VISIT course on constructing an effective survey

Slide 1: Welcome to VISIT's basic-level training session on what constitutes an effective survey. As explained in our introductory material on social science, designing surveys is more complicated than you might think.

Slide 2: The goals of this course are pretty simple. We want to try to cram highlights from a semester long research-methods class into about a thirty minute presentation. We'll start with a brief review of social science research, and follow that with a discussion on the Institutional Review Board – what it is, and why it's there. Next, we'll discuss statistical-based social science research and at the end we'll present an example of a well-designed survey on warning response. This course should not only help you formulate your own surveys, but will facilitate an overall understanding of what constitutes a good survey when you run across one in the literature.

Slide 3: The goal of any social science research is to understand individual and societal behaviors. There are two basic types of study:

- **Descriptive studies**
  - Which include things like interviews, questionnaires, or natural observations
  - This type of study can only establish correlations, it cannot establish cause and effect. Cause and effect can only be established using:
- **Experimental studies**
  - This type of study works by incorporating manipulations.
  - Manipulation means purposefully varying one or more questions within the survey, then giving different participants different versions. Usually there are 2-3 variations. The purpose is to see if answers to any of the other questions in the survey seem to be systematically affected by these variations.
  - This type of study is very complicated to design, and a lot things can go wrong[1], so we won't be addressing experimental studies in this course.

Slide 4: What is a survey?

- A survey is nothing more than a way for us to gather information about individuals and groups through a series of questions.
- The thing to be careful about is that questionnaires are based on self-reported information. So you have ask, "are the answers true?" Only the participants know for sure, and sometimes even they don't know. There can be many problems, such as memory lapses, guessing, emotional factors, not understanding the question, personal image problems, and so on. By constructing a survey carefully one can circumvent a lot of these issues.

[1] unethical manipulation of people, IRB problems, confounding variables

- There are a couple of ways to administer a survey
    - Structured, one-on-one interviews
        - These are time intensive and have fewer participants, but you can follow up to clarify misunderstandings with follow-up questions
    - Questionnaires.
        - These are anonymous, easier to use, and usually yield more participants. However, remember that follow-up is not possible, so question design is critical.

Slide 5: The Institutional Review Board

- During the middle part of the 20th century, medical and psychological studies pushed the envelope with experiments that began to have more and more potential to cause harm to participants[2].
    - The National Research Act of 1974 resulted in the Belmont Report of 1978 which defined ethical principles for human subjects in medical and psychological studies. The main focus was a respect for persons, combined with beneficence, and justice.
- This all led to new regulations and pre-review of all human-based research and the so-called Institutional Review Board, or IRB
    - By the way … IRB approval is required for all research that receives support, directly or indirectly, from the United States government. So gaining IRB approval is your first step if your you're designing a survey. If the study is a simple survey, it's usually a piece of cake. Regional offices typically handle this approval process for NWS offices.

Slide 6: Okay, so now on to the main topic.

- The first step in survey design should always be to <u>formalize the goal</u>. Let's say your office recently issued a winter storm warning and most people hadn't reacted the way you'd hoped. You might decide to put together a survey to try to find out why.
- So what is the goal of the survey? Write it down. Let's say your first cut is "To find out why people don't respond to severe winter storm warnings." But this is obviously too broad brush, and much too difficult a question for a single survey. Narrowing it down to a more reasonable scope we might say something like, "To find out why people in our warning area didn't respond to our recent winter storm warning."
- Later we'll show an example survey that had a very clear goal – to learn why a lot of people in Phoenix, AZ didn't respond to severe heat wave warnings during a summer in which there were several deaths. This goal led to the first questions on that survey.

---

[2] Examples include Project MKULTRA which was a 1950s, through early 60s series of classified mind control studies organized by the CIA. There was the Milgram obedience experiment (1963), and the Stanford prison experiment (1971), and even sexual studies like those done by Kinsey or by Masters and Johnson. Things like that.

Slide 7:

- Okay so if you wanted to find out why people didn't respond properly to heat warnings, you first need to find out whether or not participants were even aware of heat warnings. You can't "ignore" something you didn't hear about. A well-stated goal will help you focus on salient points. The rest of the first seven questions go on to ask about sources of information, any recommendations for action they remember receiving and so forth. But those who hadn't heard heat warnings were instructed to skip the next six follow-up questions. We'll look some of those questions a little later.

**** So let's get on to the survey design ****

Slide 8: There are a lot of factors involved in designing a scientifically-based question set. First, survey designers need to be almost hyper-vigilant when structuring questions that are conducive to later statistical analysis. Second, there is a veritable litany of potential biases to avoid, third there is the problem of formatting questions that are completely clear and finally, you need to be mindful of the study's limitations. Let's look at each of these separately.

Slide 9: Designing for Analysis

- A statistical plan is key to later analysis. With properly designed question and answer sets, analysis becomes simpler. Sometimes not paying enough attention to this step might invalidate important parts of the study.
- Making sure questions are quantifiable and comparable is tough, because in the case of social science, we're basically talking about *nuanced* responses to social/psychological situations. That is, on a survey we assign values to nuances, and this can be relatively arbitrary. A 5-point Likert question such as, "Do you believe that weather warnings are accurate enough for you to act upon?" with choices; strongly believe, somewhat believe, don't know, somewhat skeptical, strongly skeptical can be *assigned* numbers 1-5. This is fine as long as you keep in mind what's being done. And remember … the answer to this 5-level question can only be statistically compared with other Likert questions that have 5-point-scaled questions, too. The same levels of nuance, if you will. To compare this factor with the physical sciences, think of it as accuracy of the instrument. Picture trying to compare a set of temperature measurements from thermometers with a +/- 0.125C accuracy with another set using instruments with a +/- 2C accuracy. It could be done, but wouldn't it be better if they all had the same accuracy?

Slide 10: Designing for Analysis (2)

- Questions need to be clear and concise. We've all seen questions on surveys that leave you wondering what the researcher really wants. That kind of question distracts participants. Also, several studies have shown that concise questions tend to yield more honest answers.
- Open-ended questions (that is, asking the participant to write out their own answers or opinions). Remember – you want all of your questions to be part of later analysis. Before analysis, someone will have to design a coding scheme that somehow, someway

assigns numbers to those written answers. Open-ended questions can very be time consuming to handle during analysis and add an extra level of arbitrary ranking.

- Finally, most psychological factors, (e.g., resilience, anxiety, skepticism, etc.) have had measures established for them in the social science literature. You can find those easily enough using a search engine. If such a measure exists, it is generally expected that the researcher will find it and use the established question set, since it's been shown to actually measure what you're hoping to measure. Also, it makes life easier.

Slide 11: Formulating questions

- Questions with ambiguities and/or biases can ruin all or part of the survey's results.
- Keeping your questions clear is key. A good survey avoids lengthy, questions with confusing language or jargon. Remember … be concise.
- Also, a good survey doesn't leave participants wondering what you want to know. Questions like, "What do you do when you hear weather warnings from various sources?" will leave participants wondering more about what it is you want to know, rather than thinking about their answer. A better question would be, "When you hear a winter storm warning how do you typically react?, followed by a list of specific choices, such as – Ignore it, Look outside, Make some preparations, Cancel most plans and prepare to deal with the weather, and so on. This lets the participant know specifically what you're getting at.
- Watch out for leading, or loaded questions. Obviously, a question like, "Should responsible people always heed weather warnings?" is a leading question. A question like, "Why do you ignore severe weather warnings?" assumes that the responder does that. It is a loaded question. Multiple people should always review the question set to look for any such errors before the survey is ever conducted.

Slide 12: You want to be careful of biases

- There are many common biases that can creep into a question set and potentially ruin the results. Here are a few to think about.
- Sampling bias – Selecting your participants could affect results.
    - You might get very different answers if you choose college students versus senior citizens versus people going to work.
        - It is probably not possible to get a completely representative sample, but you'll need to get as close as you can if the survey results are to be helpful
- Volunteer bias – What about the people who didn't volunteer?
- Self-reporting bias – People sometimes practice image management (they don't want to look bad), they may have faulty memories, emotional issues, may not understand the question, practicing playful sabotage, response bias (such as they never chose extremes), and so forth.
- Cultural bias – different cultures often think about the same problem differently
- Experimenter bias – questions inadvertently phrased to correspond with what you want to hear. These usually turn up as loaded or leading questions.

Slide 13: A few more miscellaneous biases and then we'll move along.

- Extraneous variables – time of day, season, current weather, etc.
  - If one group taking a survey about tornado warning response is polled in the spring, and another in winter, will the season be a factor?
- Vague questions – Do you experience bad weather regularly?
- Limited options – Do you consider this product; excellent, very good, fair (versus excellent, good, fair, not very good, poor). If you don't want to hear the full truth, why do the survey?
- Double barrel questions – How satisfied are you with weather forecasts and severe weather warnings? Okay, so enough said. On to our example.

Slide 14: An actual survey.

- It's finally, time to look at an actual survey, to see if we can apply what we've just learned to judge how well the survey was designed.
  - The survey reference: Kalkstein, A.J., & Sheridan, S.C. (2007). The social impacts of the heat-health watch/warning system in Phoenix, Arizona: Assessing the perceived risk and response of the public. International Journal of Biometeorology, 52, 43–55. doi 10.1007/s00484-006-0073-4
  - This study examines risk perception and warning response to NWS heat warnings in Phoenix, Arizona during the summer of 2005
  - The underlying reason for the study was that;
    - Throughout the summer of 2005, an excessively hot summer, people didn't seem to respond properly to heat warnings. A total of 18 people in the greater Phoenix area died.
    - Researchers felt that a better understanding might help reduce future fatalities
  - So the goal of the study was simple: To understand why residents of Phoenix, Arizona didn't respond as hoped to heat warnings for situations that were particularly dangerous.

Slide 15: Materials and Methods for this study

- This was a descriptive study which used a paper questionnaire distributed at 8 different shopping malls in and around greater Phoenix.
  - The intent was to attract a diverse range of participants with a range of ages and ethnicities that represented the Phoenix area.
  - But notice – it's only looking at mall shoppers who are willing to take the time to fill out a form. Not that that's bad, you only need to keep it in mind.
- The questionnaire consisted of 17 questions, some with multiple parts
- Researchers estimated that the survey could be completed in about 5 minutes, though I would suspect the time was more like 10 minutes, or so.
- In the end, there were a total of 201 participants

- **Here we display the file Heat Study.pdf**, and go questions at the end of article.
    - The first two questions were presented to establish whether the participants even knew that a warning had been issued. Actually 195/201 said that they had been aware. Those who were aware were then they asked some questions about these warnings and their reactions. Those who weren't were instructed to skip the next several questions.
    - In this survey, questions 3,4,5,6, 9, and 11 with all designed with 4 levels. They can legitimately be compared. For example, if you said you took the warning somewhat seriously or very seriously (Q3) with whether you worry about it (Q4), and whether you changed behaviors as a result (Q5,6) or whether you perceive the heat to be dangerous (Q12), your answers all could be compared directly. The comparison would be statistically strong.
    - Other questions in this study cannot be directly compared. Question 8 could have had 4 levels in order to compare false alarm impressions with taking the warning seriously, but it didn't. You can still make a comparison, but a statistician would find the correlation weaker, because the answers are nuanced differently.
    - Questions 13-17 were demographic in nature. Demographics are necessary. You need to understand who was sampled to give you an idea of whether the sample represents the actual target audience. Plus, there are always interesting opinion differences between different genders, age groups, and so forth that could help you understand reactions a little better.
    - Anyway, you can take a look at the survey and review the question set with some of the things we've talked about in mind. It has its strengths and its weaknesses, but all in all, it's a pretty good questionnaire.

Now return to Slide 16 in the PowerPoint presentation

Slide 16: Let's look quickly some results.

- First, as noted, it turned out that nearly everyone was aware of the warnings (195/201).
    - 63% of participants in this study were female. But notice that more females than males (as a percentage) reported an awareness of the existing warning system
    - Also, there seems to be a pretty good awareness across ethnicities
    - Overall, the study covered a wide range of ages. Notice that even though the youngest age group (18-29) accounts for the lowest percentage of those who were aware of official warnings, there were still 2/3 who were.

Slide 17: Here are some even more interesting results.

- Of the 169 participants who reported being aware of the heat warnings, and who actually completed the action questions in the survey, about half reported they did nothing about it. This graph shows that out of those who ignored the warning, the most common reason for doing so was that they felt they were used to heat, so they didn't feel like they needed extra precautions. So there was no perceived risk.

- This answer actually resonates with me. I've surveyed Front Range Colorado residents and found the same sort of response for winter storm warnings.

Slide 18: Suggestions by Kalkstein and Sheridan

- This study showed that a large percentage of residents who ignore heat warnings, do so because "it is always hot in the summer in Phoenix." The authors pointed out that if the warning is issued for a *particularly dangerous heat event*, the unusual nature needs to be *emphasized*. Make it special, but if and only if it is special.
- Provide clear instructions on what to do about it
    - Authors suggest that, since research has shown that people are more likely to act if they have clear instructions, residents should be urged in straightforward language to drink more fluids, seek air-conditioned locations, and participate in other mitigating actions during the warning period
- The authors also suggested that city and health officials need to become more involved in getting the public to respond by perhaps
    - Distributing water, or
    - Making sure that homeless shelters are air conditioned
- Handing out water bottles would be pretty expensive, so some sort of pilot study might be called for here.

Slide 19: The important things to remember from this course;

- Every study must have a clear purpose and a well-stated goal.
- The ideal survey will try to find a representative sample. It needs to be distributed so as to reach people of all ages, economic ranges, ethnicities and so forth that truly represent the population you are trying to understand. And if you're looking at someone else's survey results, did their survey accomplish that lofty goal.

Slide 20: Surveys must be crafted in a scientifically sound manner.

- That is, good surveys have question sets that can be statistically analyzed. Remember; numerical values in social science research represent levels of nuance, arbitrarily assigned. Keep Likert levels the same whenever possible.
- Questions must be clear, concise and inclusive.
- Always collect demographics such as gender, ethnicity, age, etc. That is also part of building a representative sample.
- And always watch carefully for those many biases and ambiguities.
- The example case is a pretty good model to look at if you decide to conduct a survey and should help you recognize a good survey when you see one.

END OF COURSE

– – – – – – – – – – – – –