# Statement of the Problem Transcript

## Extended Audio Descriptions List

[Survey lifecycle chart details](#_Survey_Lifecycle_chart)

[Formula details](#_Formula_details:)

[AUDIO LOGO]

**[AD]:** OPRE Methods Meeting

**[AD]:** Dr. Paula Daneri, Office of Planning, Research, and Evaluation.

**Paula Daneri:** Now, I'm pleased to present a session that will introduce what we mean by unit missingness. Why is unit missingness in survey research a challenge, exactly? Our presenters are Mr. Brad Edwards from Westat, and Dr. Emilia Peytcheva from RTI International. Following the statement of the problem session you're about to hear, we'll hear presentations on sampling and recruitment from Dr. Mike Brick, Dr. Courtney Kennedy, and Dr. Brady West. We'll then moderate a Q&A session with your questions.

So please feel free to enter your questions into the chat as you think of them. Brad and Emilia, I'll turn it over to you.

**Brad Edwards:** Thanks, Paula.

**[AD: Dr. Brad Edwards, Westat]**

I'm delighted to be here for this kickoff session of this conference. I'd like to just start by talking amusing for a minute about this word "missingness". It's not the kind of word you hear in everyday conversation, right?

A state of being missing. It doesn't mean zero, it doesn't mean not applicable, it doesn't even quite mean blank. It means something should be there isn't there, right? And this is a particular issue in our field in survey research, and one that is the touchstone of this whole conference.

So Dr. Supplee framed this well, reason why we're concerned about surveys.

OPRE does a lot of them. You do a lot of them. They're used a lot in evaluation research, for general studies of the population. But this is an interesting point to be talking about the problem of missingness in surveys. Because there's surveys have been going through something I would call an identity crisis. After a very long time of success, surveys, the survey field finds itself kind of questioning methods and looking for new methods ways to address problems that have grown to be pretty large.

And I describe a vicious cycle here starting with our population becoming more diverse but also more polarized. There's an increasing demand for information. There are a lot more surveys done today than were yesterday or 20 years ago. But along with that, there's this huge amount of disinformation that seems to be flooding all kinds of media. Trust, distrust in government, distrust in elites has been growing.

There's a lot of evidence of that over the last 30, 40 years. Most institutions have been in decline to almost disappearing from a scale of trust. And trust is kind of the bedrock of survey research of interviewing, of asking people questions. It sort of falls apart if the person who's answering the questions doesn't really trust the questioner, right?

It causes them not to participate at all if they don't trust. So distrust really is poison to the whole process. And then just in general, there's an increasing reluctance to participate in surveys. So this kind of is creating this vicious cycle that we're all searching for ways to get out of ... to solve.

**[AD]:** Outline. Meeting motivation, objectives. Setting the stage, focusing on your roles. The case for response rates, relevance to OPRE. History of surveys, history of me. Where are we now?

**Brad Edwards:** So this is just an outline of what we're going to be talking ... what I'm going to be talking about. And I want to take you through this in about 10 minutes, so without further ado...

So in this session, Emilia and I want to introduce the concept of missingness. Talk about what it means in terms of representation, talk about response rate, the basics of that metric and just in general survey error and quality in this context, in particular, focusing on relationships between response rates and bias and sort of get you ready for the more detailed sessions that come follow us today and tomorrow.

So for the rest of the day today, we're going to try to help you tap research methodologists' expertise on missingness, hear some of their concerns about the future, and give you some tools that might help you as you're talking with other members of your survey teams, with contractors you're dealing with, with clients, about response rate issues.

So I want to do kind of just a very quick history of surveys, as I understand it. The beginning, I think of as going back to the 1930s, so 90 years ago or so, when two things came together and made for a powerful foundation for what became the field of survey research, the theory of probability sampling of rigorous selection of a sample from a larger population, following in ways that give every member of that population a chance of being selected, a measurable chance of being selected, combined with uniform, standardized way of collecting the data, the protocol for collecting it, and the actual questions themselves being asked in the same way of every member in that sample. Those two things really sparked things in the 1930s. Gallup, other organizations became, you know, almost daily news stories, people sort of hanging with bated breath on what the latest polls showed. And that continued well through the '50s and spawned various, you know, subsets of survey research like multi-population research, multicultural and multinational research. The golden age, I think of as the 1960s through the '80s, maybe, when the survey world really mushroomed. Federal funding of surveys increased a lot,

There were major social programs that were launched in the '60s that required a lot of evaluation, very large scale survey work. Computing was starting to take off during that same period. Telephone ... random digit dialing emerged as a very popular mode in the '70s, and then I think of the middle age as following that in the '90s, the early 2000s.

That's kind of my story at this point. I would not claim to still be in middle age. But I think many people would say the survey world also has reached the end of its middle age where you see the issue of response rates of representation really coming to the fore, costs coming to the fore, and people sort of looking wistfully back maybe to that golden age when things seemed simpler.

A lot of this is a story of data collection modes. In the early days, mail was a very popular mode ... '30s, '40s, '50s. People really responded, the vast majority of them, to mail questionnaires. But as advertising caught on to that and started sending out ... flooding mailboxes with junk mail, response rates started plummeting for that mode. But I mentioned telephone coming into the fore. Everybody used to answer their telephone and had landlines so it was a very easy way to sample. And a lot of telephone surveys were getting response rates in the '80s. That started declining some 20, 25 years ago. People stopped answering their phones. Face-to-face interviewing generally was known as the gold standard, in large part because it was able to maintain pretty high response rates by knocking on people's doors and continuing to knock until people answered and invited them in.

Well, that doesn't happen anymore either. So I've been participating in conferences billed as you know, 'is face-to-face interviewing still the gold standard or is there any future to face-to-face interviewing?' because as you can imagine, it's the most expensive mode to collect data. So costs in general and declining participation are causing lots of people ... funding organizations to say, really, should we continue to be doing surveys if the costs keep climbing and the ... you know, you're including fewer and fewer people. Does it really make sense? We need other sources of getting data.

So response rates have been a very easy to understand quality metric and have stood in for how representative a survey has been. They don't allow you to assess whether the sample frame you're drawing, the frame you're drawing the sample from really covers the whole population. But once you get a sample, the response rate is a handy metric for indicating just in a general way how representative things are.

Equity is an important aspect of that. Looking at different groups, have you really gotten similar response rates from the groups that are of interest to you? Are you giving every member, every group, the maximum chance to participate in the survey? And this can be especially important, I think, in evaluation research, because you're often comparing different groups, and to make a valid comparison, you'd like to have roughly the same level of participation within each of those groups.

And furthermore, I think as opposed to the survey world in general, evaluations often are designed to make decisions, and to make an informed decision, you're really affecting people's lives more or less directly with the results of your research. So you really want an equitable ... equitable approach that gives, you know, every group that you're looking at a good chance of participating.

**[AD]:** Where are we now?

**Brad Edwards:** I talked about high costs and the cost pressure and competition in the field. If you're talking about contractors, I work for a contractor. There's a lot of pressure to bid as low as possible in order to win work, but that obviously creates a tension with quality. So, you know, when you talk about response rates, represent ... representativeness that has a lot to do with quality.

So how do you balance those things? One solution that a lot of surveys move to is multiple modes where you take advantage of the positive features of one mode, and then another mode. You can increase response rates that way. You can reduce costs, but you're adding complexity to the operation. It takes more sophistication, better systems to do... choose to operate in several modes at once.

Each kind has different requirements, different ways of assessing how well they're doing and improving their yields. So there are a lot of tradeoffs in designing a survey that uses multiple modes. Some are better for some things than others. Most of my experience is in face-to-face data collection, and I was used to seeing generally minority groups have higher response rates than the general population.

But I found myself in a meeting a few years ago with people whose main background was in web surveys, and it's just the opposite. Higher educated, non-minority people participate at higher rates on the web, so it makes a difference what mode you're using, which order you use them in. Typically, you might want to start with the least expensive mode and then move, for people who don't respond to that, to another mode that's a little more expensive and keep working things to the highest response rate you can get the best results for the amount of money that you have.

Nonprobability samples are becoming a lot more common, and studies that take advantage of ... of other sources of data than surveys, like administrative data sources are also, you know, coming to the fore. And there are a couple of sessions that focus on that later. So a lot of the importance of this is in thinking about options you have in designing to minimize nonresponse and making the best choices there, and that's a lot of what Emilia is going to be touching on next.

**[AD]:** Dr. Emilia Peytcheva, RTI International.

**Emilia Peytcheva:** Thank you, Brad. I want to start with a figure that shows the survey life cycle, and highlight only two things here. And this is coming from Groves and colleagues in the Survey Methodology book.

Just something that would help us distinguish what is unit nonresponse and what is item nonresponse because when we talk about missingness, we really need to define which one of those two missingness we are talking about.

So as Brad already mentioned, we would be focused on unit nonresponse. So we are staying on the representation side of this graph, which is the right-hand side, and we will define this in a second. But we will not be talking about the measurement side of the graph, which is item nonresponse. It is a type of measurement error despite its name.

[AD]: Survey lifecycle chart details available in transcript.

### Survey Lifecycle chart details:

Two process flows (measurement and representation) both result in a survey statistic. The measurement flow begins with the construct (mu sub i) which leads to measurement (cap Y sub i). Validity happens between these two steps. After measurement, there’s time for measurement error and leads to response (y sub i). Processing error is between response and edited response (y sub i p) and the final result is the survey statistic (y bar sub p r w). The representation flow begins with the target population (cap Y bar) and leads to sampling frame (cap Y bar sub cap C). Between these two steps is the coverage error. Sampling frame leads to the sample (y bar sub s) with the sampling error in between. Sample leads to respondents (y bar sub r) with nonresponse error in between. An adjustment error is between respondents and postsurvey adjustments (y bar sub r w) and the final result is the survey statistic (y bar p r w).

Source: Groves et al. (2004) Survey Methodology.

[Back to list.](#_top)

**Emilia Peytcheva:** So I will just go ahead and define each one of them, and then we'll proceed only focusing on unit nonresponse. So we define item nonresponse as a failure to obtain a response to a survey question. This can happen because the item is particularly sensitive for certain respondents or because respondents are very tired and certain items come toward the end of the survey, and they are already lacking motivation to continue.

We can also have skip logic errors in our survey. Fortunately, we have ways to reduce item nonresponse through various questionnaire design features. Skip instructions would be a good example of that or introducing probing protocols, whether this is in a web survey, when we introduce objects and we ask respondents to please answer this question in order to proceed to the next or whether these are instructions to interviewers. We typically correct item nonresponse through imputation, and we will learn more about that as we focus on unit nonresponse because that's also one of the methods that we use to correct unit nonresponse. So with all that, you can forget about item nonresponse and we'll just start talking about unit nonresponse only.

So, unit nonresponse is defined as failure to obtain any survey information from a sampling unit. So, imagine multiplied item nonresponse in a way, because every single question does not have a response. But again, we need to make the distinction that unit nonresponse is really staying on the representation side of our survey lifecycle while item nonresponse is still on the measurement side. We have various reasons why people don't participate in surveys, thus various components for unit nonresponse.

We can have unit nonresponse that is due to our inability to get in touch with a particular sampled person. We can get in touch with them, but they can refuse to participate, or they can be unable to give us an interview just because they don't speak English, or the language that we are offering, or maybe they have some mental disability.

So for most of these, we wouldn't even know whether the selected sample respondent is eligible to participate and not or not.

We have some ambiguity when it comes to what is unit missingness and item missingness that's why we needed to define them in advance. Sometimes when we have excessive item missingness, we need to make a decision whether we will define a particular case as unit missingness, which means that if a particular respondent answered five questions out of 55, we would need to make a decision to count them as a person who we ... who did not complete an interview.

And same applies for breakoffs, especially when they happen at the very beginning of a survey. Fortunately, again, we have certain ways to correct for unit nonresponse, not ideal, but we do this through weighting and imputation and we will learn about those methods throughout the duration of this meeting.

Brad already talked about response rates, so I would flip it and talk about nonresponse rates. We will define nonresponse rate as the proportion of estimated eligible sample members who do not complete the survey.

And one very important distinction that I want to make is that nonresponse rate is not nonresponse error. When we talk about nonresponse error, we typically talk about bias and variance, and usually we will define the mean sampling error as the squared bias plus the variance, which is the formula that we have here.

What is different is that we will talk about the different components during the meeting. We will talk about bias and variance. When we talk about nonresponse rate or response rate, it's something that does not quite give us the big picture, whether we have nonresponse error in our surveys.

It's very simple to compute and easy to interpret. It's an overall measure for the whole survey.

We have standardized computation of response rates that can be found on the AAPOR website…

**[AD]:** w-w-w dot a-a-p-o-r dot o-r-g

**Emilia Peytcheva:** …but the reason we're still using nonresponse or response rates instead of talking about biases and variances or mean squared error is mostly because we really don't have an easy way of always estimating bias, for example, and we don't have better alternative measures that would give us that ease of, "oh, we can talk about representativeness of our little sample" and so on.

Here's how we define nonresponse bias. We define nonresponse bias as the difference in the respondent mean and the full sample mean. And typically, we would estimate this as the difference between respondent and nonrespondent mean multiplied by the nonresponse rate.

So, two important components here. The nonresponse rate is definitely something that we need to pay attention to. So essentially saying if we have low response rates does it mean that we necessarily have bias.

As you can see from this formula, we do have to account for the difference between respondents and nonrespondents. So we may have a really low response rate, but if there is no difference between those who respond and those who we are missing from the survey, then our bias will be close to zero and vice versa, of course.

**[AD]:** Formula details available in transcript.

### Formula details:

Formula is written as Cap E open bracket y bar sub r minus y bar sub s close bracket equals cap E open bracket m sub s over n sub s open paren y bar sub r minus y bar sub m close paren close bracket.

Full sample mean is represented in the formula by y bar sub s. Respondent mean is y bar sub r. Nonrespondent mean is y bar sub m. Number of sample elements in full sample is n sub s and the number of nonrespondent sample elements is m sub s.

[Back to list.](#_top)

**Emilia Peytcheva:** This is a figure that essentially ... a graph that represents the relationship between nonresponse rate and nonresponse bias or lack of a relationship, if you look at the estimates. So this represents close to 1,000 respondent means plotted by the response rate for the particular survey from which they come.

If we are to fit a regression line, it would most likely be close to horizontal. So this essentially tells us that there is no particular difference between response rates and nonresponse bias, in this case, nonresponse rate.

Another interesting thing from this figure is to examine the variability of biases within a study. So we see that we have biases that vary within a study more than they vary between a study. And that's really important because essentially it tells us that we should not focus on one particular measure when it comes to improving the survey quality.

If we're after bias, we should probably think of various ways how to reduce biases across different survey estimates rather than how to increase response rates, for example. So we should ... our focus should not be at the overall survey level, it should be at the survey estimate level.

**[AD]:** Plot chart: Percentage absolute relative nonresponse bias of 959 respondent means by nonresponse rate of the 59 surveys in which they were estimated. Percentages are from 0 to 100. Nonresponse rates are from 0 to 80.

Source: Groves, R.M. and E. Peytcheva (2008). “The Impact of Nonresponse Rates on Nonresponse Bias: a Meta-Analysis.” Public opinion quarterly 72(2): 167-189.

**Emilia Peytcheva:** So just to summarize the lack of relationship between response rates and nonresponse bias. Nonresponse rate is really a weak predictor of nonresponse bias.

Something that came out of this study as well, is that the magnitude of bias also depended on how exactly nonresponse bias was estimated. So we have various ways of estimating nonresponse bias, and they turn out to be meaningful when it comes to the magnitude of bias.

I mentioned already that bias varies more across statistics within a study rather than between studies, which really suggests that for each study, we should focus on how we can minimize bias for that particular study, because bias is estimate-specific.

So, the lack of association between response rates and bias should not necessarily mean that we should not strive for high response rates. We should because it would mean that we have a smaller chance likely for inducing nonresponse bias in certain estimates.

But we also have to keep in mind that higher response rates do not necessarily mean that we don't have nonresponse bias.

And, in fact, there are studies that demonstrate that if you achieve higher response rate through a certain design feature, you may induce higher nonresponse bias, in particular estimates, if this design feature disproportionately appeals to a group that you're already overrepresenting.

**[AD]:** Source: Merkle and Edelman, 2009.

**Emilia Peytcheva:** So we do have design features to reduce nonresponse in advance, kind of to avert nonresponse. We offer incentive surveys. We know that manipulating survey sponsorship also gives us a chance to increase or decrease response rates. We have full control over the number and timing of our call attempts, how long data collection would be (sometimes influences response rates), our attractiveness of our contacting materials, interviewer training, modes of data collection (Brad already talked about those), how long the survey is in general, because it is the perceived burden that respondents have to decide to participate or not in the survey, and then who we select from our household, whether we want to go with everybody in the household or we just want to select one person.

We also have ways to compensate for unit nonresponse if it happens during data collection. So one and we will be discussing pretty much all of these different methods during this two-day meeting.

One of the approaches is double sampling for nonresponse when we subselect among our nonrespondents and go after them with a more expensive protocol.

Other active management tools during data collection are responsive and adaptive designs, which we will hear about. We have different weighting procedures and different models for adjustment.

And just wrapping up ... Dr. Supplee already mentioned the scope of the meeting, so we are going back to the very first point of our representativeness of the survey life cycle. We will start with sampling and then address unit response in terms of how we measure, how we reduce, and how we adjust for it, and then move to administrative data usage, and finish with reflections.

And thank you, and I hope you find these meetings over the next two days really stimulating.

[AUDIO LOGO]