ISBN: 978-93-90847-39-6

2. Structural Equation Modelling- A Thorough Insight

Dr. M. Dhanabhakyam

Professor and Dean,
Department of Commerce,
Bharathair University,
Coimbatore.

Sowmya G.

Senior Research Fellow,
Full-time PhD Research Scholar,
Department of Commerce,
Bharathiar Univerisity,
Coimbatore.

2.1 Introduction:

Structural equation modeling (SEM) has become the most sorted method of data analysis in the modern day social science research. Structural equation modeling is one of the best alternative method in which the researchers are exponentially creating and testing relationship and testing the validity of the most complex and sophisticated models. Through adopting a SEM approach the researchers worldwide are trying to reach the new paradigms of relationship and hypothesis that was new to the world until recent. The SEM is basically a graphical technique in which the user will be developing and redeveloping a series of diagrams to explore the relationship between the variables. This is a most robust technique as this is the technique based on the covariance between the constructs used in the study. In this technique the researcher will be able to test the relationship between more variables at a time. Sometimes the researcher will be able to understand the complex mechanisms such as simple mediation, serial mediation, moderation, mediated –moderation, moderated-mediation and alike through a robust SEM.

This is a graphic based and diagrammatic approach which simplifies the presentation of dynamic interaction between the variables. The viewer or the user can easily verify the diagram for more relevant information. Moreover the result of structural equation modeling throws various kind of information which helps the researcher to verify the statistical validity of the data being used and the results being obtained. In simple words, SEM offers a technique in which the complex interaction between the variables can be simplified and more robust and unexplored relationship between the constructs can be developed. This chapter offers a comprehensive understanding about the various kinds of techniques, methodology used and the prospects of structural equation modeling in the ambits of doing social science research. This chapter further details the various forms of data validity that must be checked and assessed by the researcher before proceeding with structural analysis.

2.2 Benefits of Doing Structural Equation Modeling:

People always prefer diagrammatic representation over theoretical presentation. There is always an extra point of presenting information through diagram than representing it in words. Diagrams always caught the attention of the readers and have great impact in providing meaningful information to the desired stakeholders. The following are the major benefits of doing a structural equation modeling to the researchers.

- This presents the model in a diagrammatic way which simplifies the underlying ideas more effectively.
- This is the most robust mechanism to communicate the conceptual intention of the researcher.
- When the intention of the researcher is to explore the dynamics of interaction between the variables, the SEM is the only rescue as it checks the relationships simultaneously.
- This offers a more simplistic way to conceptualize and operationalizes an idea in the form of diagram.
- Most of the structural equation modeling techniques is covariance based; hence the validity of the obtained result cannot be questioned easily.
- The results of SEM indicate the statistical fitness of the model being developed. This reflects the robustness of the measures being tested and the results being obtained. The SEM provides a series of goodness of fit indices and badness of fit indices that the researcher can interpret to verify the statistical goodness of the obtained results.
- Through testing for SEM it allows the complex model building such as mediation and moderation and assesses the relationship between numerous variables simultaneously. Mediation and moderation is considered to be the most happening interaction between the variables which can be easily included and verified through SEM.
- The SEM canvas offers various drawing tools that can be used to make the obtained results more attractive and comprehensive.
- There are numerous software's that can be used to develop SEM with the help of collected data. Some of the platforms include Amos developed and powered by IBM, Smart PLS, and Lisrel etc.

This chapter is based on the structural equation modeling being undertaken in AMOS which is promoted by the IBM group. AMOS originally stands for analysis of moment structure which analyzes and interprets the critical relationship between the variables. There are many advantages of preferring AMOS over other software's. Amos is software which word basically on the assumptions of covariance between the variables. The covariance act a robust measure that never compromise on the validity of the result obtained. Also, AMOS is much low cost when compared to other platforms and the probabilities of software getting crashed during the model building is comparatively low for this software. There is an added advantage that AMOS supports massive bootstrapping which is seldom provided by the other platforms or has bootstrapping limitations attached with. Also, AMOS offers better path diagram while creating and testing the mediation and moderation between the variables. The authors objectively opines that AMOS is not just a mean software for running SEM but it is the most effective, simple and reliable platform for testing the structural relationship between the variables. This chapter details the various points to be considered and the drawing rules in the AMOS canvas while developing a SEM.

Before proceeding and testing the SEM the researcher must complete confirmatory factor analysis. The various aspects of conducting a confirmatory factor analysis is given below in the next section.

All you need to know about confirmatory factor analysis:

The goodness of the structural model developed by the researcher lies in the confirmatory factor analysis and the various tests of validity being examined by him/her.

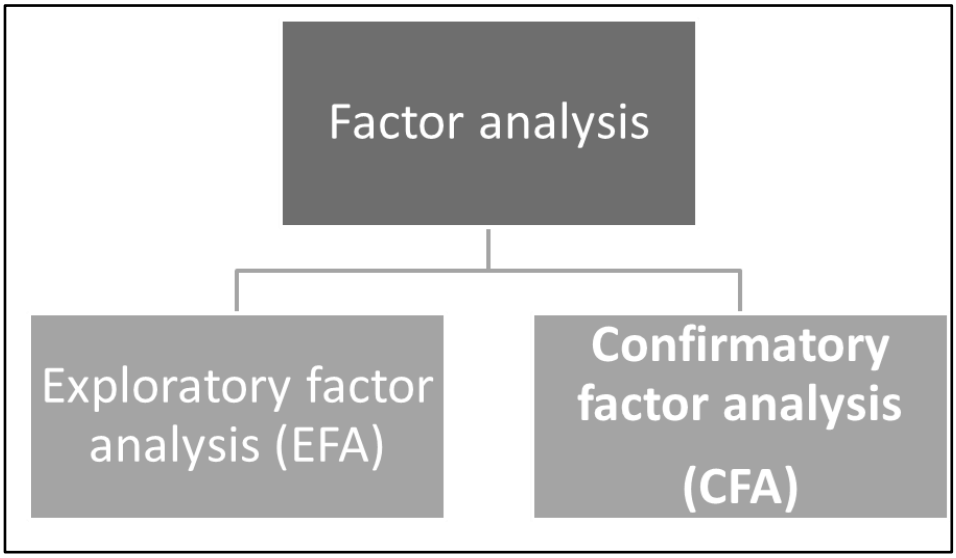

Figure 2.1: Types of Factor Analysis

Factor analysis is a technique that is widely used to explore the underlying factors from a set of observable variables.

It is basically a data reduction technique which is used to extract the hidden factors or unobservable variables from a set of known factors.

There are two types of factor analysis- exploratory factor analysis and confirmatory factor analysis. Exploratory factor analysis or EFA is concerned with exploring and extracting factors from a set of observed items or statements.

It is the first and foremost step in factor analysis. Once the factors are explored and named it is the duty of the researcher to confirm the validity of the factors.

Ensuring the statistical validity of the constructs are necessary step before proceeding with the structural analysis. It is said that ta structural model without a proper measurement model is nothing but a building without a string base which can crumble at any time.

So as a matter of due diligence it is the responsibility of the researcher o confirm the various forms of validity before proceeding with the structural model.

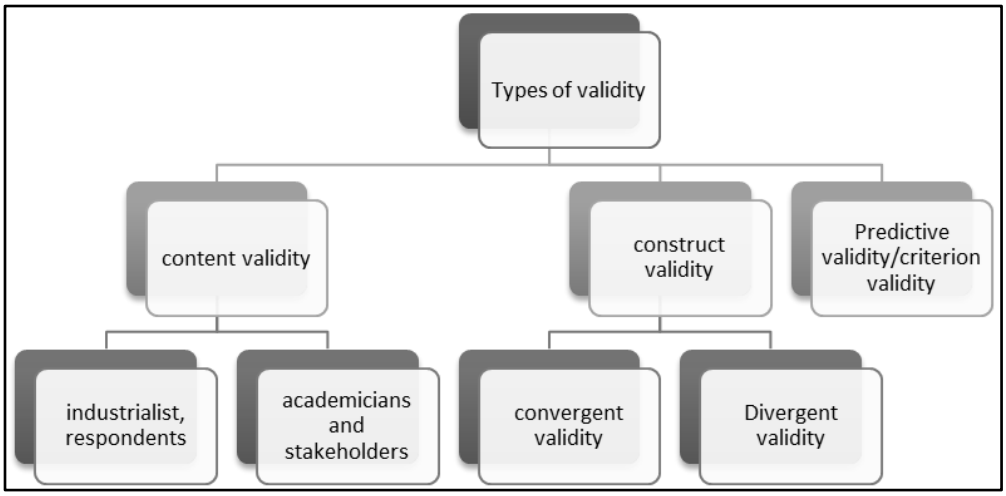

Figure 2.2: Types of Validity Tested in Confirmatory Factor Analysis

The validity of the questionnaire can be measured in multiple ways. Before moving to the different forms of validity it is very essential to understand what validity is. Validity of a research instrument or a construct denotes that the scale measures what it intends to be measured. In other words it measures the items measures the constructs which it is supposedly to measure and it does not measure something else. The different forms of validity are explained below:

A. Content Validity:

This is also known as the face validity of the measures or items being used in the questionnaire. As the name indicate this form of validity measure the facial appearance of a research instrument at a prima facie look. It is generally done by circulating the research instrument among the HR practitioners and academicians for their opinion. On a general basis the researcher supplies the structured questionnaires to a group of academicians in similar specialization and HR managers and industrialist of the study area.

The opinion suggested by them can be incorporated in the research instrument to make it more useful and inclusive. Content validity is a non-statistical measure which is based on the opinion of the stakeholders. It should be duly noted that the opinion of the researcher must be final in designing and redesigning the questionnaire by incorporating the necessary suggestions.

B. Criterion Validity or Predictive Validity:

This predicts the ability of the variable to impact another variable. This shows the predictive capacity of the item being involved in the study. A study with good criterion validity is said to have good predictive capacity. The predictive capacity of the constructs can be measured using the robustness of the r-square value. A study with significant effect size is considered to have good criterion or predictive validity. This shows the extension capacity of the study in the future purposes.

C. Construct Validity:

This is one of the most significant forms of validity. This form of validity test the goodness of each construct entered in the study. This reflects the overall goodness of the construct under the study. Construct validity can be ensured in multiple ways-

- Convergent Validity:

Convergent validity reflects the internal consistency of the item being used in the study. It reflects that all the statement entered in the study reflects and measures the intended construct. For e.g. if job satisfaction is being measured by 8 items and job commitment is being measured by 7 items ensuring convergent validity proves that none of the item measuring job satisfaction should measure job commitment and none of the item measuring job commitment should either measure job satisfaction. This further means that the items should converge together to measure the underlying construct not the other or distinct construct. The convergent validity is ensured through average variance extracted. For a construct to ensure convergent validity the average variance extracted must be greater than 0.5.

- Divergent validity or discriminant validity:

In contrast with the convergent validity this describes the distinctness of each construct entered in the study. While it is said that items should converge, the constructs should diverge from one another. The divergent validity proves that each construct should be separate from one another and should have distinct purpose in the study. The test of divergent or discriminant validity is ensured through the square root of average variance extracted and inter correlation between the variables. It must be noted that the intercorrelation between every variable must be less than the square root of average variance extracted for each factor.

D. Unidimensionality:

This is another measure of validity which measures the unidirectional role of each item towards the measurement model. It delivers the individual role of each item in the measurement model. For an item to be considered unidimensionality the factor loading of the item must be greater than 0.40. Sometimes the unidimensionality of some of the item will be less than 0.40. It is the discretion of the researcher to retain such item in the measurement model. It is always better to delete such an item with lower factor loading, but again the relevance of the item in the measurement model plays a key role i deciding the inclusion or deletion in the measurement model.

The various forms of validity have been illustrated below:

Imagine a researcher intends to measure two constructs namely job satisfaction and job commitment. Job satisfaction is measured by 6 items taken from previous scales, whereas job commitment is measured using 5 statements. Job satisfaction is labeled as JS1, JS2 etc and job commitment is labeled as JC1, JC2 etc.

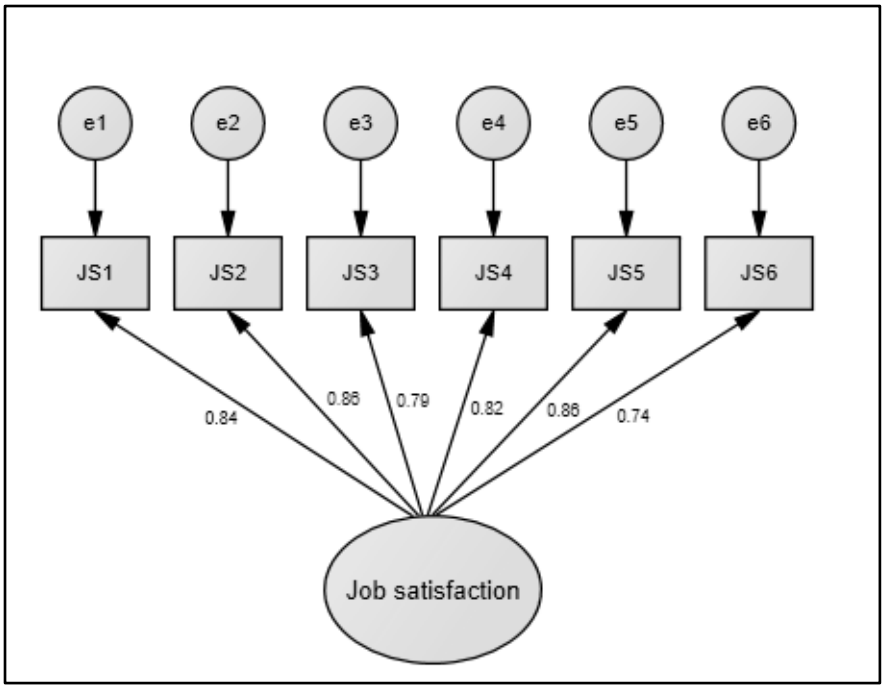

Figure 2.3:

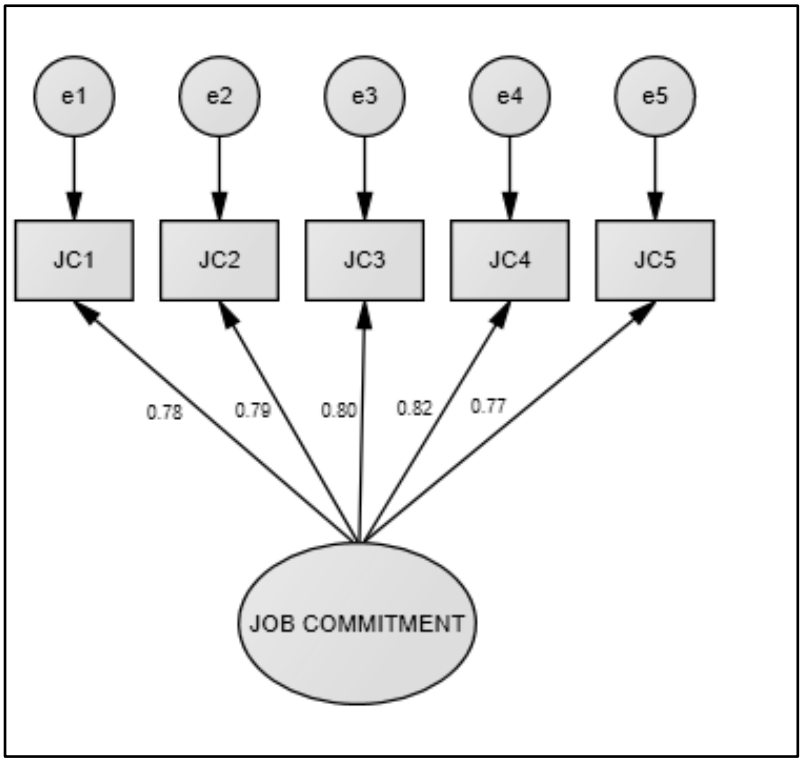

Figure 2.4:

The above figures namely figure 2.3 and figure 2.4 represents the basic measurement model representing job commitment and job satisfaction. The regression coefficients in the pathway show the factor loading of the items.

The factor loading indicates the contribution of each item towards the construct. It can be noted that in the above figures all the factor loadings are above 0.40. Hence it can be concluded that all the items have unidimensionality in measurement.

Next speaking about convergent validity, average variance extracted is the measure that is used to measure the convergent validity. Average variance extracted is the average of the squares of each factor loading of each construct.

Here it can be noted that the AVE value for each of the construct is greater 0.50. Hence there is sufficient convergent validity.

Table 2.1: Showing discriminant validity

Variables	Job Satisfaction	Job Commitment
Job satisfaction	**0.819**	
Job commitment	0.213	**0.741**

The above table shows the measures of discriminant validity. The diagonal values show the square root of average variance extracted. The values down to the diagonal values show the intercorrelation between the variables.

Here it is noted that the inter correlation between the variables are far below the square root of average variance extracted. Hence the constructs are considered to be distinct and hereby the divergent validity is ensured.

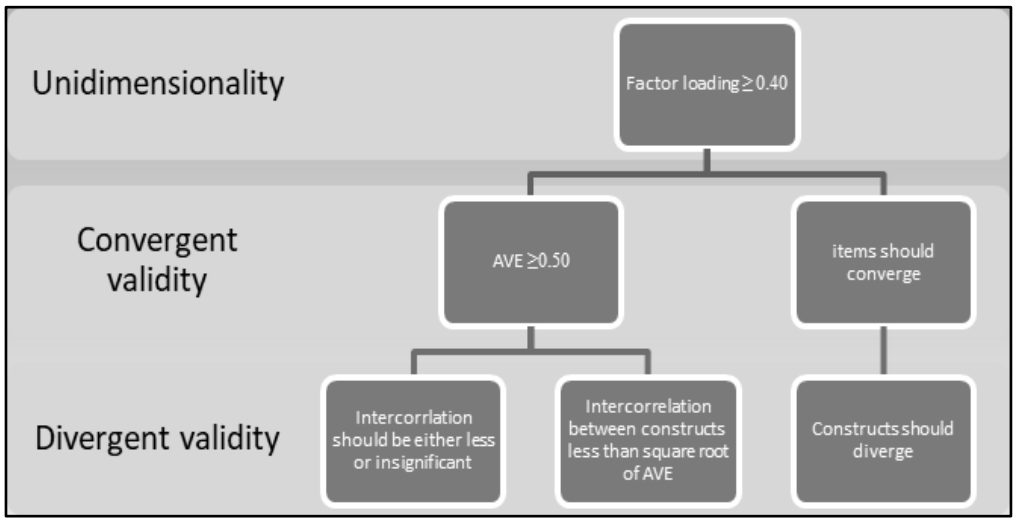

Figure 2.5: Conditions to Satisfy the Test of Validity

The drawing canvas of structural equation modeling- know your tools

Figure 2.6: Representing the Tools of AMOS

The above table shows the basic tools available in the IBM AMOS that will assist the researcher in drawing the structural model. It must be noted that the researcher must be friendly with using the various tools available for drawing.

It must be noted that the tools will be available in a comfortable left hand side whereas the drawing canvas will be available on the right hand side whose size can be adjusted according to the needs.

- The rectangle shows the unobserved variable or the factor
- The ellipse shows the observed variable or the item

- The combination of both ellipse and rectangle shows the entire object. If one clicks on that the measurement model will be on the canvas. With each additional click the new item will be added in the canvas
- The single pointed indicator shows the path
- The double pointed indicator shows the covariance connector
- Then there is the option to add the error term.
- Then there option to select one object at a time, all objects at a time and to deselect all the objects at a time
- The shape of the object can be altered
- The parameters can be moved or can be duplicate
- The eraser help to wipe out the unnecessary objects and mistakes if any in the canvas
- There is option to bring data into the canvas
- The path can be magnified using the telescope available and can be zoomed out
- After drawing the model only the researcher should bring the concerned data into the canvas.
- Remember to save the diagram from time to time
- Always remember not to click on the canvas frequently and multiple times. By clicking on the canvas frequently sometimes the software may get crashed or disrupted.

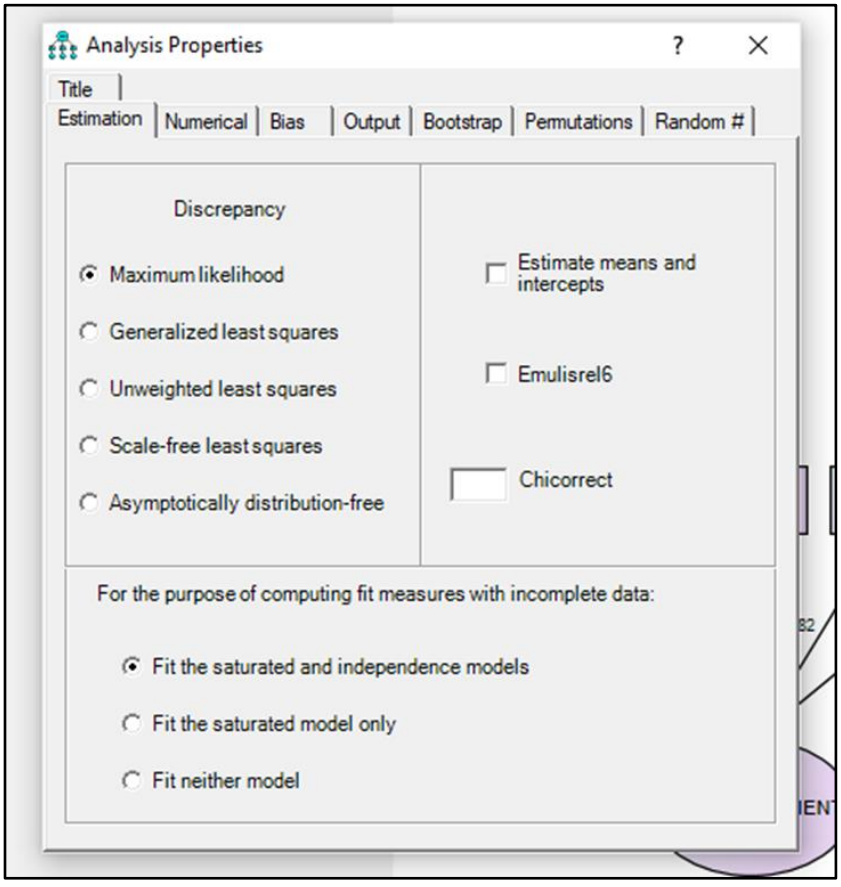

Figure 2.7: A glimpse of the Analysis Properties

All the analysis that you wish to do in AMOS comes under the option analysis properties. In the analysis properties tab there are many options available to choose from. Normally the data work under maximum likelihood option. The researcher has the discretion and should have thorough research methodology to choose from the alternatives. It should be noted that for all these options the shortcut keys are also available. The researcher has to be very much judicial in analyzing the tools, specific methodology, the exact data type of their data, the missing values if any etc for a robust data analysis procedure.

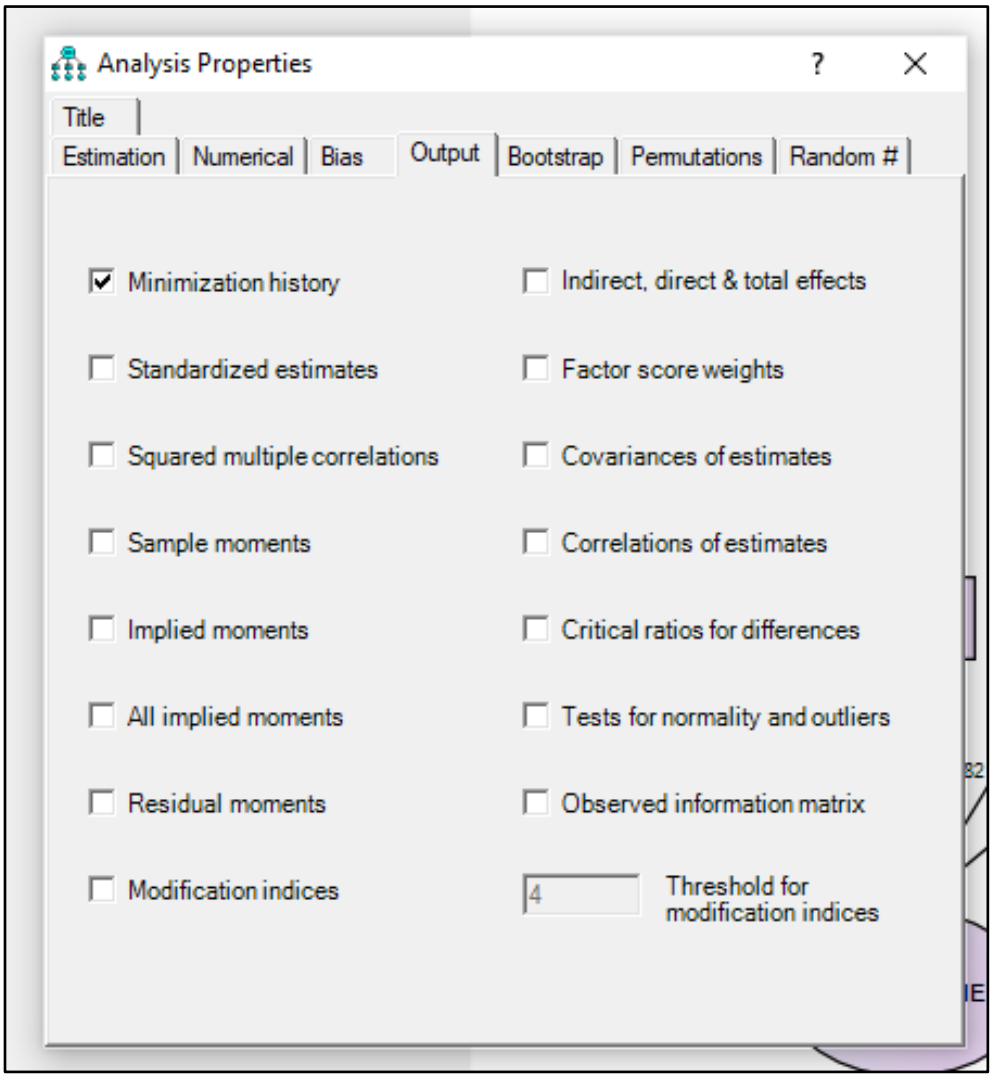

Figure 2.8: A Glimpses of Various Options in the Output Tab

The above screen shows the glimpse of the various outputs available to the researcher to select from. There are no robust rules to select the options. The researcher must be clear about the options he/she wants to get presented in a AMOS output. The researcher must be cautious that whatever output he wants to shows in the work and should not bombard with too much of information.

It must be noted that all the information shown in the thesis must be carefully interpreted otherwise the meaning of the entire model will be altered and the model will turn stale or useless for the stakeholders. It should be noted that there must be thorough justification even for using the modification indices the researcher must go with the proper research methodology in selecting the options for output. Not too much or not too low information should be portrayed by the researcher.

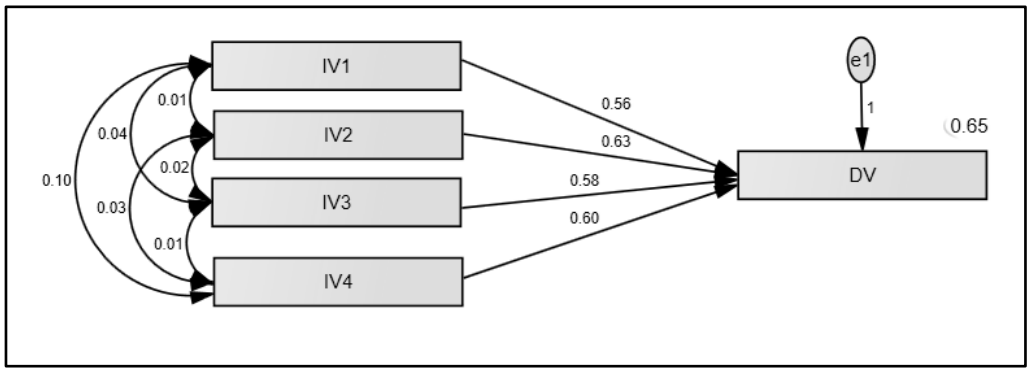

Figure 2.9: Structural Equation Modeling

The above table shows the simple result of a structural equation modeling. The Iv represents the independent variables and the DV represents the dependent variable. The e represents the error term which represents the variation that is not caused by the combined impact of the independent variable. In the above figure 2.9 it may be noted that the independent variable has a error term which explains the possibilities of error variation in the model. It should be noted that always the dependent variable must have an error term without which the entire model won't run in analysis.

The path between the independent and the dependent variable must be connected by a single headed arrow whereas the double headed arrows must be used by the researcher to connect the covariance.

The covariance must be connected for all the independent variables. Error in connecting covariance will have serious impact on the overall model performance. Once the model is drawn and the data has been loaded the researcher should select the options for running the analysis and should select which options to be included in the output generation.

- The output of the structural equation modeling can be interpreted in the following way-
- Any change in IV1 will lead to 56% change in the dependent variable
- Any change in IV2 will lead to 63% change in the dependent variable
- Any change in IV3 will lead to 58% changes in the dependent variable
- Any change in IV4 will lead to 60% changes in the dependent variable
- The R square value shows that 65% of variation in the dependent variable is explained by the change in the independent variable.
- It must be noted that the significance of all the impact and the relationship should be considered while interpreting the results and the diagram.

2.3 Understanding the Fit Indices of Structural Equation Modeling:

The structural equation modeling output throws a series of indices that the researcher must use to interpret the output. There are several indices available which will be used by the researcher to interpret the statistical fitness of the data. The various fit indices available can be termed as badness of fit indices and the goodness of fit indices. The researcher should carefully determine the fit indices to be reported as there are no robust measures to show which the fit indices to be shown are and which are the indices that need not be disclosed.

As a robust measure it must be noted that the goodness of fit indices should be more and more and the badness of fit indices must be very low below the threshold limits. Some of the goodness of fit indices include GFI, AGFI, NFI, TLI, CFI and some of the badness of fit indices includes measures such as RMSEA, RMR, SRMR etc.

The learners must note that there is no agreed consensus on what are the fit indices that have to be disclosed in the structural equation modeling. The researcher should judiciously select the fit indices and their threshold limits as desired by the propounders of structural equation modeling.

2.3.1 Identifying the Moderation and Mediation in a Structural Model:

One of the major purpose for which the researchers use structural equation modeling is its possibilities for testing moderation and mediation through this model. Mediation is known as intervening effect whereas moderation is called as interaction effect. Let us understand first about the test of mediation. A variable is said to be in mediation when the presence of such a variable alters the nature or direction of the preexisting relationship. In other words a variable is said to be mediating, when it affects the course of direct relationship between the independent and dependent variable. Learners should not that there are three types of mediation

A. Full Mediation: When the relationship between the variable have been fully affected or altered by the presence of the third variable it is called as full mediation. For eg; imagine a situation where X and Y are not at all related but the presence of a mediating variable M creates a significant relationship between X and Y. Then the M is considered to be a full mediator. Also the converse can be true. There will be significant relationship between X and Y and the presence of M makes this relationship either insignificant or makes it move in opposite direction. Here also the M is considered to be a significant full mediator. The occurrence of full mediation is a really myth situation in reality.

B. Partial Mediation: This is a more realistic situation in a structural model. In partial mediation the already existing relationship between the variables are getting affected by the presence of the mediator. In other words, when the existing relationship between the variables X and Y is being altered by the presence of mediator M it is considered to be partially mediated. Partial mediation between the variables can again occur in positive and negative manner. When the M increases the relationship between X and Y it is said to be positive partial mediation and if M diminishes the relationship between X and Y it is said to be negative partial mediation.

C. No Mediation: This is the most difficult situation in which the presence of mediation between the variables does not alter the relationship between the variables. The mediation is considered to be absent when the presence of a variable in the model will not touch the existing relationship between the variables. If there is no mediation between the variables, the researcher should make a thorough review of the literature to make sure about the possibility of including or excluding such a variable in the mediation model. The learners should always keep in mind that the presence of mediation should be determined based in the statistical significance of the result and the inclusion of a variable as a mediator or intervening variable should be based on the literature. The researchers should not set any arbitrary variable as a mediator and should run the model. Everything provided in the research should have enough support from the literature. The test of moderation is another important purpose of doing a structural equation modeling. The moderator or a condition is a variable that causes the relationship to happen between the variables. It is a variable or the presence of the conditions which affects the nature of relationship in varied ways. For e.g. Age sometimes act as a moderator in impacting the relationship between the salary and job satisfaction. Here the moderator is not a passer or intervention for the relationship to happen but it acts as a condition for which the relationship can be happen or seize to exist.

Figure 2.10: Theoretical Model for the Test of Mediation

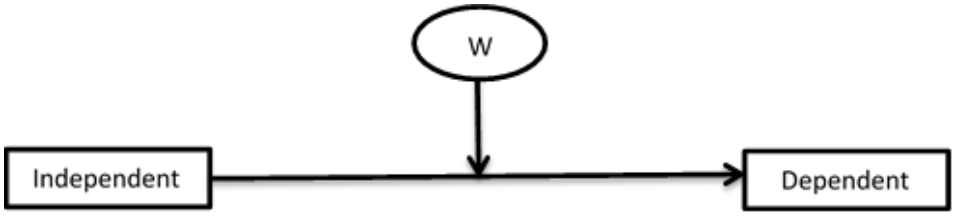

Figure 2.11: Theoretical Model for the Test of Moderation

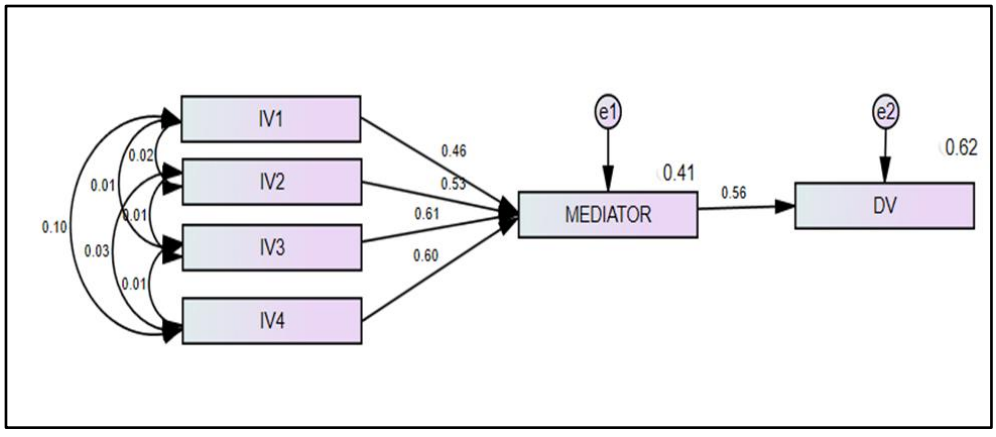

Figure 2.10: Illustrating a Structural Mediation Model

The above figures show the theoretical model difference between the test of mediation and moderation. It can be well noted that mediation act as an intervention and is placed in between the dependent and independent variables. This is the significance of mediation, it act as an intervening variable between the independent and dependent variable. It can be seen from the theoretical model of the moderation that it acts as a condition to carryover the relationship between the variables.

Here there is no intervention happening but a carry forward of the conditional effect between the variables. Both moderation and mediation have significant role in building and testing the relationship between the variables.

2.3.2 Points to Keep in Mind While Working on Structural Equation Modeling:

- There should be proper theoretical support for hypothesizing a structural model
- The mediation and moderation should be supported by proper reviews and justification
- All the variables entering the structural model should be measured on a continuous scale
- The results of confirmatory factor analysis is essential to prove the validity of the results obtained through the structural model
- The structural model should not be considered in the absence of confirmatory factor analysis
- The researcher should make use of all the available tools to make the appearance of the structural model more attractive.
- The researcher can test more complex models involving serial mediation, parallel mediation, conditional indirect effect and conditional interaction effect through the structural equation modeling.
- The possibilities of testing and verifying the unexplored relationship between the variables can be attempted through the structural equation modeling.
- The fit indices should be properly reported along with the necessary threshold limits.
- The researcher must be careful about the various outputs that must be shown in the structural output.
- Every structural relationship included in the model should be based on careful formulation of hypothesis and should ensure sufficient contribution to the literature and practical applications.

2.4 Conclusion:

This chapter had provided thorough insights about the various aspects of structural equation modeling. The researcher had discussed the various aspects of SEM, different types of validity, necessary conditions to satisfy the validity of the data, how to draw a structural diagram in the Amos canvas.

The authors have provided enough information about the various options available in developing and analyzing a SEM diagram. Further the researcher has also detailed the interpretation procedures for a SEM diagram. The knowledge gained through this chapter can be used by the researchers and budding scholars to using SEM as an efficient tool and platform to test complex relationships and sophisticated models.

2.5 References:

1. Weston, R., & Gore Jr, P. A. (2006). A brief guide to structural equation modeling. *The counseling psychologist*, 34(5), 719-751.
2. Thompson, B. (2000). Ten Commandments of structural equation modeling. In *US Dept of Education, Office of Special Education Programs (OSEP) Project Directors' Conference, 1998, Washington, DC, US; A previous version of this chapter was presented at the aforementioned conference and at the same annual conference held in 1999.* American Psychological Association.
3. Savalei, V., and Bentler, P. M. (2010). Structural equation modeling. *The Corsini encyclopedia of psychology*, 1-3.
4. Nachtigall, C., Kroehne, U., Funke, F., & Steyer, R. (2003). Pros and cons of structural equation modeling. *Methods Psychological Research Online*, 8(2), 1-22.
5. Lei, P. W., and Wu, Q. (2007). Introduction to structural equation modeling: Issues and practical considerations. *Educational Measurement: issues and practice*, *26* (3), 33-43.
6. Hoyle, R. H. (1995). The structural equation modeling approach: Basic concepts and fundamental issues.
7. Chou, C. P., and Bentler, P. M. (1995). Estimates and tests in structural equation modeling.