THE EFFECTS OF QUESTION WORDING AND FRAMING

When asked whether their country's nuclear weapons made them feel "safe," 40 percent of the respondents in a 1986 British Gallup poll said yes, and 50 percent said no (the remaining 10 percent had no definite opinion). But when another pollster used the word "safer" rather than "safe," these percentages reversed: 50 percent of the respondents said that nuclear weapons made them feel safer, and 36 percent said that nuclear weapons made them feel less safe (Lelyveld, 1986, October 5).

These results illustrate the importance of subtle changes in wording. Sometimes changes of only a word or two—either in the response alternatives or the question itself—can profoundly affect how people answer a question (Borrelli, Lockerbie, & Niemi, 1987). Consider, for example, the mysterious case of the Missing Middle Category.

A TRICKY ASSIGNMENT

The year is 1969. You, a Machiavellian pollster par excellence, have infiltrated the Gallup polling organization. Secretly, your mission is to conduct a survey showing that the American public wants a speedy troop withdrawal from the Vietnam War. How will you attain your objective?

In June of that year, Gallup told respondents that President Richard Nixon had "ordered the withdrawal of 25,000 troops from Vietnam in the next three months." Respondents were then asked whether "troops should be withdrawn at a faster or a slower rate" ("same as now" was not explicitly included in the response alternatives, but if respondents spontaneously declared that the present rate was about right, interviewers accepted this answer). Nearly half of the respondents (42 percent) said "faster," 16 percent said "slower," and 29 percent volunteered "same as now" (Converse & Schuman, 1970, June).

You release these results to the press, and the next day newspapers across the country run the headline: AMERICANS FAVOR SPEEDIER TROOP WITHDRAWAL. Sitting back in your chair, you toast the successful completion of your mission. Or so you think.

Soon after the Gallup poll, Harris conducted a similar survey using the following question: "In general, do you feel the pace at which the President is withdrawing troops is too fast, too slow, or about right?" In response to this question, approximately half of the respondents (49 percent) felt that the troop withdrawal was proceeding at the right rate. Only 29 percent said they preferred a speedier withdrawal, and 6 percent said the withdrawal was too fast (Converse & Schuman, 1970, June).

Picking up a newspaper, you now read the following headline: AMERICANS FAVOR STATUS QUO; EARLIER POLL IN ERROR. "Earlier poll in ERROR?" you say to yourself. Surely the poll conducted by Gallup is not in error. But if not, then which poll is correct?

Instead of asking which poll is correct—cases can be made for both—a more useful question is to ask what the results from each one mean. Gallup used a "forced choice" question that did not include an explicit middle category. As a consequence, people with very slight leanings came out for either faster or slower troop withdrawals. Questions without middle categories are often useful for assessing general inclinations, but to be entirely accurate, the Gallup headline should have read: WHEN FORCED TO CHOOSE BETWEEN FASTER AND SLOWER TROOP WITHDRAWALS, AMERICANS FAVOR A SPEEDIER WITHDRAWAL. Results from the Gallup poll do not indicate the percentage of Americans who were actually dissatisfied with the rate of troop withdrawals.

In contrast, the Harris survey explicitly offered respondents a middle category. The effect of adding a middle category is much the same as adding a "no opinion/don't know" filter; usually, 10 to 40 percent of the respondents shift their answers to this category, but the marginal proportions stay roughly the same (in this case, the proportion saying "too fast" or "too slow"). The disadvantage of including a middle category is that it is a "safe" answer, and in some instances, may therefore act as a magnet for pseudo-opinions.

OPEN PRIORITIES

Just as the inclusion of a middle category encourages people to choose that category, the inclusion of other response alternatives often increases the number of respondents who choose those alternatives. This pattern was clearly documented in a study published by Howard Schuman and Jacqueline Scott (1987).

Schuman and Scott asked respondents in a national survey to name "the most important problem facing this country today." As the first column of Table 6.1 shows, when respondents were asked this question in an "open" format (in which they answered in their own words), only 2 percent spontaneously named the energy shortage, the quality of public...
A QUESTION OF PRIORITIES

<table>
<thead>
<tr>
<th>Problems</th>
<th>Open Question</th>
<th>Closed Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of public schools</td>
<td>1</td>
<td>32</td>
</tr>
<tr>
<td>Pollution</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>Legalized abortion</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Energy shortage</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>All other responses</td>
<td>98</td>
<td>40</td>
</tr>
</tbody>
</table>

Note: These data come from a survey by Howard Schuman and Jacqueline Scott (1987) in which 178 respondents were asked the open question “What do you think is the most important problem facing this country today?” and 171 respondents were asked the closed question “Which of the following do you think is the most important problem facing this country today—the energy shortage, the quality of public schools, legalized abortion, pollution—or, if you prefer, you may name a different problem as most important.”

Even answers to the most mundane questions can be affected by the number and type of response alternatives. For instance, one study found that reported television usage varied as a function of the response categories that were provided in the question (Schwarz, Hippler, Deutsch, & Strack, 1985). In this experiment, respondents used one of two scales to indicate how much television they watched per day. Both scales were divided into six half-hour increments, but the first scale began with “up to a half hour” and ended with “more than two and a half hours,” whereas the second scale began with “up to two hours” and ended with “more than four and a half hours.” Of the respondents who were given the first scale, 84 percent reported watching television two and a half hours or less per day. In contrast, of the respondents who used the second scale, only 63 percent gave this answer.

According to Norbert Schwarz and his colleagues (1985, pp. 388–389): “Response scales are not simply ‘measurement devices’ that respondents use to report their behaviors. Rather . . . respondents may use the range of behaviors described in the response alternatives as a frame of reference in estimating and evaluating their own behavior.” Structured response alternatives—like questions—are never perfectly neutral, and they often convey an implicit range of acceptable answers (see Table 6.2 for some other questions that implicitly suggest appropriate responses). As the next section shows, response biases may also arise because certain answers are more socially desirable than others.

SOCIAL DESIRABILITY

In 1982, Richard Wirthlin, who used to be in charge of opinion polling for President Reagan, found that 58 percent of the respondents in a national survey agreed with this statement: “A freeze in nuclear weapons should be opposed because it would do nothing to reduce the danger of the thousands of nuclear warheads already in place and would leave the Soviet Union in a position of nuclear superiority.” But then, several minutes later in the very same polling interview, 56 percent agreed with the statement: “A freeze in nuclear weapons should be favored because it would begin a much-needed process to stop everyone
in the world from building nuclear weapons now and reduce the possibility of nuclear war in the future.” In fact, 27 percent of the respondents actually endorsed both statements, in what Wirthlin called “the most singular inconsistency on any question we’ve ever asked” (Clymer, 1982, May 6).

Why the inconsistency? In the absence of a firm opinion concerning a nuclear freeze, respondents may have tried to give the most “socially desirable” answer. The first question equated support for a nuclear freeze with Russian nuclear superiority, and the second question associated it with world peace. Very few Americans are in favor of Russian nuclear superiority or opposed to world peace.

When people do not have deep convictions about an issue, they often respond to “catch phrases” that point them in a socially desirable direction. For example, U.S. News & World Report published the results of a poll in which 58 percent of the respondents favored aid to Nicaraguan rebels “to prevent Communist influence from spreading,” but only 24 percent favored assistance to “the people trying to overthrow the government of Nicaragua” (Budiansky, 1988, July 11). To most Americans, “preventing Communism” is a commendable goal and “overthrowing governments” is not.

Schuman and Presser (1981) found similar results from a pair of opinion polls conducted in 1978. In the first poll, respondents were asked: “If a situation like Vietnam were to develop in another part of the world, do you think the United States should or should not send troops?” In response to this question, only 17 percent of those surveyed felt that the United States should send troops. When the threat of Communism was explicitly mentioned, however, support for sending troops doubled; 37 percent of the respondents supported sending troops when asked: “If a situation like Vietnam were to develop in another part of the world, do you think the United States should or should not send troops to stop a Communist takeover?”

Back in 1940, Hadley Cantril found essentially the same pattern of results in two polls on the U.S. involvement in World War II. Only 13 percent of Cantril’s respondents agreed that “the U.S. should do more than it is now doing to help England and France,” but this figure climbed to 22 percent when a second set of respondents were asked: “Do you think the U.S. should do more than it is now doing to help England and France in their fight against Hitler?” At the time, helping in the “fight against Hitler” was seen as more socially desirable than simply helping England and France.

TO ALLOW OR FORBID?

In another early pair of experimental polls, Rugg (1941) asked respondents one of two questions:

- Do you think that the United States should allow public speeches against democracy?
- Do you think that the United States should forbid public speeches against democracy?

Even though these questions seem as though they are asking the same thing, Rugg found that they led to very different responses. When people were asked whether the United States should allow public speeches against democracy, 62 percent said no. On the other hand, when people were asked whether the United States should forbid speeches against democracy, only 46 percent of all respondents said yes (the logical equivalent of not allowing such speeches). Of those respondents who expressed an opinion, about 20 percent more said they would “not allow” antidemocratic speeches than would “forbid” them!

Schuman and Presser (1981) repeated Rugg’s experiment three times in the 1970s, and on each occasion they found results similar to those observed by Rugg (1941). Thus, in a total of four survey administrations spanning more than 30 years, the “forbid” version of the question led to substantially more support for free speech than did the “allow” version (presumably because of the loss in freedom associated with “forbidding” something). The only significant difference between Schuman and Presser’s results and Rugg’s results was that Americans in the 1970s were relatively more tolerant of free speech. In both forms of the question, roughly 30 percent more people supported free speech in the 1970s than in the 1940s.

The allow-forbid distinction has also been observed with respect to a number of other topics. For example, Hans-J. Hippler and Norbert Schwarz (1986) found differences between “not allowing” and “forbidding” peep shows, X-rated films, and the use of salt to melt snow on the highways. Other polls have found that only 29 percent of those surveyed were in favor of a constitutional amendment “prohibiting abortions,” but 50 percent were in favor of a constitutional amendment “protecting the life of the unborn” (Budiansky, 1988, July 11).

FRAMING

As the foregoing results suggest, people respond differently to losses (e.g., prohibitions) than to gains (e.g., allowances). The importance of this difference was not widely appreciated, however, until psychologists Amos Tversky and Daniel Kahneman popularized the notion of “framing.” According to Tversky and Kahneman (1981, p. 453), a decision frame is “the decision maker’s conception of the acts, outcomes, and contingencies associated with a particular choice.” Tversky and Kahneman (1981) proposed that decision frames are partly controlled by the formulation of the problem, and partly controlled by the norms, habits,
and characteristics of the decision maker. What Tversky and Kahneman focused on in their research, though, was the effect of formulating problems in different ways.

To see how powerful these effects can be, consider the following pair of monetary bets (taken from an experiment by Tversky and Kahneman). In Decision 1, you must choose between Alternative A and Alternative B:

**Alternative A:** A sure gain of $240  
**Alternative B:** A 25 percent chance to gain $1000, and a 75 percent chance to gain nothing

What would you choose? (See Item #25 of the Reader Survey for your answer.) After noting your preference, consider Decision 2:

**Alternative C:** A sure loss of $750  
**Alternative D:** A 75 percent chance to lose $1000, and a 25 percent chance to lose nothing

What would you choose this time? (See Item #26 of the Reader Survey.)

Tversky and Kahneman (1981) found that Alternative A was preferred in Decision 1. Eighty-four percent of their subjects chose Alternative A over Alternative B. As mentioned in Chapter 5, people tend to be risk averse when gains are at stake (they prefer “a bird in the hand” to “two in the bush”).

In Decision 2, however, people usually prefer to gamble. Eighty-seven percent of Tversky and Kahneman’s subjects chose Alternative D. When losses are at stake, people tend to be risk seeking. Indeed, this pattern of risk seeking and risk aversion was so common that 73 percent of Tversky and Kahneman’s respondents chose Alternatives A and D, and only 3 percent chose Alternatives B and C.

The interesting thing about this problem is that choosing B and C turns out to be better than choosing A and D. You can see this by simply adding the chosen alternatives together (see Figure 6.1). By choosing A and D, a sure gain of $240 is added to a 75 percent chance of losing $1000 and a 25 percent chance of losing nothing. This becomes a 75 percent chance of losing $760 and a 25 percent chance of gaining $240. By choosing B and C, a sure loss of $750 is subtracted from a 25 percent chance of gaining $1000 and a 75 percent chance of gaining nothing.

This leaves a 25 percent chance of gaining $250 and a 75 percent chance of losing $750.

In other words, the combined choices are:

**A and D:** A 75 percent chance of losing $760 and a 25 percent chance of gaining $240  
**B and C:** A 75 percent chance of losing $750 and a 25 percent chance of gaining $250

If you choose A and D, you stand to lose more or gain less than if you choose B and C. Regardless of whether you lose or gain, choosing B and C is better than choosing A and D.
A BITTER PILL TO SWALLOW

The effects of framing are particularly important in medical decision making. The earliest and most famous example of a framing problem related to medicine is Tversky and Kahneman’s (1981) “Asian disease” question:

Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

- If Program A is adopted, 200 people will be saved.
- If Program B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved.

Which of the two programs would you favor?

Tversky and Kahneman found that 72 percent of the people given this frame were risk averse; they preferred to save 200 lives for sure rather than gamble on saving a larger number of lives.

By changing the frame, however, Tversky and Kahneman were able to elicit a very different pattern of preferences. They gave a second set of respondents the same problem, except that the consequences were now described as follows:

- If Program C is adopted 400 people will die.
- If Program D is adopted there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die.

Even though Programs C and D in this frame are numerically equivalent to Programs A and B in the former frame, 78 percent of the respondents now became risk seeking; they preferred to gamble rather than accept a sure loss of 400 lives.

Similar results have been found in other studies of medical decision making (cf. Wilson, Kaplan, & Schneiderman, 1987). For example, an experiment published in the New England Journal of Medicine found that framing influenced physicians’ decisions about how to treat lung cancer (McNeil, Pauker, Sox, & Tversky, 1982). The participants in this study were 424 radiologists, 491 graduate students who had completed coursework in statistics and decision theory, and 238 ambulatory patients with different chronic problems. All 1153 participants were presented with summary information on two forms of treatment for lung cancer (surgery and radiation therapy). Once they finished reading this summary information, they were asked which of the two treatments they would prefer.

In approximately half the cases, the summary information was framed in terms of the cumulative probability of survival after a particular amount of time (e.g., a 68 percent chance of living for more than one year). In the other cases, the summary information was cast in terms of mortality (e.g., a 32 percent chance of dying by the end of one year). Because the danger of dying during or immediately after surgery is the major disadvantage of treating lung cancer surgically, the experimenters hypothesized that surgery would be selected more frequently when the summary information was framed in terms of the probability of living than when it was framed in terms of the probability of dying.

And in fact, this is exactly what they found. Surgery was preferred to radiation therapy 75 percent of the time in the survival frame, but only 58 percent of the time in the mortality frame. This pattern was found among physicians and graduate students as well as patients. Thus, even in the case of life-and-death decisions made by highly trained specialists, framing can significantly influence the choices that are made.

DEDUCTIVE LOGIC

In a thought-provoking essay on “economic reasoning,” Harvard economist Thomas Schelling (1981) presented an interesting case in which “reframing” a choice can lead to surprising conclusions. Schelling began his discussion by noting that the U.S. income tax laws at that time allowed families to deduct $1000 for each child (i.e., to subtract $1000 per child in the process of converting gross income to taxable income). The amount of this deduction was fixed; it did not depend on the total income a family earned.

Schelling raised the question, though, whether it might not be better to allow wealthy families to take larger tax deductions for each child. After all, wealthy families spend more on their children than do poor families, and the amount of money it costs for wealthy families to have children is much higher than the amount it costs poor families.

Most people would object to such a policy. “There is no reason,” goes the argument, “to extend further privileges to the rich—and certainly not at public expense.”

But consider (says Schelling) what would happen if the tax laws were reformulated. Suppose that instead of deducting the cost of children from a tax schedule based on childless families, a “childless premium” were added to a tax schedule that assumed the typical family had two or three children. In other words, suppose a childless family were charged “extra” taxes (instead of giving families with children a tax break). Should a poor family without children pay just as high a premium as a wealthy family without children?

In this case, it is tempting to argue that rich families should pay more than poor families. Because rich families spend more on their children than do poor families, it stands to reason that a rich family without children can afford to pay higher premiums than a poor family without children. The problem is, however, that this argument directly contradicts the earlier argument against allowing wealthy people to take larger...
deductions for their children. In the words of Schelling (1981, p. 55): “Since the same income tax can be formulated either as a base schedule for the childless couple with an adjustment for children, or as a base schedule for the family with children plus an adjustment for childlessness, it should not make any difference which way we do it.”

But of course it does make a difference, and as long as there are deductions for children, people will argue—with good reason—that these deductions should be fixed in size.

**PSYCHOLOGICAL ACCOUNTING**

Not only do decision makers frame choices—they frame the outcomes of their choices. Tversky and Kahneman (1981) dubbed this process “psychological accounting.” Psychological accounting has to do with whether an outcome is framed in terms of the direct consequences of an act (what Tversky and Kahneman called a “minimal account”), or whether an outcome is evaluated with respect to a previous balance (a more “inclusive” account).

The following pair of problems (taken from an experiment by Tversky & Kahneman, 1981) illustrates psychological accounting:

**Problem 1.** Imagine that you have decided to see a play where admission is $10 per ticket. As you enter the theater you discover that you have lost a $10 bill. Would you still pay $10 for a ticket for the play?

Of the 183 people who Tversky and Kahneman asked, 88 percent said they would still buy a ticket to see the play. Most respondents did not link the loss of $10 with the purchase of a ticket; instead, they charged the loss to a separate account.

Now consider the second problem:

**Problem 2.** Imagine that you have decided to see a play and paid the admission price of $10 per ticket. As you enter the theater you discover that you have lost the ticket. The seat was not marked and the ticket cannot be recovered. Would you pay $10 for another ticket?

Of the 200 people who Tversky and Kahneman asked, only 46 percent said they would purchase another ticket. Even though the loss of a ticket is financially equivalent to the loss of $10, most respondents apparently added the cost of a second ticket to the cost of the original ticket—and viewed $20 as excessive.

Here is another example of psychological accounting from Tversky and Kahneman:

Imagine that you are about to purchase a jacket for $125, and a calculator for $15. The calculator salesman informs you that the calculator you wish to buy is on sale for $10 at the other branch of the store, located 20 minutes’ drive away. Would you make the trip to the other store?

In response to this question, 68 percent of Tversky and Kahneman’s subjects said they would be willing to make the 20-minute drive to save $5.

But what if the calculator cost $125 and the jacket cost $15? Tversky and Kahneman presented this problem to 88 people:

Imagine that you are about to purchase a jacket for $15, and a calculator for $125. The calculator salesman informs you that the calculator you wish to buy is on sale for $120 at the other branch of the store, located 20 minutes’ drive away. Would you make the trip to the other store?

Now only 29 percent of Tversky and Kahneman’s subjects said they would make the drive. When the calculator was valued at $125, more than two-thirds of the respondents said they would not drive 20 minutes to save $5, even though this price reduction is no different than the savings mentioned in the earlier problem.

Why? According to Tversky and Kahneman (1981), respondents frame the first problem in terms of a minimal account in which the 20-minute drive is weighed directly against $5 in savings. In the second problem, though, respondents include the entire purchase price in the account, and as a consequence, $5 seems rather negligible. More formally, George Quattrone and Amos Tversky (1988) have proposed that people make choices according to a “ratio-difference principle” in which the impact of any fixed positive difference between two amounts increases with their ratio. Thus, the difference between a price of $20 and $15 yields a ratio of 1.33, which is larger and more influential than the rather negligible ratio of 1.04 between $125 and $120. Richard Thaler (1985) has found similar framing effects in his research on the role of psychological accounting in consumer choices.

These findings, and the findings mentioned earlier in this chapter, do not mean that a clever salesperson or pollster can elicit any response from the public. Rather, they indicate only that question wording and framing often make a substantial difference, and that it pays to be aware of their effects. Although these observations may seem self-evident, they have been—as the next three chapters show—a long time in coming.

**CONCLUSION**

When people learn the results of a public opinion poll, they rarely think about whether the questions were filtered, whether a middle category was included among the response alternatives, whether the problem was framed in terms of gains or losses, and so forth. The natural tendency most people have is to assume that the survey respondents simply answered in keeping with their beliefs. A great deal of research suggests,
however, that framing and question wording can significantly affect how people respond.

Before relying on results from survey research and other studies of judgment and decision making, it is important to consider how people's answers would have changed as a function of factors such as:

- The order in which the questions were presented
- The context in which the questions appeared
- Whether the question format was open or closed
- Whether the questions were filtered
- Whether the questions contained catch phrases
- The range of suggested response alternatives
- The order in which response alternatives were presented
- Whether middle categories were provided
- Whether problems were framed in terms of gains or losses

If you suspect that changes in these factors would have led to changes in how people answered, you should qualify your interpretation of the results until variations in wording can be tested. As Paul Slovic, Dale Griffin, and Amos Tversky (1990, p. 25) have argued: “If [the results from multiple procedures] are consistent, we may have some basis for trusting the judgment; if they are not, further analysis is required.” Because judgments are so easily influenced by question wording and framing, the safest course of action is to elicit them in a variety of ways and compare the results.