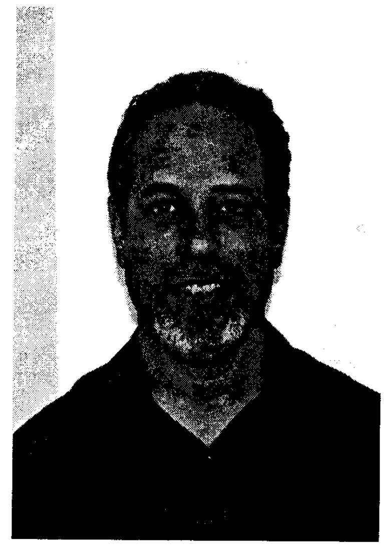
Looking for Mr. Good Data

By Scott W. Flexo



Scott Flexo received his Ph.D. in Political Science, with a concentration in Political Psychology, from the University of California, Irvine in 1992. Scott is founding partner of Flexo Hiner & Partners, Inc., a full-service marketing research company located in Long Beach, California. Scott teaches marketing and consumer behavior at University of California, San Diego, and Political Science at Fullerton College in Orange County, California.

The recent explosion in new data collection methods, including bar scanners, IVR, disk-by-mail surveys, and online research, serves business' insatiable need for more information by making it easier to collect and store vast amounts of consumer data. Unfortunately, market forces that drive business to demand more data, faster and at a lower cost, can readily clash with a practitioner's commitment to the fundamental principles of survey research design. This sets up the "practitioners dilemma." Should one compromise on research quality issues and deliver data faster and cheaper in order to meet the demands of the contemporary marketplace?

The answer, of course, is no. Though the market is putting greater demands on the research practitioner, there is no reason to compromise the design process. Indeed, a closer look at the survey process will reveal several methods for extending the value of research while remaining committed to sound principles of research design.

Deductive and Inductive Analysis

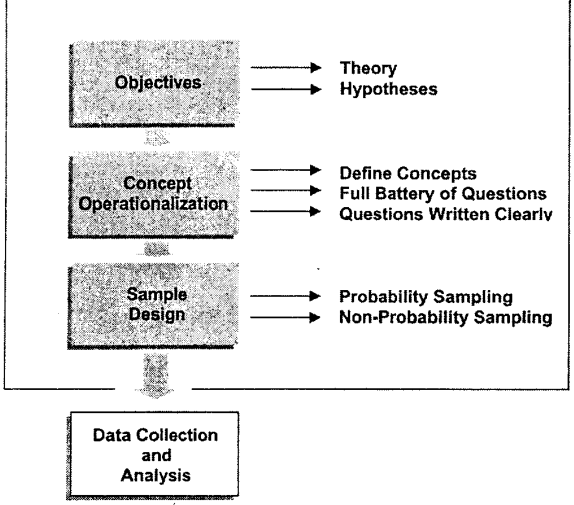
The traditional and formal objective of survey research is to test hypotheses derived from theories of consumer behavior. This is commonly referred to as deductive analysis, which goes from the general, a theory, to the particular, the data. It is a building process whereby theory informs hypotheses, and hypotheses inform data collection methods. This process assures consistency across all elements of the survey process. By continually referring back to the prior steps in the process, one can guard against systematic error that may arise from the inherent, subjective biases of the researcher (Figure 1).

The opposite approach is inductive analysis (sometimes termed "data mining"). Inductive analysis goes from the particular, the data, to the general, the theory (or meaning). Typical operating procedure in an inductive analysis is to take an existing data set and plow through it, looking for statistically significant relationships. The analyst then develops conclusions that explain the statistical findings.

Analyzing an existing data set is one way to deliver research faster and at better value. This is because inductive analysis allows the typically huge investments (of both money and time) of data collection to be amortized

Figure 1

THE SURVEY DESIGN PROCESS



across time, as the data is continually re-used and re-analyzed. Savings are even greater if the data were collected during the normal course of a business transaction such as with bar scanners or online forms.

While inductive analysis may make it possible to save time and money, it does not come without its tradeoffs. Inductive analyses carry far greater risk of bias and error. This is because the inductive analyst is essentially working backwards: attributing meaning to statistical relationships rather than using statistics as tools to test hypotheses. Unlike theory, statistics cannot create meaningful relationships. Though interesting patterns can be found in nearly every data set, statistical analysis cannot make the relationships between variables in a data set more or less true.

There is a second source of potential error related specifically to the statistical procedures that are part of inductive analysis. With a sufficiently large data set, common in inductive analysis, virtually any difference can be deemed statistically significant, since the power of statistical tests increase with sample size. As the size of the data set increases, statistical tests become more sensitive to smaller and smaller differences (or similarities) between variables. This often disregarded aspect of statistical testing increases the risk that the inductive analyst will attribute meaning to a relationship when there is actually none.

Clearly Define and Operationalize Concepts

Both deductive and inductive analyses depend on good data. Good data comes from clearly defined concepts, operationalized as well-written questions.

First, theoretical concepts must be deconstructed and the connotative meanings of essential terms must be clearly delineated in order to write effective questions. For instance, a client may be interested in why people choose to shop online, instead of in a traditional "bricks and mortar" store. "Benefits theory" states that people choose actions that offer a positive consequence or benefit. Based on this theory, we might hypothesize that people choose to shop online because the online shopping experience offers positive benefits that cannot be obtained from traditional channels. Though, at first glance, the concepts in this hypothesis may seem straightforward, they actually require further specification. What are "benefits," and what is meant by the "online shopping experience?"

Once defined, the concepts must be put to work. A questionnaire should be comprehensive, containing the entire range of items needed to completely address the hypotheses and study objectives. Questions must be written in neutral language that all respondents can easily understand. Writing good questions is, at times, hard work. However, good questions beget good data.

This may all seem obvious. However, time and financial pressure could conceivably shortcut this critical step in the design process. For instance, it is easier and less time consuming to simply use open-ended questions in place of good concept operationalization. Though quick to write, open-ended questions and answers do not provide sufficient depth and range of response. They are an incomplete substitute to good concept operationalization.

Instead of open-ended questions, there are many other ways to achieve complete concept operationalization on time and within a budget. For instance, the number of questions that may be asked in a given time frame can be significantly increased by split sampling or using a trade-off design. It is often possible that a theory of consumer behavior is so clearly and completely operationalized that it holds value for multiple clients.

This would lead to a syndicated or multi-sponsored research project where costs are shared across several clients. As noted above, a well-conceived and comprehensive data set can also be used more than once. It can be "mined" for additional patterns that may not have been obvious during the original research (though care should be taken that questions are not re-conceptualized in the re-analysis so meanings are different from the original intent). This would extend the life of a data set adding value over time.

Determine the Population to be Studied and How it Should be Sampled

Though there are many ways to gain efficiencies and still adhere to a robust design process, probability sampling is one principle that should not be compromised, except in dire circumstances. Unfortunately, more and more data is being collected via "convenience" sampling frames, restricted to readily accessible units that can be collected inexpensively on the Internet (such as people clicking on banner advertising, responding to unsolicited email, or signing-up for an ISP).

Since Internet research appears to spawn non-probability studies, let's briefly review why probability sampling is so critical. With a probability sample, each element of the population (or sampling frame) has a known, non-zero, and typically equal chance of being included in the sample. This means that different probability samples chosen from the same population would yield relatively the same responses, and statistical theory can be used to calculate margins of error, confidence intervals, and other survey estimators. Data can be weighted to remove non-response bias. Unique distributions can be assumed allowing for advanced statistical analysis like correlation and linear regression.

Non-probability samples do not have these properties, since all members are self-selected. The chance of a population element being sampled is unknown. No element of random selection is involved. A non-probability sample is a sample chosen *without any basis in statistical theory*. This means that margins of error, confidence intervals and other critical survey values cannot be appropriately estimated.

Because it lacks basis in statistical theory, a non-probability sample cannot be statistically compared to anything. For instance, comparing results of a non-probability sample to the U.S. Census may show that the non-probability sample works well. However, there is no statistical theory that can be used to guarantee that additional samples will continue to do so.

Furthermore, a non-probability sample cannot be turned into a probability sample. Drawing a probability sample from a non-probability-sampling frame (such as a panel of people who "click" on banner ads) will not correct the problems. The sample will merely include all the same errors that infect the original non-probability-sampling frame.

Moreover, weighting demographic variables will not correct problems with a non-probability sample. Statistical theory allows for the weighting of a sample to the U.S. Census or other known population proportion, in order to remove non-response and other forms of non-sampling error. However, data weighting only applies to probability samples. Weighting a non-probability sample makes no theoretical sense. Doing it may work once, even twice, but there is no guarantee that it will work repeatedly with other samples in exactly the same way.

Admittedly, there are instances when time and cost concerns must take precedence and a non-probability sample is necessary. Nevertheless, there are important rules that should be followed when using non-probability samples so that clients will not get confused. At a minimum, a research report must state directly that a non-probability sample was used. Most importantly, research conclusions drawn from a non-probability sample should not be projected, or implied to be projectable, to a larger population. Conclusions only represent the responses of the people interviewed.

A Brief Case Study

A brief case study may help illustrate these points. Many of our e-Commerce clients wanted information that could be used to reduce customer acquisition costs and increase repeat sales. They especially wanted consumer information that would help them do a better job targeting scarce resources to improve the return on their marketing, communications and website development investment. In response, we designed a syndicated e-Commerce market research study that answered these specific questions at a cost far below what each client might have paid for a custom project.

Following the flow of the survey design process as illustrated in Figure #1, we began the study by turning to a well-known theoretical perspective called "benefit segmentation." First articulated by Russell Haley in 1968, benefit segmentation divides the market into different subgroups of consumers based on benefits sought from a given product, brand, industry or channel. The central aspect governing benefits based research is that consumers enter the marketplace to address particular needs. Consumers are able to achieve desired benefits by electing certain behaviors such as what, where, when and how to shop. Indeed, the act of shopping itself is merely the behavior of a consumer attempting to match product or brand attributes to the benefits they desire. As such, Haley notes, "...benefits which people are seeking in consuming a given product are the basic reasons for the existence of true market segments."

We considered benefits research to be most appropriate because it would give our clients a more complete understanding of why people choose to shop online rather than in a traditional "bricks and mortar" store. Prioritizing the reasons why people choose online shopping would help our clients develop motivating sales propositions and compelling web sites consistent with the needs of online shoppers. Knowing how reasons differ across the online shopping population would help our clients better manage and target scarce resources.

Benefit segmentation is a particularly useful theory for developing hypotheses about online shopping behavior. The Internet offers consumers a new choice in how and where to shop. Those choosing to shop online do so because the channel offers unique, motivating benefits that cannot be obtained from shopping in a traditional way. In addition, different people choose to shop online for different reasons. If true, this final hypothesis would allow us to segment the market, based on the different benefits consumers seek from online shopping.

We next set about defining the concepts and creating the questionnaire. We needed to clearly operationalize the theory, hypotheses and objectives with questions that would cover *every* aspect of why people choose to shop online. We also wanted to be certain that questions would be easily understood by all respondents.

The questions were designed through a process that included focus groups, internal ideation sessions, client feedback, and extensive pre-testing. Our first cut yielded over 100 questions about the benefits sought from online shopping. We next built a conceptual factor structure from the 100 questions, so we would be able to eliminate redundant questions and those which did not clearly operationalize the theory. It also allowed us to focus on the language of each individual question, and re-write it if necessary, so they would be easily understood by respondents. At the end of this process, we yielded 60 benefit items that became part of the questionnaire.

Benefit segmentation is about explaining why people behave as they do. Therefore, in addition to the benefit items, we added several hundred questions designed to capture the entire online shopping behavior sequence, from advertising awareness to desired methods of payment. We assessed respondent shopping behaviors across more than 40 product and service categories. We included a comprehensive battery of questions designed to elicit respondent media habits, lifestyles and demographics that would be used to profile the benefit segments. At the end of the process, we had a matrix of over 300 questions that were to be measured across 3,000 respondents.

Next, we needed to create the optimal sampling design, one that would yield a representative sample of the online shopping population at the lowest possible cost. Because of the length of the questionnaire, and the sensitivity of some of the questions, we opted for a phone-mail methodology. We contemplated an online survey, but our experience has shown that non-response increases dramatically with online surveys longer than 10 minutes. (We also contemplated doing the entire survey over the telephone, but the length of the questionnaire would have led to respondent fatigue and the sensitive nature of some of the questions might have led to response bias.)

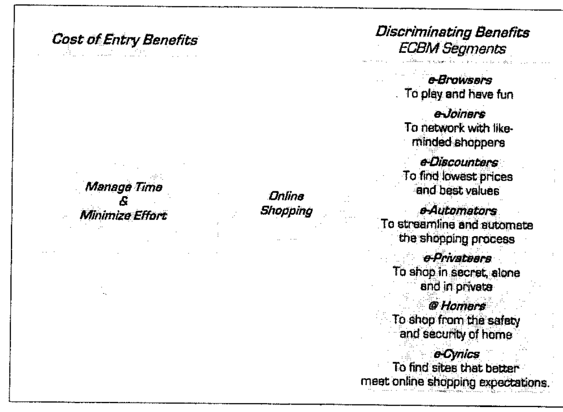
We wanted to use a probability sample. We rejected outright the non-probability panel data built from banner ads and unsolicited email. Rather, we took a traditional and somewhat more costly approach, drawing our sample randomly from all listed and unlisted telephone num bers in the U.S. population. We used the telephone call to screen for those who own a computer and use it for online shopping either for themselves or as gifts for someone else (in order to eliminate business to business online purchases). We used the telephone survey to collect basic demographic and online shopping behavior information from all respondents who qualified. At the end of the telephone survey, we solicited agreement to do the mail portion of the survey.

This sampling design included three points of known non-response error. The first was the telephone call itself. Though we did everything possible to reduce non-response error, we had to assume that those who agreed to the telephone interview were different from those that could not be reached or did not agree to be interviewed. To remove response error we weighted telephone responses so demographic and behavioral data conformed to our population of interest, the U.S. population of people shopping online. Sampling theory allows us to do this because the original sampling frame was a probability sample drawn from all households in the U.S.

The next two areas of non-response bias derived from respondents who completed the telephone interview, but declined to do the mail survey, and from those who did not send back a completed mail survey. Again, we had to assume that people who agreed and sent back a completed mail survey were different from those who did not agree to the mail survey or did not send back a survey. We applied additional weights to the mail returns in order to address these particular sources of non-response bias.

The work that went into applying a strong theoretical perspective, developing hypothesis, thoroughly defining and operationalizing the concepts, and creating a probability sampling design yielded significant results. We called the survey the e-Commerce Benefits Monitor (ECBM). The survey yielded two general sets of benefits that completely explain why people choose to shop online, and seven specific online shopping "benefits" based segments (Figure 2). Each online shopping segment was fully profiled across the large number of Internet navigation and usage questions, online shopping behavior questions, and media habits, lifestyle, and demographic questions included in the survey. We believe ECBM to be the first, and only, truly comprehensive investigation of *why* people choose to shop online.

Figure 2
Why America Shops Online



As a syndicated study, the cost of ECBM is shared across multiple clients, allowing us to cost effectively answer our client's questions. In addition to meeting the immediate objectives, the huge ECBM data set is continually "mined" for new insight. This can be done because we spent time in the design phase fully operationalizing each of the theoretical concepts. We also created a segmentation "short-form," or battery of 9-12 questions (reduced from the original list of 60) that can be used to classify online consumers into each of the seven unique shopping segments, easily and inexpensively. This further extends the value of the ECBM database. Putting the short form online, or using it in other research projects, our clients can apply the entire ECBM database to their own unique and particular circumstance.

All this helps illustrate that we believe it is possible to deliver very comprehensive, timely and inexpensive research without compromising the design process. Research practitioners are being called upon to generate more information, faster and at less cost. Because of these market forces, remaining committed to fundamental research principles is becoming increasingly difficult. There are however, many creative ways to deliver cost effective research without compromising design principles. In spite of all the pressures, as an industry, we must remain vigilant and committed to a strong, reliable survey design process. Only in this way can we discriminate between good and bad data, and deliver significant, meaningful, and cost efficient marketing research to our clients. ◆