

Research for Marketing Decisions

Vaibhav Chawla

Department of Management Studies

Indian Institute of Technology Madras

Week - 06

Lecture - 30

Questionnaire Design: Framing Common Demographic and Screener Questions

so where were we in the last class if you remember we were discussing the questionnaire design right so there were a few important things that we still did not discuss in the questionnaire design that we are going to discuss right away so let's start with basic questions like if you have to write if you have to ask for the income of what is the question you will let's say you want to ask for the income of the respondent personal income or income of the respondent how you would ask what you how you would frame the question similarly for age tell me in a survey you want to ask for the income and the age gender how you will ask this question simple question every MBA student would be required to do at some point of time and hopefully very very soon as soon as you get into a job doing something like this just floating a form right so how would you ask for the income okay let's say it is the survey about the skin care cream for women right you want to ask the income you want to ask the age sorry or let me reverse the table let me say that for the men for adult men the razor company The shaving razor company, they are conducting a survey for a particular problem and they want to ask their income.

How would you write the income for the adult men? What is the question that you will frame? Let's say adult working men, right? Let us not talk about students. We are, let's remove the students, adult working men.

We want to ask their income. What should I, how should I ask? Sorry? No, no, that is there. Just focus on this question.

Income, how you will ask? Let's say, I will write please tick the category that in which your income falls right one ways I gave the categories the other ways I can directly ask

income what is the correct way first way is incomplete unless I write the categories So one is there are problems here, right? Which I taught in the previous class also.

So when you say please take the category in which your personal annual income in rupees So there were so many problems. So which means this straight away asking is a problem. Now how we will make the categories. So for example one way is.

0 to. 3 lakh. three lakh one two six lakh six lakh one two nine or let me say six lakh one two fifteen lakh 15 lakh one two 25 lakh 25 let me say here one is this way of writing right other ways This is one way, another way of writing.

There is third way. There is another way. I am not saying these all are correct. You have to tell me what is correct. Another way to write will be

and so on what is the way you would want to use so who are the one who say first one is the best way raise your hands please one two no one only ok How many they say the second is the best way, this one? Please raise your hands. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. So, let us say 18 around.

How many say the third one is the best way? 1, 2, 3, 4, 5, 6, 7. How many say last one? Less than 3, above 3 but less than 6, above 6 but less than 9. 0 let's see the problems with each one of them now this one 0 to 3 lakh 3 lakh 1 to 6 lakh 6 lakh 1 to 15 15 to 25 the intervals between in each category they are not equal when the intervals are not equal

There is a problem. What is the problem? We cannot code them while data entry as 1, 2, 3, 4, 5 because this difference we cannot say one is more than other by one unit. Data entry coding which is coding of the answers becomes a problem because these intervals they are not constant. There is 3 lakh difference here.

3 lakh difference here. But here the difference is 10 lakh. 9 lakh. 9 lakh. Here the difference is 10 lakh.

Here the difference is 3. So since the interval difference is not consistent, we cannot code them as the interval scale. And there comes the problem. So one is that try to keep the intervals for if not for first and last answer but in between answers the interval should be same. Even if you can include first one that is good.

So last one you cannot keep the intervals same as that is fine. Even the first one sometimes we would not be able to keep the but the middle options you should be able to keep the this interval difference same for the all the middle options if you can do it for

first option that is also good enough because it becomes problem in coding now let's look at this one 0 to 3 lakh 3 lakh 1 to 6 lakh 6 lakh 1 to 9 lakh 9 lakh 1 to 12 lakh and about 12 lakh it is good right problem is people would be looking at the zeros whether it is lakhs crores what is that If it is send over mail where you are not there to tell them it is in lax. Sometime the problem could come but here the intervals are same.

This is one of the ways of doing it. When you think your population is educated that they would be able to know that this is in lax. And good part of it is intervals are consistent. Now if you look at this one. Less than 3 lakhs, between 3 to 6, between 6 to 9.

Now sometimes, now somebody who has an income of 3 lakhs or 6 lakhs or 9 lakhs, they could sometime mark here or there. Generally people should mark on the, generally people would say my income is higher, why should I mark on the lower category? But you don't know. So there also comes the problem. And if you look at the last one, less than 3 lakhs.

Above 3 lakhs but less than 6 lakhs. Above 6 lakhs but less than 9 lakhs. This is intervals are also same. Language is also correct. Very clear.

Right? This is one of the best ways of doing it. Provided your population would be able to, the sample drawn from the population should be able to understand, should be able to make sense of above 3 lakhs but less than 6 lakhs. Because it becomes little, little complex to read.

The options if they are, yeah. Sir, here whose income is 6 lakhs exactly, he has not any point of view. Equal and above 6 lakhs but less than 9 lakhs, yeah. So, so you have to see what is the so here above 3 lakhs less than 3 lakhs here equal to or above 3 lakhs equal to or above 3 lakhs here also equal to or above 6 lakhs that will become the best one but you see it will become very complicated to read so that is one problem

So, one can use this one, this one or this one, any of these three. Now, if you look at these answer categories, there are two characteristics. One is categories are mutually exclusive. Which means they don't overlap. We have to keep the categories in such a way that there is no overlap.

There is no overlap which means there is no confusion. A respondent can fall into any one of the categories but not more than one. That is mutually exclusive. So when you make the questionnaires, your answer should be mutually exclusive and exclusive.

Collectively exhaustive, which means beyond them all the answers should come under these answer options.

Anybody and everybody who is your respondent should be able to mark an answer among your categories. There cannot be any answer beyond them. That is called collectively exhaustive. There are two basic characteristics when we are asking for questions. Mutually exclusive, collectively exhaustive.

These two are basic things when you are making a question. In addition to whatever we have done in the previous class, remember that intervals in income you should keep same in between all the answer options and if possible for the first option as well. Write in the simple way that your respondents would be able to understand and mutually exclusive, collectively exhaustive. that's what if your population would be able to let's say there itself when you say please take the category in which your income in which your personal income falls in rupees right you may also write in bracket the numbers below are in lakhs so that zeroes If you are giving this particular one, if you think reading zeros will be a problem, you can say the numbers below, sorry, or that could also create a problem, but then they would feel that it is 0 to 3 lakh lakhs.

So... See, if your population is educated, you know, they would be able to figure out, then use this one because there is exactly no overlap. Even the end point of the category, there is no overlap. But generally, you will see something like this more times. Which, at the end point, there, you know, people generally would tick on the higher side.

Between 3 to 6, somebody who is a 6th They would not know whether they will mark in the second or third. So there's little ambiguity is there. There we would have to then if we give the this way of answering we will have to then accept those ambiguities. Otherwise this is the best ways and this one also.

Although the individual options are difficult to they're lengthy and difficult to read if you can make it easy. But the characteristics that you have to keep in mind is mutually exclusive collectively exhaustive. Now let's do for age. Now, age, let's say, skin care cream for adult women. That company is doing a survey.

Can you write a question for age and tell me? Like this. Adult women. These are basic questions. If you don't write them correctly, then there is no future.

The difference should be, you can't go to every category. Yes, many surveys would do that. They would follow that income categories. But the problem then becomes you cannot code them as 1, 2, 3, 4, 5. So,

first of all you would have to question whether the income brackets that have been made are they representative of people with similar lifestyle because your focus in marketing is about lifestyle attitude the way they live they spend their time they spend their money if you see that these income brackets somebody who is 10 lakh and above to 50 lakhs right do they have similar lifestyle Do they spend their time and money at similar kind of places? Generally, when you have such broad income categories, they may not be a correct representative of their lifestyle. So, better is to break it. Some companies do not break and they also do it.

But it depends upon whether you see with respect to your... not even with respect to your product category. If you see that the lifestyle of the people changes every 10 years of their income, every 5 lakh or 10 lakhs, then you better keep the categories of your own where the differences are 10 lakhs or 5 lakhs or something. If you see 20 lakhs difference, there is no change in lifestyle and so on, then you can keep it same. Age.

Somebody has to tell me. So what did you write? What is the statement? No, no, no. Tell me what is the statement you have written.

Please. Take the category. Of age. In which. You.

Belong. That is one way of writing. Please take the category of age in which you belong. In which units? In which units?

In years. Right. Now somebody who is 26.5. Whether they will write 27 or 26. Let us say in.

Completed years or years and months. Whatever you see is appropriate. So you have to be very very specific. There is no room for ambiguity. Or you can say your personal age.

Generally age nobody would actually you know. But if you see there will be people you can write your age. Your own age in completed years as on today. As on today, you can even write today's date, whatever. Then?

Institutions use this as on today or as on this date. Yeah, as on today you can write because some may say, okay, end of the year, shall I or whatever. Then you write the age categories. So, what is the first one, adult woman? Okay, equal to?

But below? 28. Okay. What is the next one? Equal to or above 28 but below? Ahh Sorry 30?

38. Equal to or above 38 but below 48. Equal to or above 48 but below 58. equal to or above 58 but below 68 and you can write 60 equal 68 and above this is one way to write which is correct anybody else who has written the this is all mutually exclusive and collectively exhaustive right somebody who is 150 years old also would be able to mark 68 and above right Somebody who just on that day turned 18 would be able to mark you.

So collectively exhaustive. Whoever belongs to your population would be able to mark it. And there are no two answers of age for the same person. So mutually exclusive, collectively exhaustive. So is it absolutely necessary to have equal data?

Wherever possible, yes. If not possible at all, then leave it. then you can still mark it as 1 2 3 4 but they will not be the interval scale they will become more of ordinal type of scale but if possible you can try it keeping the intervals equal but for this case age between less than 35 there can be more categories of people so it is more appropriate to yes so if you see that 18 to 35, they might have the similar kind of lifestyle needs related to skincare cream. Then you do it like that, 18 to 35.

And then you are saying if 35 to 45 instead of 35 to, so the categories are not equal, then what you will do? Then you can code still 1, 2, 3, 4, 5, treat them as ordinal scale. So there is a screener questionnaire. We will come to that screener questionnaire as well. One has to use the screener questionnaire.

So that very good question. Screener questionnaire, which means those who are not 18, they will be filtered out. So we will look at the screener questionnaire as well. So if let's say it is, as he said, 18 to 35, 35 to 45, 45 to 55, 55 to 70, unequal. unequal are there so here let's say 58 55 to 75 so then also you can code is a 1 3 1 2 3 4 5 whatever comes here but they will be treated as the ordinal scale no not really but what will happen is the number of statistical operations that you will be able to use on that data will be then limited

Can't we ask age directly? That's what. You have to be... In marketing, we look at the customers. We don't look at our convenience. If you do a girl 18 plus, you will ask what is your age?

What are you going to get? Will you get the answer? They will give you approximate age. Some might give them, you know, 40. They would be 40.

They will mark 18. What you will do? They can do in, but there is a bracket. 18 to, let's say, 18 to 35. So, you cannot predict 35, 18.

You cannot make a judgment. They are safe if they mark in the correct category. Likewise, income for men. It can be between 15 to 50 lakhs. Let's say 15 to 25.

It can be 25, it can be 15. I don't tell you. I mark the category. So that is one way to make people safe when they answer. If you are making people unsafe, they will not answer the right option.

So one of the ways, first of all, you yourself have to appear non-judgmental. if you are looking like a judgmental person then they the answers might be a problem so that is what in field training the staff that collect data if let's say face to face that kind of field training is required whereas in email survey you know such trainings are you require trainings in making this survey and floating it around but not the face face to face Very good question. Now width of the interval again the experts in that industry would have an idea that the needs changes every 5 years or 10 years of this particular you know target group. Then accordingly you would have to.

There is no right answer to it. It depends upon somebody who is an expert can provide you that, okay, we see that, let's say from the market data, you have data about the customers and what they are buying for the existing skincare creams. You would be able to figure out what particular group is buying, age group is buying similar or different creams. It is possible that age might not be a differentiator.

That is also possible. Another one is One is screener questionnaire. This I took from Professor Michael Hyman slides because very interesting screener questionnaires were given. So this is a screener questionnaire for soft drink.

They were doing a survey for soft drinks, soft drink consumption. So this is a screener questionnaire so that as he asked that if the Respondent, if the respondent is less than, in case of adult women, respondent is less than 18, how would you would find out? Very good question. Then we would use the screener questionnaire.

And in screener questionnaire, we will ask certain questions. First of all, we will ask, do you use skincare creams? Is your age above 18? Have you participated in similar type of surveys before? We don't want such people because people who have participated in some surveys, they might, you know, you don't want people who might, who have, you

know, good amount of idea how to fill up the surveys and who can, you know, make their inferences about what is required to be filled in the survey.

So you want people who have not much participated in surveys, although it's not a very essential requirement. But in your screener questionnaire, you use the questions to filter out the respondents who you don't want. Like in this case, there are questions like, do you know me at all? If you know the interviewer, then you will not be asked to. So the familiarity with the interviewer also is a problem.

if you think familiarity would create bias discarded are you the lady of the house yes or no if say in in between all these questions if even one of your screener criteria is not fulfilled then you can at that point of time either terminate the survey or what you can do is in this case there are if you look at this particular survey which is on screen when the service start there are identification questions for the type of risk identification question which is your identify what are the identification questions name a gender and all these personal questions then type of respondent there are three types of respondents here current users of the product aware but non trials and trials and rejecters so this is for a particular product this company is conducting a survey and let's say they want to understand the likes and dislikes of the current users they have this screener questionnaire here to find out the type of respondent they want to know for the current users what are the likes and dislikes so there would be some respondents who are aware but never tried you cannot have them as on giving answers to current to the survey for brand likes and dislikes of current users but you can actually give them a different question are about why they did not try so that you know why some people in the target group are not trying whereas for tryers and rejecters those who have tried and rejected again you cannot give the same questionnaire about likes and dislikes of the current users because they are not the current users for them you might use a different questionnaire about what they did not like why did they rejected

So there will be some questionnaire difference between these three surveys. So screener questionnaire help you screen out either filter out the respondent that you don't want or you have you make multiple surveys like in this case a company wanted to understand product a product company wanted to understand the likes and dislikes of the current users. So for the current users they would have this survey. But what about, let's say, who know the brand but have never tried? If you want the responses from them so that you can win them, you can give a different survey about why they did not try.

Whereas for the others who have tried but rejected, you can give a different survey about finding out why they rejected. Because that is also important to you. But this one is majorly important about the current brand current users brand likes and dislikes that is given that is okay. But also try to make different questioner for two other sets because that is important for a particular brand right. So screener questioners are also used.

Yeah so this is one of the ways of asking questions if you see the first question. The question is there. And then answer options and with answer options 1, 2, 3 these are for you to write the to enter the data. You will enter the data with these codes 1, 2, 3, 4, 5, 6 which represents a particular ascending or descending order. Generally you will keep as ascending order as the

Some extent of that particular characteristic increases. So in this survey, it is your wish whether you want to mention these codes or you don't want to mention. Generally, they mention. So you can mention it like this. So in this question, they haven't said that in a day, in a week or in a month.

I think that... what i'm trying to show you here is how just to represent the questionnaire how just to write this so if you see the respondents are going to be confused with these codes then don't have them but have them definitely in the likert scale strongly agree disagree strongly agree five let's say and strongly disagree five even there also if you want to not have them you need not have them but in your copy in your copy which you will use for your reference have the codes also there so that you remember what code to enter data entry otherwise will become a problem instructions is the once you are floating a survey the first page should have the information about Who is the agency who is conducting this survey? And what is this survey about?

And how much time it will take of the respondent? And the responses will be kept anonymous and only with the team. Nobody will be identified ever through this survey and their voluntary consent to fill the survey. Once do you get all that, then you move next. Then instructions are there.

If you have some instructions, you can add the instructions like this. But in the first page, the introduction, the cover letter, if it is there, put the cover letter saying who is the agency you represent or either you are an independent researcher not representing any company. or not belonging to any consultancy agency not sponsored by any other any company what is your purpose of doing research why you are asking a particular type of respondent then take their consent tell them anonymous tell them there are no right and

wrong answers and so on and then you can take the consent and then move to the instructions if there are instructions otherwise directly you can start the question now this was all about the questionnaire