Research Methods in Health Promotion
Dr. Arista Lahiri
Dr. B.C. Roy Multi-Speciality Medical Research Centre,
Indian Institute of Technology Kharagpur
Week 04
Lecture 20: Measurements in Health promotion

Hello. Welcome back and in this lecture we will be discussing on the measurements regarding health promotion. Now, in our discussion in our pursuit regarding research methods in health promotion, we have been discussing about the different research methods techniques and we will also be discussing regarding the different methods of intervention delivery development etcetera etcetera. But in between what is most important over here for any research method is how do we measure the components or how exactly do we I mean you measure the variables that you are interested in because otherwise without the measurements your interpretation of the results they they they will hardly be possible. So, measurements are basically important to understand specifically in the behavioral context how exactly are we acquiring the exact ah phenomena because ah we will be discussing later on also that the behavior all these aspects they are not directly tangible or you cannot directly weigh them or just measure the height it is not similar. So, in this context the the measurement particularly is very much important because this ensures that that whatever you have studied or the variables that you have ah you have elicited through your research they do carry certain meanings and the associations or the the relationship between the variables they have certain important inferences for you ok. So, in this lecture basically we will be covering the matrix of measurement exactly what are the the components that we measure how do we measure then we will cover a little bit about the scales and the index.

Next the variable measurement techniques. So, how exactly are we going to measure the variables and what are the different techniques that we already have in our hands. So, let us start our discussion the first point is the matrix of measurement also in some other text books you will find that the the same thing is written under the scales of measurement. Basically the matrix of measurement means how do we measure a particular variable.

For example, if you have age as a particular variable and and for example, you have asked a participant what is the age what is your age in this scenario what happens the participant ah may respond in 1 2 3 4 like this or ah you know the participant the age of the participant may be say 18 months that means, 1.5 years like this it it all depends. So, this is a continuous measurement and you see over here you have the numbers. So, similarly what other measurements you can make you can just simply go on ah in a room and simply calculate the number of people who who have ah you know who have ah a black jacket with them simple or you can simply calculate the number of males and females in a particular room. Now what

happens over here is they just give you a certain number a certain number of the male and female entity.

So, like this the the different measurements that we make it varies based on these matrix. For example, you know when we have a criteria as we mentioned number of males and females. So, here the criteria is the different category of gender. Here what happens we call it a nominal measurement why see what is written for nominal data based on categories that cannot be placed into an order ah based on preference or any other criteria. So, you cannot really order male and female for example, you cannot order the genders they are simply the names ah of the different sexual attributes that is the definition of gender.

So, male and female that is why we call them nominal it comes from noun ok. Next is we we were unable to order them in the nominal ah measurement, but in ordinal measurement what happens the data based on categories that can be placed into an order of preference. For example, if you have ah ah you know ah a Likert scale Likert scale may be like this say 1 to 5 and 1 corresponds to strongly disagree, then disagree, then neutral, then agree, then ah then strongly agree like this. See what happens over here you have an order it is a spectrum from strongly disagree to strongly agree. Also in other situations if you consider ah the visual analog scale for pain say for example, it rates the the pain response from 0 to 10.

Now from 0 to 10 means 0 is lower and 10 is higher although here the important part is 0 1 2 3 4 ah up to 10 it may not actually mean a particular value of of of 10. So, the basic essence over here is the difference between the two categories. For example, if you consider the pain scale now you have 0 1 2 and 3 I am not going into details just for 1 0 1 2 and 3 these these 4 ah these 4 responses you have ah gathered. Now the difference between the response ah participant who responded 3 in the visual analog scale for pain and ah the difference between 2 to 1 and the difference between 1 to 0 you cannot really say that the differences they are all the same you really cannot because it depends on the perception of the individual. Similarly for the the strongly agree strongly disagree neutral agree and strongly agree scale you cannot say that the difference between strongly agree and agree is similar to agree and neutral like this.

So, this is this is why this is called a simple ordinal scale. Here you have an order you you have an order of preference, but you cannot comment that the difference between the two two orders or the different between the two categories they are same ok. So, this is called the ordinal measurement. Next comes the interval measurement here obviously, you have ah the category you have the order of preference, but what you also have here is the difference. Say for example, you have measured ah in the visual analog scale they have standardized that ok.

So, the difference between 10 minus 9 is 1 and 9 minus 8 is again 1 and these two ones they are comparable in that scenario this becomes an interval scale where you can really compare these differences. The difference between 3 to 2 and the difference between 2 to 1 they they are now equal then you call it an interval scale. Data typically are continuous they are not based on categories typically this is the criteria for interval and ratio both of these scales you do not have categories like strongly agree disagree no it is not like that rather you have a continuous data from say one to 100 or like this and the separation between points on the continuum is always equal. So, that means, 3 to 1 the here from 1 to 3 the separation is at point 2. Now the difference between 3 minus 2 is 1 and 2 minus 1 is 1 and and this really makes sense only then this data will be called an interval data ok.

Now we come to perhaps the most robust and most commonly used scale ah in in in every research is the ratio scale. It is basically again a betterment over the interval scale because the interval scale it it may not have a true 0 value to it. That means, whenever you are you are calculating something even though this 0 in the visual and analog scale although the value over here is 0 and the difference between 1 to 0 and ah 1 2 to 1 they are all the same, but this 0 does not mean that there is absolute absence of pain right or absolute absence of any other attribute. So, that means, this is not an absolute 0. So, that is where the ratio scale comes in it is one step better than the interval data because they also have a true 0 point a feature not described by the interval data.

Now a typical example for ratio scale is considered the Kelvin scale for measurement of temperature because the true 0 in the Kelvin scale does really mean that it it is it is absolute absence of heat. So, so that is the basic idea behind using a ratio scale. Now for our everyday purpose what we can consider the typical example may be like age group you consider from 0 to say 50. Now it is again continuous now that 0 again the question is what should that 0 mean ah that 0 should mean that absolute absence of age well that really does not make any sense what is absolute absence of age. Rather what you can consider if you have to consider age as a as a ratio scale and it in fact, is is that 0 means it the the person has the age of the person is not countable at this moment that may be because the person has not born yet ok.

So, this is the basic idea behind using the ratio scale and we usually in our health promotion research we usually ah prefer to use the ordinal scale for different measurements and the interval and ratio scale they come in when we when we consider the cumulative measurements for say particular construct. So, we will be coming to all those issues ah later on in this lecture. As we were speaking that the matrix they they help us in measuring the constructs in our health promotion research. So, basically what are these constructs? As you can see what we have mentioned because the health promotion research in that point with the health behaviors and their underlying psychological and psychosocial influences research questions are frequently focused on intangibles such as self esteem, depression, self efficacy, attitude, perception and beliefs. That means, in our health promotion research what we have is certain intangible things

what you cannot directly go on and measure that is the basic problem because for height and weight you do know how to measure weight and you you know how to measure height.

You have the equipment for measurement of weight you also have the equipment for measurement of height that is what we call the tangible or directly tangible items. But here what happens is the constructs or typically the psychological or psychosocial factors the health behaviors that we want to quantify you cannot directly touch them or you cannot directly understand them and you obviously, do not have equipment like a weighing scale or a stadiometer to measure these kind of behaviors that is why we call them intangible because you can only feel them, but you cannot really touch them it is as simple as that. And for these issues like self esteem or depression psychological and psychiatric factor self efficacy attitudes these are the different constructs these are the different intangible items or sometimes what we call them as the concepts not constructs because at this moment we can only understand what what is a self esteem what is depression, but we really cannot measure because we have not described the measurement of all of these. So, when we cannot measure them we we consider them as the concepts. So, that is what we have mentioned the these intangible items are better viewed as concepts because we cannot directly understand them we have to devise certain other indirect measures to to measure them or to basically touch them because we we hardly have any other instrument physically or tangible instrument to measure these different psychological phenomena.

Now these concepts now that we understand say self esteem what do you mean by self esteem whether I am able to do any particular item or not able to perform a particular behavior or not we know that this is the the self esteem. Now in order to measure self esteem self esteem is now a concept because we understand what is self esteem now in order we have to measure this now in order to measure that self esteem we have to provide certain operational definition how exactly what should be the exact measurement for this. The operational definitions they may not be universal in some instances, but you you can really formalize or utilize them for your research that is why they are called the operational definition because you are operationalizing your concept into a measurable item or measurable variable. So, after operationalization these concepts they now become the constructs. So, once we have only understood what do we mean by this behavioral phenomena it is called a concept now we are able to measure these the typical behavioral phenomena it is now called a construct right.

So, the questions or items should be designed to to distinct effect indicators of the construct. The the questions that you have say for example, you take the example of a questionnaire this is a questionnaire you want to measure suppose you want to measure a self-efficacy that means, whether I will be able to do a particular thing or not you have 1 2 3 4 5 say 5 questions. Now this is how you operationalize the concept of self-efficacy into the construct of self-efficacy here the I questions or items that you have they should be designed. So, that they have distinct effect indicators on the constructs so that means, now these items they become an indicator

indicator of say the self-efficacy you can conceptualize a self-efficacy in terms of 5 different domains just for example, now these 5 questions each of these questions they will represent these individual domains. So, now these questions are considered as indicator for those distinct domains that are used to operationalize the concept into construct and they should be distinct.

Now the concept of distinct questions and whether they are able to measure the same construct or not these issues they all come into the purview of whether your question is valid or reliable and whether they are valid and reliable both. So, these parts are basically discussed later on when we discuss reliability and validity of of quantitative study tools. Now the effect indicators they signify that the each question or item relating to the concept infer some effect or influence on an observable behavior and taps into the construct. So, that means, the self efficacy you are not able to directly see rather you can observe the different domains that you have identified over here and these question they directly tap into these domains so that they ultimately have a bearing on the measurement of this of this overall construct. Now as our discussion progresses I hope you are able to understand till this point that a particular construct is usually made up with different domains or you know with different indicators.

So, multiple indicators that is what we have mentioned multiple indicators to assess one construct here your 5 questions that is 5 indicators are measuring the construct of self efficacy. It is very much rare to find one question or one one indicator to directly measure one construct usually we do not come across these kind of questions or these kind of phenomena in health behavior or health promotion research because you really cannot measure the psychological phenomena which you cannot directly directly observe through a single question you have to observe or you have to measure it from different angles to actually gather the idea. Now consider this this different domains of questions of measurement of construct in such a way that it is a 3 dimensional graph. In a 3 dimensional space if you have to identify this this particular point say this is the point P and it is located in a 3 dimensional space this is the x axis say this is the y and this is the z. Now, if you have to specify the location of P you must have the distances from all of these axis.

So, if you want to consider P as the construct over here this is your construct this is the concept that you have you have operationalized through measurements in x axis y axis and z axis what happens over here is you have 3 different measurements and you have 3 different ah points or ah or lengths that the construct now can be specified ok this is this point. Why I am telling you this because the scenario is same you really cannot observe this construct directly rather if you have this distance how far it is from this this that that you get from your different questions then you can have an idea ok. So, this construct can be located in this area or this is the measurement for this construct this is how and this is the basic reason why we have multiple questions to basically operationalize a single concept into a construct and measure that construct. Next our issue is discussion of scale and index remember indicator and index they are not the same. So, previously in the last ah slide we discussed regarding the indicators the

indicators they multiple indicators they build a particular construct and for these indicators you have typical questions for them ok and what happens with the with the index is we will come in the next slide is it is in fact, in itself a complex thing.

So, just remember that indicator and index they are different. Now what is the scale? Scale means you have a few questions maybe you have a measurement of these questions in built with them and finally, through that questionnaire you are particularly again observing a particular construct. So, let us see what we have mentioned over here the self efficacy represents a single distinct constructs we were discussing with the example of self efficacy and that can be measured by using multiple effect indicators. We also discussed that the different domains can be can be elicited through the different questions and these domains or these indicators can ultimately measure the self efficacy. Now the measure of self efficacy is called a scale a measure composed of theoretically correlated items measuring the same constructs.

Now if you remember in the last slide we mentioned that the distinct effect you you need to have right now the distinct effect yet say these questions these were the questions right these were the questions. Now these questions will have a distinct effect on a measurement of self efficacy or they will distinctly define the different aspects of self efficacy, but they will also be correlated among themselves. That means, for example, if your question regarding self efficacy is whether you are able to perform this behavior is one question a particular behavior of of a whole construct. Then if you have another question like this what is your age then conceptually although you cannot you can really correlate them in a statistical way, but for measurement of self efficacy these two questions they are not correlated and they are also not having a distinct in interference on the measurement of self efficacy. So, that will not form a scale, but if the measurement of self efficacy is say it is composed of measurement of five behaviors whether the individual or the respondent is able to able to perform all the five behaviors or not.

Now you are asking the participant regarding whether you you you feel that you are you can perform regular physical activity next whether you feel that you can perform eating a healthy diet like this. Now physical activity diet they can be considered as correlated items. Now that becomes a scale you have a correlated items and you also have the items you can really even score them and after after that score you can get a total. Although you know the total the cutoffs and the interpretation all of those things you need to validate you need to find out its reliability, but those are the things reliability and validity will be discussing in a later part of later lecture. Here the example of a scale is given as geriatric depression scale.

For example, if you have the GDS short form the the 15 questions it has they are correlated with each other. In fact, we did a study on rural population regarding the validity and reliability

of GDS Bengali version among the participants and we found out that all of those questions were highly correlated with each other and they they measure the construct in a very good way. What happens is they have they they are the yes no responses no means 0 yes means once like that. So, you have a score and finally, you have a total score. Now the whole thing is considered a scale because you are able to quantify the depression that the elderly people they are having.

So, the idea of having a scale is again to quantify not only to understand whether this is present or not also to quantify ok. So, this is what we have discussed in this point operational definition of the construct or the concept that means, it is now operationalized into a construct then we have created questions or items that measure the newly defined construct the different domains we have identified for measurement of self efficacy, we identify the different questions and test each test the items as an entire scale. That means, for example, over here you have 15 items you consider them as a scale and you just simply test whether they are measuring depression or not and whether the items ah fulfill the criteria of being in a scale or not. Now since we now have understood this concept of scale that means, we are now able to understand how do we measure a single construct quantify a single construct this brings us to the next question of index where we will discuss what if we have multiple constructs in our hand. Now an index it refers to a measure in which the items are considered causal indicators because they themselves define the construct and influence the value of the construct ok.

So, here the index means it has certain components inbuilt into it which finally, is causally related to the the outcome. Consider this example of quality of life the WHO quality of life index it has different domains in it. Now these different domains you can consider them as different constructs of measurement of ah quality of life. What happens is for example, the the the social ah engagement component if you consider it has a causal relation with the quality of life because see if a social engagement of a participant is improved then the quality of life will also be improved. So, it has a causal relationship and these are the building blocks for the index of quality of life.

That means, whenever you have the constructs that are causally influencing the outcome and you you ah take all the constructs together it becomes an index. The items in an index typically are heterogeneous and may not necessarily be correlated with each other. Because in the in the quality of life example see you have 6 domains. Now the questions in those 6 domains they may be correlated with each other, but if you take one question from one domain and other question from the other domain you may find that they are ah not coherent enough or they are not correlated enough. So, this brings in a substantial amount of heterogeneity and ah in statistical terms we may consider ah the different dimensions for ah validating this questionnaires ok.

Again the the the discussion on validity and reliability we will have at a later point in this course. So, I hope the difference between a scale and index is clear. Just to put it simply you may have ah a single scale that can identify a single construct, but for an index you you may have multiple scales as well for example, for measurement of multiple constructs. Now this brings us to the next topic of discussion how exactly do we measure the different behaviors or the different constructs in our health promotion. We now know that we have scales and index, but how exactly do we implement these ah these tools and ultimately exactly measure the components.

Now first is the self-report version. This is one of the most used technique of measurement of the behaviors in health promotion research although it has certain major issues. For example, self-report means you are asking the participant and you are just taking the the response from the participant and noting it down and you are ah taking the response in its face value, you are not cross checking it or you are neither you are validating the response. Now what happens is it is usually clouded by certain biases if there are certain ah systematic deviations in in the measurement because if you ask a participant regarding his or her ah beliefs regarding smoking you may find there are certain social desirability bias that means, since the since the society is against smoking the participant will say no no I am I can I really can quit and or I have quit smoking like this although that may not be the exact scenario. So, the response bias in the scenario will be considered social desirability bias because this is based or influenced by the social desirability to certain ah behaviors. There can be inaccurate recall bias participant or repeatedly occurring even for example, an individual is having ah consider the example of a child who is having recurrent diarrhea.

Now you are you are studying certain ah hygiene practices or hygiene behaviors among the parents ah of the children in a particular area and in that case if you ah if you take the number of diarrheal episodes of the child or when the diarrheal episodes occurred for a particular child as one of the variables it may suffer from inaccurate recall bias because that is a recurrent phenomena for for a particular family and that particular family or the mother of the child may inaccurate recall say the the the date of the first episode or the date of a second episode. Now what I am trying to make you understand is in order to avoid inaccurate recall bias whenever you are doing some measurement by self report you put certain prominent ah behaviors. So, that the people are able to recall them. For example, age at menarche if you ask any any female regarding age at menarche the the respondent will obviously, be able to tell you the particular age, but if you consider ah when did you have ah say the the diarrheal episode say the fourth diarrheal episode of this child in this in this diarrhea example the mother may ah may not be able to recall, but if you ask when did the diarrheal episode start that means, you are asking the first diarrheal episode. Obviously the mother of the parents they will be able to tell you oh no that was the date when the episode started like this.

So, you have to mark certain prominent things then there may be certain selective recall bias. For example, if a particular behavior is very much of importance to a particular individual that individual when asked about those behaviors of many other behavior will selectively recall those important behaviors only. So, that is the selective recall bias you will obviously, have these kind of issues with the self report. Now this is another table that I have given to you basically this is a pyramid which shows as you go up the accuracy level of self report ah is is higher. At the baseline is the infrequent behavior and low salience that is low importance and frequent behavior with low importance.

So, that means, if the particular behavior is of low importance then I mean irrespective of whether that behavior is frequent or infrequent the response accuracy will not be very much high. One step further is frequent behavior with high salience for example, the episode diarrheal illness in a child that is a very much important behavior important issue for the parents and that is also frequent. So, although the parent is not able to accurately recall the episode of say third or fourth episode, but can you know simply state when the episode started because again that is of high importance. Highest is infrequent behavior and high salience. So, here the accuracy level is very high very much important to the individuals and also that particular behavior or the practice is not that much frequent among them.

Next we consider the direct observations here simply you just go to the community ah or you go to the people and directly observe what they are doing. We may have certain checklist you have a list of what are the points you need to observe you just stick or cross stick or cross like this. So, these are direct observation, but you as a researcher is observing among the participants. Usually the direct observations are done for you know the practice related items. Often there can be record reviews certain variables like the income inequality, social capital and employment rates all these these are often available through the public records.

So, you can review these records and get data from by reviewing the record what happens with the record review is it gives the whole research a particular context. So, it is important to understand the contextual nature of the inquiry. If you have certain economic bearing to to your research then it is obviously, useful if you can understand the employment rates or the average spending patterns. Now, for these you may have the reports available from where you can gather the data that is called the record review you have a record and you are reviewing the record. And the best method of measurement is considered as the biological assessment.

Usually biological assessment is done for the outcome variables because obviously, as it goes with the with the best method it is very much tedious to do and is also very much costly in certain instances. Usually the different markers are measured through the biological assessment, measurement of actual biomarkers they indicate certain things. For example, you may consider ah measurement of stress through a scale, but the measurement of level of

cortisol it it provides a better indication of measurement of ah measurement of stress. Typically whenever you are studying stress as a particular outcome variable it is obviously, better if you have the resources to measure the level of cortisol. Also ah the measurement of ah the the different assessment techniques basically they differ based on the availability of resources and also based on the ah based on the research question the rigor of the methods proposed like this.

Lastly we come to the discussion of mediating and moderating variables. Now that we understand how we measure the different variables, the mediating and moderating variables they give you an idea of which concept or which construct or any other external variable is basically influencing the relationship between the predictor variable and the outcome. Say this is a particular behavior which is related to a practice of ah certain things. Now this is your construct and ah this is you have measured through scales and this is you have observed the practice. It may so happen you may have one variable v 1 over here which is directly influencing the relationship.

Now if it is so that because of the presence of the variable v 1 which is again ah which is basically changing this relationship, we call it the moderating variable loosely I am telling you because it is moderating or changing the nature of the relationship. It may so happen that you have a variable v 2 over here where it is basically mediating the effect b leads to v 2 and this v 2 ultimately leads to the practice. So, these are all the issues that you need to consider whenever you are proposing your measurement ah technique for your health promotion research and ideally all of these issues which variables ah how you can measure the variables, how you are planning to operationalize the thing, the measurement of the of the different constructs, the scales, the different mediating or moderating factor, even the confounding factors that is a very common thing you have to specify at the protocol stage. So that it is clear ah on how you are going to collect the data and how you are going to analyze the data ok. That is what we have discussed we have discussed the different matrix of measurement the nominal ordinal interval and ratio measurement.

We also stated that it is hardly possible to take a single question to measure a particular construct. We noted that a scale can measure a construct and an index it usually is a complex thing where you may have different constructs that are causally related to the final outcome measurement. And lastly we also discussed about the different assessment techniques where we also stated that the self report is the commonest assessment technique for our health promotion research, but it also has certain problems on its own. So, these are your resources you can go through them. that is it for this lecture. Thank you.