

Master's Series on Field Research

A series of interviews with major figures in field research conducted in the early 1980s
by Peter Blanck

Transcript of an interview with Jim Davis



Burton Blatt Institute
SYRACUSE UNIVERSITY

Peter Blanck: Talking about survey techniques, and how it's useful in field research and the state of the art today, and where it's developed from. A very general question.

Jim Davis: If you draw a line from here to here, at one end you would have experiments. In the laboratory, very carefully controlled. At the other end of the line you would have the very informal field researcher of the anthropological type, many of whom I guess you'll be talking to. The survey researcher is half way in the middle. And to understand what it is, you have to see its differences between the other two types. Now, start with the laboratory. The obvious advantage of the survey is that you get out of the laboratory. And what that means really is the sampling. The people you actually get in your net. In a survey if you want to spend enough time and money, you can cover any logical group. If you want to study the handicapped you can get a national representative sample of the handicapped, or older people, or younger people, or blacks or whites, or Greek Orthodox, or you name it. And we know enough to know that you can get very misleading results if you work with a non-representative sample. Classic examples are of class bias. That is if you just collar people, or you talk to your friends. You're more likely to have people who are better educated than the average and since education makes a lot of difference you are going to have misleading results.

But there are more subtle ones. The famous study of marriage announcements in *The New York Times*, which turned out to be very biased because they were done during Lent or not during Lent, and some religion did or did not announce its marriages. If you are operating in a laboratory you generally get people from around from wherever you are. So doing laboratory experiments here at Harvard University; we would have people from Boston, and people from San Diego would be very badly represented in the studies. Survey can get out of the laboratory, and with enough time and effort have a fair that is accurate representation of any group. But there is another aspect to the representation. A survey will broaden the range of the variables we studied. There is a statistical principal that if you shorten the range of your variables you are going to lower your correlations. For instance, if you are studying age at Harvard University, Harvard students will range from eighteen to twenty-two, something like that; a four or five year range. If you are studying age in general population it goes from zero to ninety-nine. So in a survey we are looking at the effect of age, you can get an enormous range in age; whereas, if you are working with college students, or prison records, or workers in a factory, there won't be so much range. So compared with the laboratory experiment the survey has the distinct advantage of enabling you to get representative groups. There is no bias in the selection. You don't have too many of this and too many of that. It also gets a healthy spread on any sort of variables.

If you are interested in religion you are going to get all the religions.

If you are interested in T.V. watching, you are going to get whatever range there is in the audience. This is the classical argument for surveys. Now, at the other end the modern professional area probability survey is quite different from the classical field work in that is well two ways: it is much more rigid and organized, and it is a very complicated social enterprise. Now as a solo field-worker, I can put on my pith-helmet and go out in the subway and sit in a bar and talk to people and get all sorts of insights and ideas. And there is usually no great control over what I am going to do. Next, a he simply says something; I follow it up and say "Oh, tell me more about that," and so forth and so on. Surveys are actually extremely rigid and there is the image of the individual respondents talking to Dr. Gallop and saying "Well Dr. Gallop here is what I

think of the president's policy." The point in fact, the respondent that talks to one of perhaps one-hundred or two-hundred, part-time workers who are chosen, not for their personality, but for their slavish care in following instructions literally. And in the ideal survey, each of these one-hundred interviewers would do exactly the same thing. It's ideally, each question would be read from a schedule with exactly the same stress on this word or that word, and one of the real horrors of the professional survey is the creative interviewer who would say, "Well, I'll do question twenty-two first because it might work better on this respondent."

So the whole thing is very far from spontaneous, very far from creative. And the survey interviewer is not supposed to be a particularly charming or effervescent sort of a person; rather we encourage a business-like sort of manner for two reasons. Uniformity, so that each person deals with a standard sort of stimulus, psychologists would call it. And also because we are afraid of leading or biasing the respondents. Because if I say "What is your party preference?" And you say "Republican." And I say "Well that's wonderful!" then you're much more likely in your subsequent questions to give a pro-republican kind of response because you say "Oh, that person was pleased I'm republican, they're probably republican." In that sort of a way, and for a certain fraction of the interviews, we call back in a week or ten days, frankly, to make sure it took place, but more than that to check up on the quality. And some of our very best interviewers, the interviewee can barely remember the person.

Oh yes, there was a woman here last week and she asked me about so and so, but does not remember that they wore a red hat, or were a striking blond, or whatever. So highly standardized and very, very rigid, along with the interview schedule, the interviewer gets a "gynormous" mimeographed thing called "Specs," for specifications. And they tell them exactly what is meant by each question, how to handle if the respondent says "What do you mean by fiscal policy?" the interviewer is instructed to say "By fiscal policy I mean what the government is doing with the banking system" or whatever. The purpose, of course, is not to destroy creativity but to have scientific control.

So, while the survey interview does take place outside the laboratory, in the home, real people talking to real people, which is a distinct advantage in social science. It is quite different from the informal participant observation, anthropological kind of interview because it is very, very carefully controlled, specified, organized. The other aspect of it, which really follows from this organization, is it's a very large, socially organized enterprise. I couldn't give you an exact number, but the typical national survey of this sort done by National Opinion Research Center, where I work, or the Michigan Survey Center, or Gallop, or Roper, or Harris, or anybody; I would suspect that several hundred people will be involved in the sense of charging their time to the project and using its funds from start to go. And it is a very divided division of labor. And no one has the skills to do all of this. In the solo field work pretty much you have the lone pilot out there flying through the fog, this is more like the 747 with the computers and the crews and the ground control and so forth and so on, and different people do different things. The interviewer at the end has nothing to do with making up the questions, has nothing to do with drawing the sample. The interviewer is instructed how to choose the respondents, the sampling people and most of these organizations have a sampling department, which is highly technical. The sampling department has nothing to do much with the questions or with the actual interviews or seeing the people, and the data processing people, who are another group, never talk to a live person; but, work with

piles of interviews that are delivered. So it's a very large organization with a very sharp division of labor and large number of very specialized people.

The net result is that I know very little about all the things I'm talking about in the literal sense. I've been involved in national surveys for twenty-five or thirty years, I have never drawn a sample. I would no more draw a sample than I would say Mass, or scuba-dive, or whatever, they are highly qualified people who do that. The study director, whoever is in charge of it, has to know enough about these other specialties to try to get them to do it right and get it coordinated and try to keep it within the budget, but you don't actually have to know how to do any or all of these things; you have to coordinate a group of people. So the natural result of all this is that almost like cyclotrons for physicists, centers have grown up where surveys are actually conducted. And a big change in the nature of social research, in the last twenty-five or thirty years, has been that individual professors or investigator do not run surveys out of their hip pockets anymore, but subcontract the work to professional organizations.

Peter Blanck: Do you want to talk more specifically. Give a definition of NORC and maybe your association with it as an example of these types of.

Jim Davis: Okay, I mentioned N-O-R-C, which is known to everybody in the profession as NORC. It's the National Opinion Research Center. It is the oldest nonprofit survey center in the world. Surveys began essentially in the middle '30s with the Gallup poll commercially, which grew out of political polling but a lot media research, radio, and the beginnings of T.V. Other surveys grew out of the Federal Government, which was interested, say, in predicting famous example, predicting crops, how many soy beans are we going to produce next year and the Department of Agriculture developed surveys to interview farmers and find out how many soy beans they were going to have.

NORC was founded in the early '40s with the mission of providing sort of a nonprofit analog to the commercial polls that had grown up. The original notion was to do surveys for labor unions, and charities, and little people. The net result is somewhat ironic. Because of this very large scale in-division of labor, professional surveys are extremely expensive. A high-quality national survey these days in the early '80s can run up to half a million dollars without being extraordinary in any way. Because of that, labor unions and little people in charitable organizations cannot afford surveys. And the only people who can afford surveys are large federal agencies. So most of those university-affiliated survey organizations, there are several: NORC is affiliated with the University of Chicago, in Chicago, Illinois; the University of Michigan in Ann Arbor has a survey research center which is the General Motors of the industry, luckily there is no Japanese competition; Temple University has a national survey organization.

The nonprofit survey organizations are generally affiliated with the universities because they draw on so much skill and technology. In order to carry out a survey you need a sampling statistician, you need consultants in the specific contents; if it's an economic survey you have to talk with economists about how to measure income or wealth, or whatever. You need graduate students for cheap labor, and so forth and so on. So that most non-profit survey organizations are connected with universities, but they are funded by outside sponsors, called clients in the trade. The larger ones--almost all these days--from the Federal Government, so it is very interesting flow of ideas and money, from the university sector to Washington, D.C., and back and around.

Peter Blanck: I want to slowly push you to be more and more technical. One of the prevalent things that they use on the outside is mail surveys, which they send out in bulk; the types of things that only the people in back of the Business School actually end up tabulating, the graduate students. If you could talk to use direct interview versus mail survey as different techniques, which may apply to different types of questions you'd want answered, or maybe not.

Jim Davis: There are at least three main approaches to collecting survey data. The statistical theory doesn't care whether it's by mail or telephone or personal interview, but the actual mode of operation is quite different. You have the personal interview where somebody knocks on the door, sits down, goes through the long questionnaire. This has obvious advantages. The main one is that if I ask you something and you say [gibberish] and I say, "Can you explain a little more about [gibberish]?" Whereas, if you're filling out a questionnaire and you put some weird inscrutable thing, I will never get it unscrewed. So the personal interview has the distinct advantage, that the interviewer even through operating under this very tight leash, can probe confusing things.

Well in some examples you can ask the person to check something. A lot of people do not know their ages. If you ask them their age, they say "Well, I'm not sure," the interviewer can say "Well is there a family bible or something you can look it up in?" So a personal interview has that distinct advantage. It is also extremely expensive. And there is a lot of technology associated with surveys at the computer-end, but the main costs are in data gathering, and there is just no way an interviewer can interview fast. An hour interview still takes an hour just as it did in 1930 and interviewers are getting paid somewhat more than they got in 1930. So using that as a standard, an alternate form is the direct mail interview, where you send somebody a form, they fill it out, and they mail it back. It has a number of very distinct advantages. Not always the obvious one. First of all it's much cheaper. Cost per case is a lot less. Secondly, there is some evidence although studies say this and studies say that; that for hot topics--potentially embarrassing or confidential things-- people are more likely to be candid when they are sitting there alone checking it off than talking to somebody. Say it's a question about my sex habits, or income tax cheating, or do I drive over the fifty-five mile an hour limit? Now you look like an honest citizen. So if you said "Do you drive over the fifty-five miles per hour?" I would say "hardly ever;" whereas, in fact, I always do. And when filling out a form I'd be more likely to tell the truth. It, direct mail, has another advantage that the mailman can get in places that the interviewer can't. Everybody gets their mail. A survey interviewer may be turned back at the door in say a very fancy high-rise apartment. The greatest obstacle to social science progress is the doorman, a good doorman, bad doorman not a problem. But good doormen keeps people out of high-rise apartments. But the mail man is always let in so you can mail to anybody. The disadvantages if direct mail are the obvious one that you can't, if the answer is inscrutable or illegible or whatever, there you are. Although there are many combinations; for instance, you can have a direct mail survey and call up people if you have their phone number. Sometimes you have a personal interview and then leave a longer questionnaire part with people, more sophisticated people of mixed modes.

Having said these wonderful things about the direct mail survey it has the notorious drawback of very, very low completion rates. The completion rates for a personal

interview survey--well, this is highly controversial seem to be slipping these days; but with proper time, and effort, and funds, 70 percent completion is quite reasonable, and actually it should be 80 percent or 90 percent. Now direct mail 50 percent is remarkable, and often you're down to 5, 10, 20 or 30. Now statistically there is a big difference between the number of cases and the completion rate. If given the choice of the same amount of funds, I may be able to collect many more cases by a mailing. I can mail thousands of them out and get hundreds of them back; whereas, I might only be able to afford a few hundred interviews.

You can get many more cases. But there is always the question of the representativeness, particularly if you have a very, very small return rate. This depends an enormous amount on the topic. A general rule of thumb is that "Involved people are more likely to respond." So if I mail out a questionnaire on political attitudes or something, people who are active in politics are probably more likely to answer; whereas, people who are totally indifferent to politics. And a reasonable rule of thumb would be--not that republicans or democrats are more likely to return but "Political partisans are more likely to return a political questionnaire than the apathetic." So direct mail surveys tend to overestimate involvement, interest, concern, or whatever.

The country is made up of millions of people who really don't care much about whatever you doing your survey on. That can be very important. Particularly, say, in marketing or something like that. People who sell dog food presumably need to know how many people don't have dogs. If you send out a direct mail questionnaire about dog food I suspect that people who don't own dogs are quite likely not to answer. So, that's direct mail. There is a third, former survey, which has become rehabilitated. This is the telephone survey and has an interesting intellectual history.

In the middle '30s there was a famous magazine called the *Literary Digest*. It presented a public opinion poll, I guess, of its readers every month on timely topics. And they polled voters in, I believe, the 1936 election, and elected, I guess that would be Landon . . . Hoover, Landon, Willkie. Their poll showed Landon winning. The ballot boxes showed Roosevelt in a tremendous landslide. Now what has happened was that they had gotten their sample from telephone books. In the 1930s there was a big social class bias in telephone owning. Just as these days, if you see a Mercedes parked out in front of a house you say "Well, that's an up-scale house." In 1936, if you had a telephone you had an up-scale house and up-scale people voted for republicans, still do. Because of that, and more particularly since the *Literary Digest* went out of business shortly afterwards, polls based on telephones has a very, very bad reputation because of bias, which turns out to be essentially social class bias.

However, times are changing and these days, as best we know, about 90 percent of the population has a telephone. The pockets without telephones are very clear and they are very important. The rural south and, in particular the black rural south, is relatively low on telephones. But the north, and within the north there isn't much race difference between blacks and whites in the north, outside of the rural south virtually all American households have telephones. These days a telephone sample is not automatically biased in terms of social class. The telephone is even better than the post man in terms of getting there. That is, the postman could get it stuck in the slot in the door, and the telephone actually rings and somebody answers it. Telephone samples have been rehabilitated and are now, among sophisticates, considered quite sheik because of the

ability if the telephone to get into high-rises and to actually reach people. The problem of non-response in surveys is not people saying "take your survey and go away," but people who you never find. People who work odd hours, or traveling salesmen who are never home, and so forth and so on. And they are easier to hound down by telephone. Now the step beyond that, the really cutting edge technology and it's used every day by some of the most sophisticated market researchers, thousands and thousands of interviews is the computerized telephone interview where an interviewer with a headset and T.V. scope will sit at the T.V. scope. And the scope will type out the question, there is no printed questionnaire. You read what it says, you say [gibberish] and you type something into the thing. Now, when you do that, the machine can be programmed to make very sophisticated checks and adjustments in the questionnaire.

Questionnaires often have branches. Take an obvious one, if you're asking people about finances for college you probably wouldn't want to ask if they don't have any children. If they have only young children, you want to ask if they are planning ahead. If they have older children you want to find out if they are in college, and if they are not in college, you know a very complicated sort of thing. The computer can figure out all these things and simply present the appropriate question, and the interviewer does not have to say "Now wait a minute, your children are under sixteen, so I go to page thirty-five and have to do this." This is called CATI, C-A-T-I, Computer Assisted Telephone Interviewing.

And, as I said, is sort of the current fad in research and it allows enormous flexibility and this control that I was talking about. Because you can see although the computerized interview can be highly flexible, the interviewer has zero choice at all the interviewer simply reads what it says at the scope. Now, there have been elaborate studies of the difference between telephone interviewing and personal face-to-face interviewing. And the difference is fairly obvious. You get a lot more material face-to-face. Simply, a telephone interview with the same question will generate many fewer words of response. Somehow, if there is a face there, you tend to say more to it. If there is a voice there, you say "da-da da-da;" if there is a person there you say "da-da da-da, oh and da-da." Now of course that works both ways; if you paying interviewers by the hour, sometimes the shorter interview is an advantage.

There is a lot of fancy research on this, and it is not clear whether people are more honest, or less honest, or whatever. There is the argument that people are more likely well, there is this famous finding as I recall that Catholics are more likely to admit that they don't go to Mass regularly in a telephone than in face-to-face back to what I was saying earlier about people are trying to make-nice with the interviewer. On the other hand, it's very hard for the telephone interviewer to see that the person is nervous, or evasive, or whatever, and they know they get a lot less material. So, you have these various modes and the . . . it used to be that they were divided into good and bad. Now I think a more sophisticated view is that each has its advantages and disadvantages and very frequently an efficient, successful project will combine them.

Peter Blanck: A related topic, I was watching a Nova program, you've probably seen it, on "America Computer Intrusion into Our Lives," and they showed a telephone survey. And the question I'm asking is, "Where, as a survey researcher, as researcher like yourself, do you draw the line between intrusion and between good investigation?"

It seems like a difficult question, your laying on these telephone, computer surveys onto these people, they really have no choice.

Jim Davis: The ethics of surveys have been debated, discussed, thought through enormously in the last few years. To begin with the obvious, the vital question is disclosure. The fundamental rock bottom ethical requirement of professional research is that I not disclose that you said you haven't been to Mass for five years, or you always drive over the speed limit, or that you have three children and their names are Agnes, Louise, and Charlie.

So, the major protection of the respondent is that the results will not be disclosed. And there is a paradox that, as far as any one knows, there has never been a violation at this level. I don't know of a case . . . once in a hundred years, inevitably by the laws of probability, you get a famous person in your sample. I once had a survey in which a famous movie star was drawn and I read his or her questionnaire very, very carefully, and so forth. But I didn't announce, and very careful, even, say, when teaching a class. Even now, not to say who it was and that this famous movie star was a republican or took tap-dance lessons.

Now the paradox is that the only threats to this confidentiality that anybody knows about come from the government. And a very complicated set of legal problems, in that, on occasion zealous district attorneys will subpoena questionnaires. Classic example is a famous study called the "Income Maintenance Experiment" in which families were given money to see what would happen, sort of a test of the negative income tax. And a district attorney subpoenaed some of the questionnaire to find out if there was any welfare fraud involved. By and large the survey organizations have been successful in preventing this by negotiation, but there is no legal protection. Because if I ask you all sorts of naughty things; I cannot, I can give you my personal promise that I will not reveal them, I cannot say and I can guarantee that the district attorney of Middlesex County will not get their hands of them.

So in some areas, say for instance drug research, a lot of surveys on drug usage and drug research, and very difficult problems of assurance of confidentiality. So the first ethical matter is confidentiality. And this in everybody's experience is no problem. Then you have the question of consent. Now survey people argue that surveys almost intrinsically have consent. Our interviewers, very seldom, knock people down sit on their chests and say "I'm not going to let you up until you tell me what you think about the Federal Reserve Bank." And at any time the respondent may break off the interview, some do, the interviewers are instructed to make a diplomatic attempt to maintain them. But if people really don't want to tell you anything they don't have to. There is no way to compel them, which is quite different than some social science experiments where you can get data on people without their knowledge. So, in a sense surveys are less problematical in this ethical area than say some forms of observation. If you're working in a large office and I unobtrusively count how many times you go to the water cooler; you may not know that I am a researcher. And I might be getting information potentially harmful to you: that you're at the water cooler half the time, and never working, and somebody else is a hard worker. But if I ask you how often you go to the water cooler and you don't want to tell me, or you want to lie to me or whatever, it is your prerogative. So the matter of consent, surveys have a built-in self-protection, which is quite different from the field observation where I go to the bar, and you're

dumb enough to start talking about how your goof-off on the job; there is an ethical problem. Or, in the experiment where I manipulate the heat in the room to make you thirsty so you go to the water cooler.

Where having said that of the ethical issues, and we're very concerned about them, anonymity is the big one and seems to be working well, that consent really is built in the nature of the process. The tricky question is that of the concept of informed consent, as it is not ethical to say I want your consent to be in some research. I've got to tell you what the research is, or you may be signing up for something that is not to your own interest. This can be a problem. Say I am studying the effects of early family experience on bigotry. If I say to you "Well, I'm interested in studying bigots and you seem to be a very bigoted person, now will you participate in my study?" I may lower my completion rate, but also may bias the study. If you know the study is about religious bigotry then you may start thinking about religious bigotry, or whatever, and it may shade your answer. So the question of informed consent is a very delicate one. Can you tell people. Actually with the experience we have they don't want to know that much. And they do not care that the "Department for the Improvement of National Roads and Curbs" is sponsoring question thirty-two, and the "Ministry for the Development of Preschool Architecture" is sponsoring this or that.

People are generally less interested than you think. They just think it's something moderately interesting and they want to get it over with. The question of informed consent is a subtle one. But I would say the remaining ethical question on this, which is a very subtle one I don't think has an easy answer--is whether groups and outfits have a right to anonymity. Take a classical example. You do a study of cheating at universities. It would be totally unethical to say that Susie Jones is a cheater. Her questionnaire says that she cheats; no question of whether it's unethical. But to then say that Athabaskan State University has the highest rate of cheating of any university in the country is a very delicate matter because you don't really have at the Athabaskan University's permission. You get permission from a bunch of people in the study, and the question; for instance, research on different ethnic and racial groups often make some look better than others.

You can't get informed consent from all the Texans to do a study of differences between Texans and Oklahomans or something like that. So this is the remaining issue of the identification of groups and collectivity that is really very, very hard to solve.

Peter Blanck: Let's switch gears a little bit and talk about the science of questionnaire and survey design. I think Bob Rosenthal talked about that a little. But if you have a view on that, general techniques; how a beginning survey researcher would think about how to generate questions, obviously depending on the topic of the issues they're interested in.

Jim Davis: You usually distinguish between the question and the design. The design is the sample of people you're going to compare. So, you're going to compare people who are very active in politics and not very active in politics. Controlling for reach in social class, or something like that. That can be worked out with a sampling statistician pretty much with a slide rule, except they don't have slide rules any more.

Question writing is very much an art. And we know very little about it except that a change of a word or two will make an enormous difference. A famous example is a question on happiness, and there are two versions: there is a happiness question that Gallop people used, and there is a happiness question that the NORC Michigan people used. And version A says "Generally speaking, are you very happy, fairly happy, pretty happy?" Version B says "Generally sp" -- Whoops Version A says "Generally speaking, are you very happy, fairly happy, not too happy" Version B says "Are you very happy, pretty happy, not too happy?" So the difference between "pretty" and "fairly," which doesn't sound like much of a difference I had to stop and remember it-- will produce a 15 percent difference in the proportion saying they are happy.

So, very slight differences in question wording can make enormous differences in the answers. We know that. What we don't know is which was the better way. One gives you more happiness, one gives you less. We know that. What we don't know is: Which is the correct answer? And it's a trap which is almost impossible to escape since the textbook answer is to compare it to a true yard-stick. For instance, in voting, if we have a question "Did you vote for this or that candidate?" we can check it against the election returns to see if we are getting about the right proportions of democrats or republicans. But there are no happiness returns.

So we cannot go to the "National Happiness Register" and see whether version A or version B gives the better answer. What we do have to do are to work with our items over and over again, and learn from the way they behave, what they mean, and how they work. About the only rule of questionnaire writing is to "avoid creativity." By and large you can find a question that somebody else has used, then you should use it again, if only to compare your results with the results other people use. This is particularly important when you're doing studies over time. This happens to be an area where I am working. Look at social trends and you have to repeat your questions word for word or you can't tell whether any change you get is from change or from your difference in question wording. But, question wording is the most important part of the survey. The things I talked about--response rate, sampling, interviewer effects, and so forth--have all been studied. There's thousands of studies of them. But their affects are perhaps one-tenth of the effect of the wording of the question. And although we're doing increasing research on it we're learning more and more it is basically, I wouldn't say and art, but very much a craft.

Peter Blanck: So just as the interviewer can bias the interview by the way he phrases things or the emphasis he gives, you can easily bias things by the way you write the question.

Jim Davis: Oh yes. Anyone who wishes can change the level of approval. The classical example is in foreign relations. There is a classical example of a question in the Vietnam War. "Do you approve of what we are doing in Vietnam?" "Do you approve of what we are doing in Vietnam to stop the communists?" The difference between those two questions changed approval of American policy in Vietnam by 30, 40, 50 percent or something like that.

You can always produce a level of approval you want. However, at the same time it's blatant what you are doing, so you really can't get away with anything. And what really is more important is not the level. It's like . . . well, I will get different quote levels of

education by asking are a college graduate or asking if you are a high school graduate. Well, that's not a real difference it's simply that in one case I've asked for an extreme cut and the other case I've asked for a more middling cut. A lot of question wording simply moves a question from sorting people into very high versus everybody else, or above average versus below average. That's no problem. Where the question wording is really problematic and very subtle is where we change a relationship. Sometimes you can write a question that will look just like another question, but better educated people will be more likely to say "yes;" and the other version of the question that looks very much the same, the better educated people are not more likely to say "yes." Now that is scary.

That's when question wording affects the relations between variables that it really becomes problematic, and we're just beginning to study that. The other related matter is we're just beginning to understand how the order of questions and the arrangement of questions makes a real difference. Take an obvious one, if you ask people about federal spending after you've asked them a whole series of specifics, "What about this program, what about that program, and what about the other program?" After you sort of reviewed all the programs, they are more in favor of spending; whereas, just out of the blue you asked "How do you feel about federal spending?" people are less favorable to it. That is where you put the question in the order can have a big effect on this distribution. We're just beginning to understand these things, which are crucially important.

Peter Blanck: Okay, if you could talk a little more technically about sampling procedures in general, the technicalities involved.

Jim Davis: The actual procedure in drawing a sample is called -- for most surveys, there are exceptions -- are called "Area Probability Sampling." The easiest sort of sampling is when you have a list.

If I want to sample the students at Harvard -- except for questions of confidentiality, Harvard would not let me have a list of all the students I would get a copy of the Harvard telephone directory and in effect by using dice pick every fifth name, or whatever. No problem at all. But there is no list of the American population. In effect, who is in the American population changes momentarily. People are born, people are dying, people move in, and people move out. So when you're dealing with a population where there is no convenient list, everybody has to be someplace. So here is a map of the U.S. and we know from censuses that approximately X fraction are here; in particular, in large metropolitan areas.

So we sort of have a lottery and choose metropolitan areas giving them as many tickets in the lottery as they have people, and we end up with maybe thirty, forty or fifty towns or communities. Within each then we carry out another lottery and pick areas of town. Now this will differ: sometimes it is something called a census tract, sometimes its ward, or precinct, or something like that. Then, eventually, we get down to a smallish level, a group of blocks, and at that point staff are sent out to conduct a sort of a mini census. They go through this five or ten block area and find out exactly how many people are living there.

Because original drawings were based very heavily on census materials, but whatever census it was, it was, it no longer is. The American population is mobile enough that -- well this is now the beginning of 1982 the results of the 1980 census, which are not out yet, are no longer true. The number of people living in Cambridge, Massachusetts, is not whatever it was in 1980, or if that has tended to cancel out, the number of people of this block certainly are different.

So we have an enterprise called "listing," and the interviewers go out and determine how many people live on that block and really take it apart. They look in garages, and basements and roomers, and so forth and so on. Then there is a final lottery among those people and an interviewer will be given a list of, assignments vary, maybe ten to twenty names and told to get interviews with Theophilus P. Schmertz at 1222 North Maple Street. The economics of it as you can see that the very expensive part in this listing at the end and you only have to do that in a very few places. You don't have to list every garage in the United States. The net result is, however, a very statistical effect called "clustering." And, which simply means that people in the same geographical area tend to be more similar; you don't find many farmers living in urban Boston.

So if you pick a neighborhood in downtown Boston you aren't going to have a lot of farmers. You pick a neighborhood in central Kansas you're going to get a lot of farmers. People who live in the same geographical area tend to be rather similar. This doesn't change their odds. The odds that a farmer in rural Kansas will get drawn in your study are proportional to the number of farmers in rural Kansas. But it does mean at the end that the data requires some statistical adjustment. It's ignored in all statistics textbooks. Correcting for something called "design effects." But the net result, the economics of it, as you can afford to collect so many more cases that these studies are very cost efficient, which then means, oddly enough, that sample designing, and it's all done on computer these days, is really an economic rather than a, quote, "scientific decision." You tell the sampling designer how much money you have to spend and how much precision you must have. You can buy as much or as little precision as you want; basically, in terms of the number of cases. If you interview a small number of cases, you can have big confidence-band of error around your estimate. If you have a huge number of cases, very small. You tell the man how much you can afford, how much precision you need, and he turns around the computer and out comes the sample assignments.

The net result, however, is a very complicated process and in the end you have to make some statistical adjustments to get back the proper numbers.

Peter Blanck: One final question, which we end the interview with, what do you enjoy about social research, and what do you enjoy about your branch of social research in particular? What is stimulating for you, what is fulfilling, what do you value?

Jim Davis: I enjoy making little discoveries. I was working with some data from an NORC study at home yesterday and I discovered that English Catholics that is American population, people who are Catholic but whose national origin is England, rather than say France or Germany or whatever have very different attitudes and opinions in social standing than other groups.

And I'm probably the only person in the world who knows that. It's not very important, it's not a world shaking fact. But when you work with data about the whole United States, you learn endless, fascinating facts and properties in this very complicated society we live in. So that a survey takes a long time; it takes a year to two years from start to end. Most is tedium, and the part where you have to cut the budget is just hell. What is fun is when the interviews are finally in, and the data are all cleaned up, and you start to make your tabulations and you see "Hey, oh my, this increases," or "You know that 22 percent of the population do this?" or "English Catholics are very different."

Actually something quite relevant in the same data set we're finding nobody really looked at very seriously if you take the black population born in the north, living in the north; black population born in the south, living in the south; and the black population born in the south living in the north; they are as different as Jews and Protestants. They are three very different social groupings in the black population; whether they're old northerners, old southerners, new migrants. And discovering these things is an enormous amount of fun.

Then once you've discovered them, they have to be analyzed formally and written up, and that's tedium again. So, for me at least, the fun is in the middle of actually seeing interesting results come out of the data. And I guess we have enough fun in that ten or fifteen minutes to justify all the tedium in the six months before and after.

Peter Blanck: Good, great,