PREVENTING CHRONIC DISEASE
PUBLIC HEALTH RESEARCH, PRACTICE, AND POLICY
VOLUME 2: NO. 1 JANUARY 2005

SPECIAL TOPICS

A Catalog of Biases in Questionnaires

Bernard C.K. Choi, PhD, Anita W.P. Pak, PhD

Suggested citation for this article: Choi BCK, Pak AWP. A catalog of biases in questionnaires. Prev Chronic Dis [serial online] 2005 Jan [*date cited*]. Available from: URL: http://www.cdc.gov/pcd/issues/2005/jan/ 04_0050.htm.

PEER REVIEWED

Abstract

Bias in questionnaires is an important issue in public health research. To collect the most accurate data from respondents, investigators must understand and be able to prevent or at least minimize bias in the design of their questionnaires. This paper identifies and categorizes 48 types of bias in questionnaires based on a review of the literature and offers an example of each type. The types are categorized according to three main sources of bias: the way a question is designed, the way the questionnaire as a whole is designed, and how the questionnaire is administered. This paper is intended to help investigators in public health understand the mechanism and dynamics of problems in questionnaire design and to provide a checklist for identifying potential bias in a questionnaire before it is administered.

Introduction

Health surveys have been and will continue to be important sources of information for evidence-based public health and medicine (1). The principal instrument for collecting data in surveys is the questionnaire. To allow the investigator to collect the most accurate data from respondents, the questionnaire must be unbiased. Bias is a pervasive problem in the design of questionnaires.

In this paper bias is defined as a "deviation of results or inferences from the truth, or processes leading to such a deviation" (2). Questionnaire bias is a result of unanticipated communication barriers between the investigator and respondents that yield inaccurate results. Bias may arise from the way individual questions are designed, the way the questionnaire as a whole is designed, and how the questionnaire is administered or completed.

Based on a review of the literature, this paper identifies and categorizes 48 common types of bias in questionnaires and provides an example of each type in addition to a brief explanatory comment. This paper goes beyond the general lists of biases that previous authors have provided (3-6) by cataloging the types of bias according to their source (Table). Organizing the types of bias in this way keeps related problems together and makes analysis of a questionnaire easier. The catalog below is meant to help public health investigators understand the mechanism and dynamics of problems in questionnaire design and to provide a checklist for identifying bias in a questionnaire before it is used as a survey instrument.

Types of Bias in Question Design

Problems with wording

Ambiguous question. Ambiguous questions lead respondents to understand the question differently than was intended and, therefore, to answer a different question than was intended (1).

Example: *Is your work made more difficult because you are expecting a baby?*

This question is ambiguous. A "no" answer may mean, "No, I'm not expecting a baby," or "No, my work is not made more difficult" (7).

Complex question. Complex and lengthy questions should be avoided in a questionnaire.

Example: *Has it happened to you that over a long period of time, when you neither practiced abstinence, nor used birth control, you did not conceive?*

This question, which was used in a survey on family planning conducted for the Royal Commission on Population (8), is vague ("a long period of time"), too formal ("Has it happened to you that"), and complex, because of its length and use of the neither/nor construction.

Double-barrelled question (also known as two questions in one). Questions that are made up of two or more questions make it difficult for the respondent to know which part of the question to answer and for the investigator to know which part of the question the respondent actually answered (9).

Example: *Do you agree that acquired immunodeficiency syndrome (AIDS) can be transmitted by shaking hands with a person with AIDS or through other means of physical contact?*

A "no" answer may mean not by shaking hands, or not through other means of physical contact, or both (1).

Short question. Short questions may not be as accurately answered as questions that are longer. A question that is short may come across as abrupt in an interview situation. Questions that include a transition to the next topic give respondents more time to gather their thoughts and also more clues to use in formulating their responses (7).

Example: *Have you had bad sore throats?*

Now a question about bad sore throats. We're looking for information about these. Have you had bad sore throats?

The first question has been found to be less accurately answered than the second question when compared with information obtained from the respondents' physicians (10). The second question includes an introduction that sets up the query.

Technical jargon. Technical jargon and the profession's technical terms may not be understood by the general public and should be avoided.

Example: *What was your age at menarche?*

What was your age when your menstrual periods started?

The technical term in the first question may not be understood by many women, so it is preferable to ask the same question in more common terms, as in the second question (9).

Uncommon word. Uncommon and difficult words should be avoided in questionnaires.

Example: Gowers (11) and Day (12) have produced lists of words that can be replaced by simpler alternatives. For example:

Uncommon	Common
Assist	Help
Consider	Think
Effectuate	Cause
Elucidate	Explain
Employ	Use
Initiate	Begin/Start
Major	Important/Main
Perform	Do
Quantify	Measure
Require	Want/Need
Reside	Live
State	Say
Sufficient	Enough
Terminate	End
Ultimate	Last
Utilize	Use

Use common words in questionnaires, especially questionnaires targeted for the general population, to avoid misunderstanding.

Vague word. Vague words in vague questions encourage vague answers (1).

Example: *How often do you exercise?*
[ ] Regularly
[ ] Occasionally
How often do you exercise?

[ ] twice a week or more often
[ ] once a week
[ ] less than once a week

The first question is vague because "occasionally" and "regularly" are not defined. The meaning can easily be made more precise, as in the second question.

Missing or inadequate data for intended purpose

Belief vs behavior (also known as hypothetical question or personalized question). Questions that ask the respondent about a belief (hypothetical) can yield quite different answers than questions that ask the respondent about his or her behaviors (personalized) (9).

Example: *Do you think that it is a good idea to have everyone's chest regularly checked by X-ray?*

Have you ever had yours checked?

The two questions generated different results. Ninety-six percent of the respondents answered "yes" to the first question, but only 54% answered "yes" to the second (13). The answers to both questions may be accurate even though the results are different. The investigator must determine whether the purpose of the question is to collect data regarding a belief or a behavior and design the question accordingly.

Starting time. Failure to identify a common starting time for exposure or illness may lead to bias (3).

Example: *In the last 12 months, have you had an accident causing head injury?*

Because a survey is normally conducted over an extended period, the time frame of "last 12 months" will vary depending on the date of the interview. The data obtained therefore cannot be used to estimate incidence rates. The following question is better and will provide a common time frame:

From January 1 to December 31 of last year, did you have an accident causing head injury?

Data degradation. It is better to collect accurate, continuous data at source instead of degraded data. Once degraded data have been collected, it is impossible to recover the original continuous data or to change cut-off criteria for categories (7).

Example: *What is your birth date?*

What is your age in years?

Which age category do you belong to?

For information on age, the first question is the best because it can provide accurate continuous data, followed by the second question. The third question is the least desirable because data are degraded (1).

Insensitive measure. When outcome measures make it impossible to detect clinically significant changes or differences, Type II errors occur (3).

Example: *How important is health to you, on a scale of 1 to 3?*
(Unimportant) 1 - 2 - 3 (Important)

How important is health to you, on a scale of 1 to 10?
(Unimportant) 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 (Important)

The first question may not have sufficient discriminating power to differentiate the respondents because of the limited categories. The second question may be better.

Faulty scale

Forced choice (also known as insufficient category). Questions that provide too few categories can force respondents to choose imprecisely among limited options (7,9).

Example: *Do you agree?* Yes [ ] No [ ]

Do you agree? Yes [ ] No [ ] Don't Know [ ]

The first question, which does not have a "don't know" category, may produce a bias because respondents who have no opinion are forced to select an answer that may or may not reflect their true feelings. The second question is recommended.

Missing interval. Missing intervals in response choic-

es can cause confusion.

Example: *How often does the computer system go down?*
[ ] Less than once per month
[ ] Once per month
[ ] Once per week
[ ] More than once per week

Respondents do not have a place to put "once every two weeks." The following response categories are recommended:

How often does the computer system go down?
[ ] Less than once per month
[ ] Once per month to once per week
[ ] More than once per week

Overlapping interval. Overlapping intervals in response choices can cause confusion (9).

Example: *How many cigarettes do you smoke per day?*
[ ] None [ ] 5 or less [ ] 5-25 [ ] 25 or more

Respondents smoking exactly 5 or 25 cigarettes per day do not know in which category to place themselves. The following question is more appropriate:

How many cigarettes do you smoke per day?
[ ] None [ ] 4 or less [ ] 5-24 [ ] 25 or more

Scale format. An even or an odd number of categories in the scale for the respondents to choose from may produce different results.

Example: *Do you agree?* (Agree) 1 – 2 – 3 (Disagree)
Do you agree? (Agree) 1 – 2 – 3 – 4 (Disagree)

The first question, with an odd number of categories, tends to result in neutral answers (i.e., 2), and the second question, with an even number of categories, tends to force respondents to take sides (1). The two approaches produce different results, but there is no general consensus as to which one is better.

Leading questions

Framing. Some questions may be framed in such a manner that respondents choose an inaccurate answer.
Example: *Which operation would you prefer?*
[ ] An operation that has a 5% mortality.
[ ] An operation in which 90% of the patients will survive.

Patients scheduled for surgery may choose the second option when they see or hear the words "90%" and "survive," but in fact a 90% survival rate (or 10% mortality) is worse than a 5% mortality (1).

Leading question. Different wording of the same question can guide or direct respondents toward a different answer (1,7).

Example 1: *Do you do physical exercise, such as cycling?*

This is a leading question because it will likely lead the respondent to focus only on cycling.

Example 2: *Don't you agree that . . . ?*

This negatively worded question leads respondents to answer no (14). The preferred phrasing is, "Do you agree or disagree that . . . ?"

Mind-set. The mind-set of the respondent can affect his or her perception of questions and therefore can affect answers.

Example: *1. How many cigarettes do you smoke per week?*
2. How many cigars do you smoke per week?
3. How many beers do you drink per month?

The change in wording from "per week" to "per month" can result in wrong answers for the third question above, because of the possible mind-set of the respondents.

Intrusiveness

Reporting (also known as self-report response). A respondent may selectively suppress information, such as past history of sexually transmitted disease (2,15).

Example: *In the past five years, have you engaged in anal intercourse, that is, rectal intercourse?*

This question is so direct and up-front that many people may refuse to answer. The following question may reduce reporting bias by deliberately loading the question to suggest that others also engage in the behavior (1): *People*

practice many different sexual activities, and some people practice things that other people do not. In the past five years, have you engaged in anal intercourse, that is, rectal intercourse?

Sensitive question. Sensitive questions, such as age, personal or household incomes, sexual orientation, or marital status, may elicit inaccurate answers and may also affect the interviewer-interviewee relationship so that all subsequent answers can be affected.

Example: *How old are you?*

In what year were you born?

The first question, which is direct, tends to result in a high percentage of refusals to answer. The second question tends to yield fairly accurate responses (1).

Inconsistency

Case definition. Definition of cases based on different versions of the International Classification of Disease (ICD) codes, for example, or first-ever cases vs recurrent cases, may change over time or across regions, resulting in inaccurate trends and geographic comparisons (16).

Example: *How many bladder cancer cases do you see in a year?*

How many histologically confirmed bladder cancer cases do you see in a year?

The use of two different case definitions can present problems when comparing results.

Change of scale. If the measurement scale for a quantity changes in different surveys, the results may not be comparable.

Example: *Compared to other persons your age, would you say your health is excellent, good, fair, or poor?*

Would you say your health in general is excellent, very good, good, fair, or poor?

The first question, which was used in the 1985 National Center for Health Statistics National Health Interview Survey (NCHS-NHIS) (1), has four categories of health, and the second question (1995 NCHS-NHIS) (1) has five categories. Therefore, the categories may not mean exactly the same in the two surveys and will cause a problem for comparison over time.

Change of wording. If the precise wording of a question changes in different surveys, the results may not be comparable (7).

Example: *Compared to other persons your age, would you say your health is excellent, good, fair, or poor?*

Would you say your health in general is excellent, very good, good, fair, or poor?

The first question (1985 NCHS-NHIS) (1) and the second question (1995 NCHS-NHIS) (1) use different wording, namely, "compared to other persons your age" vs "in general." This may guide respondents to evaluate their health in a different context.

Diagnostic vogue. The same illness may receive different diagnostic labels at different points in space or time (3).

Example: *Do you have bronchitis?*

Do you have emphysema?

The terms "bronchitis" and "emphysema" are used in Great Britain and in North America, respectively, to refer to the same disease (3). It is therefore important to use the term that is appropriate in space and time.

Types of Bias in Questionnaire Design

Formatting problem

Horizontal response format. In self-administered questionnaires, horizontal vs vertical format of the response choices can affect the answers (17).

Example: *Your health is:*
Excellent ... [ ] Good ... [ ] Fair ... [ ] Poor ... [ ]

Your health is:
Excellent . . . . [ ]
Good . . . . . [ ]

Fair . . . . . [ ]
Poor . . . . . [ ]

The horizontal response format (first example) can cause confusion among the respondents because of poor spacing and may result in the wrong answers being checked or circled. The vertical response format (second example) has been suggested as better for listing response options (17).

Juxtaposed scale (also known as questionnaire format). Juxtaposed scales, a type of self-report response scale that asks respondents to give multiple responses to one item, may elicit different responses than separate scales (18).

Example*: 1. Indicate how important and how satisfied you are with each of the following using a scale of 1 to 5:*

(Unimportant) 1 - 2 - 3 - 4 - 5 (Important)
(Dissatisfied) (Satisfied)

	Importance	Satisfaction
a. Your family	_ _ _ _ _ _ _ _	_ _ _ _ _ _ _ _ _ _ _
b. Your career	_ _ _ _ _ _ _ _	_ _ _ _ _ _ _ _ _ _ _
c. Your marriage	_ _ _ _ _ _ _ _	_ _ _ _ _ _ _ _ _ _ _

The above question is in a juxtaposed scale format. The advantage is that it can force respondents to think and compare the importance and satisfaction for each item because they are side by side. However, this questionnaire format has been shown to cause confusion among respondents who are less educated, in which case the following question, with parts A and B in a separate scale format, may be preferred (18):

1A. Indicate how important each of the following is to you using a scale of 1 to 5:

(Unimportant) 1 - 2 - 3 - 4 - 5 (Important)

	Importance
a. Your family	_____
b. Your career	_____
c. Your marriage	_____

1B. Indicate how satisfied you are with each of the following using a scale of 1 to 5:

(Dissatisfied) 1 - 2 - 3 - 4 - 5 (Satisfied)

	Satisfaction
a. Your family	_____
b. Your career	_____
c. Your marriage	_____

Left alignment and right alignment. Alignment of the response choices to the left or right side of the possible responses can produce a bias (17).

Example: *Your health is:*
excellent . . .[ ]
good . . . . . .[ ]
fair . . . . . .[ ]
poor . . . . . .[ ]

Your health is:
[ ] excellent
[ ] good
[ ] fair
[ ] poor

It has been suggested that placing the response choices to the right side of (i.e., after) the list of possible responses will result in fewer errors on the part of interviewers in a personal or telephone interview. This facilitates subsequent data input directly from the questionnaire. For mailed and other self-administered questionnaires, placing the response choices to the left of (i.e., before) the possible responses makes it easier for the respondent to circle or check them (17).

Questionnaire too long

No-saying (also known as nay-saying) and **yes-saying** (also known as yea-saying). Some respondents tend to answer no to all questions or to answer yes to all questions (1).

Example*: What are the reasons why you do not exercise daily?*

	Yes	No
It takes too much time	[X]	[ ]
There is not enough time in my day	[X]	[ ]
There is no equipment at home	[X]	[ ]
There are no community resources	[X]	[ ]
I feel I am not trained to do it	[X]	[ ]
I feel I do not want to do it	[X]	[ ]
I am too tired	[X]	[ ]
It is too difficult	[X]	[ ]

In the above example, the respondent chooses yes for all

items. One way to reduce the no- or yes-saying bias is to use both positive and negative statements about the same issue in a battery of items to break the pattern (1), as in the following example:

	Yes	No
1. People with AIDS deserve to have the disease	[ ]	[X]
2. People with AIDS should be given more help	[X]	[ ]

Open question (also known as open-ended question). Open-ended questions can result in data with differential quality (14). Also, respondents are likely to be unwilling to take the time to answer them.

Example: *What kind of physical exercise do you do?*

This open-ended question presents a difficult recording task. The interviewer must decide whether to record everything that the respondent says, record only what the interviewer considers relevant, or paraphrase the respondent's answer. However, in some circumstances, open-ended questions are more appropriate than close-ended questions, particularly in surveys of knowledge and attitudes, and can yield a wealth of information through appropriate qualitative methods such as content analysis (7).

Response fatigue. Questionnaires that are too long can induce fatigue among respondents and result in uniform and inaccurate answers (19).

Example: *Now we would like to move on to our Question no. 618, concerning the health of your pet fish...*

Personal interviews usually last 50 to 90 minutes; telephone interviews typically last 30 to 60 minutes; self-administered questionnaires typically take 10 to 20 minutes to complete (19). From field experience, interviewers and respondents report these times to be acceptable, and common sense suggests that much longer times are not feasible. Respondents are unable to concentrate and give correct answers in a lengthy interview, especially if the topics are not of interest. Toward the end of a lengthy session, respondents tend to say all yes or all no or refuse to answer all remaining questions (1).

Flawed questionnaire structure

Skipping question. Skipping questions may lead to the loss of important information because of logical errors in the flow of questions.

Example: *1. Are you self-employed?*
[ ] Yes
[ ] No (Go to question 8)

2. Do you smoke?
[ ] Yes
[ ] No

3. . . .

8. Do you use a cellular telephone?
[ ] Yes
[ ] No

The above questions, because of errors in the skipping sequence, will not collect smoking information for those who are not self-employed. Pretesting of the survey instrument should prevent such a bias.

Types of Bias in Administration of Questionnaire

Interviewer not objective

Interviewer. Bias can be caused by an interviewer's subconscious or even conscious gathering of selective data (2,4), which can result from inter-interviewer or intrainterviewer errors (4).

Example: *Do you smoke?* Yes [ ] No [ ]

If an interviewer knows that the respondent does not have a smoking-related disease, and therefore is unlikely to be a smoker, he or she may rephrase the question and ask instead, "You don't smoke, do you?" This is a leading question and is likely to lead to a negative answer (14). Proper interviewer training is needed to prevent such biases.

Nonblinding. When an interviewer is not blind to the study hypotheses, he or she may consciously gather selective data (20).

Example: *1. Do you have lung cancer?* Yes [ ] No [ ]

2. Do you smoke? Yes [ ] No [ ]

The first question reveals to the interviewer the disease status of the respondent, and this may affect the way he or she asks or records the answer for the second question. Besides providing interviewer training, it is also important to ensure that interviewers are blind to the study hypotheses.

Respondent's subconscious reaction

End aversion (also known as central tendency). Respondents usually avoid ends of scales in their answers. They tend to try to be conservative and wish to be in the middle (7).

Example: *Do you agree?*
[ ] Strongly agree
[ ] Agree
[ ] Disagree
[ ] Strongly disagree

Respondents are more likely to check "Agree" or "Disagree" than "Strongly agree" or "Strongly disagree" (1).

Positive satisfaction (also known as positive skew). Questions on satisfaction may cause problems.

Example:	Yes	No
1. Are you satisfied with your family?	[X]	[ ]
2. Are you satisfied with your career?	[X]	[ ]
3. Are you satisfied with your marriage?	[X]	[ ]

Respondents tend to give positive answers when answering questions on satisfaction (1).

Respondent's conscious reaction

Faking bad (also known as hello-goodbye effect). Respondents try to appear sick to qualify for support (1).

Example: *Which of the following symptoms do you have?*

Respondents tend to check more types of symptoms than they have (1).

Faking good (also known as social desirability, obsequiousness). Respondents may systematically alter questionnaire responses in the direction they perceive to be desired by the investigator (3). Socially undesirable answers tend to be under-reported (7).

Example: *Did you smoke during your pregnancy?*
Yes [ ] No [ ]

Mothers tend to answer no even if they smoked during pregnancy (1).

Unacceptable disease. Socially unacceptable disorders (e.g., sexually transmitted diseases, suicide, insanity) tend to be underreported (1).

Example: *Do you have a sexually transmitted disease?*

Ask these questions toward the end of the questionnaire so that they will not affect other questions. Also consider using anonymous, mailed questionnaires instead of face-to-face interviews.

Unacceptable exposure. Socially unacceptable exposures (e.g., smoking, drug abuse) tend to be underreported (1).

Example: *Do you now smoke cigarettes every day?*

A direct and intruding question like the one above may result in reporting inaccuracy. Instead, when asking about undesirable behaviors, it is better to ask whether the person had ever engaged in the behavior in the past before asking about current practices, because past events are less threatening (21). For example:

1. Have you smoked at least 100 cigarettes in your entire life?

2. Last year, were you smoking cigarettes every day?

3. Do you now smoke cigarettes every day?

Unacceptability. Measurements which hurt, embarrass, invade privacy, or require excessive commitment may be systematically refused or evaded (3).

Example (22): *We would now require two urine specimens from you. The first specimen will be collected over a 24-hour period, part of which will be while you are in your natural working environment, probably toward the end of the work week, such as on a Friday. The second specimen*

will be taken over another 24-hour period while you are at home, out of the work environment for at least 24 hours, such as on a Sunday. During collection, keep all urine samples refrigerated, in the refrigerator at home, or by the portable thermos bottle and ice-packs at work. When finished, please call the taxi company with the instruction sheet to deliver the samples to the laboratory.

Avoid measurements by intrusive means, or consider using incentives to increase participation rate.

Underlying cause (also known as rumination). Cases may ruminate about possible causes for their illness and thus exhibit different recall of prior exposures from those of controls (3).

Example: *Did you have skull x-rays in the past five years?*

In a case-control study of childhood brain tumors, a significantly elevated risk was reported by cases for skull x-rays compared to controls (23). It is not known whether this was a true effect of x-rays on brain tumors or of cases' thinking that x-rays were the cause of their illness.

Respondent's learning

Learning. Completing a questionnaire can be a learning experience for the respondent about the hypotheses and expected answers in a study.

Example (24):

1. Which of the following investigations would you order for a patient of yours with asthma-like symptoms?
[ ] spirometry
[ ] lung volumes, diffusing capacity
[ ] peak expiratory flow rate
[ ] chest X-ray

2. Under what conditions would you order spirometry for a patient?

Having thought about prior questions (such as the first question) can affect the respondent's answer to subsequent questions (e.g., the second question) through the learning process as the questionnaire is completed. To avoid learning bias, it may be necessary to randomize the order of the questions for different respondents.

Hypothesis guessing. Respondents may systematically alter questionnaire responses when, during the process of answering the questionnaire, they think they know the study hypothesis.

Example:	**Yes**	**No**
1. Does your child have headaches?	[ ]	[ ]
2. Does your child play with battery-operated toys?	[ ]	[ ]
3. Does your child play with batteries?	[ ]	[ ]
4. How many and which types of batteries do you have at home?		

The respondents, perceiving that the study is about headache and battery use, may overreport the number of batteries if they have a child with headaches.

Respondent's inaccurate recall

Primacy and **recency**. Depending on the type of questionnaire (interviewer-administered questionnaires or self-administered questionnaires), respondents may choose answers differently.

Example: (24): *Which of the following types of doctors did you see in the past year?*
[ ] family doctor
[ ] pediatrician
[ ] lung doctor/internist
[ ] allergy doctor/immunologist
[ ] emergency room doctor
[ ] some other kind

Research has indicated that in mailed surveys, respondents may tend to choose the first few response options on the list (primacy bias), though in telephone or personal interview surveys, they are more likely to respond in favor of the later categories (recency bias) (25,26). These effects can be minimized by reducing the number of categories presented to respondents and by randomizing the order of categories in survey instruments.

Proxy respondent (also known as surrogate data). For deceased cases or surviving cases (e.g., brain tumors) whose ability to recall details is defective, soliciting information from proxies (e.g., spouse, family members) may result in differential data accuracy. In general, it is not advisable to ask someone to answer attitudinal, knowledge, or behavior questions for others (1).
Example: *1. What is your wife's occupation?*

2. Please tell me how afraid your wife is of getting cancer?

Not at all afraid				Extremely afraid
1	2	3	4	5

The first question is appropriate but the second question is not.

Recall. This type of bias is because of differences in accuracy or completeness of recall prior to major events or experiences (3).

Example: *How many diagnostic x-ray examinations did you have when you were pregnant?*

It was found that mothers whose children have had leukemia were more likely than mothers of healthy children to remember details of diagnostic x-ray examinations to which these children were exposed *in utero* (2).

Telescope. Respondents usually recall an event in the distant past as happening more recently (1). This is a form of recall bias.

Example: *In an interview in May, an event which was thought to have occurred in March actually happened in November of the previous year.*

Telescope bias can be reduced by the bounded recall procedure in which respondents are interviewed at the beginning and end of the time period referenced in a survey questionnaire (1). The first interview would serve to identify events that occurred before the interview period so that they could clearly be eliminated if the respondent later reported that they occurred during the period between the first and second interviews. However, this procedure must ask people about the same thing twice.

Cultural differences

Cultural. The culture of the respondents can affect their perception of questions and therefore their answers (27).

Example: *What is your gross monthly income?*

The culture in North America and Europe is to think in terms of annual income. For the above question, it is inevitable that some respondents will put down a figure representing annual, not monthly, income. This question would be appropriate for a survey in Asia, however, since the culture there is to report monthly income. Pretesting the survey instrument should minimize this bias.

Discussion

Questionnaire bias is an important subject given the less than optimal questionnaires that are produced in the health research field. This paper can serve as a resource for health researchers and practitioners using questionnaires. It provides a catalog of types of bias that can be used as a checklist for identifying potential problems when designing and administering questionnaires.

This paper focuses on biases specific to questionnaires (design and administration). It does not cover such biases as sampling and selection biases (5). Nor is it within the scope of this paper to discuss such general practices as survey development, interviewer training, or for that matter, how to best conduct an adequate survey. For example, inadequate survey design may sometimes result in biases in sample selection, such as by not having the questionnaire translated in all necessary languages (language barrier bias), by restricting the survey to those subjects with telephones (telephone sampling bias) (5), or by selecting only those born close to the date of interview (next birthday bias) (28). There may be other common errors in survey development that may cause interpretational problems, including asking for family history of a disease (family history bias) (29), not having pilot surveys to pretest the questionnaires (lack of pretest bias) (14), or using telephone interviews where visual aids cannot be used to illustrate the questions (telephone interview bias) (19). Furthermore, different kinds of study methods, such as mailed questionnaires, personal interviews, telephone interviews, Web surveys, routine data and registries, surveillance systems, and focus groups, should be used depending on the nature of the study to avoid bias (wrong instrument bias, also known as wrong study method bias) (1,7). For a simple survey of an educated section of the population (e.g., a professional group) concerning a subject of interest to its members, a mailed questionnaire might be appropriate. On the other hand, a survey of the general population on detailed and complicated information would almost certainly call for a personal interview. These are questionnaire problems that affect or are related to study designs and, strictly speaking, are not questionnaire

biases. They are therefore not included in our catalog of questionnaire biases.

Author Information

Corresponding author: Bernard C.K. Choi, PhD, Department of Public Health Sciences, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada, Department of Epidemiology and Community Medicine, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada, Centre for Chronic Disease Prevention and Control, Public Health Agency of Canada, Room A147, 120 Colonnade Road, PL# 6701A, Ottawa, Ontario, Canada K1A 1B4. Telephone: 613-957-1074. Email: Bernard_Choi@phac-aspc.gc.ca.

Author affiliations: Anita W.P. Pak, PhD, Office of Institutional Research, University of Ottawa, Ottawa, Ontario, Canada.

References

1. Aday LA. Designing and conducting health surveys. 2nd ed. San Francisco (CA): Jossey-Bass; 1996.
2. Last JM. A dictionary of epidemiology. New York (NY): Oxford University Press; 2001.
3. Sackett DL. Bias in analytic research. J Chron Dis 1979;32:51-63.
4. Choi BC, Noseworthy AL. Classification, direction, and prevention of bias in epidemiologic research. J Occup Med 1992;34:265-71.
5. Choi BC, Pak AWP. Bias, overview. In: Armitage P, Colton T, editors. Encyclopedia of biostatistics. Vol 1. Hoboken (NJ): John Wiley & Sons, Inc; 1998. p. 331-8.
6. Delgado-Rodriguez M, Llorca J. Bias. J Epidemiol Community Health 2004;58:635-41.
7. Foddy W. Constructing questions for interviews and questionnaires: theory and practice in social research. Cambridge (United Kingdom): Cambridge University Press; 1993.
8. Lewis-Faning E. Report on an enquiry into family limitation and its influence on human fertility during the past fifty years. Vol 1. London: Her Majesty's Stationery Office, Royal Commission on Population; 1949.
9. Kelsey JL, Thompson WD, Evans AS. Methods in observational epidemiology. New York: Oxford University Press; 1986.
10. Marquis KH. An experimental study of the effects of reinforcement, question length, and reinterviews on reporting selected chronic conditions in household interviews. Ann Arbor (MI): University of Michigan, Institute for Social Research, Survey Research Centre; 1969.
11. Gowers EA. The complete plain words. London: Penguin (Pelican) Books; 1954.
12. Day RA. How to write and publish a scientific paper. 5th ed. Phoenix (AZ): Oryx Press; 1998.
13. David ST. Public opinion concerning tuberculosis. Tubercle (Journal of the British Tuberculosis Association) 1952;33:78-90.
14. Schlesselman JJ. Case-control studies: design, conduct, analysis. New York: Oxford University Press; 1982.
15. Fennema JS, van Ameijden EJ, Coutinho RA, van den Hoek JA. Validity of self-reported sexually transmitted disease in a cohort of drug-using prostitutes in Amsterdam: trends from 1986 to 1992. Int J Epidemiol 1995;24:1034-41.
16. May DS, Kittner SJ. Use of medicare claims data to estimate national trends in stroke incidence, 1985-1991. Stroke 1994;25:2343-7.
17. Dillman DA. Mail and telephone surveys: the total design method. New York: John Wiley & Sons, Inc; 1978.
18. Hunt DM, Magruder S, Bolon DS. Questionnaire format bias: when are juxtaposed scales appropriate: a call for further research. Psychol Reports 1995;77:931-41.
19. Hartge P, Cahill J. Field methods in epidemiology. In: Rothman KJ, Greenland S, editors. Modern epidemiology. 2nd ed. Philadelphia (PA): Lippincott-Raven Publishers; 1998. p. 163-80.
20. Rothman KJ, Greenland S. Types of epidemiologic studies. In: Rothman KJ, Greenland S, eds. Modern epidemiology. 2nd ed. Philadelphia (PA): Lippincott-Raven Publishers; 1998. p. 67-78.
21. Sudman S, Bradburn NM. Asking questions: a practical guide to questionnaire design. San Francisco (CA): Jossey-Bass Publishers; 1982.
22. Choi BC, Connolly JG, Zhou RH. Application of urinary mutagen testing to detect workplace hazardous exposure and bladder cancer. Mutat Res 1995;341:207-16.
23. Howe GR, Burch JD, Chiarelli AM, Risch HA, Choi BC. An exploratory case-control study of brain

tumours in children. Cancer Res 1989;49:4349-52.
24. Jin R, Choi BC, Chan BT, McRae L, Li F, Cicutto L, et al. Physician asthma management practices in Canada. Can Respir J 2000;7:456-65.
25. Ayidiya SA, McClendon MJ. Response effects in mail surveys. Public Opinion Quarterly 1990;54:229-47.
26. Salant P, Dillman DA. How to conduct your own survey. New York: John Wiley & Sons, Inc; 1994.
27. van Hemert DA, Baerveldt C, Vermande M. Assessing cross-cultural item bias in questionnaires: acculturation and the measurement of social support and family cohesion for adolescents. J Cross-Cultural Psychol 2001;32:381-96.
28. Salmon CT, Nichols JS. The next-birthday method of respondent selection. Public Opinion Quarterly 1983;47:270-6.
29. Khoury MJ, Flanders WD. Bias in using family history as a risk factor in case-control studies of disease. Epidemiology 1995;6:511-9.

Table

Table. Sources of Questionnaire Bias

Source	Bias
1. Question Design	
Problems with wording	ambiguous question complex question double-barrelled question (two questions in one) short question technical jargon uncommon word vague word
Missing or inadequate data for intended purpose	belief vs behavior (hypothetical question, personalized question) starting time data degradation insensitive measure
Faulty scale	forced choice (insufficient category) missing interval overlapping interval scale format
Leading questions	framing leading question mind-set
Intrusiveness	reporting (self-report response) sensitive question
Inconsistency	case definition change of scale change of wording diagnostic vogue
2. Questionnaire Design	
Formatting problem	horizontal response format juxtaposed scale (questionnaire format) left alignment and right alignment
Questionnaire too long	no-saying (nay-saying) and yes-saying (yea-saying) open question (open-ended question) response fatigue
Flawed questionnaire structure	skipping question
3. Administration of Questionnaire	
Interviewer not objective	interviewer nonblinding
Respondent's subconscious reaction	end aversion (central tendency) positive satisfaction (positive skew)
Respondent's conscious reaction	faking bad (hello-goodbye effect) faking good (social desirability, obsequiousness) unacceptable disease unacceptable exposure unacceptability underlying cause (rumination)
Respondent's learning	learning hypothesis guessing
Respondent's inaccurate recall	primacy and recency proxy respondent (surrogate data) recall telescope
Cultural differences	cultural