Survey questions design

Hello I'm Debbie Collins and in this video we're going to look at some of the assumptions that underpin survey question design: the errors that can arise and how we can identify them.

Consider this question: how many hours did you work last week? Try to answer it now. How did you go about answering it? How did you define work? What did your last week include? Monday to Friday? Monday to Sunday or the seven days before today? How did you come up with your answer? Did you guess, give your contracted hours or try to calculate the actual hours you worked last week ? What if last week wasn't typical, for example because you were ill or on holiday? What did you do? Did you think about last week or report a typical week's working hours?

Can you think of particular types of people who might find this question difficult? Here are some groups I thought about: people looking at after the home and family would they think that this question is asking them about their work? What about people doing voluntary unpaid work? The self-employed might find this question difficult to answer: is it talking about all work that they do or only that that's paid? Will they even keep a record of how many hours they work? You may have other thoughts.

We need to be very clear about what information we want - the conceptual clarity of the question - and in what form we want it (the type of measure we need). Implicit in any form of measurement is that there is a standard procedure or tool which is used to obtain data. By applying the standard we can be certain that observed differences are real differences and not an artefact of differences in the way the data were obtained.

In survey research the standard procedure is a survey interview and the tool or instrument is the questionnaire. The measurements obtained are the respondents' answers to survey questions, as this quote from Fowler and Mangione illustrates on the slide. However this assumes that the wording of questions constitutes a complete and adequate script such that when respondents hear or read the question they will be fully prepared and able to answer it. However standardizing question wording may not ensure that questions produce valid, reliable, sensitive, and unbiased data as the hours work question we looked at illustrates. In fact survey researchers have known for many years that despite attempting to standardize questions and procedures errors still occur. These measurement errors can be classified in different ways. On the slide they have been grouped into two broad categories: those connected with question errors resulting from how the

question is phrased and the assumptions it contains; and administration errors, for example resulting from how the question is read, errors in recording an answer, for example respondents typing in the wrong key on the keypad. Oksenberg, Cannell and Kalton came up with a different error typology focusing on the nature of problems encountered by respondents and survey interviewers. They classified these errors in terms of: comprehension problems resulting from the use of vocabulary, sentence structure, or understanding of the nature of the task and the rules about how to respond; validity problems resulting from problems with respondents interpreting the same question in different ways or in the same way as each other but in a way that's different from that intended by the researcher; processing difficulties with respondents being unwilling or unable to retrieve the information required to answer the question; and pronunciation or communications difficulties.

There are other ways of thinking about measurement error. Another approach to thinking about measurement error is to think about the response process and what can go wrong. Perhaps the most often referred to response process model in social research is Roger Tourangeau's four-stage model. The model proposes four stages of cognitive processing involved in responding to a survey question. The stages are shown on the slide. Comprehension involves people thinking about what the question is asking: what do specific words and terms and phrases mean to the respondent when they're presented with them. The next stage is retrieval from memory: what types of information need to be recalled. For example in the hours worked example I needed to recall how many hours I'd worked in the last week, however that information may not be accessible so I may need to make a judgement about what to do when that information is not available and I have to think about how motivated I am to try to answer the question accurately and thoughtfully. If the information is not accessible then what estimation strategy will I use? Will I guess, will I think about my contracted hours or will I think about my contracted hours and then use an anchor and adjust strategy where I think about whether last week was similar or different.

Finally, I need to respond to the question. One of the factors that will influence my response is whether I feel the question is sensitive and I may not want to divulge the truthful answer that I have in my mind, I may want to modify it to appear in line with social norms. I then have to map my in mind answer to the answer options provided: does it fit or do I need to reformat my answer to fit into the answer options provided?

Some steps in this model may be conscious others may be unconscious and respondents may be go back and forth between the stages. You might want to pause this video now and reflect on how you could use this model to look at the response process to the hours worked question we discussed earlier.

Some methodologists have argued that the four stage model implies a set processing sequence in a linear fashion and ignores social, cultural and contextual factors that can influence response behaviour, such as who is funding the research or what's happening in the news. Alternative models have been proposed that explicitly allow for back-and-forth between response stages such as Willis, Royston and Bercini's Flexible Processing Model or for contextual factors such as Cannell, Miller and Oksenberg's Process Theory.

As survey question designers we can use these theories and models to help us when writing survey questions, to assess the potential sources of measurement error that might arise. If you want to find out more about these different response models then go to the resources that go with this video. This final slide illustrates the range of question testing methods and techniques available that we can use in developing survey questions. In the next video we will look at cognitive interviewing. For more information on other methods look at the resources pages.