PATH ANALYSIS VS. SEM A PRACTICAL APPROACH USING THE MODEL OF HIVAIDS

Chapter · January 2021

1 author:

Raquel A Benavides-Torres
Autonomous University of Nuevo León

UNIVERSIDAD AUTÓNOMA DE NAYARIT
ÁREA ACADÉMICA DE CIENCIAS DE LA SALUD
UNIDAD ACADÉMICA DE CIENCIAS DE LA SALUD

Tepic, Nayarit; 20 de febrero de 2021

PhD. Raquel A. Benavides-Torres

P R E S E N T E

Por medio de la presente se le informa que después de pasar por un arbitraje estricto de doble ciego, su capítulo en idioma inglés titulado “Path Analysis Vs. SEM: a practical approach using the Model of HIV/AIDS Prevention In Adolescents”, incluido en el libro “Use of Nursing Models and Theories from a Sexuality Perspective”, bajo la coordinación de Martha Ofelia Valle Solís, fue **ACEPTADO** para su publicación durante el año 2021 por la Editorial Nova Science Publishers, la cual cuenta con circulación y reconocimiento a nivel internacional y esta indexada en Scopus.

A T E N T A M E N T E

Dra. Martha Ofelia Valle Solís

rtha Ofelia Valle Solis, DSN
Editor

Use of Nursing Models and Theories from a Sexuality Perspective

nova
science publishers

ISBN 978-1-53619-180-6

9 781536 191806

www.novapublishers.com

Use of Nursing Models and Theories from a Sexuality Perspective • Solis

NOVA

Martha Ofelia Valle Solis, DSN
Editor

Nova Medicine & Health

Use of Nursing Models and Theories from a Sexuality Perspective

Nursing - Issues, Problems and Challenges

NOVA

Library of Congress Cataloging-in-Publication Data

ISBN: 978-1-53619-180-6

Published by Nova Science Publishers, Inc. † New York

CONTENTS

In: Use of Nursing Models and Theories … ISBN: 978-1-53619-180-6
Editor: Martha Ofelia Valle Solis

Chapter 13

PATH ANALYSIS VS. SEM: A PRACTICAL APPROACH USING THE MODEL OF HIV/AIDS PREVENTION IN ADOLESCENTS

***Raquel A. Benavides-Torres*[*], PhD**
School of Nursing, and Center for Research and Development in Health Sciences (CIDICS), Universidad Autónoma de Nuevo León, Monterrey, Nuevo León, México

ABSTRACT

Conceptual models are very important for building up the scientific body of knowledge in nursing. In the development of new models, researchers use path analysis and structural equation modeling, therefore this paper used the Model of HIV/AIDS prevention in Adolescents to evaluate the application of both techniques. The principal use of path analysis is to test the fit of correlations against two more causal models that are the focus of the researcher. Based on formal and informal theory, previous research, and logic, the researcher can develop the path model. An extension of the general linear model is SEM, which is a more powerful technique with the similar purpose of conducting multiple regressions. This is a more potent technique than path

[*] Corrersponding Author's E-mail: rabenavi@gmail.com.mx; raquel.benavidestr@uanl.edu.mx.

analysis because it takes into account the modeling of third factor explains 11.22% of the total variance. Cronbach's Alpha was .84. The methodology used for the translation and adaptation of this instrument was adequate.

Keywords: multiple regression, path analysis, structural equation modeling, model development

INTRODUCTION

Conceptual models in nursing are recognized as important and useful products for building up the scientific body of knowledge in nursing (Fitzpatrick & Whall, 1996). The growing interest in the development of new models to understand this phenomenon's importance in nursing science implies that the study of its construct and concepts can only be measured using indicators of causality (Aggleton & Chalmers, 2000). Statistic analysis of relationships among variables is one of the main focuses of empirical research in the health sciences. The model most frequently used is multiple regression. First used by Pearson (1980), its general purpose is to aid in learning more about relationships between several independent or predictor variables and a dependent or criterion variable. In the development of new models, researchers sometimes use inappropriate statistical techniques when analyzing their data (Williamson, 2003). The two most frequently used methods in the literature are path analysis and structural equation modeling. Therefore, this paper will assess and evaluate the application of path analysis and structural equation modeling. In this paper we will also use concepts from the model of HIV/AIDS prevention in Mexican adolescents (MAPA; Benavides, 2006). MAPA has four major concepts: individual, microsystem, and psychological influences that are directly or indirectly related to sexual health behaviors regarding HIV/AIDS prevention. These concepts will be used as exemplars for path analysis and structural equation modeling.

PATH ANALYSIS

Path analysis makes many aspects of multiple regression more understandable and is a better choice for explaining the analysis of non-experimental data (Keith, 2006). The principal use of path analysis is to test the fit of correlations against two more causal models that are the focus of the researcher (Pedhazur, 1982). The first step in this process is to conduct a review on the variables that have been shown to be associated with the dependent variable. To illustrate this association, the researcher should create a path model.

A path model is a diagram that illustrates independent, intermediary, and dependent variables with arrows. Based on formal and informal theory, previous research, and logic, the researcher can develop the path model (Wright, 1921; 1934).

Using variables from MAPA, we can create a path model which considers the following variables: sexual behaviors as regards HIV/AIDS prevention, parent–child communication about sex, and HIV/AIDS knowledge. This exemplification of path analysis only included three variables from MAPA, but it is important to mention that the researcher can include as many variables as they find to be related to the outcome.

The second step in the model is to check the status of the model. This step includes running the correlations between the variables and deciding on the paths to be solved. Therefore, we can say that in this case the researcher wants to determine the effects of "parent–child communication about sex" on "sexual behaviors regarding HIV/AIDS prevention". However, the researcher may be worried that HIV/AIDS knowledge might affect the above-mentioned variables. In this path model, we consider what is called *weak causal ordering* (Belnap, Perloff, & Xu, 2001), meaning that the path from communication to sexual behaviors does not affirm that one is a direct cause of the other. Therefore, we can assume that knowledge mediates this relationship. For this to be true, we need to use the following rule: "The correlation between two variables X and Z is equal to the sum of the products of all paths for each possible tracing between X and Z. These tracings include all possible routes between X and Z, with the exceptions that the same variables are not entered twice per tracing and a variable does not both enter and exit through an arrowhead" (Kenny, 1979, p.30). Thus, the correlation between HIV/AIDS knowledge and sexual behaviors will equal to "C" plus the product of "A" and "B". The model may be interpreted as demonstrating the effects of "HIV/AIDS knowledge" and "parent–child communication about sex" on "sexual behaviors regarding HIV/AIDS prevention" along with the effects of the two independent variables.

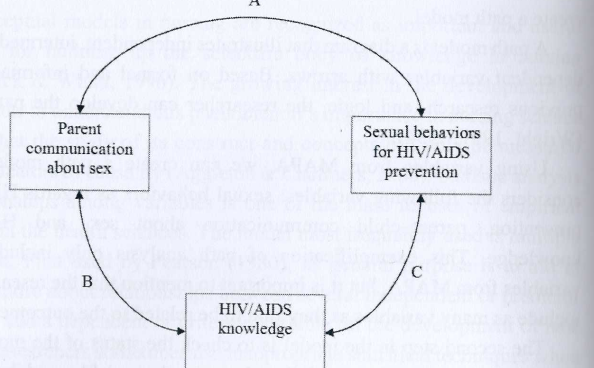

Figure 1. Path analysis with three variables from MAPA[1].

The third step involves measuring the variables in the model. After drawing the model and obtaining correlations between the variables, the researcher may make use of a multiple regression to show the standardized regression coefficients in the analysis (β). Using a statistical package such as SPSS (Statistical Package for Social Science), we can regress "sexual behaviors regarding HIV/AIDS prevention" on "parent–child communication about sex" and "HIV/AIDS knowledge". The information in the

[1] Dependent variable: sexual behaviors for HIV/AIDS prevention; Independent variables: parent communication about sex and HIV/AIDS knowledge; A, B, and C are the correlation coefficient between the variables. Note: using curved lines represent correlations.

output can then be used to interpret the results. First, we need to check whether the model is statistically significant using the F value; if so, we can interpret the results for each variable (Keith, 2006). Using evidence from the literature, we can assume that the β of the path analysis may have a low or moderate effect on sexual behaviors surrounding HIV/AIDS prevention because MAPA conceptually shows that more variables influence the outcome variable. Therefore, it is important to consider all the disturbance variables that are not included in the model presented in Figure 1. Those effects from the variables that are unmeasured (disturbances) are the residuals from a multiple regression (Wright, 1960a). The calculation of unmeasured effects should be included in the path model (Figure 2). In this path, we can observe that, at this step, we can have arrows that show the effect (β) of one variable on the other variable.

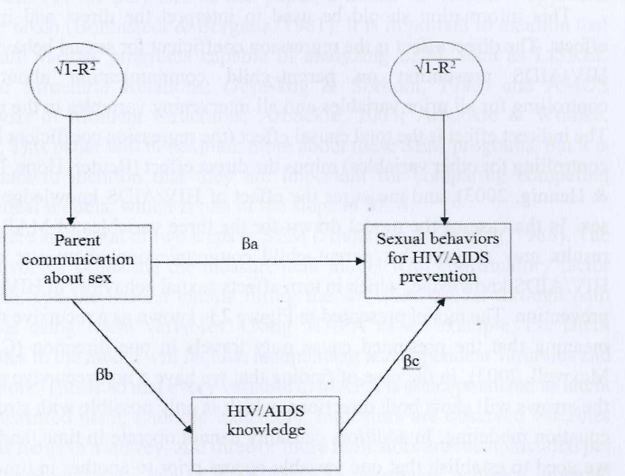

Figure 2. Path analysis using simultaneous multiples and disturbances[2].

[2] βa, βb and βc = standardized coefficient for the regression; the circles indicated disturbances;

The last step is to estimate the model. At this point, simultaneous multiple regressions were used. In the first regression, both independent variables are included, and the R^2 from this first analysis is used to calculate the residuals or disturbances on sexual behaviors regarding HIV/AIDS prevention. In the second regression, only the analysis for HIV/AIDS knowledge was regressed on the outcome. In this case, the R^2 for this regression was used to calculate the residuals or disturbances on parent–child communication about sex. Those disturbances or residuals are variables that the researcher wishes to include in the path model, but there are no measures included in the data (Wright, 1960b). After obtaining all the results from the regression analysis, we need to consider the non-standardized (b), standardized coefficients (β), variance (R^2), and the values for the disturbances.

This information should be used to interpret the direct and indirect effects. The direct effect is the regression coefficient for sexual behaviors in HIV/AIDS prevention on parent–child communication about sex, controlling for all prior variables and all intervening variables in the model. The indirect effect is the total causal effect (the regression coefficient before controlling for other variables) minus the direct effect (Reuter, Hope, Netter, & Hennig, 2003), and measures the effect of HIV/AIDS knowledge about sex. In the case of the model drawn for the three variables of MAPA, the results may suggest that parent–child communication about sex affects HIV/AIDS knowledge, which in turn affects sexual behaviors in HIV/AIDS prevention. The model presented in Figure 2 is known as a recursive model, meaning that the presumed cause only travels in one direction (Cole & Maxwell, 2003), In the case of finding that we have a non-recursive model, the arrows will show both directions, which is only possible with structural equation modeling. In addition, causality cannot operate in time, and so, if we need to establish that one variable occurs prior to another in time, it is easier to draw a path, yet this is impossible to do with simple path analysis (Sheder, 1995). Therefore, the following part of this paper will focus on an explanation of structure equation modeling, which is a more advanced method that allows for this type of analysis.

SRTUCTURE EQUATION MODELING (SEM)

An extension of the general linear model is SEM (Structure Equation Modeling), which is a more powerful technique with the similar purpose of conducting multiple regressions (Arminger, Clogg, & Sobel, 1995). This is a more potent technique than path analysis because it takes into account the modeling of interactions, nonlinearities, correlated independents, measurement error, correlated error terms, multiple latent independents, and one or more latent dependents with multiple indicators (Fan, Thompson, & Wang, 1999). This procedure is recognized as confirmatory rather than exploratory with three different approaches: a strictly confirmatory approach, an alternative model approach, and a model development approach. For the purposes of this paper, a model development approach will be taken (Bohmstedt & Borgatta, 1981). It is important to mention that there are various programs capable of analyzing SEM, such as LISREL (Linear Structural Relations; (Jöreskog & Sörbom, 1996) and AMOS (Analysis of Moment Structures; Arbuckle, 2003; Arbuckle & Wothke, 1999). This paper will not explain more about these SEM programs, but it is important to mention that they are important for comparing competing theoretical models, which is one of the steps in SEM.

There are a total of two steps in SEM (Anderson & Gerbing, 1988). The first involves validating the measurement model with confirmatory factor analysis, and the second entails fitting the structural model through path analysis using latent variables. Using MAPA as an example, the latent variables in the model will include independent and dependent variables and mediators. This is so that every variable in MAPA is conceptualized as latent and measured using multiple indicators. Indicators are observed variables such as items in a survey, and three or more indicators are recommended per latent variable, since this is more likely to accurately estimate modeled errors (Kline, 1998). Knowing all these terms, we can then simulate a path analysis using the four main concepts in the model: individual influences, microsystem influences, psychological influences, and sexual behavior in HIV/AIDS prevention.

The last step is to estimate the model. At this point, simultaneous multiple regressions were used. In the first regression, both independent variables are included, and the R^2 from this first analysis is used to calculate the residuals or disturbances on sexual behaviors regarding HIV/AIDS prevention. In the second regression, only the analysis for HIV/AIDS knowledge was regressed on the outcome. In this case, the R^2 for this regression was used to calculate the residuals or disturbances on parent–child communication about sex. Those disturbances or residuals are variables that the researcher wishes to include in the path model, but there are no measures included in the data (Wright, 1960b). After obtaining all the results from the regression analysis, we need to consider the non-standardized (b), standardized coefficients (β), variance (R^2), and the values for the disturbances.

This information should be used to interpret the direct and indirect effects. The direct effect is the regression coefficient for sexual behaviors in HIV/AIDS prevention on parent–child communication about sex, controlling for all prior variables and all intervening variables in the model. The indirect effect is the total causal effect (the regression coefficient before controlling for other variables) minus the direct effect (Reuter, Hope, Netter, & Hennig, 2003), and measures the effect of HIV/AIDS knowledge about sex. In the case of the model drawn for the three variables of MAPA, the results may suggest that parent–child communication about sex affects HIV/AIDS knowledge, which in turn affects sexual behaviors in HIV/AIDS prevention. The model presented in Figure 2 is known as a recursive model, meaning that the presumed cause only travels in one direction (Cole & Maxwell, 2003), In the case of finding that we have a non-recursive model, the arrows will show both directions, which is only possible with structural equation modeling. In addition, causality cannot operate in time, and so, if we need to establish that one variable occurs prior to another in time, it is easier to draw a path, yet this is impossible to do with simple path analysis (Sheder, 1995). Therefore, the following part of this paper will focus on an explanation of structure equation modeling, which is a more advanced method that allows for this type of analysis.

Srtucture Equation Modeling (SEM)

An extension of the general linear model is SEM (Structure Equation Modeling), which is a more powerful technique with the similar purpose of conducting multiple regressions (Arminger, Clogg, & Sobel, 1995). This is a more potent technique than path analysis because it takes into account the modeling of interactions, nonlinearities, correlated independents, measurement error, correlated error terms, multiple latent independents, and one or more latent dependents with multiple indicators (Fan, Thompson, & Wang, 1999). This procedure is recognized as confirmatory rather than exploratory with three different approaches: a strictly confirmatory approach, an alternative model approach, and a model development approach. For the purposes of this paper, a model development approach will be taken (Bohmstedt & Borgatta, 1981). It is important to mention that there are various programs capable of analyzing SEM, such as LISREL (Linear Structural Relations; (Jöreskog & Sörbom, 1996) and AMOS (Analysis of Moment Structures; Arbuckle, 2003; Arbuckle & Wothke, 1999). This paper will not explain more about these SEM programs, but it is important to mention that they are important for comparing competing theoretical models, which is one of the steps in SEM.

There are a total of two steps in SEM (Anderson & Gerbing, 1988). The first involves validating the measurement model with confirmatory factor analysis, and the second entails fitting the structural model through path analysis using latent variables. Using MAPA as an example, the latent variables in the model will include independent and dependent variables and mediators. This is so that every variable in MAPA is conceptualized as latent and measured using multiple indicators. Indicators are observed variables such as items in a survey, and three or more indicators are recommended per latent variable, since this is more likely to accurately estimate modeled errors (Kline, 1998). Knowing all these terms, we can then simulate a path analysis using the four main concepts in the model: individual influences, microsystem influences, psychological influences, and sexual behavior in HIV/AIDS prevention.

The first step is the measurement of the model using confirmatory factor analysis (CFA). The main focus of CFA is to confirm that the indicators short themselves as factors (Maruyama, 1997). In other words, CFA places test items that are highly correlated with one another as one factor and those that are less correlated as a separate factor. The use of previous research and relevant theory is used to confirm the factors or construct the underlying measures (Keith, 2006). Similarly, as in the first step in path analysis, the researcher should now map out the variables of interest into a path. The simulated model for CFA using MAPA will show which subtests are decided on for measuring each latent variable. In this case, individual influences, microsystem influences, psychological influences, and sexual behaviors in HIV/AIDS prevention.

The initial model for MAPA has four latent variables and 12 subsets using the actual measures or measured variables. The arrows in the MAPA figure below show the causal assumptions underlying testing models. For example, in the case of individual influences of adolescents, the CFA will determine which of the three scores out of HIV/AIDS knowledge, being of female gender, or age, has the primary influence on the latent variable. Continuing with this example, every measure, as mentioned before, has a score for unreliability, which is represented in the above figure by "u". After obtaining this path, we can then proceed to use CFA to assess the role of the measurement error in the model, to validate a multifactorial model, and to determine group effects on the factors (Chen, Bollen, Paxton, Curran, & Kirby, 2001). This model is often called a "null model", in which the covariances in the covariance matrix for the latent variables are all summed to zero (Silvia & MacCallum, 1988). The standardized covariances (correlations) among the factors are used when the researcher wishes to compare competing models. In addition, there is the possibility that two of the latent variables are highly correlated, so both factors can be used as one when constructing a new model (Hu, & Bentler, 1999). For example, if individual and microsystem influences are highly correlated, the researcher may want to put them together to form a new model with three latent variables, but the model with two combined latent variables will have six measures. If the fit of the measurement model is found to be acceptable ($\beta >$ 0.05, $p < 0.05$), then the researcher may proceed to the second step (Bohmstedt & Borgatta, 1981).

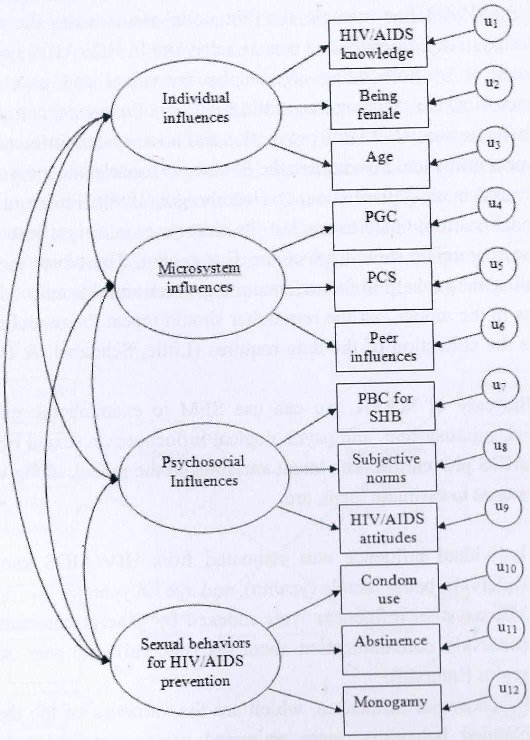

Figure 3. Full latent variables from MAPA in a SEM model.[3].

[3] The curve lines represent relationship among the latent variables; PBC = perceive behavioral control; SHB = sexual health behaviors; u = unreliability

The second step consists of testing the structural model by comparing its fits with those of different structural models. Here, the researcher should focus on whether the model as it is drawn is correctly specified (Bohmstedt & Borgatta, 1981). For example, the first model assumes that the only way that microsystem influences and sexual behaviors in HIV/AIDS prevention are related is by both being affected by individual and psychological influences. However, if using longitudinal data, another model can show that sexual behaviors in HIV/AIDS prevention and microsystem influences have a reciprocal influence. To compare the fit of those models, the most common statistic used is chi-square (Lee, & Hershberger, 1990). For example, the first model has more parsimony, but the chi-square is insignificant. In this case, the researcher should reject the first model. Therefore, the use of longitudinal data is helpful for understanding which variable antecedes other variables in the model, but the researcher should repeat the model as many times as the collection of the data requires (Little, Schnabel, & Baumert, 2000).

In the case of MAPA, we can use SEM to examine the effects of individual, microsystem, and psychological influences on sexual behaviors in HIV/AIDS prevention. The latent variables in the model, along with the variables used to estimate them, are:

11. Individual influence was estimated from HIV/AIDS knowledge (interval), being female (yes/no), and age (in years).
12. Microsystem influences were indexed by general communication (interval), communication about sex (interval), and peer influence scales (interval).
13. Psychosocial influences, which are the variables of the theory of planned behaviors, were estimated using perceived behavioral control (interval), subjective norms (interval), and attitudes (interval) about sexual behaviors in HIV/AIDS prevention.
14. Sexual behaviors regarding HIV/AIDS prevention was measured with three additional scales on condom use, abstinence, and monogamy. All these are interval scales.

At this point, when there is more than one measure for each variable, the use of composites is very helpful. This is based on an assumption from multiple regression that implies the prediction of outcome variables from the optimally weighted combinations of the measured variables (Berk, 2003). Similarly, in the case of SEM, all the latent variables are optimally weighted combinations of the measured variables (Keith, 2006). If this is needed, it is important to do it before using an SEM program to reverse the scores of the measured variables, because this will help in the final interpretation. Both SEM programs mentioned above used measured variables (indicators) to run the model, not latent variables. This explains why they are called "latent", because they are estimations from the measured variables. For example, in the case of the latent variable *microsystem influences*, the program used the scores of general communication, communication about sex, and peer influence scales.

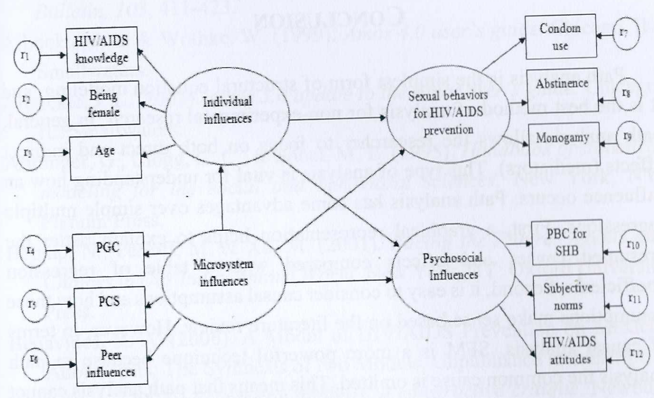

Figure 4. Full latent variables from MAPA in an SEM model[4].

For the model presented in Figure 4, "r" is the unique error variance of each measured variable as r_1 all the way through r_{12}. This representation

[4] PGC = parent general communication; PCS = parent communication sex; PBC = perceived behavioral control; SHB = sexual health behavior; r = unique error variance.

includes the error that every measurement necessarily has. For example, the greater the error in the scale of parent–child communication about sex, the less the adolescents' score on that measure is a result of true variation and the less reliable the measurement. Because of this, the error effect is very important for SEM, because we can underestimate the true effect that a variable has on sexual behaviors in HIV/AIDS prevention. This means that, when researchers do not take into account the error effect, they reduce the effect of one variable on another (Teàs, 2000). Please note that, when we talk about an error term, we are not referring to the disturbance terms or residuals as discussed before in path analysis, which reflects the unexplained variances in the latent variable. In Figure 4, however, we did not include disturbances (residuals), so we can assume that those will be present when running the SEM program.

CONCLUSION

Path analysis is the simplest form of structural equation modeling, and it is the best method of analysis for non-experimental research. In general, path analysis allows the researcher to focus on both direct and indirect effects (mediators). This type of analysis is vital for understanding how an influence occurs. Path analysis has some advantages over simple multiple regressions. First, a graphical representation helps to explain better the presumed causes and effects compared with a table of regression coefficients. Second, it is easy to consider causal assumptions and how those assumptions make sense based on the literature review. However, in terms of causal analysis, SEM is a more powerful technique because in path analysis the common cause is omitted. This means that path analysis cannot determine whether a variable affects both the presumed cause and the presumed effect. Here, the answer depends on the power that SEM has when using longitudinal data to explain what antecedes the outcome variable. Nevertheless, assumptions about both models have not been mentioned here; these can be found in various books by authors such as Munro (2001, p. 355–404), who provide theoretical and statistical assumptions. In conclusion,

there are more advantages in using SEM over path analysis: 1) more flexible assumptions about multicollinearity, 2) having multiple indicators per latent variable for reducing measurement error (confirmatory factor analysis), 3) a graphical representation of the modeling interface, 4) modeling multiple dependent and mediator variables, and 5) the ability to handle difficult data (abnormal or incomplete data).

REFERENCES

Aggleton, P. and Chalmers, H. (2000) *Nursing models and nursing practice*. (2nd ed.) Basingstoke: Macmillan.

Anderson, J. C. & Gerbing, D. W. (1988). Structural equation modeling in practice: a review and recommended two-step approach. *Psychological Bulletin, 103*, 411-423.

Arbuckle, J. L. & Wothke, W. (1999). *Amos 4.0 user's guide.* Chicago, IL: Smallwaters.

Arbuckle, J. L. (2003). *Amos 5.0 update to the Amos user's guide.* Chicago, IL: Smallwaters.

Arminger, G., Clogg, C. C., & Sobel, M. E. (1995). *Handbook of statistical modeling for the social and behavioral sciences.* New York, NY: Plenum Press.

Belnap, N., Perloff, M, & Xu, M. (2001). *Facing the Future: Agents and Choices in Our Indeterminist World.* New York, NY: Oxford University Press.

Benavides, R. A. (2006). A Model of HIV/AIDS Prevention on Mexican Adolescents: The Synthesis of two Models. Unpublished manuscript.

Berk, R. A. (2003). *Regression analysis, a constructive critique*. Newbury Park, CA: Sage Publications.

Bohmstedt G. W. & Borgatta, E. F. (1981). *Social Measurement.* Thousand Oaks, CA: Sage Publications.

Chen, F., Bollen, K. A., Paxton, P., Curran, P., & Kirby, J. (2001). Improper solutions in structural equation models: Causes, consequences, and strategies. *Sociological Methods and Research 29,* 468-508

Cole, D. A. & Maxwell, S. E. (2003). Testing Mediational Models With Longitudinal Data: Questions and Tips in the Use of Structural Equation Modeling, *Journal of Abnormal Psychology, 112* (4), pp.

Fan, X., Thompson, B., & Wang, L. (1999). Effects of sample size, estimation method, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling, 6,* 56-83.

Fitzpatrick J. J. & Whall A. L. (1996). *Conceptual models of nursing: Analysis and application* (3rd ed.). Stamford, CT: Appleton & Lange.

Hu, L. & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling 6*, 1-55.

Jöreskog, K. G. & Sörbom, D. (1988). *A program for multivariate data screening and data summarization. User's Guide (2nd Ed.).* Chicago, IL: Scientific Software International.

Keith, T. (2006). *Multiple regression and beyond*. Boston, MA: Pearson Education.

Kenny, D. A. (1979). *Correlation and causality.* New York, NY: Wiley.

Kline, Rex B. (1998). *Principles and practice of structural equation modeling*. NY: Guilford Press.

Lee, S. & Hershberger, L. (1990). A simple rule for generating equivalent models in covariance structure modeling. *Multivariate Behavioral Research, 25,* 313-334.

Little, T. D., Schnabel, K. U., &Baumert, J. (2000) *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples*. Mahwah, NJ: Lawrence Erlbaum Associates.

Maruyama, G. M. (1997). *Basics of structual equation modeling.* Thousand Oaks, CA: Sage Publications

Munro, B. H. (2001). *Statistical methods for health care research* (4th ed). Philadelphia, PA: Lippincott Williams & Wilkins.

Pearson, K. (1908). On the generalized probable error in multiple normal correlation. *Biometrika*, *6*, 59-68.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research,* (2nd ed). NY: Holt. Chapter 15 (pp. 577-635).

Reuter, M., Hope, M., Netter, P., & Hennig, J. (2003). Structural equation modeling (SEM) vs. configural frequency analysis (CFA): An empirical approach. *Psychology Science, 45*, 280-297.

Sheder, J. (1995). On inferring causation from correlation. *Measurement and Evaluation in Counseling and Development,* 27, 564-571.

Silvia, E. & MacCallum, R. C. (1988). Some factors affecting the success of specification searches in covariance structure modeling. *Mutlivariate Behavioral Research 23,* 297-326.

Teas, R. K. (2000). Path and LISREL. *Marketing Research, 12*, 20-22.

Williamson G. R. (2003). Misrepresenting random sampling? A systematic review of research papers. *Journal of Advanced Nursing.* 44, 278-288.

Wright S. (1934). *The method of path coefficients. Annals of Mathematical Statistics,* 5, 161-215.

Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, 20, 557-585.

Wright, S. (1960a). The treatment of reciprocal interaction, with and without lag, in path analysis. *Biometrics*, 16, 423-445.

Wright, S. (1960b). Path coefficients and path regressions: Alternative or complementary concepts? *Biometrics*, 16. 189-202.