Marketing Research and Analysis - II
Prof. Jogendra Kumar Nayak
Department of Management Studies
Indian Institute of Technology Roorkee

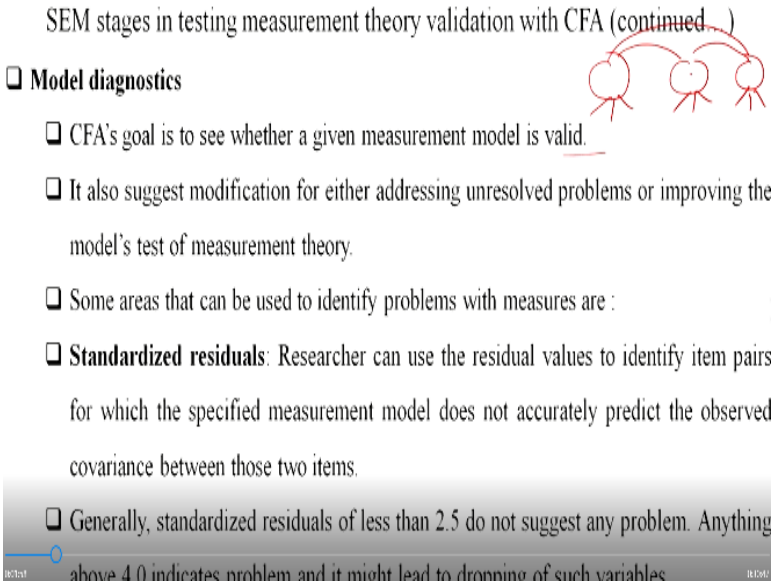
Lecture –56
Confirmatory Factor Analysis in SPSS-III

Welcome everyone to the class of marketing analysis and researches. So in the last lecture we were discussing about confirmatory factor analysis. Confirmatory factor analysis is a technique where we understood which is based on a theory right. It is unlike the exploratory factor analysis where the statistics decides which factors or constructs to be developed. But in the Confirmatory factor analysis we are developing a model on basis of certain theory.

This theory might be maybe a scale developed earlier by somebody or you know some literature that we have studied or through some even sometimes you know some discussion with some experts in the field. So you have developed a model and you feel it should you know behave in a manner that you are thinking. So to test that we do a confirmatory factor analysis okay.

So in the last lecture while we ended we I had said suppose you have developed a model and now the model is not showing a very a fit model. So how to check that and how what should you do in that case so model diagnostics is the one which will be starting the discussion today.

(Refer Slide Time: 01:39)



CFA goal is to see whether it given measurement model is valid or not right. So that means you have made a model so we know this is a covariance structure so right. So we are trying to see and each one has got let us say at least 3 variables okay now were saying whether this model that we are drawn is a valid model or not. Now how do you know whether its a valid model or not to do that we have some cut off values right.

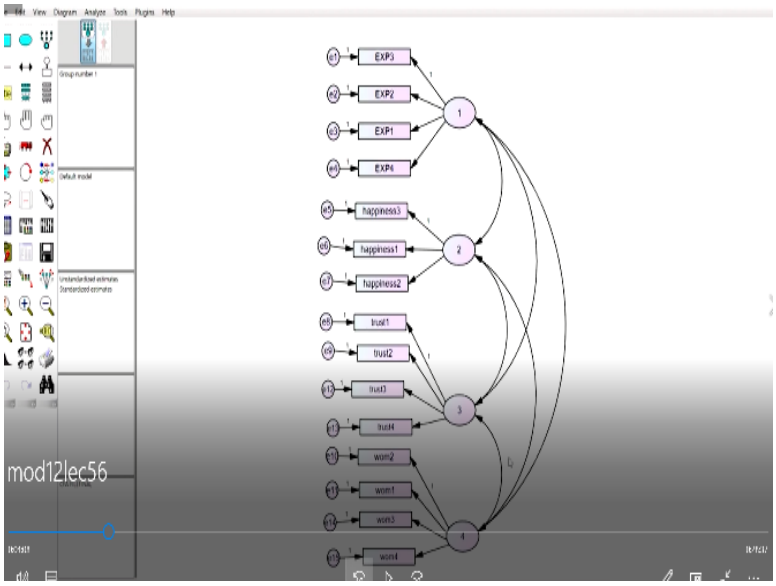
Which you will say as the indicis now this indices will help us in in finding out whether the model is a fit model or not. So we will see it also suggests modification for either addressing unresolved problems or improving the models test of measurement theory. So suppose you have while checking the fitting the says you found that the modern fitness is low right in such condition can you improve the model yes.

Now generally what people do they are when they you know when the model is not coming fit so they panic and they tried to do something which is you know unscientific. So you need not worry about I will simply help you to understand how scientifically you can improve the model okay. Some areas that can be used to identify problems with measures are for example standardized residuals.

So what are the standardized residuals I will explain researcher use the residual values right. Residuals this word you have already used in regression also when you were talking about the error types. Identify item pairs for which the specified measurement model does not accurately predict the observed covariance between these two items. So whenever let us see let us take this.

Generally, the standardized residuals of less than 2.5 do not suggest any problem right. But anything above 2.5 and specifically above 4 indicates serious problem and might lead to dropping of such variables that means what. When you have a model let us say, let us go to a model right so I had opened a model this is the model I had brought.

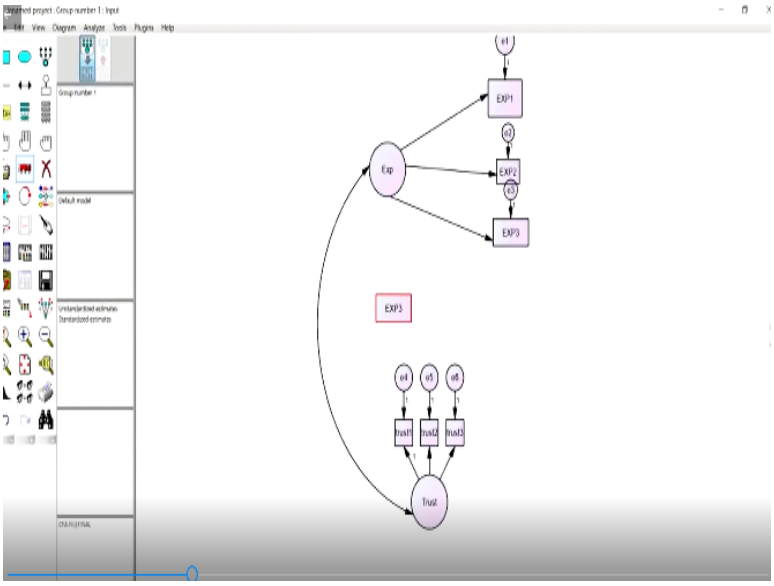
(Refer Slide Time: 03:43)



So suppose this is the model now I am saying there are four constructs 1 2 3 and 4 the first one is linked to experience so there are four you know indicators of experience, second is three indicators of happiness the 3rd one is having four indicators of trust and the 4th one is having four indicators of a word of mouth. So when we are, this is the CFA model is a covariance model here there is no definite relationship. Now what do you mean by definite relationship?

Where we say that something moves from, when there is a direction A moves to B so this is nondirectional correct so there are moving both sides the covariance so in such condition we will learn this and see whether the model is valid or not and then can be improved. But before that for some people who are extremely new to you know structural equation modelling and confirmative factor analysis. Then let me explain how to draw a model first of all right. So I have brought a blank model.

(Refer Slide Time: 04:41)



So you see what I am doing is suppose how to draw a construct. So there are several ways this is a construct right so this is how you draw a constructs this is the indicator right so let us say I have three indicators right I have three indicators linked to this so I say this is how I draw right. So this is how I draw right and each indicator will have an error term right so if indicator has an error term.

So this is you know indicator has got an error term. So suppose instead of that we can we do it a little differently yes we can do it. So this is how you draw it so in the same thing I have drawn it so easily this facility has been provided here. Now once you just click on it so that this icon is gone now you draw here now suppose this is the construct and the rest is something you want to right now whatever you want to write.

So I am writing trust okay trust so this is my construct right. Now what is a variable how do I bring the variables now to do that I will explain you just go to this file now remember you need to have your data file ready. Now the data file is the file which from where you collect import the data. Now the file name I am giving is let us say the CFA file right. So let me take it out open okay.

Now once I have done with it. I will go to view variables and data set and now there are trust. So trust 1 so I am pulling trust 1 here, I am pulling trust 2 here, I am pulling trust 3 here. Okay once

I have done now suppose you want to also do the same for this one let us say. This is let us say experience I am taking experience 1, experience 2, experience 3 okay and this one I will name it as experience okay so this is experience.

So let me draw it for you experience okay this is all I am writing. Now since this is a covariance structure so I will draw a covariance model right CFA is covariance. So now this is what now this if you see there are some unobserved, these are error terms they have to be given name so just go to this place name unobserved variables. So each error term has been given name correct now you can run it okay.

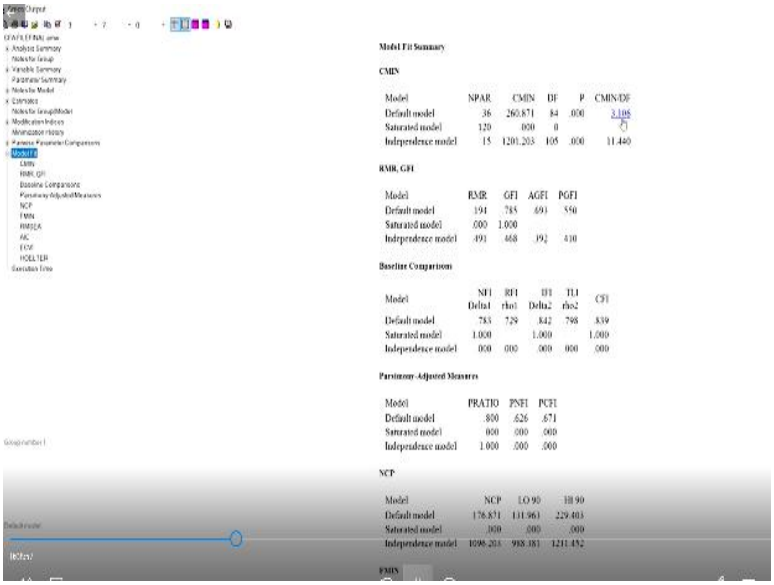
So this is how you draw the model so my interest is only to show you how to draw the model you want to make any changes to all those icons are there right so copy so you want to copy something you can copy for example I want to copy this and I want to paste it here right. So I want to or I want to just move something from this place to this place I can drag it here. So these are all the icons which the more you play.

You will realize more and more you will understand more and more okay. So I am stopping it with this okay. Now I will go to my model of interest now this one so already my data file is connected okay so what I am going to do I am going to first run this model and see whether this model is a valid model or not. Now how do I run now this icon you can see is what it says about running right.

So but one thing before you run you will need to go to the output file so this is a maximum likelihood you know condition now go to the output now I need certain for example there is residual elements now if you go back to this file what did it say I need these standardized residuals. Generally, people get confused of where do we find this okay so this is the residual moments right so you have done it.

Now I have run it already I think so let me leave it no issues. So okay so this is my output file okay now first let me check my model.

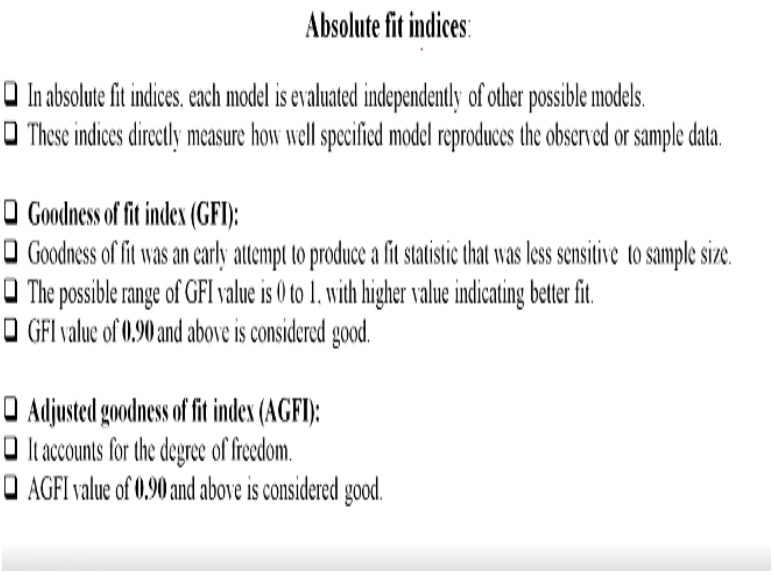
(Refer Slide Time: 08:48)



Now before you understand you see the values and let me tell you there has there is a chi square value because this model is based on the chi-square technique so this chi square minimum by degree of freedom this value is should be around you know less than 3 preferably less than 3 and if it is more than 3 or 4 and especially more than 5 then it is the model is a weak model right. So the lower the chi squared by degree of freedom the better it is okay.

Now this GFI, AGFI, NFI, IFI, TLI right these are all different indices. So CFI so these indices if you if you go back to my slide I will show you.

(Refer Slide Time: 09:29)

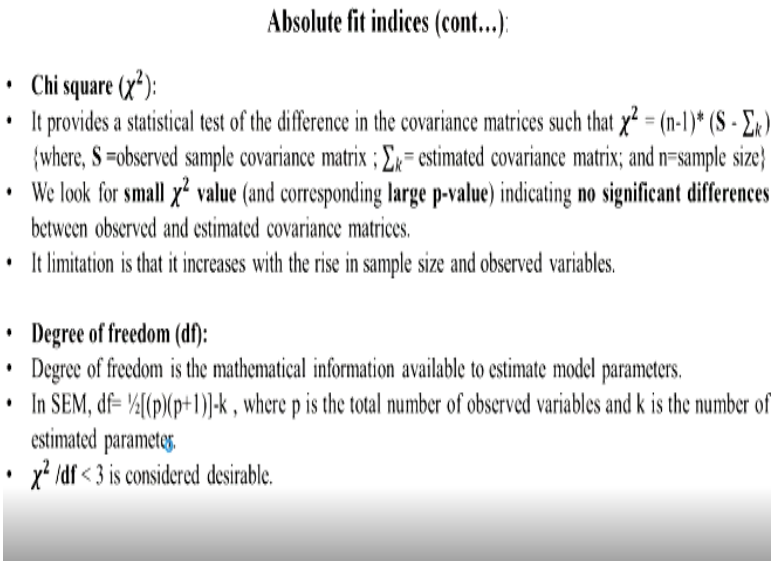


This in this indices this is an indices absolute fit indices right. So this absolute fit indices are for example goodness of fit indices right. Each model when you test a model at least from each indices you should take 1 right. So for example the goodness of fit indices this was an attempt to produce a fit statistic that was less sensitive to sample size. Now what is the normal person for somebody who is new you understand that higher the value.

The higher the better is a fit of the model so the cut off value has been kept has 0.9 so GFI is above 0.9 it is considered to be good right. The adjusted goodness of fit similarly this should also be 0.9 or above right close to 0.9. But remember do not ever worry because sometimes this 0.9 although it may be a very you know stringent it is a measure it is a cut off value but this might if you get something on 0.85 87 also you need not worry.

If your sample size is very large it is sensitive to sample size so if you get a slightly lesser you know fitness value, then also you should not worry. Because if you have a large value or sample size it will be very sensitive and this might come low right. So you need not bother about it but you have to check your sample size right.

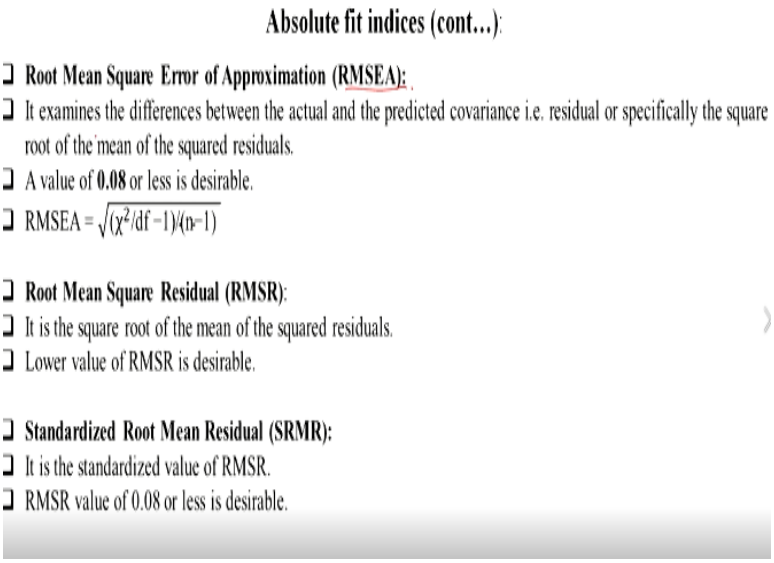
(Refer Slide Time: 10:51)



Similarly, the chi-square this value right which tells about the difference between the estimated covariance matrix and the observed covariance matrix right. So this value also we are trying to

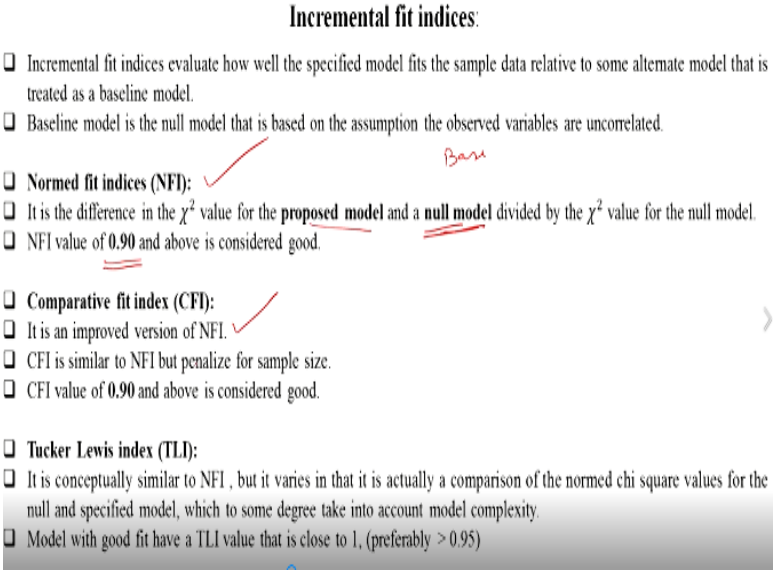
check the chi-square by degrees of freedom okay. So this should be preferably < 3 as I have mentioned here right.

(Refer Slide Time: 11:10)



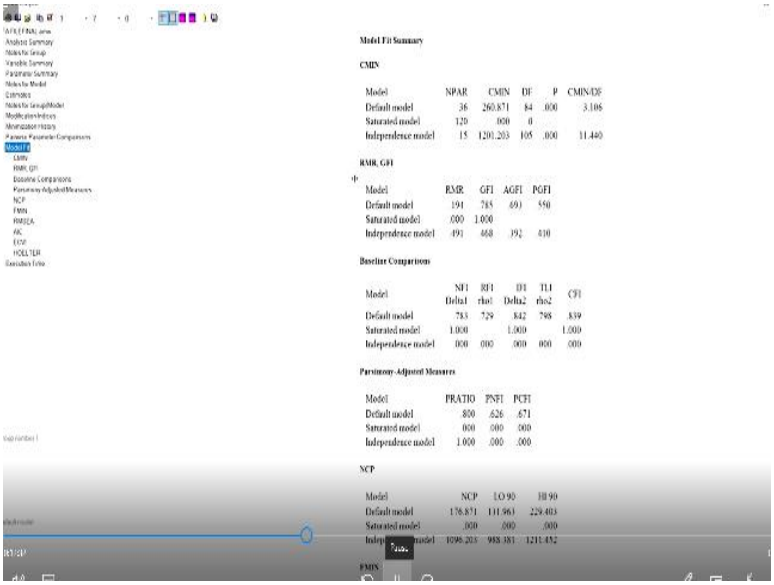
RMSEA this is another indicis right, it examines the difference between the actual and the predicted covariance model right and the value of 0.08 or less is desirable. Okay you can go back to the formula check the slides later and read it slowly how did the formula develop. But let me at this point let me explain you its meaning. The root mean square residual is another such indices which is also has a cut off value of about let us say 0.05 okay. So if it is less than this it is better the model is fit right so these are different fit indices okay.

(Refer Slide Time: 11:52)



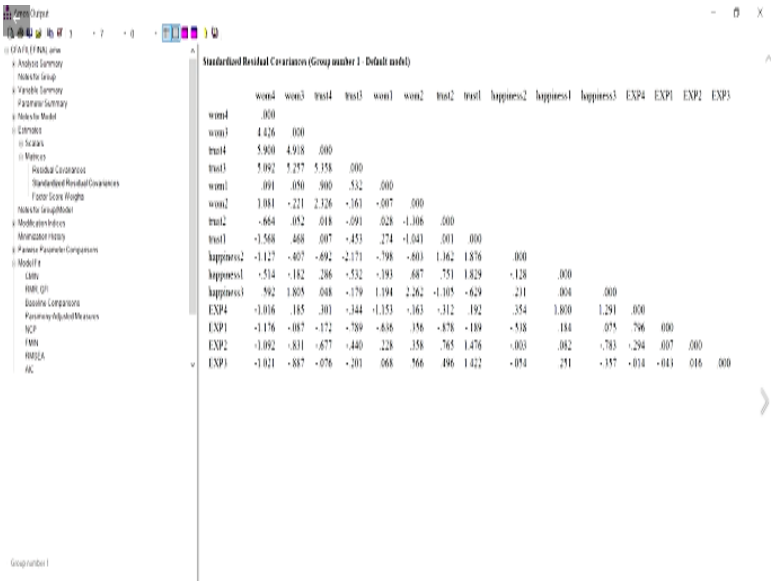
NFI it is a difference in the chi square value of the proposed model and the null model the baseline model you must have you will see when you use Amos this is called a baseline model. This baseline model is a null model okay divided by the chi square value for the null model. And NFI also should be about 0.90. Similarly, the CFI is an improved version of the NFI right. But CFI similar to NFI but it penalizes for the sample size. The higher sample size more sensitive and it should be 0.9. So these are several Tucker Lewis index.

(Refer Slide Time: 12:24)



So these are some several you know fit indices okay so now let me show you what we have got. Now we were here in the Amos output now we said it should be < you know 0.3 correct. Now let us see can we make some improvements in the model now and you see a the NFI is 0.783 GFI is 0.785 CFI is 0.839. So generally if we see the cut off the models cut off values are low right now how do we improve. Now let us go to the estimates right.

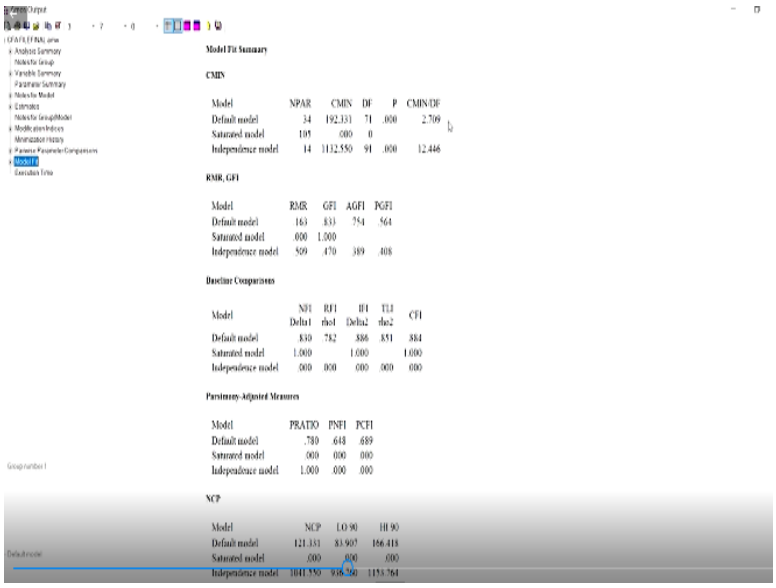
(Refer Slide Time: 12:58)



So there are two things scalar and Matrices now go to the matrices right now when you go to this you we want this standardized residual variances. Now let me drag it a bit now look at this you know this is a covariance matrix. Now if you see look at any value try to find out the highest value which is the one which is where the value is more than 4 and above right so 4 and above I think wom 4 right.

This one has a very large impact right the higher the values with the other when its the covariance values are very high. So 4.426, 5.092 similarly trust 3 also if you see its all about 5. So why not first start with this trust 3 so we go to the model trust 3 so we first start you know deleting this trust 3. So I am deleting trust 3 now one by one we will go and well run this model again you remember the chi square degree of freedom 3 point something. 1 or something was there I will run it again.

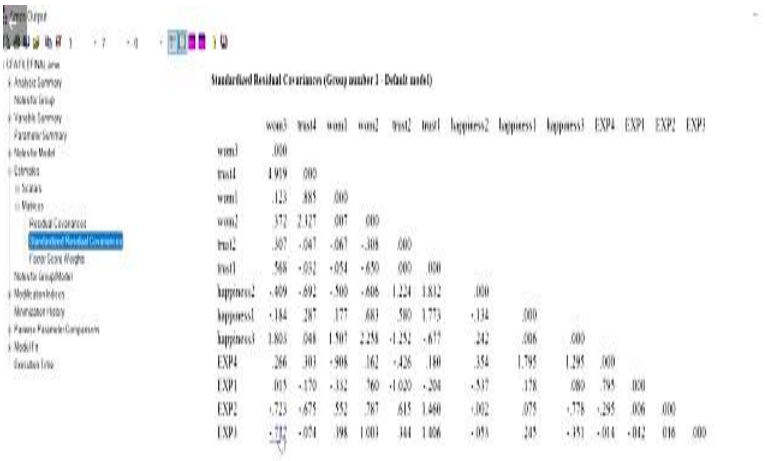
(Refer Slide Time: 14:08)



Now let us go to the output let us check the model fit now you will see there has been a significant change 2.709 and have the indices improved GFI 0.833, CFI 0.844 which is better than the earlier right. Now let us see is there any other variables similarly you know responsible for the downfall of this model weakness of the model. Yeah now let us go to this wom 4 right so we go to the model again.

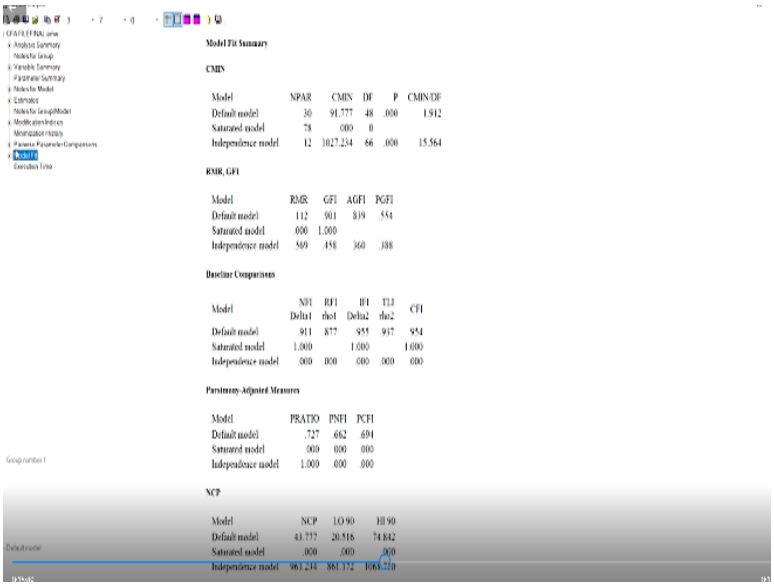
And we will see if we can if we remove this wom4 what is the change right now we will run the model again okay. So has there been any improvement 2.7 now 2.2 so from 3.1 to 2.7 now 2.2 and look at the GFI, the AGFI, the you know CFI and if you can see there has been significant changes right now let us see can we can we you know is there any other such covariance you know the relationships which are you know keep making the model weak? So again let us go to the estimates and check finally the last for the last time.

(Refer Slide Time: 15:27)



There is trust 4 which is also you know above a 4 so we can play with this trust 4 right. So I am removing trust 4 okay so we will have only maybe two values of trust. Now finally we will run.

(Refer Slide Time: 15:46)



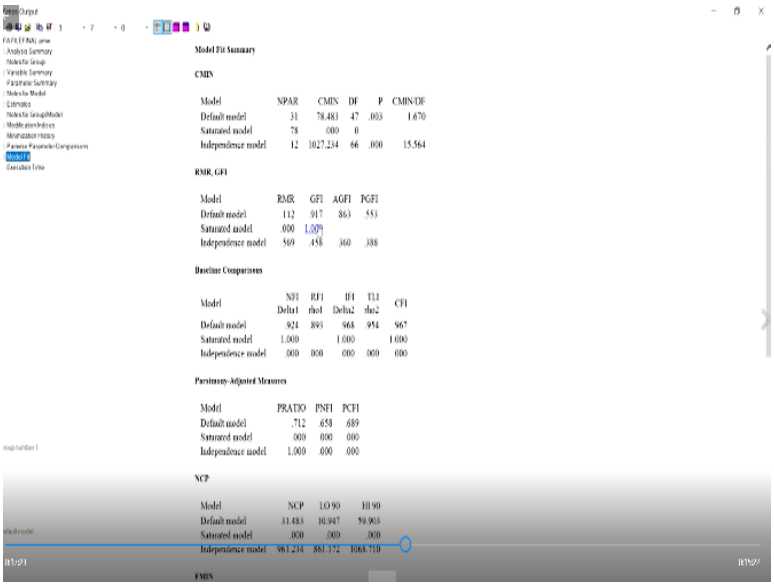
So model fit so now it is 1.912 so there is a significant improvement the GFI values have become above 0.9 so which is good enough. Suppose now you want to check again is this some value variable which is still you know responsible now we can see it there is nothing above point 4 the cut off values are in fact 2.5 and above 2.5 and in between 2.5 to 4 itis then you may treat it if you do not treat it also no issues.

But above 4, you should surely treat it that is an item for deletion right now there is no such problem now still if you want to see make an improvement suppose this would have been the table and you see our model fit is there say I want this this would not have been 1.9 something like 3.9 or something or 2.9 then can I make some improvement and yes what do you do you go to the modification indices right.

And when you would go to the modification indices see within the same construct which are the items or which are the error terms which have a relationship you should never this is theoretically wrong to make draw a covariance between a construct and error you should not do it. Between two error terms and that to within the same construct you should first check for it. So 12.830 is it in the same construct.

Let us see e3 and e4 yes so these two there is the if we make a covariance it says that there will be an improvement in the model right. Okay now let us run it was 1.9 now let us see is there any improvement in the model.

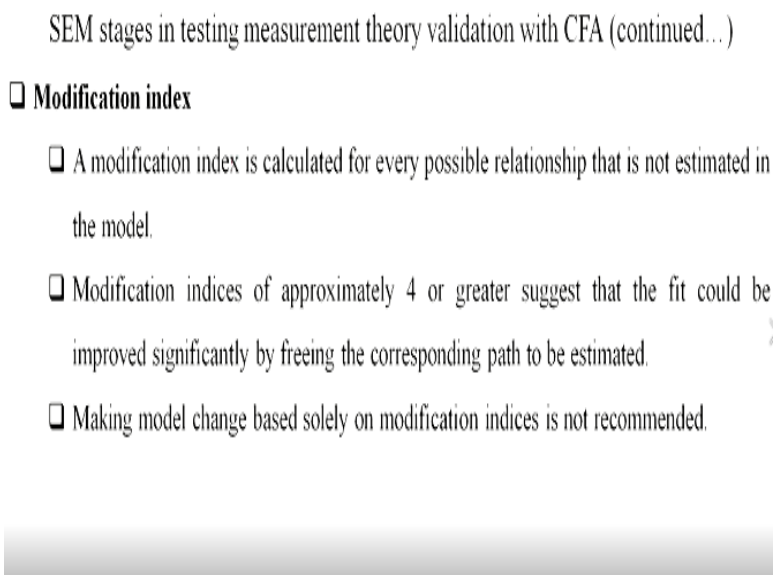
(Refer Slide Time: 17:22)



1.670 look at all the values right and this is an improvement so you can do further also you can again check for modification indices and see whether there is any you know any such relationship e2, e5, e3 so this is so fine right so this is not to be connected so e2, e5 is it within the same here no. So that means there is no more improvement possible in the model so we stop

here so now this is your final CFA model okay so as I said so this is this helps you to identify which items to indicators to delete okay.

(Refer Slide Time: 18:04)



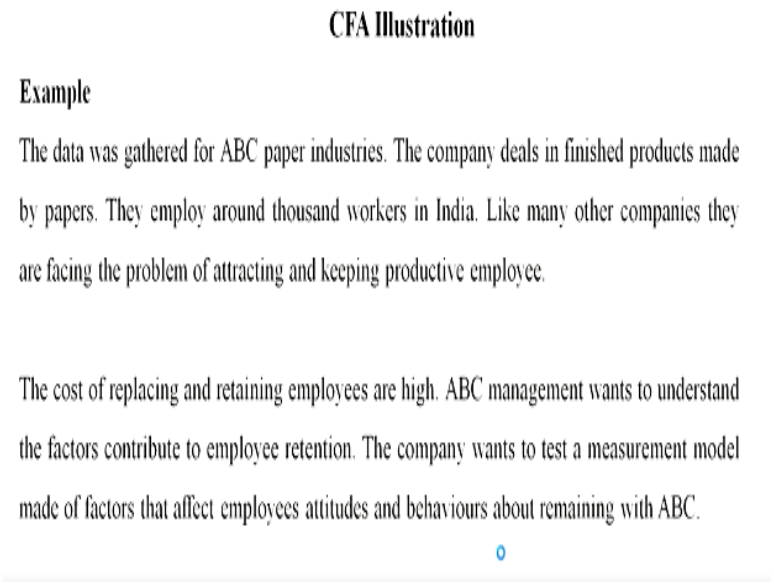
Similarly, the modification index which I just did it and showed you is calculated for every possible relationship that is not estimated in the model and modification index of approximately 4 or greater suggests that the fit could be improved significantly by freeing the corresponding path to be estimated right. So making model change basically solely on modification indices is not recommended.

So you should not first try to see the residuals and after that you can go for the modification indices and I have given you some conditions like for example do not try to draw a relationship between error terms of 2 different constructs and do not do it too many things at one time right that means two error terms together at the same time do not do that go for one by one the highest which has the highest value.

So one by one and you should not also theoretical it is also wrong to you know connect a construct with error term because the software will do it but then theoretically that it is wrong actually you should not be interpreting that way so this is what we have understood. Now there is one more example I have brought for you to because after this once you are thorough with the

CFA then well get into this. After the measurement model is found to be fit we will get into the structural model okay.

(Refer Slide Time: 19:25)

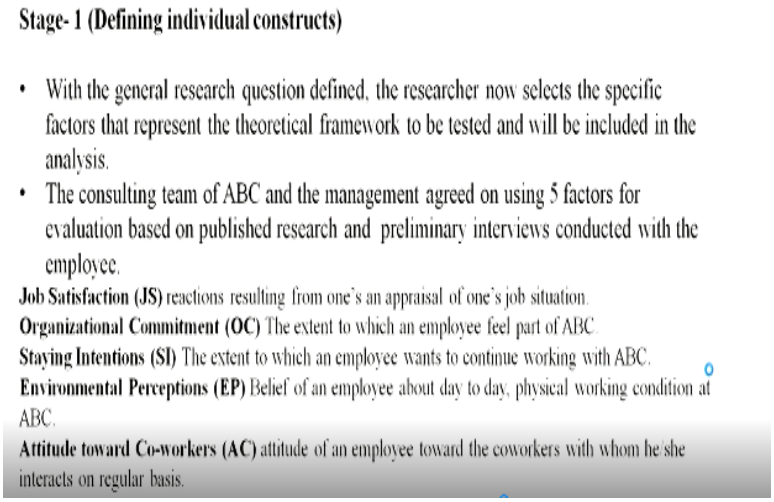


So here this is a case which I have brought this is a case of ABC paper industries the company deals in finished products made by papers. They employ around thousand workers like many of the companies they are facing the problem attracting and keeping productive employee. The cost of replacing and retaining employee is high. ABC management wants to understand the factors which contribute to employee retention.

So it wants to test a measurement model made of factors made of factors that affect the employee's attitudes and behaviours about staying with the company.

(Refer Slide Time: 19:58)

CFA Illustration

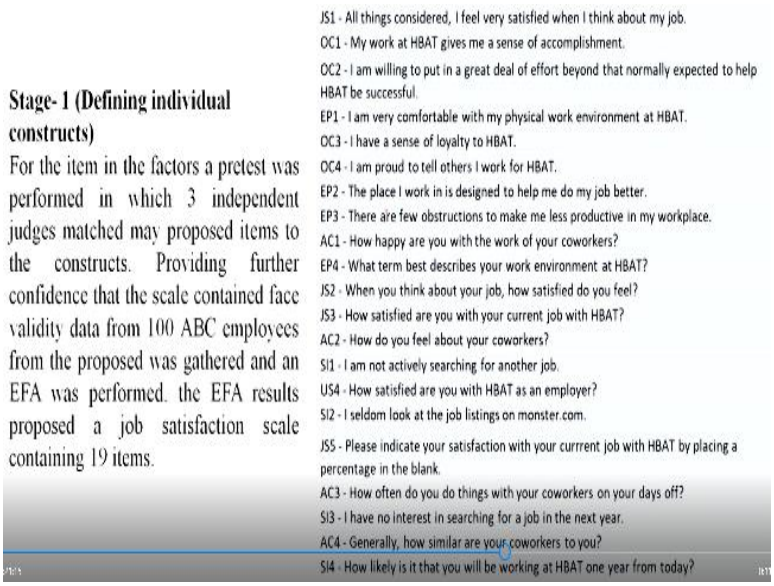


So first we will define the individual constructs so let us do that with the general research question defined, the researcher now selects the specific factors that represents the theoretical framework to be tested and will be included right. So what are the constructs now they have seen job satisfaction, organizational commitment, staying intention and environmental perceptions attitude towards co-workers.

These are some of the variables or the constructs that would be a part of this study, job satisfaction reactions resulting from once an appraisal of one's job situation. Commitment that is an extent to which an employee feels part of the company. Staying intention an extent to which an employee wants to continue working with ABC. See remember in the measurement model we are not saying anything is dependent or independent.

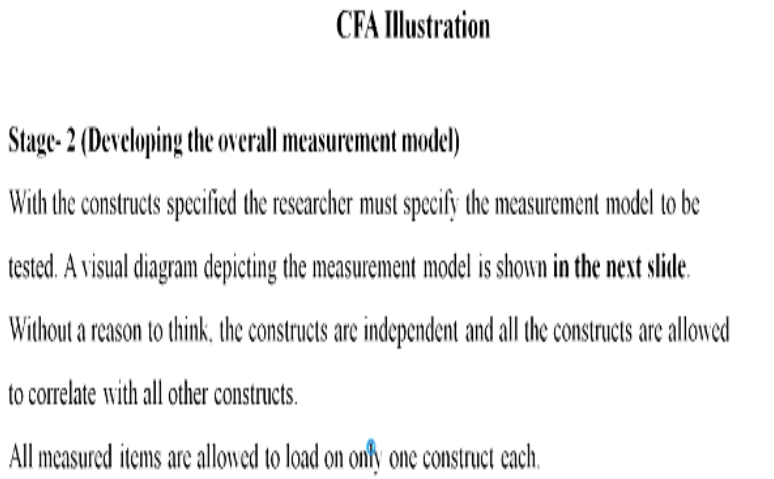
It is just a covariance model every there is no direction as such right. So these are the five constructs we are having. So job satisfaction JS, organizational commitment OC, staying intention SI, and Environmental perceptions EP right that means belief of an employee about day to day physical working condition right and attitude towards co-worker AC.

(Refer Slide Time: 21:13)



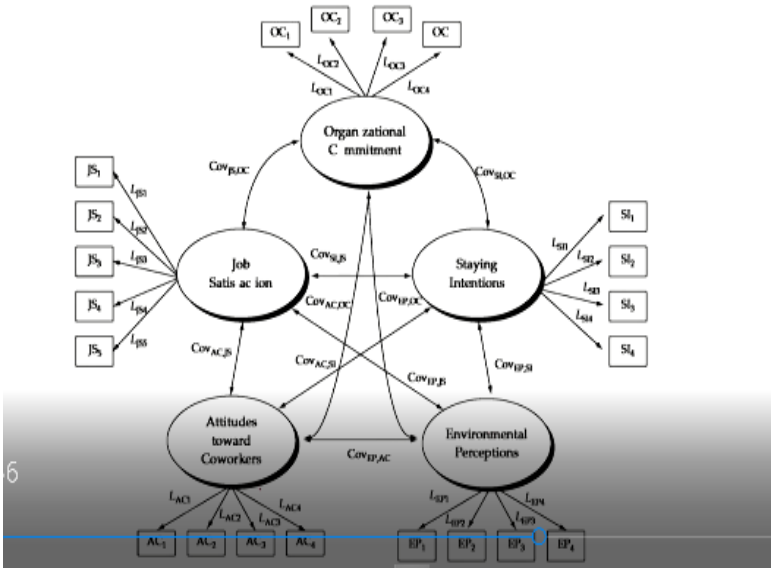
Each you see each constructs has several indicators so job satisfaction has JS1, JS2, JS3 right there are right you can check from here you can read all this right.

(Refer Slide Time: 21:28)



So then we do the develop the measurement model so with the constructs specified, the researcher must specify the measurement model to be tested right okay.

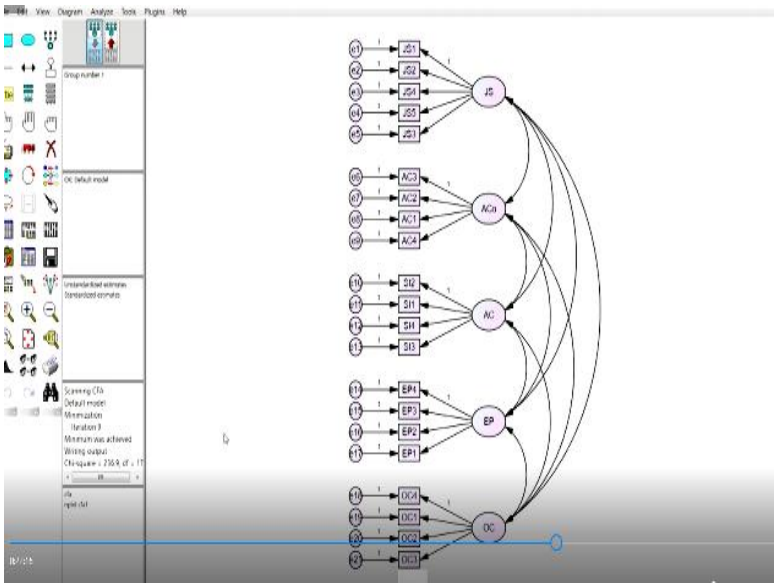
(Refer Slide Time: 21:36)



So this is the measurement model okay so job satisfaction, organizational commitment, staying intention attitude towards co-workers, environmental perceptions and you will see they are all correlated right. So we have a theoretical understanding how they are correlated okay so this is all you can see the correlation right its a covariance structure and each of the indicators are connected with the and you can see these are reflective constructs right.

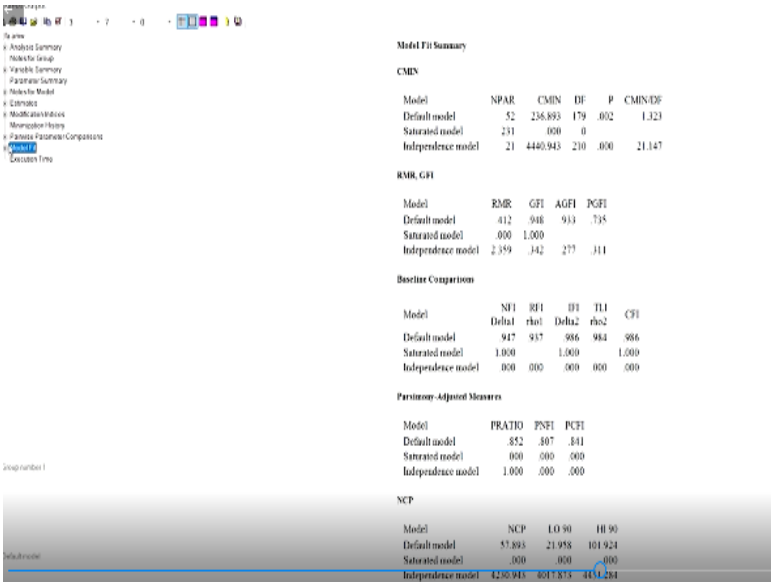
So that means the latent construct is explaining the variables okay and each variable what does it mean suppose SI 1 what does it mean you can check here.

(Refer Slide Time: 22:10)



Now let us run this we have this file in fact with us so we have this right so I already connected to the data. Now you try to understand first what we do we will take we have already run so it is already there now let us run the model right I hope you can draw the model at the moment by using this icons okay.

(Refer Slide Time: 22:32)



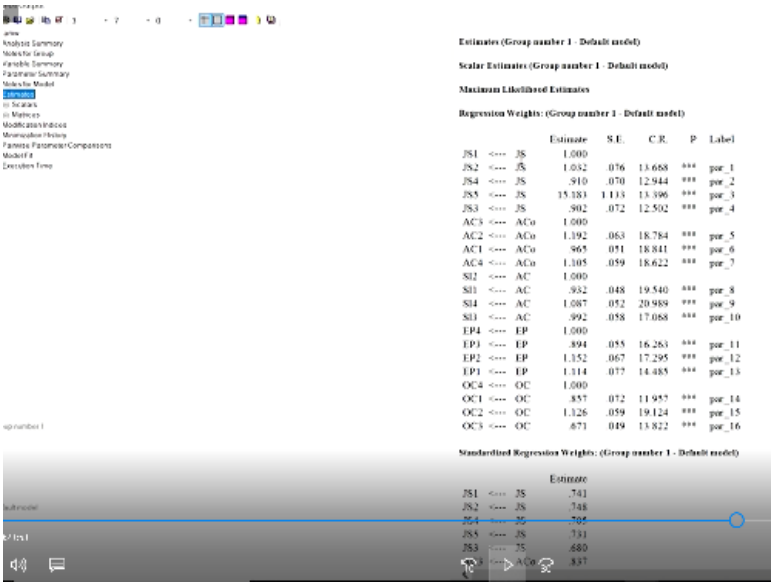
Now first let us check the model fit the model is because there is already we have tested and tried and tested already this model is coming every clear and very robust model look at the you know CMIN by degrees of freedom the chi square by degrees of freedom it is 1.323 right and look at the GFI the goodness of fit indices this is what you need to report correct when you write your you know paper research paper.

You need to write the chi-square by degrees of freedom was this much the GFI value the AGFI value you can take one from each right so and the RMR should be <0.5 as I had said and baseline you can say out of NFI, RFI, IFI, CFI you may take one or two of these. So all are above 0.9, right so once we have done now there is no need for improvement but the model is by default we are saying the model is an improved model.

That means we have we can say the model is a is a strong model okay that means what does it mean that difference between the observed and the estimated model is very lucky right. Now

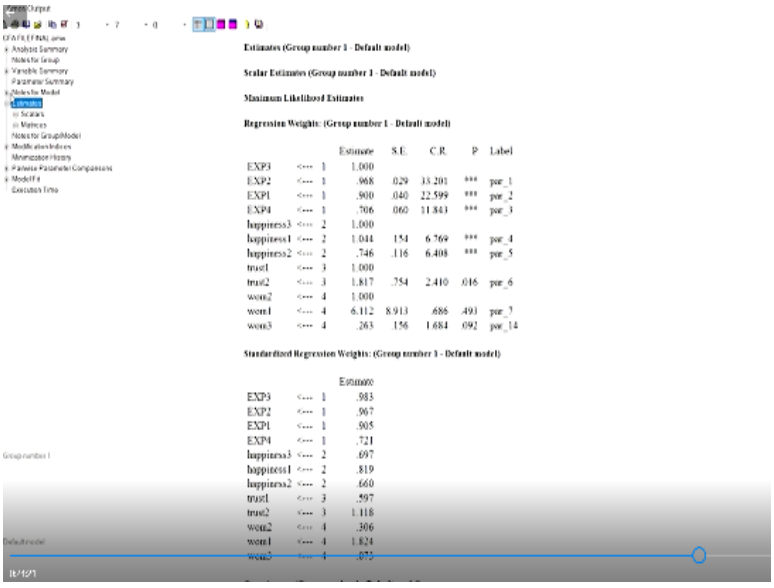
suppose you would have needed some improvement to check then what would have you would have done go to the estimates right now this is an important.

(Refer Slide Time: 23:52)



Now you can although this is less important but this is important for us this side. Now is this is there any insignificant thing any insignificant relationship no. If there would have been insignificant relationship, then we would have deleted. Now let me let me go back to the earlier the one slide which I had drawn okay now let us run this okay.

(Refer Slide Time: 24:21)

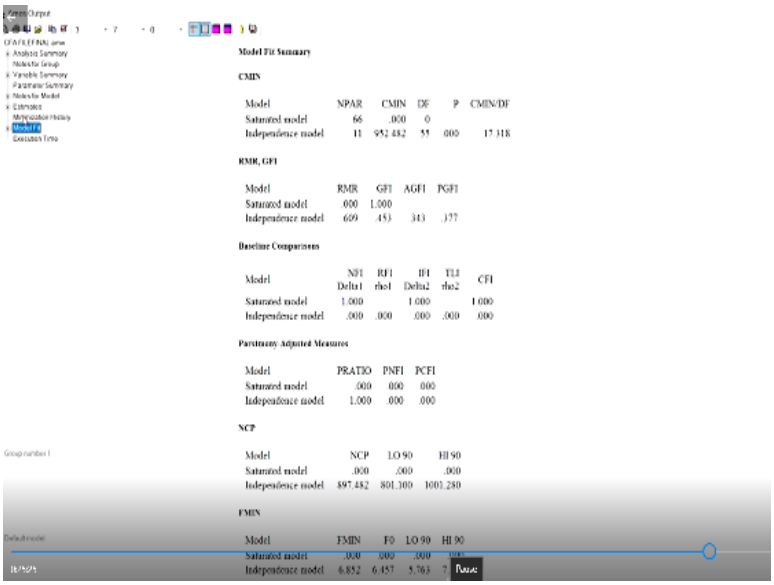


So go to the estimates now you can see some of them is okay. We have already done it otherwise maybe the ones which you deleted that we deleted there we deleted some residual errors would

have also shown here maybe they would have been insignificant at this you know this place it would have for example you see the relationship between wom1 right and the 4th construct wom1 and 4 right. So 4 and wom1 this is still coming in significant that means if you delete it.

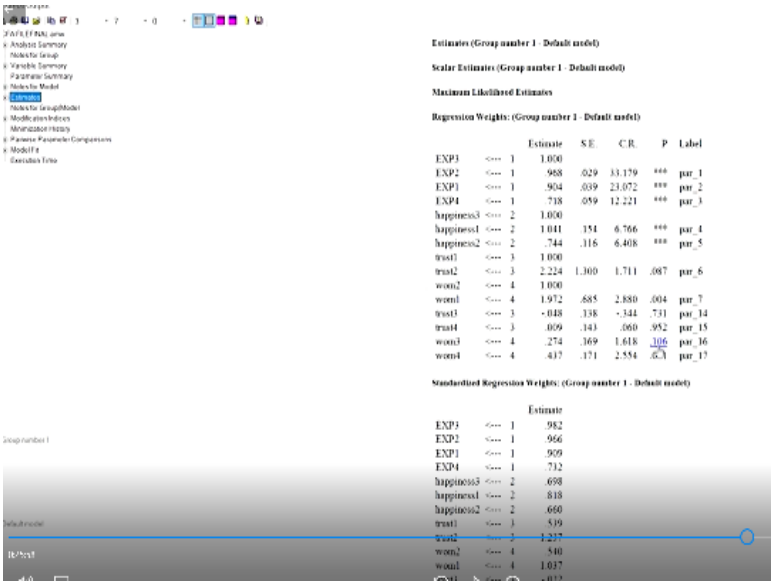
Now look at the model fit first go to the output right now look at the model fit so 1.67 and whatever you are getting now if I deleted now this is a good thing suddenly we got it struck my mind. So I think wom1 0.493 wom 1 right. So let us remove this so suppose I remove this right and now let me run the model again now you will see now go to the model fit must have improved.

(Refer Slide Time: 25:24)



Okay the problem here is one what has happened this is good that this problem has come to you. Now everywhere there is a parameter the fixed and every parameter has to be explained right so there is one more way I can show you where you can see whether you know some of the indicators are fitting or not. So as I explained you can run the model go to the estimates right so in the estimates first you look for the model fit and normally so it is how it has there.

(Refer Slide Time: 25:53)



Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
EXP3	<---	1	1.000				
EXP2	<---	1	.988	.029	33.179	***	par_1
EXP1	<---	1	.904	.039	23.072	***	par_2
EXP4	<---	1	.718	.059	12.221	***	par_3
happiness3	<---	2	1.000				
happiness1	<---	2	1.041	.151	6.766	***	par_4
happiness2	<---	2	.744	.116	6.408	***	par_5
trust1	<---	3	1.000				
trust2	<---	3	2.224	1.300	1.711	.087	par_6
worn2	<---	4	1.000				
worn1	<---	4	1.972	.685	2.880	.004	par_7
trust3	<---	3	-.048	.138	-.344	.731	par_14
trust4	<---	3	.009	.143	.060	.952	par_15
worn3	<---	4	.274	.169	1.618	.106	par_16
worn4	<---	4	.417	.171	2.554	.011	par_17

Standardized Regression Weights: (Group number 1 - Default model)

			Estimate
EXP3	<---	1	.982
EXP2	<---	1	.966
EXP1	<---	1	.909
EXP4	<---	1	.712
happiness3	<---	2	.698
happiness1	<---	2	.818
happiness2	<---	2	.660
trust1	<---	3	.519
trust2	<---	3	1.217
worn2	<---	4	.540
worn1	<---	4	1.037

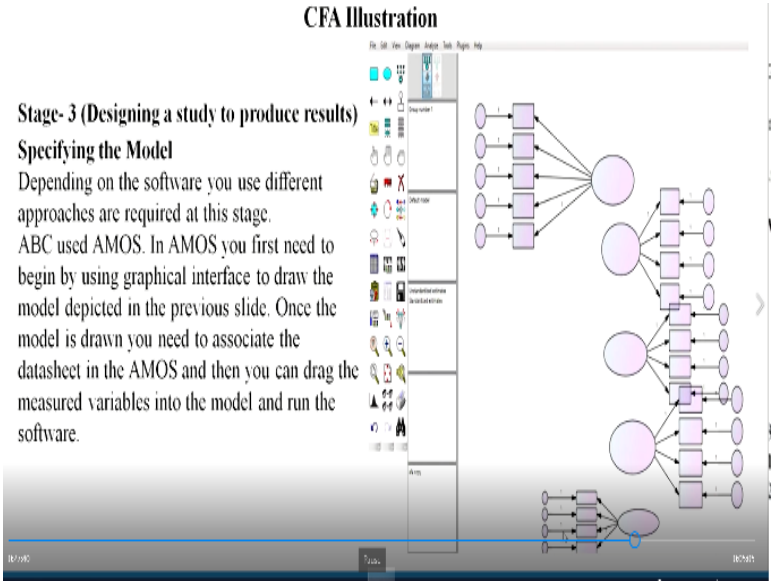
Now go to the estimates now I said if you remember that the estimates need to be significant right but here this is very interesting what is the what is the if you see this one trust the construct 3 with the trust 4 right this is a 0.952 right. So 0.95 this is not significant right so we have a cut off value of let us say 0.01 so it is much above it so that means with this null hypothesis cannot be rejected which says that there is no relationship that is to be cannot be rejected okay.

So now so when you find this what you can do is go to this and drop this model trust 3 I think it was trust 3 right so what we will do is trust 3 you can delete it from here and keeping other things the same you can re-run the model right. So this is also a good way to understand how to improve your model now let us see the estimates there are a few for example trust 4 is also there so you can check now this data every data set is different.

So when you do this so you can understand where to how to improve so I said through the estimates you can go and check for significance. I also explain to you how through the residuals you can go for you know something which is above 4 or something and try to delete those variables and if there is no problem in that still the model is not improving then you can also see for modification indices right.

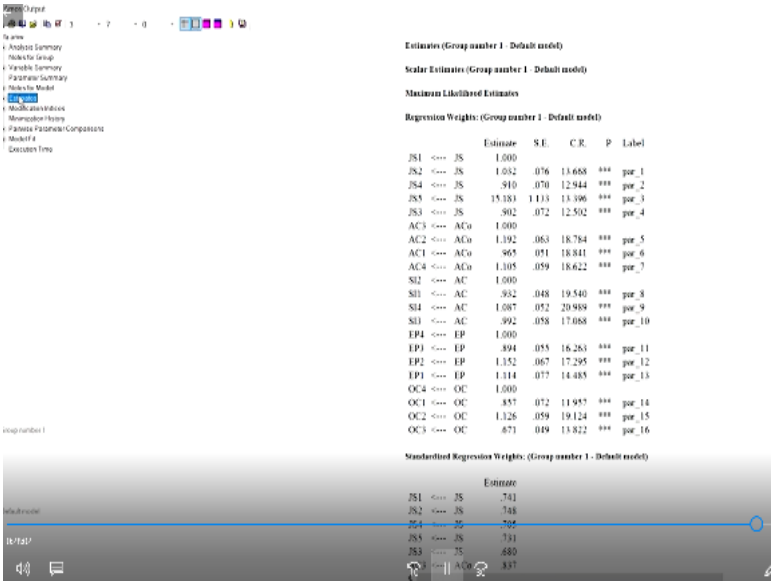
So these are the 3 things that you can check to improve your model if suppose still it does not going correct improving then do not worry then you have to go back to the field and check your data again right so now coming to the slide so we were here.

(Refer Slide Time: 27:39)



So as I was working on this you know job satisfaction this is okay so this file we were working so I think yes this is the one so we had run this and we were checking is you know model first so the model was fit everything was good. So this model says okay now this is everything is good right.

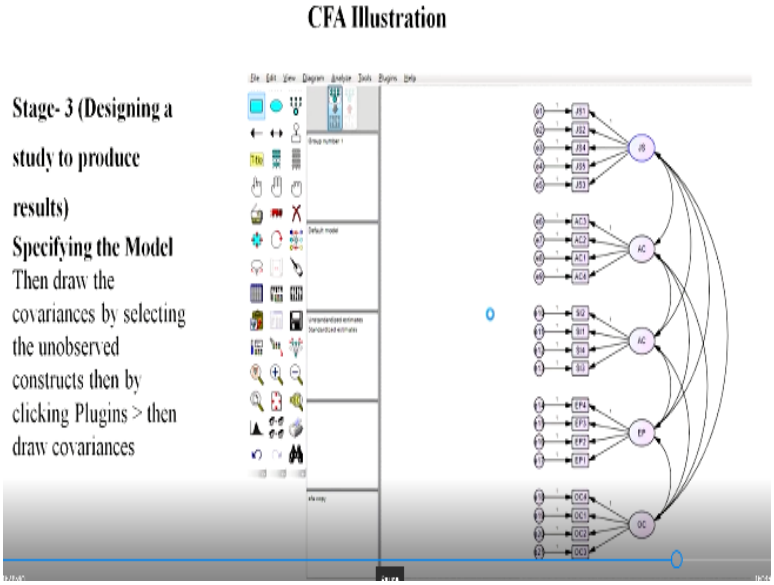
(Refer Slide Time: 27:58)



Now estimates let us look at the estimates now although it is not a directional study like a structural equation modelling or measured as structural model. In the measurement model we are not worried but still this helps us to identify whether any of this you know indicators are giving a poor result to the entire model. So had it been a not significant case then it is an item for deletion as I had just shown to you.

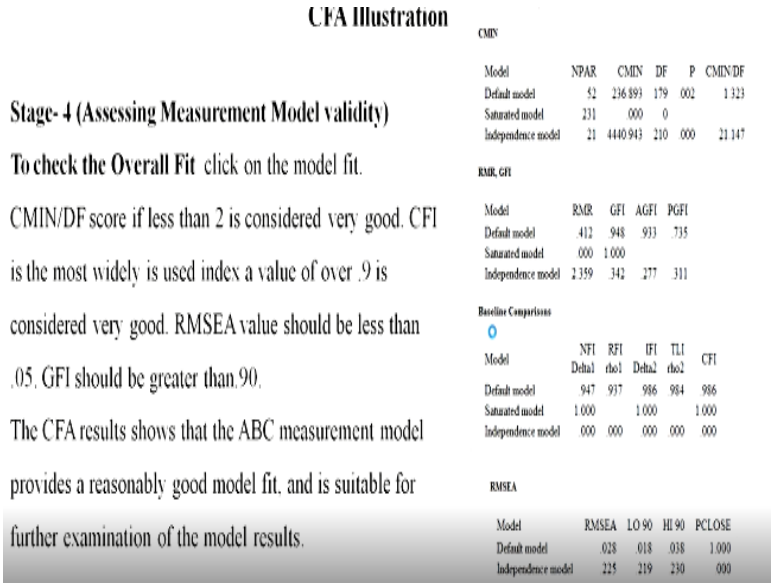
But here in this case there is no such problem that means now if you go back to the model that is how you do this right so you have drawn all these things.

(Refer Slide Time: 28:35)



And now the model has been drawn and these things you can check later on how to take the values into here and finally we will come to the scores right.

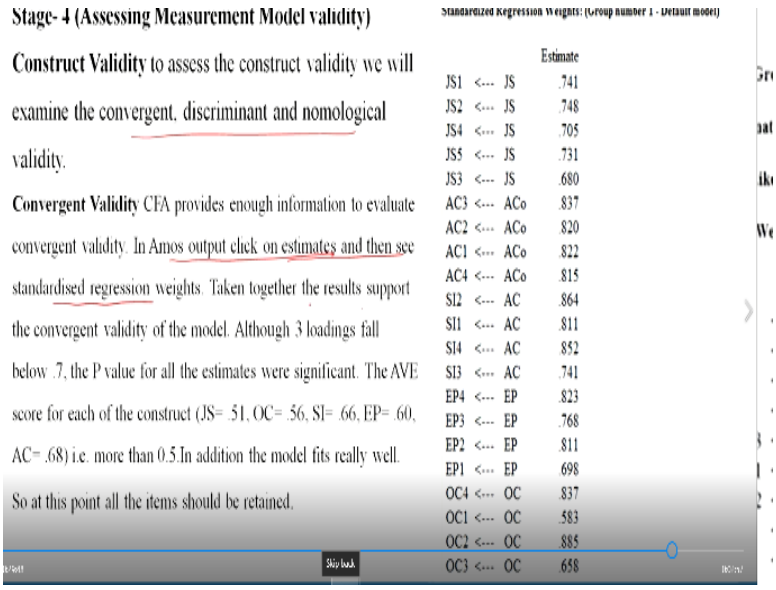
(Refer Slide Time: 28:48)



Okay this is how the model will look like you see 1.323 and the GFI so here this is how you write to check the overall fit click on the model thus a chi square degree of freedom is <2 it is considered very good CFI is over 0.9 RMSEA is <0.5 GFI should be greater than 0.90 so this result shows that the ABC measurement model provides a reasonably good model fit and is suitable for further examination of the model results.

So this is only one part that measurement model says that the model is good and it is can be tested in the further relationship as in the structural model okay. Construct validity, to assess the construct validity we will examine the convergent discriminant and nomological validity.

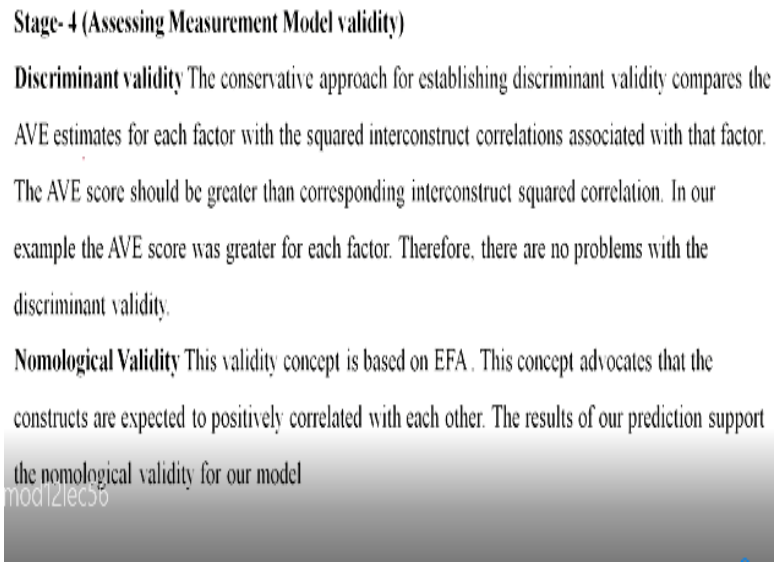
(Refer Slide Time: 29:34)



So convergent validity as I have already explained right let us see here. In Amos output click on the estimates and then see the standardized regression weights. Taken together the results support the convergent validity right. Although 3 loadings fall below 0.7 so there are you see point they should be within 0.7 at least it should be above 0.5 so nothing is below 0.5 here right there just everything is more than 0.5.

In addition, the model fits really well because the chi-square by degrees of database is 1.323 right now how will you check for discriminant validity.

(Refer Slide Time: 30:15)



The conservative approach for establishing discriminant validity compares the AVE estimates which I had shown you in the diagonal it should be the AVE scores if you take the constructs C1, C2 so C1, C2, C3 and you take the AVE loadings right. So this loading AVE scores should be more than the square of the correlation among the among the loadings right. So the AVE scores should be greater than the corresponding interconstruct squared correlation. In our example the AVE score was greater for each factor therefore there are no problems okay.

(Refer Slide Time: 30:54)

Summary

Four stage of CFA has completed. The chi square is significant above the .01 level. Both the CFA and RMSEA appear quite good. Overall, the fit stats suggest that the estimated model reproduces the sample covariance matrix reasonably well. Further evidence suggest a good construct validity. Thus, ABC paper mills can be fairly confident at this point that measures behave as they should in terms of the unidimensionality of the five measures and in the way the construct relate to other measures.

So this is how finally you write the CFA has completed the chi-square is significant above the 0.01 level. Both the CFA and RMSEA appear quite good because of the you have to show the values right. Overall the fit stats suggest that the estimated model reproduces the sample covariance matrix reasonably well. This is the most important finding further evidence suggest a good construct validity.

Thus ABC paper mills can be fairly confident at this point that the measures behave as they should in terms of the unidimensionality of the five measures, five measures are the five constructs. And in a way the construct relate to each other but here you have not said which construct affects the other in which way because no direction has been provided right. So only we say that the model is a valid model and the constructs relate to each other.

So this is all you do in the CFA correct so in the next lecture, I will explain to you once within taking same example how to explain the relationship or a provide a direction and check whether there is a relationship exists and if it exists how does it exist okay. So this is all for the day do we have. Thank you so much.