In the following report, Hanover Research outlines best practices in questionnaire design from selected prominent sources.
# Table of Contents

Executive Summary and Key Findings ................................................................. 3

Introduction ....................................................................................................... 3

Key Findings .................................................................................................... 3

Best Practices in Questionnaire Design ............................................................ 4

Purpose of a Survey ......................................................................................... 4

Question Types ............................................................................................... 6
- Closed Questions ......................................................................................... 7
- Open-Ended Questions .............................................................................. 9

Questionnaire Format .................................................................................... 10
- Question Order ......................................................................................... 10
- Question Wording ..................................................................................... 12

Check List ..................................................................................................... 13
EXECUTIVE SUMMARY AND KEY FINDINGS

INTRODUCTION

In this report Hanover Research outlines best practices in questionnaire design from selected prominent sources: (1) the UNESCO International Institute for Educational Planning’s *Quantitative research methods in educational planning* series, Module 8: Questionnaire Design; and (2) the Pew Research Center for the People & the Press. Additional information was collected from (3) the Research Methods Knowledge Base.

KEY FINDINGS

- **There are appropriate uses for different question formats that rely primarily on the type of data required to address the hypothesis being tested with that question.** Closed questions and open-ended questions each have benefits and limitations that must be considered. Similarly, a survey may not be the best means to capture the desired information. Other options include interviews, focus groups, and direct observation of activities or environments.

- **The details of a survey can have unintended effects on the way respondents answer each item.** While these effects cannot be completely eliminated, careful consideration of question type, question wording, and question order can help minimize the survey designer’s influence on the respondent. Randomization, alternative formats, and pre-testing items can help the survey designer understand internal influences on the respondent.

- **Surveys should exhibit word economy and item efficiency.** While internal test/re-test for target concepts is an appropriate approach, overall the survey should be short, easy to understand, and minimize the amount of conscious thought the respondent must use to provide answers. This will help the survey produce valid and reliable results that accurately measure the topics being tested.
BEST PRACTICES IN QUESTIONNAIRE DESIGN

PURPOSE OF A SURVEY

Surveys are one means of data collection. In the case of student exit surveys, the population is presumably defined, literate, and accessible based on previous interactions with the researching institution: their college or university. However, additional considerations should be made regarding the types of information that are hoped to be gained from the research. The UNESCO module categorizes the types of information that can be gathered from a survey into three groups, detailed as follows:

In the field of educational planning, the information that is collected [from a questionnaire] can be classified broadly into: (a) inputs to education (such as school resources or various background characteristics of schools, teachers or students), (b) learning and teaching processes, and (c) the outcomes of education (such as pupil achievement, attitudes towards school, and measures of school efficiency such as survival rates etc.).1

The Research Methods Knowledge Base brings up several other sets of considerations, including question issues, content issues, and bias issues. Below we reproduce their guiding questions to test the relevance of a survey for the given research purpose:

- **Question Issues**
  - **What types of questions can be asked?** Are you going to be asking personal questions? Are you going to need to get lots of detail in the responses? Can you anticipate the most frequent or important types of responses and develop reasonable closed-ended questions?
  - **How complex will the questions be?** Sometimes you are dealing with a complex subject or topic. The questions you want to ask are going to have multiple parts. You may need to branch to sub-questions.
  - **Will screening questions be needed?** A screening question may be needed to determine whether the respondent is qualified to answer your question of interest. For instance, you wouldn’t want to ask someone their opinions about a specific computer program without first “screening” them to find out whether they have any experience using the program. Sometimes you have to screen on several variables (e.g., age, gender, experience). The more complicated the screening, the less likely it is that you can rely on paper-and-pencil instruments without confusing the respondent.
  - **Can question sequence be controlled?** Is your survey one where you can construct in advance a reasonable sequence of questions? Or, are you doing an initial exploratory study where you may need to ask lots of follow-up questions that you can’t easily anticipate?

---

Will lengthy questions be asked? If your subject matter is complicated, you may need to give the respondent some detailed background for a question. Can you reasonably expect your respondent to sit still long enough in a phone interview to ask your question?

Will long response scales be used? If you are asking people about the different computer equipment they use, you may have to have a lengthy response list (CD-ROM drive, floppy drive, mouse, touch pad, modem, network connection, external speakers, etc.). Clearly, it may be difficult to ask about each of these in a short phone interview.

Content Issues

- Can the respondents be expected to know about the issue? If the respondent does not keep up with the news (e.g., by reading the newspaper, watching television news, or talking with others), they may not even know about the news issue you want to ask them about. Or, if you want to do a study of family finances and you are talking to the spouse who doesn’t pay the bills on a regular basis, they may not have the information to answer your questions.

- Will respondent need to consult records? Even if the respondent understands what you’re asking about, you may need to allow them to consult their records in order to get an accurate answer. For instance, if you ask them how much money they spent on food in the past month, they may need to look up their personal check and credit card records. In this case, you don’t want to be involved in an interview where they would have to go look things up while they keep you waiting (they wouldn’t be comfortable with that).

Bias Issues

- Can social desirability be avoided? Respondents generally want to "look good" in the eyes of others. None of us likes to look like we don't know an answer. We don't want to say anything that would be embarrassing. If you ask people about information that may put them in this kind of position, they may not tell you the truth, or they may "spin" the response so that it makes them look better. This may be more of a problem in an interview situation where they are face-to-face or on the phone with a live interviewer.

- Can interviewer distortion and subversion be controlled? Interviewers may distort an interview as well. They may not ask questions that make them uncomfortable. They may not listen carefully to respondents on topics for which they have strong opinions. They may make the judgment that they already know what the respondent would say to a question based on their prior responses, even though that may not be true.

- Can false respondents be avoided? With mail surveys it may be difficult to know who actually responded. Did the head of household complete the survey or someone else? Did the CEO actually give the responses or instead pass the task off to a subordinate? Is the person you’re speaking with on the phone actually who they say they are? At least with personal interviews, you have a reasonable chance of knowing who you are speaking with. In mail surveys or phone interviews, this may not be the case.2

Within a given survey, it is important to keep these broad questions in mind as not all items may be appropriately addressed in or suited for a survey instrument.

---

**Question Types**

The general question types are *closed* and *open-ended*. The UNESCO module mentions a third type of question, the *contingency* question, but this is only differentiated from the other two types in its function, not its format. The type or format of a question can have a powerful impact on the observed results. Consider this example from the Pew Research Center for the People & the Press:

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.³

In this example, we see that the format of the question influences the frequency of responses and the set of responses received. It also indicates more generally that respondents can be influenced by the presence or absence of pre-existing options. The following figure summarizes the question types.

---

CLOSED QUESTIONS

Closed questions have a finite set of responses from which the respondent can choose one or more. The advantages and disadvantages of this question type are summarized below:

<table>
<thead>
<tr>
<th>ADVANTAGES</th>
<th>DISADVANTAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>▪ the respondent is restricted to a finite (and therefore more manageable) set of responses,</td>
<td>▪ they can introduce bias, either by forcing the respondent to choose between given alternatives or by offering alternatives that otherwise would not have come to mind,</td>
</tr>
<tr>
<td>▪ they are easy and quick to answer,</td>
<td>▪ they do not allow for creativity or for the respondent to develop ideas,</td>
</tr>
<tr>
<td>▪ they have response categories that are easy to code, and</td>
<td>▪ they do not permit the respondent to qualify the chosen response or express a more complex or subtle meaning,</td>
</tr>
<tr>
<td>▪ they permit the inclusion of more variables in a research study because the format enables the respondent to answer more questions in the same time required to answer fewer open-ended questions.</td>
<td>▪ they can introduce bias, where there is a tendency for the respondent to tick systematically either the first or last category, to select what may be considered as the most socially desirable response alternative, or to answer all items in a list in the same way, and</td>
</tr>
<tr>
<td></td>
<td>▪ they require skill to write because response categories need to be appropriate, and mutually exclusive.</td>
</tr>
</tbody>
</table>

Source: UNESCO

As noted in the table above, issues of bias are strong with closed questions because the respondent is presented with what the researcher feels are the best and exhaustible list of responses relative to a given question, often summarized in just four or five categories.

In scaled answers, such as level of agreement (disagree, agree), or frequency (never, rarely, sometimes), “it is important to set the time points defining the scale so that they make sense in relation to the specific activity of interest, and to the purpose for which the data are collected. [...] The categories [should be] defined so that they do not overlap, and are not too distant from one another.”

---

with, defined similarly to other options in the scale, and clearly distinguished from the other options as well.

Additionally, it is important to recognize differences in question types. For example, there is a difference between *ranking* (asking for ordering of a fixed list) and *rating* (asking for judgment on each item in a list, separately). The UNESCO module emphasizes, "The advantages of rank order methods are basically that it is easy for respondents to understand the instructions, and the questions force discrimination among objects. One of the disadvantages is that forced responses may not yield a real degree of preference or attitude, but rather information that the respondent prefers one object over another."6

These concepts of ordered options and ranking-vs-rating lead to another type of closed-question problem: that of measuring attitudes. The UNESCO module summarizes the issues around surveying attitudes as follows:

The **main difficulties in measuring attitudes** are that (a) the object of an attitude can range from the very specific to the very general, (b) attitudes are not static, and (c) attitudes are both shaped and changed by socio-demographic circumstances and life experiences. [...] The most popular approach to attitude measurement has been via the use of attitude scales. Attitude scales usually consist of a number of attitude statements which are presented to respondents with a request that they should indicate whether they agree or disagree. Scaling techniques are deployed to order respondents along some underlying attitudinal continuum.7

Additionally, the authors recommend randomization to minimize the effect of repetition – agreeing with one statement because one agreed to the previous statement.8 The Pew Research Center summarizes issues around the design of rating questions as follows:

- **Error of proximity**: the tendency to rate items similarly because they are near to each other in the questionnaire
- **Central tendency error**: the tendency to rate most items in the middle category (when the middle category is offered). Such respondents either dislike extreme positions, or lack knowledge.
- **Error of leniency**: the tendency to give high ratings to most items by liking or agreeing with everything.
- **Error of severity**: the opposite to the error of leniency: respondents who dislike, or disagree, with most items.
- **Halo effect error**: the tendency to rate a particular statement according to how respondents feel about it in general.9

---

6 Ibid., p. 60.
7 Ibid., p. 61.
8 Ibid., p. 65.
9 Ibid., p. 66.
Finally, the most common type of scaled question, the Likert-type question, requires certain assumptions on the part of the survey designer. These are as follows:

- That there is a continuous underlying dimension which is assessed by total scores on the attitude scale and that each item contributes in a meaningful way to the measurement of this dimension.
- That a more favorable attitude will produce a higher expected score, and vice-versa.
- That items are of equal value in that they each provide a replicated assessment of the dimension measured by the total score on the scale.\(^{10}\)

It is important to keep these assumptions and potential effects in mind when crafting the wording and placement of each question. Additional information on question order is provided in the next section, “Questionnaire Format.”

**OPEN-ENDED QUESTIONS**

Open-ended questions have a potentially infinite set of responses, allowing the respondent a relatively un-restricted space for generating thoughts and experiences. The advantages and disadvantages of this question type are summarized below:

<table>
<thead>
<tr>
<th>ADVANTAGES</th>
<th>DISADVANTAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>- they allow respondents to express their ideas spontaneously in their own language,</td>
<td>- they may be difficult to answer and even more difficult to analyze,</td>
</tr>
<tr>
<td>- they are less likely to suggest or guide the answer than closed questions because they are free from the format effects associated with closed questions, and</td>
<td>- they require effort and time on behalf of the respondent,</td>
</tr>
<tr>
<td>- they can add new information when there is very little existing information available about a topic.</td>
<td>- they require the development of a system of coded categories with which to classify the responses,</td>
</tr>
</tbody>
</table>

| Source: UNESCO\(^{11}\) |

In surveys the number of open-ended questions should be minimal, if at all present. A survey with many open-ended questions indicates that another data collection method (interview, focus group, observation) is likely more appropriate. However, there are cases when open-ended questions are appropriate in a survey. Examples would include defining

\(^{10}\) Ibid., pp. 66-67.
\(^{11}\) Ibid., pp. 32-33.
“Other” categories to check the exhaustiveness of a given list of answer options in a closed question, or providing a final opportunity for feedback at the end of the survey.

The UNESCO module cautions that “[c]are should be taken in writing open-ended questions so as to avoid formats that elicit a dichotomous yes/no or agree/disagree response. In addition, the wording of questions should seek to reduce the possibility of eliciting responses that are aligned along very different dimensions and therefore cannot be systematically coded.” That is, open-ended questions should clearly require phrasal responses, and should define the scope of responses tightly (e.g., “Please comment on the efficiency of food service in the dining hall.” versus “What do you think about food service in the dining hall.”)

**QUESTIONNAIRE FORMAT**

In addition to question-specific considerations, the entire survey instrument (or questionnaire) must be treated carefully. Considerations can be summarized in two concepts: question order, and question wording. (Additional considerations such as style will not be covered in this report.) The first sub-section identifies general guidelines for knitting together questions within a larger survey framework. The second sub-section provides guidelines on developing an appropriate overall tone for the survey instrument.

**QUESTION ORDER**

As indicated in the previous sections, the context for a specific question within the overall survey instrument can impact how respondents answer the item. The concept of question order suggests ways to manage the kinds of interaction effects that arise from asking respondents multiple questions in a survey. Overall, reports the UNESCO module, there are two patterns of question sequence, described below:

- **funnel sequence**
- **inverted funnel sequence**

The characteristic of the funnel sequence is that each question is related to the previous question and has a progressively narrower scope. The first question can be either open format, or multiple choice. It should be very broad, and is used to ascertain something about the respondent’s frame of reference on a topic. This ordering pattern is particularly useful when there is a need to prevent further specific questions from biasing the initial overall view of the respondent.

In the inverted funnel sequence, specific questions on a topic are asked first, and these eventually lead to a more general question. This sequence requires the respondent to think through his or her attitude before reaching an overall evaluation on the more general question. Such a question order is particularly appropriate when there is reason to believe that respondents have neither a strong feeling about a topic, nor a previously formulated view.

---

12 Ibid., p. 33.
13 Ibid., p. 69.
14 Ibid., p. 70.
The Pew Research Center explains that “asking a more specific question before a more general question” about the same topic “can result in a contrast effect.” That is, if someone is asked to evaluate a specific program, then the overall experience, s/he may exclude the specific program from his/her evaluation of the overall experience.15

Beyond these two patterns – one which places broad questions first, the other which places them second – there are additional considerations for placing particular items. For example, where do demographic questions go? How does one ensure maximum response rates for the questions highlighting the items of major interest to the research? The UNESCO module suggests the following general guidelines:

- Non-sensitive demographic questions should be placed at the beginning of the questionnaire because they are easy to answer, non-threatening, and tend to put the respondent at ease.
- Items of major interest to the research study should be placed next since there is greater probability of the respondent answering or completing the first section of the questionnaire.
- Sensitive items that cover controversial topics should be placed last so that potential resentment that may be provoked by these items does not influence responses to other questions.
- Items on the same topic should be grouped together. However, care should also be taken to prevent one item influencing responses to later items.
- Items with similar response formats should be grouped together when several different response formats are being used within a questionnaire.
- Section titles should be used to help the respondent focus on the area of interest.16

The Pew Research Center explains further that question types can interact with each other. For example, they write, “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.” On the other hand, they continue, “For closed-ended 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.”17 This suggests caution when placing open-ended questions into the survey instrument.

However, it is important to consider during instrument revisions that “modifying [the] context of the question” by changing order or phrasing “could call into question any observed changes over time.”18

18 Ibid.
**QUESTION WORDING**

Many of the question-wording suggestions have been addressed in the previous sections considering specific question types. However, regardless of question type, there are a few general guidelines for determining the overall tone and diction of a survey. The Pew Research Center indicates the following key considerations for question wording:

- **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).

- **Ask only one question at a time.** Questions that ask respondents to evaluate more than one concept (known as double-barreled questions) [...] are difficult for respondents to answer and often lead to responses that are difficult to interpret.

- **Use simple and concrete language.** [...] 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.

- **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. [...] Several experiments have shown that there is much greater public support for expanding ‘assistance to the poor’ than for expanding ‘welfare.’

- **Consider the acquiescence bias,** which suggests that some respondents are more likely to be agreeable to offered statements – i.e., to suggest everything is positive – than 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.19

The UNESCO module offers a similar list, advising survey designers to:

- **Keep the vocabulary simple**
  - If a rare or technical term has to be used, then its meaning should be explained.
  - Acronyms and abbreviations should always be spelled out in the questionnaire. Do not assume that respondents will or should know what an acronym represents.
  - When a general term is used (e.g., “systematically”, “illustrations”, “frequently”), concrete examples should be given to clarify its meaning.

- **Keep the question short**

---

Avoid double-barrelled questions
Avoid hypothetical questions
Don’t overtax the respondent’s memory
Avoid double negatives
Avoid overlapping response categories
Beware of ‘leading’ questions

**CHECK LIST**

Overall themes of survey design best practices include attention to detail, attention to the audience, and attention to the research question or hypothesis being tested via this data collection method. The authors of the UNESCO guidelines offer a “checklist for reviewing questionnaire items,” which suggests the following critical questions:

- Will the item provide data in the format required by the research questions or the hypotheses?
- Is the item unbiased?
- Will the item generate data at the level of measurement required for the analysis?
- Is there a strong likelihood that most respondents will answer the item truthfully?
- Do most respondents possess sufficient knowledge to answer the item?
- Will most respondents be willing to answer the item, or is it too threatening or too sensitive?
- Does the item avoid ‘leading’ respondents to a specific answer?
- Is the language used in the questionnaire clear and simple – so that all respondents are able to understand all of the questions?

---

21 Ibid., pp. 41-42.
PROJECT EVALUATION FORM

Hanover Research is committed to providing a work product that meets or exceeds member expectations. In keeping with that goal, we would like to hear your opinions regarding our reports. Feedback is critically important and serves as the strongest mechanism by which we tailor our research to your organization. When you have had a chance to evaluate this report, please take a moment to fill out the following questionnaire.


CAVEAT

The publisher and authors have used their best efforts in preparing this brief. The publisher and authors make no representations or warranties with respect to the accuracy or completeness of the contents of this brief and specifically disclaim any implied warranties of fitness for a particular purpose. There are no warranties which extend beyond the descriptions contained in this paragraph. No warranty may be created or extended by representatives of Hanover Research or its marketing materials. The accuracy and completeness of the information provided herein and the opinions stated herein are not guaranteed or warranted to produce any particular results, and the advice and strategies contained herein may not be suitable for every member. Neither the publisher nor the authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. Moreover, Hanover Research is not engaged in rendering legal, accounting, or other professional services. Members requiring such services are advised to consult an appropriate professional.