Research Methods in Health Promotion
Dr. Arista Lahiri
Dr. B.C. Roy Multi-Speciality Medical Research Centre,
Indian Institute of Technology Kharagpur
Week 07
Lecture 34: Validity and reliability of study tools in quantitative research (Part II)

So, in the last lecture we started our discussion regarding reliability and validity of your data collection tool. Now, we discussed a bit about the validity we said that we will be discussing four different types of validity measurement and amongst that list we discussed about the phase validity and also the content validity. Now, in this lecture we will continue our discussion on validity of the of the questionnaire or the data collection tool and we will focus on the construct and criteria validity because these are the two remaining validity measurements from our last lecture we will focus on these two validity measurement and also we will give you an idea of the approach on how to establish the psychometric validity because the psychometric validity as we discussed in the last lecture also it it entails the whole spectrum of reliability and validity whether the questionnaire is psychologically comprehended in the same way and whether it measures the same construct or not these concepts they constitute the whole spectrum of psychometric validity. So, we will give you an idea of how we approach the question of psychometric validity. Now to start our discussion we shall first focus on the construct validity question. So, the construct validity construct means we have operationalized a concept into a construct for its measurement.

Now, we want to understand whether that operationalize is actually valid or not through the concept of construct validity. See what we have mentioned over here the construct validity refers to how well you translated or transformed a concept idea or behavior that is a construct into a functioning operating reality that means, we are studying the operationalization. Now to study the operationalization as we discussed in the in the slide with content validity in the last lecture that we have a content of self efficacy beliefs. Now self efficacy beliefs they are again a construct because you cannot directly touch them or tangible you can measure them through different items that you have devised and you have a significant overlap with the perceived behavioral control of certain items.

Now operationalize means how well the items that you have framed in your questionnaire they are able to measure only the self efficacy and how well they are able to measure only the PVC, but also how well they are able to differentiate between self efficacy and PVC. See otherwise because there is a certain overlap if the questionnaires or the measurements that you have considered if they are not able to distinguish between the constructs then the measurements or the findings that you get they are basically a merged finding. So that means, the measurement of self efficacy and measurement of PVC both are basically clouded with the with the in between intervening part that means, the measurements of self efficacy in that scenario will be heavily influenced by the measurement of PVC and vice versa. So, that is what construct validity does it tries to understand and it tries to measure how well the construct has been operationalized. So, how does it differ from content validity because see in content

validity we wanted to ensure that all the items that are necessary to define this construct or to operationalize this construct is present.

We there in content validity we came up with the content and based on the expert review we understood that which content is essential for this kind of construct and which contents and which questions or items as the content of the questionnaire are not required. Now, but in construct validity we are actually testing the operationalization the items that we selected in the content validity here we are actually testing these items whether they are actually measuring self efficacy in this example or not ok. We shall discuss the construct validity in terms of convergent and discriminant validity and the discriminant validity is also called the divergent validity. So, sorry so, let us start our discussion with discriminant validity. Now if you remember we started with this construct C since we are not able to directly understand this construct or directly measure this construct say certain beliefs you cannot directly measure you only can feel it, but to directly measure it we develop certain questions like this.

Say three questions it operationalizes this construct that means, it measures this construct what happens here you are directly measuring these questions, but not the construction that is what we loosely call a latent variable also you know if we if we get a chance in some other lecture or in some other course particularly we the concept of latent variable is a huge concept which basically deals with all the variables which we cannot directly measure, but can be measured through these different operationalizations. So, here the concept that we are measuring we will be referring them as the latent variables. So, what is discriminant validity? Discriminant validity is the extent to which the latent variable discriminates see A is the latent variable over here we mentioned it in terms of C now consider this as A the latent variable A it discriminates from other latent variables for example, say B or C or D now typical example you can consider these self efficacy beliefs and the P B C if you consider these two self efficacy belief is one construct is is one latent variable in this case P B C is also another latent variable is in in this case. So, what discriminant validity does is it basically operationalizes on how the items can distinguish between these latent constructs between these two latent constructs ok. So, that is the idea behind discriminant validity it diverges ok.

Now next follow the second bullet the discriminant validity means that a latent variable is able to account for more variance in the observed variables associated with it associated with it means associated with this construct only then the measurement error or similar external unmeasured influence or the other constructs within the same conceptual framework. Now consider this figure that we discussed in the last slide see here this part is overlapping with P B C ok. Now the objective of discriminant validity is to understand whether we are able to discriminate this portion or not to put it simply the variance whenever we are measuring discriminant validity we should reduce the items from this part to reduce the variance that is attributed to this overlapping part because if the variance that is explained by this overlapping part is more that means, the construct that you haveutilized in developing the that means, the items that you have basically utilized in developing this construct they are not typically measuring self efficacy rather they are measuring somewhat in between. But in this scenario

you have developed these three questions which are solely measuring the self efficacy what happens in this scenario is the self efficacy this particular variable it is able to explain the variances or the variability in the responses of these items. So, the variance is basically the statistical term to understand the variability in responses of all the items.

To put this concept simply you will have certain external influences or certain measurement errors for each of the items that is a general rule in in any form of field based research your objective should be to minimize these variances, but still you will get certain noise. That means, the variability that you get in these items some of it will be attributed to these these errors, but for discriminant validity it should be like that the variability that you get in these items they are mostly and in majority explained by self efficacy only ok, but not the PBC. That means, although you have certain items that can relate to self efficacy and PBC both see what happens in those items is both self efficacy and PBC are contributing in explaining those items, but the items that are exclusive to self efficacy see the variability in response to those items are not really explained by PBC ok. So, that is what it means by most variance is associated with the with the construct that we are measuring and not with the construct within the same conceptual framework say within the same conceptual framework of the planned behaviors we may we may have the self efficacy belief we may have also the PBC, but the items in the questionnaire some items that are directed to measure only the self efficacy they should only measure self efficacy and not the PBC. That means, the variability in self efficacy is sufficient to explain the variability in the items or the levels of self efficacy among the different individuals is sufficient to explain the variation that you get in responses of these questions among between among the different individuals, but still PBC although it is within the same conceptual framework is not able to explain the variability in the items.

So, that is the idea behind discriminant to divergent validity I am repeating the concept of the of differentiating between the two ah between these two you know constructs. So, that you understand the divergent validity means it diverges exclusive to a particular construct ok. See it tests that the constructs that should have no relationship do in fact, not have any relationship. Although PBC and AC measures certain common ground, but they are basically not related they are mutually exclusive constructs ok. So, your measurement in this scenario should be such that that you are able to capture that ok these two constructs are different and they are not related that is why it is always essential to you know include questions from these segments which typically measure the constructs solely.

If you include questions from these overlapping segments then what will happen will not be able to prove the divergence between the constructs that the constructs are mutually exclusive why because then both the constructs simultaneously will explain the variable and that is not the basic concept of ah discriminant validity. So, you should be very much cautious on how you choose your questions from the content universe. Now, in discriminant validity we try to discriminate between the different construct we identified items that will discriminate the different constructs we understood that a particular construct will ultimately ah explain the variability in certain items only, but not the other construct which is within the same conceptual

framework so that part we understood. Now, the convergent validity is somewhat opposite of the divergent validity. What it actually says is ah I mean it refers to the degree to which two measures of construct that theoretically should be related are in fact, related ok.

See here we try to distinguish between the self efficacy and the perceived behavioral control and now here we try to relate that ok self efficacy this is self efficacy the latent variable and this is p b c. Now, conceptually we understand that they should be related somewhat although they are mutually exclusive, but they should be related in some way. Convergent validity basically captures that part which ensures that ok from this questionnaire this relationship you can infer that means, there are certain questions which the self efficacy explains also which the p b c explains. Now, through these questions you can identify the relationship between these latent variables self efficacy and the perceived behavioral control ok. So, that is the basic idea behind convergent validity divergent you differentiate the two variables convergent you just understand the relationship between the the hidden constructs or the latent constructs that you are studying ok.

The convergent validity tests the constructs that are expected to be related are in fact, related divergent it studied that are not that should not be related are actually not related and convergent it studies that should be related are actually not related. For example, you can consider that that simply the knowledge and the practice they can be related, but in I mean in actual scenario there are certain steps between transformation of knowledge into practice. So, you can in a way consider that knowledge and practice they should be mutually exclusive and in that scenario you can test for the divergent validity whether the questions that you have they typically focus on knowledge and other set of questions typically in practice or you have questions that basically you know jumbles up knowledge and practice together. But see in self efficacy and p b c you may consider typically certain what we call a certain correlation amongst these latent constructs. So, here you can utilize the convergent validity to understand what is the relationship between them ok.

This will be you know clear when we discuss this example, but before we discuss this example we have to understand how basically we understand or measure these divergent and convergent validity. The measurement technique that we use is called a factor analysis. Factor analysis means here you have the items different items present over here from these items since you do not know how to measure the constructs, but you understand that these items will be able to capture the construct. You have the latent variable L 1 say and the latent variable L 2 these are your different constructs and you have your items. You try to fit these hidden variables and explain the variability in the response of all these items that you directly measure.

Now these are called the factors that explain the variability in the items ok. How do we do that? When we try to find out the score or say in some cases we find out the z value that is also a score to this latent variable in one way we do it by factor analysis. These are the principal components that we want to we want to understand or we want to elicit. So, that is why we

use the principal component technique when we are actually performing the factor analysis we want to explore or we want to identify this from these items ok. See then we are utilizing the principal component technique to understand these different factors or latent variables that we have also we may have to use certain rotations.

These are more a statistical issues that you will face whenever you are going to run the factor analysis in any of the statistical softwares, but the idea is you should be very much clear on which factor to consider and which factor to not consider. Because see after you have run your factor analysis and you have and the software is showing that you have identified say two factors L 1 and L 2 and all the fact all these variables all these items that you studied they are mostly related to both these factors in certain way. Now how to identify which items to keep and which items to consider against the identified factor. Consider this cutoff for example, items loaded above 0.40 where typically you get an get a table like this say this is question 1 and this is question 2 and this is question 3 typically what you get typically what you get a table like this if you have explored two factors this is factor 1 and this is factor 2 and here you get in each of these cells here you get the loadings of these variables in each of these factors.

Now we will not be discussing the concept of loading and correlation, but please understand we devise the factors in the factor analysis or we decide on which factors to keep based on the loading value. So, you have your values mentioned over here this is a typical output where the loading minimum loading of 0.4 is reached you should consider that item for further research. That means, say for example, in in factor 1 the Q 1 is having a loading of 0.5, but in factor 2 the Q question 1 is having a loading of 0.

1. That means, you know you can be you can omit these questions in factor 2 no you do not consider the question 1 to be related to factor 2 so that means, question 1 directly answers to factor 1, but in some situations if this happens like question 3 is having a loading of say 0.5 in factor 1 and also again a loading of 0.5 in factor 2 then how what do you do do you infer that question 3 is is is a very good question that basically relates to factor 1 and factor 2 both and so you can keep it no this is called a cross loading. So, if you have a cross loading above 0.4 see if you have a loading of point more than 0.

4 you keep that question and you assign that question to that particular factor, but if you have a cross loading of more than 0.4 then you simply delete the question because this question belongs again to this common area because it is related to the self efficacy as well it is related to PVC as well. So, I hope the concepts are getting clearer now. Now consider this example that basically makes you understand the whole idea behind the the divergent and convergent validity. See the factor analysis results will satisfy the criteria of construct validity including both the discriminant validity and the convergent validity.

How do we explain the discriminant validity from the results of factor analysis? The loading of at least 0.4 and no cross loading of items above point 0.4. So, we are omitting this item that means, the remaining items see question 2 also is having a 0.5 loading in factor 2, but say 0.

1 loading in factor 1. So, factor 2 question 2 is not considered in factor 1 rather it is considered in factor 2. That means, these questions are distinct to distinctly they are able to identify factor 1 and factor 2. In this example question 1 typically identifies factor 1 and question 2 typically identifies factor 2 that means, this is sufficiently discriminated and how do we measure how do we understand the convergent validity? We have another term called the eigenvalues eigenvalues and at least a loading of 0.4 that load on the posited constructs. So, eigenvalues what is the concept of eigenvalue? You have a plot called a scree plot and in that scree plot we consider a cut off of 1.

What will happen? Here in the x axis you have factor 1 and factor 2 like this factor 3 is for example, if you have more than 1 factors and you have a graph where these kind of the loadings in each factors you have you get. Based on this graph what you do is you you do typically a cut off of 1 or 1.5 based on your research you decide how many factors you keep ok. Say the these 2 factors 1 and 2 they were above the cut off level, but the factor 3 was below the cut off level. Now you have decided on these factors factor 1 and factor 2 that is why you have decided to see this table of factor 1 and factor 2 only.

Here in this table you have only the factor loadings you will also have another table where instead of this loading over here you will be you will have the eigenvalues. Now the eigenvalues if the eigenvalues of more than 1 you get in in a particular for a particular factor you consider that this particular question also the loading should be at least more than 0.4. Now that means, through this scree plot and the eigenvalue table what you get is the convergent validity whether the items see these 2 factors from the scree plot what you get is the from these 2 factors they are related and that is why they are above the cut off level and you can then extract them. That means, your questionnaire is actually able to identify these 2 related factors which should have some relationship on their own and that is why they are consistent with the ultimate explanation of the construct from your study ok.

And and this particular table where you omit the question number 3 for having a cross loading of more than 0.4 helps you to discriminate between the factors you correlate between the factors you consider the eigenvalues to correlate between the factors and then again you discriminate them to have a crisp measurement of both the factors in a distinct way ok. I hope the example has made the things on how to you know measure the convergent and divergent validity clear also you can take these as the guidelines whenever you are going to test your own questionnaire. Next we come to the topic of criteria validity. So, what is a criteria validity? It is also called a concrete validity it is the extent to which I mean to which a measure is related to an outcome.

Here we are focusing now on the outcome see in construct validity we are focusing on the construct whether whether the items they are actually able to measure the construct or not and whether the constructs can be differentiated between them and whether the items they are able to correlate between the construct like this. Here in criteria validity we are focusing on the outcome, outcome means the final health related outcome that we say we want to bring change ok. So, that is the that is the difference when we are discussing the criteria validity. See we have mentioned it measures how well one measure predicts an outcome for another measure. It is not only now about the question items in the questionnaire only here we are predicting the outcome itself the outcome that we want to maybe we want to change or maybe we want to study here we are doing that through the criteria validity.

What we say is a tool has this type of validity if it is useful useful for predicting performance or behavior in another situation. Now when the situation will occur another situation or circumstances it may be in the past present or future. Based on that the criteria validity can be of three types. See concurrent validity that means, the performance of the behavior or that particular criteria that is present in the current time that is the concurrent validity. Predictive validity means from your questionnaire you are basically predicting an outcome if the questionnaire is able to predict properly the outcome that is called a predictive validity.

And post predictive validity means the criteria was set or the criteria is already there in the past. And now what you are doing is you are basically measuring the the outcome you have measured the outcome and now you what you are doing is you are again taking the outcome and comparing it with your test with with your questionnaire perhaps. So, the concept of predictive validity is quite simple it refers to when scores from the predictor measure are taken first so that means, your questionnaire you have taken the questionnaire first and then the criteria data is collected later. So, the criteria data means here the criteria data they are basically the the assessment that you that you want to predict for example, actual performance of a behavior if you want to predict that and now you are measuring the intention. So, predictive validity you can do this by comparing the intention with the actual performance of the behavior whether you are able to predict the outcome in a proper way or not.

So, that is what the predictive validity says it assesses the construct operationalizations the operationalizations ability to predict something it should theoretically be able to. See the example of intention should theoretically be able to predict practice typically in our health promotion research the intention is not directly observed or not directly measurable that is why what we do we formulate questionnaires and we try to elicit the intention. Now what predictive validity says is whatever questionnaire we have formed that questionnaire should also be able to then then identify or predict the occurrence of practice because ultimately the core construct it leads to practice. So, theoretically the questionnaire should again be able to predict practice. If the questionnaire the data that you get and later on when you observe the practice if they corroborate that means, your questionnaire has a predictive validity that means, you through your questionnaire you can properly predict actually what what action the participants are

going to do what practice will be there and based on that you really can propose certain you know decision measures you can propose certain activities that ok.

If the participants are responding in this way to the questionnaire that means, they will ultimately perform this behavior. So, better the health system should take certain actions to rectify those behaviors like this. So, it is not possible if your questionnaire is not able to predict the behavior that means, without predictive validity you really cannot propose certain decisional or forecast that you may get from your data. Concurrent validity means now it is not predicting some some behavior in future rather it is the operationalizations ability to distinguish between groups that should theoretically be able to distinguish between that means, in concurrent validity say for example, a screening test. Now, a screening test should be able to distinguish between two groups as positive and negative and you have your own questionnaire.

Now, the screening test is implemented in the same setting as the questionnaire you compare the data whether your questionnaire the findings from your questionnaire it corroborate with the results of the screening test. If it does then it has a concurrent validity simply the idea behind concurrent validity is to ensure that the distinction that is made between the groups which should be there it is actually there and concurrent means you are taking the the gold standard test or the criteria that you are you are setting to test your questionnaire in the present situation when you are actually implementing your own research questionnaire. So, that is the idea behind concurrent validity as you can understand predictive validity may take a longer duration, but the concurrent validity it will happen in the same setting. Another concept is called the post predictive validity usually the post predictive validity is not much used in in our health promotion research rather predictive and concurrent validity are much more used, but the basic understanding we have to have regarding post predictive validity the criteria here is set in the past. Say you have results from the past and now you are able to you are able to perform the questionnaire based analysis and simply just like the concurrent validity assessment you simply now corroborate the results of your current day findings and what you found out through a different criteria from the historical data or certain secondary data.

So, that criteria that that defines whether your question is appropriate or not that criteria is set in the past and you are implementing your study in the current time that gives you the idea regarding the post predictive validity. Now that we have discussed the different types of validity and also the concept of reliability we now have a clear understanding on how to you know make a questionnaire and then finally, how to revise refine the questionnaire to make it in a in a final implementable form. So, the two step process it is given over here the two step process is first you check for reliability whether it is reliable or not that is the first thing you have to do a reliability coefficient of 0.7 or greater will ensure the reliability. If no then you have to again revise after you have ensured the reliability you proceed to step 2.

What is step 2? You now measure whether it is valid or not. See remember reliability is useless if it is not combined with validity you have different types of validity all the forms of validity has its own meaning and own interpretation. The more measurable forms are content construct and criteria and the most important forms are construct and criteria often we use the predictive validity for ensuring criteria validity, but the predictive validity you take some time. After all these has been achieved and you tick off in each of these validity statements you know that your questionnaire is now valid and it can be implemented in the general population ok. So, these are the things that we basically discussed in this lecture what is the idea of construct validity and how to understand whether the operationalization of the concept idea is basically able to measure the particular construct and whether it is able to distinguish from it distinguish the construct from the other construct and also whether it is able to mean to measure the relationship between the different constructs.

Now, in convergent validity I would again repeat there are debates regarding how to interpret the convergent validity another group of opinion is that the items it should directly converge on the latent variable that you are actually measuring, but not on the other variables that is where the concept of convergent validity and divergent validity they come together and that is why they are put together as a part of the construct validity. And the criteria validity that we discussed it can be of three types like concurrent, it can be predictive or it can be post predictive it depends on when the criteria for assessment of validity is set usually we utilize the predictive and the concurrent validity during our health promotion research. We understood that all these activities that we have we are conducting is part of ensuring psychometric validity of a data collection tool. So, these are the resources basic resources that you should go through when you are studying reliability and validity of the health promotion data collection tool. Thank you for your patient hearing. Goodbye.