

White Paper: Garbage In, Garbage Out

Survey-Question Design

February 2010

David Sleeth-Keppler, PhD

Issued by:

VALS™
Strategic Business Insights
333 Ravenswood Avenue
Menlo Park, California 94025
www.strategicbusinessinsights.com/vals

CONTENTS

THE PURPOSE OF SURVEY RESEARCH.....	1
PROBLEMS WITH USING LONG QUESTIONS NEEDLESSLY.....	2
WHEN TO FORCE CHOICE OR USE “SELECT ALL THAT APPLY”	3
HOW MUCH IS “A LOT?”	5
DOUBLE BARRELS	7
SHOULD “NEUTRAL” BE AN OPTION?	9
JUST ASK?.....	10
PRETESTING.....	12
SUMMARY	13

THE PURPOSE OF SURVEY RESEARCH

Companies routinely conduct market surveys to gain insights into what the population as a whole does or what their customers in particular do, how people prioritize behaviors, and what attitudes people hold about a host of product-related issues, such as the importance of a particular brand or the brand's attributes. Companies use survey data for a variety of strategic purposes. They use survey data most frequently when making marketing decisions about new-product development, customer identification and targeting, communications and advertising, and ways to build customer satisfaction, trust, loyalty, and retention. Because each of these business decisions is critical to the financial health and well-being of an enterprise, data reliability is critical.

Although many companies rely heavily on data for decision making, few companies appear to pay special attention to the way they construct questions on their surveys. Sometimes the way in which companies ask survey questions—researchers often refer to the questions as “items”—can undermine the very purpose of the entire survey. Consequently, many companies make their most important, data-driven business decisions on the basis of distorted, incomplete, or misleading information.

As a general rule, the quality of data generated by survey research depends on the quality of the survey questions. After a survey has taken place, no amount of analytic sophistication, presentation creativity, or convincing delivery can compensate for suboptimal “ingredients”—the data that come from poorly designed survey questions. Unfortunately, at the time when companies determine actions on the basis of the data, they no longer take into consideration the quality of the initial questions. The negative effects of poor survey-question design on the insights that come from marketing surveys go unnoticed.

The main aim of this paper is to sensitize individuals developing survey questions or working with survey data to some—not all—of the potential pitfalls involving survey questions and to offer alternatives to those pitfalls. Well-crafted survey items increase the potential to derive the most accurate and useful insights in future research and strategic efforts.

PROBLEMS WITH USING LONG QUESTIONS NEEDLESSLY

Many individuals who write survey questions appear to neglect the need for simplicity. Generally, the simpler a question's wording, the better. Some exceptions apply, such as when an item is measuring a response to a particular phrase or issue that authors cannot reword because to do so would change the meaning. Often however, items are not as simple as they could be.

For example, "I like to try new things" is a better item than "I love to buy the latest cell phone to see what new features and apps it has." "I like to have a lot of excitement in my life" is preferable to "I always try to have a happy life full of excitement. Similarly, "I like to learn about science" bests "I like to read and watch TV shows about science." Writing survey items is rarely about linguistic creativity. Survey questionnaires are not poems.

To the extent that market researchers fail to use simple questions, the burden of reading and responding to items can weigh heavily on a respondent. A respondent who becomes overburdened by a questionnaire will, in an effort to finish quickly, often choose responses with very little effort, possibly at random, and may ultimately withdraw from responding altogether. Therefore, the overall goal in designing survey questions, particularly if a survey questionnaire includes many items, is to keep the burden of responding to the questionnaire as light as possible.

WHEN TO FORCE CHOICE OR USE “SELECT ALL THAT APPLY”

Which question format should one use when the aim is to quantify customer behaviors relevant to a product or service? Similar questions may seem to measure the same quantity, but depending on how the survey asks the question, the differences in the reported quantities can be important. A popular way to keep respondent burden light or to shorten the length of a survey when asking about many items in a single pass, is to use the “Select All That Apply” question format. Take the following example:

Version 1:

Do you currently have an account at any of the following social-networking sites with a registered user name and password? (SELECT ALL THAT APPLY)

1. Facebook
2. AIM
3. Friendster
4. LinkedIn
5. MySpace
6. Twitter

An alternative way to ask a question about many items, such as social-networking accounts, is to break each item out separately and create “yes/no” or forced-choice response categories for each.

Version 2:

Do you currently have a Facebook account with a registered user name and password?

1. Yes
2. No

Do you currently have an AIM account with a registered user name and password?

1. Yes
2. No

.

.

.

Do you currently have a Twitter account with a registered user name and password?

1. Yes
2. No

On the surface, one might expect the two versions to yield similar response frequencies. In fact, this expectation is incorrect. A recent convenience sample of individuals responding to a Strategic Business Insights (SBI) study found that 37% of respondents to the question in version 1 (“Select all that apply”) indicated that they owned a Twitter account. Only 26% of the same respondents indicated that they owned a Twitter account when SBI asked the question using version 2 (“yes/no”).

What might explain the difference? Perhaps SBI’s respondents answered the Select All That Apply question more mindlessly than they answered the forced-choice question, quickly scanning the list and ticking off answers without reflecting whether they participated in the service. However, when a question breaks out each item in a Select All list for separate response, respondents might respond more mindfully, eliminating the additional responses in the Select All format. In this example, a 10% difference in the penetration of a service among the respondents—evident simply as a function of question format—is not insignificant by any standard.

In some published cases of research, forced choice has yielded higher incidences of responding to an item than has Select All (but in separate samples). Unfortunately, in many cases, no way exists to validate which number is correct without comparing each sample result with other independent and reliable data sources. The bottom line is that companies ought to consider carefully the various pros and cons associated with question formats.

HOW MUCH IS “A LOT?”

Many survey-research questionnaires ask customers to quantify—that is, provide a numeric value for—many of their daily activities: “How many times a day do you check the news?” “How many meals do you eat per day?” “How many DVDs do you own?” Typically, customers answer these types of questions by one of two methods. Some companies choose open-ended question formats, in which respondents fill in a number for each question in a blank-form field. Other companies provide response categories for each question; data processors can handle these categories easily in preparing the data for analysis.

When a company decides to provide response categories, how should it choose the categories? In other words, how much is “a lot” or “a little” of a behavior? The following example illustrates why this question requires careful consideration.

Version 1:

“How many hours of television do you watch in a typical day?”

0 hours 1 hour 2 hours 3 hours More than 3 hours

Version 2:

“How many hours of television do you watch in a typical day?”

0 hours 1–2 hours 3–4 hours 5–6 hours More than 6 hours

A survey respondent who watches 4 hours of television per day and responds to question version 1 will have the impression that his or her television-viewing habit places the individual at the top of the distribution of possible viewing lengths. When answering the same question using the response scale in version 2, the same respondent will have the impression of being an average television viewer.

Research has shown that response categories in survey questions can change the way people view themselves. A respondent’s self-view can affect responses to subsequent survey questions. For example, imagine, after asking the television question, researchers ask a respondent:

“How sociable are you?”

Not at all sociable Slightly sociable Somewhat sociable Very sociable

A respondent who watches 4 hours of television a day and responds to version 1 of the question might not think himself or herself particularly sociable, because in this circumstance, the respondent’s self-view might be that he or she spends too much time watching television and not enough time with friends. By contrast, a survey taker responding to version 2 of the question might think he or she is fairly sociable, given that his or her average viewing habit appears to leave more time for friends. Note that the actual viewing behavior across the two question versions is the same. However, the scaling of a question can fundamentally affect not only the measurement of the quantity, but also how survey participants respond to seemingly unrelated questions. One way to decide about the range of values to provide in a scale is to use nationally

representative frequencies of the behavior of interest in order to calibrate the scale (select plausible high and low values from the distribution).

The way authors word the question itself can introduce yet another distortion of a measured quantity. A famous example comes from the research on eyewitness testimony, conducted in the 1970s by Elizabeth Loftus, a cognitive psychologist, and her colleagues. In one study, Loftus varied a single word in a question about a car accident that participants saw on a videotape:

Version 1:

“Approximately how fast was the green car going when it **hit** the other car?”

Version 2:

“Approximately how fast was the green car going when it **smashed** into the other car?”

Study participants who heard version 1 of the question (“**hit** the other car”) provided lower estimates of the speed of the green car than did a separate group of participants who heard version 2 of the question (“**smashed** into the other car”). Subtle differences in wording can elicit powerfully different responses from individuals.

For marketing surveys, the extremity of the chosen words to measure a reaction can also change the numeric response that a respondent provides on a response scale:

Version 1:

“How satisfied are you with ACME toothpaste?”

1	2	3	4
Not at all satisfied	Slightly satisfied	Satisfied	Very satisfied

Version 2:

“How satisfied are you with ACME toothpaste?”

1	2	3	4
Not at all satisfied	Slightly satisfied	Somewhat satisfied	Satisfied

A respondent who feels “satisfied” with ACME toothpaste would circle option 3 on version 1 of the question but option 4 on version 2 of the question. By providing a restricted range of options in comparison with options in version 1, version 2 would certainly make the numbers look more favorable for ACME toothpaste but would overestimate satisfaction among its consumers. If the goal of the research is to diagnose a brand, product, or service accurately, companies need to ensure that all meaningful response categories to distinguish consumer reactions are in the questions.

DOUBLE BARRELS

The tendency to “double barrel” a question—that is, to ask respondents about two (or more) separate issues in a single item rather than just one—can significantly inhibit the ability to interpret responses. Obvious double barrels can occur quite frequently in questionnaires, such as the one below:

“I don’t think editorials or op-eds are useful sources of information, even if written by respectable authors or published by respectable newspapers.”

In addition to being quite lengthy, the item really asks respondents to agree or disagree with a question that contains four separate items that may differentially influence respondents: “editorials,” “op-eds,” “respectable authors,” and “respectable newspapers.” If a respondent agrees with the question, he or she possibly agreed because confidence in information from “opinion pieces” is low, because confidence in editorials (but not op-eds) is low, because confidence in journalists (but not newspapers) is low, or because of any combination of agreement or disagreement possible with the four items. Thus, the main reason to avoid double-, triple-, or quadruple-barrel questions is to avoid making the response impossible to interpret.

Some double barrels are much more subtle than the one in the example above:

“I enjoy watching television as a valuable source of learning”

1	2	3	4
Disagree	Somewhat Disagree	Somewhat Agree	Agree

The item is a double barrel because the item asks about both television viewing and television’s value as a learning source. It is impossible to tell—if a respondent disagrees with the item—whether the disagreement is because the respondent does not enjoy watching television or because the respondent does not think television is particularly useful as a source of learning.

The below item lacks the double barrel and is interpretable:

“Television is a valuable source of learning”

1	2	3	4
Disagree	Somewhat Disagree	Somewhat Agree	Agree

In recent years, survey methodologists have argued that the item above could have even clearer wording. Specifically, an agree-disagree scale makes it difficult to interpret exactly how a respondent is valuing or devaluing television as a source of learning. For example, if a respondent disagrees with the above statement, he or she could think television is “not at all” valuable as a source of learning or still “somewhat” valuable.

To avoid these types of ambiguities, methodologists recommend writing survey questions that scale according to the construct of interest—in this case, “value”:

“How valuable is television as a source of learning”?

1	2	3	4	5
Not at all valuable	Slightly valuable	Somewhat valuable	Very valuable	Extremely valuable

Perhaps the most difficult-to-detect double barrels arise from the silent and often untested assumptions that many companies have about their customers. Even though the item “Television is a valuable source of learning” is not explicitly a double barrel, the writer assumes that all the survey respondents care about learning. A respondent whose underlying psychology reflects a disinterest in learning altogether might disagree with the item because learning is simply not part of the self-concept, regardless of the source of learning.

Similar assumptions about consumers can color entire sections of a questionnaire, such as the importance of brand in consumers’ decision making. A survey may ask respondents to agree or disagree with any number of items about why brands might matter. However, a company may fail to realize that for some consumers, the overall importance of brand is near the bottom of a whole list of prepurchase considerations. For example, a cosmetics company may ask 20 questions about how a brand might influence the choice to buy a certain cosmetics product, not realizing that many cosmetics consumers care much more about the color, price, and applicability of a product than about the product’s brand.

The ability to contextualize responses on survey questions within the framework of a larger list of consumer priorities often goes missing. As a result, the basis of strategic decisions can be unchecked assumptions about what consumers care about.

SHOULD “NEUTRAL” BE AN OPTION?

The following investment question includes a 5-point agree-disagree scale with a neutral midpoint where respondents have the option of registering neither agreement nor disagreement.

“Maximizing the return on my household’s investments is of great concern to me”

1	2	3	4	5
Disagree	Somewhat Disagree	Neither agree nor disagree	Somewhat agree	Agree

Regardless of the number of points on the scale, rarely are midpoints—the equivalent of “I don’t care”—desirable responses. Responses at a neutral midpoint do not inform the researcher if the respondent chose the neutral response because he or she is neutral about the attitude issue or if the respondent simply doesn’t feel qualified to answer the question.

A larger theoretical concern is whether an individual can truly be neutral about an attitude that he or she comprehends. To avoid these ambiguities, response scales should avoid a midpoint but include a “Have No Idea” or “Does Not Apply” option.

“Maximizing the return on my household’s investments is of great concern to me”

1	2	3	4	5
Disagree	Somewhat Disagree	Somewhat Agree	Agree	Have No Idea

Neutral options invite an “Acquiescent Responding” systematic error: a response pattern in which respondents select answers without giving much thought to their true, underlying feelings about an issue. Often, the goal in acquiescent responding is to finish the questionnaire quickly. If a neutral option is available, a respondent can, with little guilt, finish the survey in minutes simply by selecting the neutral option for all items.

The prevalence of attitude neutrality on an issue can sometimes be a useful piece of data in and of itself, such as when a researcher is interested in the extent to which consumers avoid taking a clear stance on an issue. Achievers—a consumer group in the U.S. VALS™ Framework—avoid registering extreme opinions on surveys; they would find a neutral option very attractive, for example. To the extent that “hiding” in survey scales is of practical or theoretical importance for understanding consumers, a neutral option would certainly be appropriate.

JUST ASK?

A classic school of thought in survey-question design is that the best approach to measure how people think, feel, or act is just to ask them. Taking a direct approach to crafting survey questions is certainly the correct approach in many circumstances. No need to beat around the bush when asking respondents to express their level of satisfaction with ACME toothpaste—the example in “How Much Is “A Lot?” As a guiding principle, “just ask” can serve as a gentle reminder to survey-question designers to use simple questions.

However, many areas of interest to companies and business researchers do not necessarily allow for a simple just-ask approach. Take the following examples of questions:

1. Have you ever stolen from your place of employment?
2. Would you pass somebody else’s work off as your own, if you had no chance of being caught?
3. If you had a choice, would you prefer to be rich or wise?
4. How much money do you make per month?
5. Which of the following options best reflect why you like the taste of ACME toothpaste?

As survey items, each of the above questions is problematic for a number of reasons. Most people are keenly aware of societal expectations for appropriate behavior and would resist admitting to transgressions against societal or normative codes of conduct, such as the ones in questions 1 and 2; this situation is especially true if the respondent is actively committing such acts. Here, the just-ask approach would simply result in a respondent’s lying.

More generally, many people have an explicit self-presentation goal that guides many responses to survey questions—even responses to much more benign questions. For example, question 3 has a built-in demand that could result in desirable responding. Even though many respondents may prefer to be rich rather than to be wise, a respondent may choose “wise” to avoid looking overly materialistic. Even a question that simply asks respondents to indicate their monthly income could result in respondents’ lying. Survey writers should not necessarily take responses to such questions at face value. The general recommendation is to build in safety checks against self-presentation on surveys and collect data using more unobtrusive measures. In less-developed countries, for example, where responses to income questions are notoriously unreliable, a survey researcher could measure ownership of durable home goods as a proxy for income. If employee ethics is a topic of concern, asking simple “lie” items such as “I have never known someone I did not like” could provide a clue to the self-presentational goals of the respondent. Agreement with the item could indicate that the respondent is systematically lying, because everyone knows at least one person whom he or she does not like.

The last question—“Which of the following options best reflect why you like the taste of ACME toothpaste?”—requires additional explanation. Most companies assume that their

customers, or survey respondents in general, come to a questionnaire with preferences already in place. If a respondent is a heavy user of a product, for example, companies assume that the respondent will be able to identify his or her particular preference for the product and have clearly defined reasons for being a heavy product user. In reality, not only are many preferences ill defined, but the reasons for existing preferences may have even poorer definition. Simply, many people have preferences without necessarily having good reasons.

Thus, one danger in survey-question design is that rather than measuring a valid justification for a product preference, the question itself creates the justification in real time, by providing plausible reasons why somebody might enjoy a product; one of the reasons may momentarily appeal to the respondent. Often, companies preselect “reasons” for product preferences that fit their marketing strategy or existing assumptions about consumers. The list of reasons may not necessarily reflect consumers’ true underlying reasons. At the extreme, the predetermined list of reasons will not reflect the fact that the respondent may have no reasons at all for a preference.

These examples serve as reminders that using survey questions or direct just-ask measurement may not be the best tool to investigate all phenomena of interest to companies. Using unobtrusive behavioral measures, diaries, open-ended responses and various other research methods might be more appropriate to answer certain questions.

PRETESTING

Finally, a few words about the importance of pretesting questionnaire items on a small sample of respondents before conducting a survey. Pretesting can often be critical to improving the overall quality of questionnaire items. In particular, pretesting is critical in terms of item comprehension and flow. Many researchers make decisions about the “interface” between the respondent and the survey instrument without testing the effectiveness of the interface. Type fonts, spacing of items, and arrangement of response options on the page all factor into the ease or difficulty with which respondents answer questionnaires: All these issues may affect the quality of the data. Pretesting allows respondents to provide feedback to the researcher about item wording and allows respondents to flag any potential ambiguities that need resolution before the questionnaire goes to the field.

SUMMARY

Gathering valid survey data that one can act upon with confidence requires a careful approach to survey-question design. When companies field surveys without paying attention to survey questions, the most important ingredients to any marketing strategy may be suboptimal. The old adage “garbage in, garbage out” is a useful reminder about the importance of survey-question design. In addition to making informed decisions about each survey questionnaire item, a fruitful approach to marketing questionnaires should also include taking a holistic view of the effort:

- How best to arrange the survey questions across the document?
- What should the various sections contain?
- Are all the relevant elements of a concept or area of interest in the survey?
- What are the company’s unmeasured assumptions about customers and how might those assumptions color survey responses?
- Finally, is a survey really the best way to address the question to begin with?

Considering the kinds of effects of and decisions about survey questions that this white paper presents can significantly increase the return on investment when conducting marketing research.

For more information about survey design or to have SBI design your next survey, contact us: telephone: +1 650 859 4600 or e-mail: info@sbi-i.com .



Understanding customers is critical to marketing.

VALS™ strengthens demographic characterizations by explaining the deeper psychological drivers of consumer behavior.

VALS is a consulting and consumer-research service. VALS consulting provides clients with tailored marketing strategies for targeting, positioning, and communications—real-world, real-time, actionable strategies.

Consumer surveys and focus groups inform our work. Current research focuses on how consumers around the world make decisions about their lives.

VALS™ Benefits

- Gain a fresh perspective by getting “into the heads” of your customers.
- Create richly textured consumer profiles or personas that include psychological motivation, attitudes, behaviors, lifestyles, and demographics.
- Understand the distinct communications styles and media preferences of your best targets.

VALS™ Applications

- Innovation: New Product Development
- Targeting: Positioning
- Communications: Channel Planning
- Geolocation
- Policy

VALS™ Research Methods

- Custom surveys
- Syndicated surveys
- Focus Groups

Ongoing training and reports strengthen persona knowledge and provide your team with a common language.

Visit our Web site at www.strategicbusinessinsights.com/vals .

About Strategic Business Insights

Strategic Business Insights (SBI) works with clients to identify and map new opportunities based on emerging technology and market insights. We combine ongoing research with consulting services to create insights that affect customers, business, and technology.

With our help, leading organizations identify coming changes faster, separate hype from reality, and create strategy with a greater awareness of what the future may bring.

SBI stands apart from our competitors in our staff's ability to integrate all pieces of the puzzle necessary to manage the future: the technology developments that will determine commercial opportunities, the forces that will define the business environment, and the shifts in consumer demand that will shape the marketplace.

SBI is the former Business Intelligence division of SRI International that has worked with clients on opportunities and change since 1958. Headquartered in Silicon Valley—with offices in Japan, the United Kingdom, and New Jersey—we have a global reach and work across a wide range of government and business sectors, including electronics, health care, energy, and financial services.



The standard in consumer financial-services research, strategy, and insight



Continuous evaluation of opportunities in technology commercialization and applications



Organizational radar to identify signals of change



Consumer insights that drive market innovation



An exploration of emerging technologies and applications



Leveraging of our research expertise and proprietary methods

- Scenario planning
- Opportunity discovery
- Strategic roadmapping
- Customer targeting

STRATEGIC BUSINESS INSIGHTS:

333 Ravenswood Avenue
Menlo Park, California 94025
Telephone: +1 650 859 4600
Fax: +1 650 859 4544

201 Washington Road
Princeton, NJ 08540-6449
Telephone: +1 609 734 2048
Fax: +1 609 734 2094

Knollys House
17 Addiscombe Road
Croydon, Surrey
CR0 6SR, England
Telephone: +44 (0) 20 8686 5555
Fax: +44 (0) 20 8760 0635

Parkside House 3F.
2, Ichibancho, Chiyoda-ku
Tokyo 102-0082, Japan
Telephone: +81 3 3222 6501
Fax: +81 3 3222 6508

Visit our Web site at
www.strategicbusinessinsights.com .