While there is no panacea for avoiding these problems, the researcher must be aware of the difficulty of finding the right number and type of response categories and might think through these categories carefully. A possible way of dealing with this challenge is to generally offer "no opinion" options, but to omit them for items on well-known issues or questions where individuals should have an opinion. For example, if you ask respondents about their self-reported health, there is no reason to assume that respondents do not know how they feel. In contrast, if you ask them knowledge questions (e.g., "who is the president of the European Commission?"), the do not know option is a viable choice, as it is highly possible that politically uninterested citizens do not know who the president of the European Commission is. Alternatively, the researcher can decide to omit the "no opinion" option and test the strength of an attitude through follow-up questions. This way, it might become evident whether the given response represents a real choice rather than a guess (Krosnick 1999: 43 f.; Weisberg et al. 1996: 89 f.). For example, a pollster could first ask an individual about her political knowledge using the categories low, rather low, middle, rather high, and very high. In a second step, she could ask "hard" questions about the function of the political system or the positioning of political parties. These "hard" questions could verify whether the respondent judges her political knowledge level accurately.

The Semantic Differential Scale The use of the semantic differential scale allows for more options than the use of a Likert Scale, which is restricted to four or five categories. Rather than having each category labeled, the semantic scale uses two bipolar adjectives at each end of the question. Semantic differential scales normally range from 7 to 10 response choices (see Tables 4.3 and 4.4).

Large Rating Scales These are scales that have a larger range than 0–10. Such scales can often have a range from 0 to 100. Within this range the respondent is free

Table 4.3 Example of a semantic differential scale with seven response choices

How satisfied are you with the services received?						
1						7
Not at all satisfied						Very satisfied

Table 4.4 Example of a semantic differential scale with ten response choices

Do you think immigrants make your country a better or a worse place to live in?									
1									10
Better place									Worse place

to choose the number that most accurately reflects her opinion. For example, researchers could ask respondents how satisfied they are with the state of the economy in their country, on a scale from 0, not satisfied at all, to 100 very satisfied.

Guttman Scale The Guttman scale represents another rather simple way of measuring ordinal variables. This scale is based on a set of items with which the respondents can agree or disagree. All items refer to the exact same variable. However, they vary in their level of "intensity" which means that even people with low agreement might still agree with the first items (questions with a low degree of intensity usually come first) while it takes high stakes for respondents to agree with the last ones. The respondent's score is the total number of items she agrees with—a high score implies a high degree of agreement with the initial questions. Since, all items of one variable are commonly listed in increasing order according to their intensity, this operationalization is based on the assumption that if a respondent agrees with any one of the asked items, she should have agreed with all previous items too. The following example clarifies the idea of "intensity" in response patterns.

Example: Do You Agree or Disagree that Abortions Should Be Permitted?

1. When the life of the woman is in danger.
2. In the case of incest or rape.
3. When the fetus appears to be unhealthy.
4. When the father does not want to have a baby.
5. When the woman cannot afford to have a baby.
6. Whenever the woman wants.

Due to the Guttman scale's basic assumption, a respondent with a score of 4 agrees with the first four items and disagrees with the last two items. The largest challenge when using a Guttman scale is to find a suitable set of items that (perfectly) match the required response patterns. In other words, there must be graduation between each response category and consensus that agreement to the fourth item is more difficult and restricted than agreement to the third item.

Table 4.5 Example of single-answer multiple-choice question

Who is your favorite politician among the four politicians listed below?		
A	Donald Trump	
B	Emmanuel Macron	
C	Angela Merkel	
D	Theresa May	

Table 4.6 Example of a multiple-answer multiple-choice question

Which medium do you use to get your political information from? (click all that apply)		
A	Television news broadcasts	
B	Radio	
C	Internet	
D	Newspaper	
E	Discussion with friends	
F	Other (please specify)	
G	I do not inform myself politically	

4.7.2 Dichotomous Survey Questions

The dichotomous survey question normally leaves respondents with two answering choices. These two choices can include personal characteristics such as male and female, or they can include questions about personal attributes such as whether somebody is married or not.

Example: Did You Participate in a Lawful Demonstration in the Past 12 Months? (Yes/No)
Please note that some questions that were asked dichotomously for decades have recently been asked in a more trichotomous fashion. The prime example is gender. For decades, survey respondents had two options to designate their gender man or woman, not considering that some individuals might not identify with either of the dichotomous choices. Thus the politically correct way now is to include a third option such as neither nor.

4.7.3 Multiple-Choice Questions

Multiple-choice questions are another form of a frequently used question type. They are easy to use, respond to, and analyze. The different categories the respondents can choose from must be mutually exclusive. Multiple-choice questions can be both single answer (i.e., the respondent can only pick one option) and multiple answer (i.e., the respondent can pick several answers) (see Tables 4.5 and 4.6).

One of the weaknesses of some type of multiple-choice questions is that the choice of responses is rather limited and sometimes not final. For example, in

Table 4.7 Example of a categorical survey question with five choices

Which category best describes your age?	
	Younger than 18
	19–34
	35–50
	51–64
	65 and higher

Table 4.8 Example of a categorical survey question with six choices

In which bracket does your net income fall?	
	Lower than $20,000
	$20,000–$39,999
	$40,000–$59,999
	$60,000–$79,999
	$80,000–$99,999
	Over 100,000

Table 4.6, somebody could receive her political information from weekly magazines, talk shows, or political parties. In the current example, we have not included these options for parsimony reasons. Yet, the option E (other) allows respondents to add other items, thus allowing for greater inclusivity. (But please note that if the respondents list too many options under the category "other," the analysis of the survey becomes more complicated).

4.7.4 Numerical Continuous Questions

Numerical continuous questions ask respondents a question that in principle allows the respondent to choose from an infinite number of response choices. Examples would be questions that ask: what is your age? What is your net income?

4.7.5 Categorical Survey Questions

Frequently, researchers split continuous numerical questions into categories. These categories frequently correspond to established categories in the literature. For example, instead of asking what your age is, you could provide age brackets distinguishing young individuals (i.e., 18 and younger), rather young individuals (19–34), middle-aged individuals (35–50), rather old individuals (51–64), and old individuals (65 and higher) (see Table 4.7). The same might apply to income. For example, we could have income brackets for every $20,000 (see Table 4.8).

Using brackets or categories might be particularly helpful for features, where respondents do not want to reveal the exact number. For example, some respondents might not want to reveal their exact age. The same might apply to income. If they are

Table 4.9 Example of a rank order question

Consecutively rank the following five parties from most popular (coded 1) to least popular (coded 5)	
	Christian Democratic Party
	Social Democratic Party
	Free Democratic Party
	Green Party
	Left Party
	Alternative for Germany

Table 4.10 Example of a matrix table question

How would you rate the following attributes of President Trump?					
	Well below average	Below average	Average	Above Average	Well above average
Honesty					
Competence					
Trustworthiness					
Charisma					

offered rather broad age or income brackets, they might be more willing to provide an answer than be probed for their exact age or income.

4.7.6 Rank-Order Questions

Sometimes researchers might be interested in rankings. Rank-order questions allow respondents to rank persons, brands, or products based on certain attributes such as the popularity of politicians or the vote choice for parties. In rank-order questions, the respondent must consecutively rank the choices from most favorite to least favorite (see Table 4.9).

4.7.7 Matrix Table Questions

Matrix-level questions are questions in tabular format. They consist of multiple questions with the same response choices. The questions are normally connected to each other, and the response choices frequently follow a scale such as a Likert scale (see Table 4.10).

4.8 Different Variables

Regardless of the type of survey questions, there are four different ways to operationalize survey questions. The four types of variables are string variables, continuous variables (interval variables), ordinal variables, and nominal variables.

A **string variable** is a variable that normally occupies the first column in a dataset. It is a non-numerical variable, which serves as the identifier. Examples are individuals in a study or countries. This variable is normally not part of any analysis.

A **continuous variable** can have, in theory, an infinite number of observations (e.g., age, income). Such a question normally follows from a continuous numerical question. To highlight, personal income levels can in theory have any value between 0 and infinity.

An **interval variable** is a specific type of variable; it is a continuous variable with equal gaps between values. For example, counting the income in steps of thousand would be an interval variable: 1000, 2000, 3000, 4000, . . .

A **nominal variable** is categorized. There is no specific order or value to the categorization (e.g., the order is arbitrary). Because there is no hierarchy in the organization of the data, this type of variable is the most difficult to present in datasets (see discussion under Sect. 4.9.1).

An example of a two-categorical nominal variable would be gender (i.e., men and women). Such a variable is also called **dichotomous** or **dummy** variable.

An example of a categorical variable with more than two categories would be religious affiliation (e.g., Protestant, Catholic, Buddhist, Muslim, etc.)

An **ordinal variable** consists of data that are categorized, and there is a clear order to the categories. Normally all scales are transformed into ordinal variables.

For example, educational experience is a typical variable to be grouped as an ordinal variable. For example, a coding in the following categories could make sense: no elementary school, elementary school graduate, high school graduate, college graduate, Master's degree, and Ph.D.

An ordinal variable can also be a variable that categorizes a variable into set intervals. Such an interval question could be: How much time does it take you to get to work in the morning? Please tick the applicable category (see Table 4.11).

What is important for such an ordinal coding of a continuous variable is that the intervals are comparable and that there is some type of linear progression.

The most frequent types of ordinal variables are scales. Representative of such scales would be the answer to the question: Please indicate how satisfied you are with Chancellor Merkel's foreign policy on a scale from 1 to 5 (see Table 4.12).

Table 4.11 Ordinal coding of the variable time it takes somebody to go to work

Less than 10 min	10-30 min	30-60 min	More than 60 min

Table 4.12 Ordinal coding of the scaled variable satisfaction with Chancellor Merkel

1 not satisfied at all	2	3	4	5 very satisfied

4.9 Coding of Different Variables in a Dataset

Except for string variables, question responses need to be transformed into numbers to be useful for data analytical purposes. Table 4.13 highlights some rules how this is normally done:

1. In the first column in Table 4.13, we have a string variable. This first variable identifies the participants of the study. In a real scientific study, the researcher would render the names anonymous and write: student 1, student 2, . . .
2. The second column is dichotomous variable (i.e., a variable with two possible response choices). Whenever researchers have a dichotomous variable, they create two categories labeling the first category 0 and the second category 1. (For any statistical analysis, it does not matter which one of the two categories you code 0 and 1, but, of course, you need to remember your coding to interpret the data correctly later.)
3. Columns three and four are continuous variables representing the self-reported income and self-reported age the respondent has given in the questionnaire.
4. The fourth column captures a three-item ordinal variable asking individuals about their satisfaction with their job. The choices respondents have are low satisfaction (coded 0), medium satisfaction (coded 1), and high satisfaction (coded 2). Normally, ordinal variables are consecutively labeled from 0 or 1 for the lowest category to how ever many categories there are. To code ordinal variables, it is important that categories are mutually exclusive (i.e., one value cannot be in two categories). It is also important that progression between the various categories is linear.

4.9.1 Coding of Nominal Variables

For some analysis such as regression analysis (see Chaps. 8 and 9), non-dichotomous categorical variables (e.g., different religious affiliations) cannot be expressed in any ordered form, because there is no natural progression in their values. For all statistical tests that assume progression between the categories, this causes a problem. To circumvent this problem for these tests, we use dummy variables to express such a variable. The rule for the creation of dummy variables is that I create one dummy variable less than I have categories. For example, if I have four categories, I create

Table 4.13 Representing string, dichotomous, continuous, and ordinal variables in a dataset

Student	Gender	Income	Age	Job satisfaction
John	1	12,954	43	2
Mary	0	33,456	67	1
James	1	98,534	54	0

Table 4.14 Representing a nominal variable in a dataset

	Dummy 1	Dummy 2	Dummy 3
Muslim	0	0	0
Christian	1	0	0
Buddhist	0	1	0
Hindu	0	0	1

three dummy variables, with the first category serving what is called the reference category. In the example in Table 4.14, the Muslim religion is the reference category against which we compare the other religions. (Since this procedure is rather complex, we will discuss the creation of dummy variables in Chap. 5 again.)

4.10 Drafting a Questionnaire: General Information

In real research, the selection of an overarching question guiding the research process and the development of a questionnaire is theory driven. In other words, social science theory should inform researchers' choices of survey questions, and a good survey should respond to a precise theoretically driven research question (Mark 1996: 15f). Yet, for the exercise in this book, where you are supposed to create your own questionnaire, you are not supposed to be an expert in a particular field. Rather, you can use your general interest, knowledge, and intuition to draft a questionnaire. Please also keep in mind that you can use your personal experience to come up with the topic and overarching research question of your sample survey. To highlight, if a researcher has been affected by poverty in his childhood, he might intuitively have an idea about the consequences of child poverty. In her case, it is also likely that she has read about this phenomenon and liked a hypothesis of one of the authors she has read (e.g., that people who suffered from poverty in their childhood are less likely to go to college). Once you have identified a topic and determined the goals and objectives of your study, think about the questions you might want to include in your survey. Locating previously conducted surveys on similar topics might be an additional step enabling you to discover examples of different research designs about the same topic.

When you select the subject of your study, it is important to take the subject's practicability into account, as well. For one, a study on the political participation of university students can be relatively easily carried through by a university professor or a student. On the other hand, studies on human behavior in war situations for instance are very difficult to be carried out (Schnell et al. 2011: 3 f.). Also keep in mind that before thinking about the survey questions, you must have identified