D. M. Akpa

Selected Topics in Medical Statistics

Written in Honour of Professor Elijah Afolabi Bamgboye,
Eminent Professor of Medical Statistics
on the commemoration of his 70th Birthday

Edited by
Adebowale S. A. (Ph.D)
Akinyemi J. O. (Ph.D)

Chapter 23.

Application of Structural Equation Modelling to Public Health Concern on Psychosocial Functioning and Quality of Life of Adolescents in Nigeria

Onoja Matthew ***Akpa***[1*], *Kayode Raphael Fowobaje*[1,2] *and Olaniyi Matthew Olutola*[1]

[1]Department of Epidemiology and Medical Statistics, College of Medicine, University of Ibadan; [2]WHO collaboration Centre, Department of Paediatrics, University College Hospital (UCH), Ibadan, Nigeria

onojamatthew@yahoo.co.uk; +2348073732015

Summary

Structural equation modelling (SEM) is an efficient statistical method for the analysis and evaluation of complex research objectives involving relationships between observed and unobserved (latent) variables. Unfortunately, its application in Public Health research has been very limited in Nigeria. This paper documents basic theories, principles and application of SEM to research in Public Health with a focus on the psychosocial functioning and quality of life of adolescents in Nigeria. The cross-sectional study was conducted among adolescents (n=983) in Benue state, Nigeria. The study presents a step-by-step approach to SEM and also describes variety of research questions that SEM can be used to answer in public health domain with empirical application of SEM on psychosocial functioning and quality of life. Results of the empirical application showed that poor psychosocial functioning negatively impacts the quality of life of adolescents in Nigeria. There is need to encourage researchers in public health to apply SEM in their research.

23.1 Introduction

Structural equation modelling (SEM) is a multivariate statistical methods used to estimate simultaneously, a given system of hypothesized relationships among observable and latent variables in order to determine whether these associations are consistent with an obtained sample data (Schumacker and Lomax, 2010). SEM combines three separate mathematical and statistical analysis methods. This includes path analysis, factor analysis, and simultaneous equation modelling (Kline, 2011).

The work of Karl Jöreskog provided bridge to earlier works in path and factors analysis (Cudeck et al., 2001). His earliest work in the development of SEM and related methods such as confirmatory factor analysis (CFA) can be linked to other previous studies on maximum likelihood and factor analysis which created basic measurement tool that is common to all SEM softwares (Kaplan, 2000). However, modern day SEM techniques has evolved beyond the study of just measurement models to become a mixture of factor analysis and path analysis (Kaplan, 2000). In the past two decades SEM has been seen as the most important contribution of statistics to the social and behavioural sciences (Schumacker and Lomax, 2010). Unfortunately, the application of SEM in public health research has been very limited especially in sub-Saharan Africa due to lack of documented guide to its application in relevant research problems in this setting.

In the present paper, we documented the basic theories of SEM, its principles and application. We also further documented how SEM could be applied in major areas of public health and an empirical application of SEM to a public health survey on psychosocial functioning and quality of life of adolescents in Nigeria.

23.2 The Principles and Theories of Structural Equation Models

23.2.1 *The principles of Structural Equation Models*

There are two common components to a SEM: the measurement model (Fig.1a) and the structural model (Fig.1b) (Schumacker and Lomax, 2010). The measurement model investigates the relationships among a set of indicator variables and a predetermined number of latent variables. Indicator variables are those collected in the researcher's measurement instrument or questionnaire, while latent variables are constructs that may not have been captured by human measurement. The association among the indicator variables and the latent variables in a model are established a priori and tested against an empirical data set to assess if the hypothesized measurement relationships match the empirical data set that have been collected. Apart from the associations analysed by measurement model, the structural part of the model also analyses a series of a priori relationships established between latent variables (Schumacker and Lomax, 2010).

The basic objective of SEM is to provide a means of estimating the relationships among the underlying constructs of a hypothesized model. This methodology differs from others statistical methods such as regression analysis and contingency table analysis in that it focuses not on the relationships among the observed variables but on those among the unobserved (latent) constructs of the substantive model (Kline, 2011; Schumacker and Lomax, 2010; Tacbachnick and Fidell, 2007).

Structural equation modelling can be conducted through five basic process as proposed by Schumacker and Lomax (2010). These include model specification, model identification, parameter estimation, model testing and model modification. Model identification is a crucial step in SEM process as it determines to the possibilities of finding unique values for the parameters of the specified model (Kline, 2011). A SEM models may be under-identified when one or more parameters of the model are not uniquely determined; in this case, the number of the unknown parameters exceeds the number of equations and there is no empirical information to allow its unique estimation and hence its estimation should not be relied upon (Kline, 2011). When a SEM model is over-identified, the model has a number of possible solutions, and the task will be to select the one that comes closest to explaining the empirical data some uncertainties or error (Peter and Russell, 2009). A model is just-identified when the model has only one unique solution that will be able to perfectly reproduce the correlation matrix - Kline, 2011). However, Hair et al., (1998) said the solution is not of interest because it has no generalizability. Identified models are the only models that can be estimated (Kline, 2011).

23.2.2 *The Theories and Procedures in Structural Equation Modeling*

In this study the procedure recommended by Schumacker and Lomax (2010) was adopted in fitting the final structural model which are: Model specification, identification, estimation, testing and modification. To achieve the first procedure, SEM involves two main components i.e the measurement model/equation (Fig.1a) and the structural model/equation (Fig. 1b). The models involve two categories of variables classified as exogenous or endogenous variables. The exogenous variables are the independent or predictive variables while endogenous or dependent variables the dependent or response variables Jorëskog, 1966. In the modes, the variable $\boldsymbol{x}$ represents the indicator variables of the exogenous latent and $\boldsymbol{y}$ the indicator variables of the endogenous latent variable. The general form of the mathematical expression of the measurement model introduced by Karl Jorëskog (Jorëskog, 1966) is represented in the matrix form as follows:

$$\boldsymbol{y}_{(p\times1)} = \boldsymbol{\Lambda}_{y(p\times m)} \times \boldsymbol{\eta}_{(m\times1)} + \boldsymbol{\varepsilon}_{(p\times1)}$$
$$\boldsymbol{x}_{(q\times1)} = \boldsymbol{\Lambda}_{x(q\times n)} \times \boldsymbol{\xi}_{(n\times1)} + \boldsymbol{\delta}_{(q\times1)}$$

Where:

- $\mathbf{y}$ = vectors of observed scores of endogenous variables
- $\boldsymbol{\eta}$ = vectors of endogenous latent constructs
- $\mathbf{x}$ = vectors of observed scores of exogenous variables
- $\boldsymbol{\xi}$ = vectors of exogenous latent constructs
- $\boldsymbol{\Lambda}_y$ = matrix of construct loadings on endogenous latent construct
- $\boldsymbol{\Lambda}_x$ = matrix of construct loadings on exogenous latent construct
- $\boldsymbol{\varepsilon}$ = vectors of random measurement errors of endogenouse variables
- $\boldsymbol{\delta}$ = vectors of random measurement errors of exogenous variables
- $\mathbf{p}$ = number of endogenous indicator variables
- $\mathbf{q}$ = number of exogenous indicator variables

Given the empirical data describing the variables $\boldsymbol{y}$ and $\boldsymbol{x}$, the measurement equations group together the correlated indicator variables to form the latent variables in $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$. This is done by assigning fixed parameters and defining unknown parameters in Λ_y and Λ_x Karl Jorëskog (Jorëskog, 1966).

The interrelationship among the latent factors or components is explained through a structural equation model expressed in matrix form as follows:

$$\eta_{(m\times1)} = \mathbf{B}_{(m\times m)} \times \eta_{(m\times1)} + \Gamma_{(m\times n)} \times \xi_{(n\times l)} + \varsigma_{(m\times1)}$$

Where:

- **η** = vectors of endogenous latent constructs
- **B** = matrix of structural parameters relating the endogenous constructs together
- **Γ** = matrix of structural parameters relating the exogenous
- constructs to the endogenous constructs
- **ξ** = vectors of exogenous latent constructs
- **ς** = vectors of disturbances representing the unexplained
- variation in the exogenous constructs
- **m** = number of endogenous latent constructs
- **n** = number of exogenous latent constructs

It is assumed that **B** has zeros in the diagonal, and **(I - B)** is required to be non-singular, ξ and ς are uncorrelated. Model evaluation of the parameters is achieved using the criteria for estimation of the solution, measure of overall fit, and detailed assessment of fit. Appropriate Parameter estimates with their respective reasonable standard errors and correlations of parameter estimates are often used to check the relevance of each variable. The R-squared (R^2) is computed for the measurement and the structural equations to account for the amount of the explained variation in the relationship. The overall model fit is then evaluated to examine how well the specified model fit the empirical dataset using some global (multiple) fit indices including the Goodness-of-fit Index (GFI), Adjusted Goodness-of-fit Index (AGFI), Comparative Fit Index (CFI), Incremental Fit Index (NFI), Tucker-Lewis Index (TLI) etc. Model information criteria such as the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and the Consistent AIC (CAIC) are used for model comparison with smaller value indicating a better fit (Daire et al., 2008; Akpa et al., 2015).

Finally, to test the hypothesis and improve the model fit, most especially for exploratory purpose, Statistical packages such as AMOS software, LISREL etc. are used to modify each fixed parameter which indicate a minimum improvement that could be obtained in the chi-square value if the parameter were fixed for estimation (Schumacker and Lomax, 2010; Kline, 2011).

23.3 Software applications in SEM

Like most statistical procedures, SEM has been built into many popular Statistical software packages. There are a variety of such packages including LISREL, Amos, CALIS (a module of SAS), EQS, Mplus and SEPATH (a module of Statistica). There are also two other free packages in R, lavaan and "sem", which can perform SEM efficiently. One of the most important features of software version of the SEM is the capability with which simple restrictions are imposed on the parameters which allow for test of the theoretical specification of the model (Peter and Russell, 2009). Any parameter in the model can be fixed either to zero or to another value or can be fixed to be equal to another parameter or set of parameters. In particular, when parameters in the structural part of the model are constrained to zero it allows for a test of the hypothesis that latent constructs vary independently of one another (Peter and Russell, 2009).

Two or more indicators can be constrained to have same loadings on common latent construct or indicator-specific errors with equal variance. Also, in multiple group analysis, parameters can be constrained to be equal across groups in either the measurement or structural model, allowing tests of whether one or more parts of the model are equivalent across groups. The aim of such analyses is to determine the extent to which a model can be generalized across population groups (Peter and Russell, 2009; Kline, 2011).

23.4 Advantages of SEM in Public Health

The desirability of SEM methodology stems from several advantages over several multivariate statistical methods such as multiple regression or path analysis.

These are:

i. Structural equation modelling analysis allows for issues related to complex prediction as well as measurement in public health to be accurately assessed (Peter and Russell, 2009).
ii. In SEM, multiple observed variables can be assessed compared to some other statistical methods that can only use a limited number of variables at a time (Schumacker and Lomax, 2010).
iii. Measurement error is accounted for during SEM analyses (Kline, 2011; Schumacker and Lomax, 2010) which could be a very useful information for assessing accuracy of measures in public health.
iv. SEM is more powerful and provides more valid and reliable measures when compared to others statistical methods (Peter and Russell, 2009).
v. When compared with multiple regression, it is possible to have more than one dependent variable in SEM and a variable can be both a dependent and independent variable (Kline, 2011; Schumacker and Lomax, 2010; Peter and Russell, 2009). This makes SEM a powerful tool for simultaneously assessing multiple outcomes and/or predictors in a single study.
vi. It is possible to simultaneously assess direct and indirect effects of variables with SEM analysis (Kline, 2011) which could be helpful in assessing confounders and effect modifiers in a public health study.
vii. Compared to path analysis, SEM can have latent variables, which are theoretical constructs not directly observed (Kline, 2011; Schumacker and Lomax, 2010). The particular characteristics that makes SEM useful for studying problems in mental public health.

23.5 Application of SEM in Public Health Research

Structural equation modelling can be used to investigate a variety of research questions in all research areas in public health discipline. A critical barrier however would be the type of study design, the type or definition of variables collected, the investigator understands of the principles and application of SEM, among others. A typical example of the use of SEM in public health investigation was reported by Amorim et al (2010); the study investigated the relationship between nutritional status, quality of home stimulation, socioeconomic conditions, and cognitive development in a sample of children from 20 to 42 months of age living in an urban area of Brazil (Santos et al, 2008). In this study, the cognitive development (consisting of three subscales) was measured using the Bayley Scales for Infant Development (Bayley N, 1993) but only the mental subscale, which includes items evaluating problem-solving, memory, habituation, numerical concepts, vocalization, generalization, social strategies and language was used. The Home Observation for Measurement of the Environment (HOME) inventory (Caldwell & Bradley, 1984) was used to investigate the characteristics of the home environment and parenting style. Nutritional status was evaluated using the z-scores of anthropometric measures including height/age, weight/age, and birth weight.

The burden of non-communicable diseases (NCD) can be attributed to their increasing prevalence, chronic nature, and their high-risk complications (Scalene et. al, 2014). Many NCD patients (especially terminal NCDs such as cancer) are faced with numerous psychosocial problems including poor quality of life, depression, anxiety and sometimes, poor treatment outcomes. Structural equation modeling (SEM) approach was used by Hengqing et. al, 2010 to model the quality of life (QOL) of cancer patients with the assumption that QOL of cancer patients is affected by psychological and physiological factors which cannot be measured directly. A system of equations SEM was used to analyze the effect of demographic factors, social support, type of cancer, and psychological factors on the QOL of cancer patients.

Also, the health status of an individual is affected by many factors including the economic, social, environment, and biological factors (Ferra, Kamarulzaman, & Abdul Aziz, 2010). The determination and the impact of these factors on the health status of an individual varies from one environment to the other. In addition, some of these factors cannot be directly observed i.e. they are latent which complicates the process of investigating the health status of an individual using these factors. Structural equation modeling is a suitable technique that can be used to investigate the interrelationship between the observable and the unobservable factors which describe the health status of individuals in such situations. For instance, in a study conducted by Boniface and Tefft (1997), SEM was used to construct an index for the measurement of health-related behaviours while Stafford et al., 2008, applied SEM techniques to model the potential determinant of environmental health;

such as the built environment, socio-relational characteristics and amenities present in the environment. Chern et al., 2002, used SEM to model the relationship between health-related expenditures on health and the indicators of health.

Among children and adolescents, health behaviours and academic achievement are two domains whose relationship is confounded by complex interrelated variables which includes school-age children characteristics, family backgrounds, social and a wide range of other unobserved individual psychological variables (Ivanovic et. al, 2004). Alfgeir et. al, 2010 tested a SEM that estimated the relationship among health behaviours, body mass index (BMI), dietary habits, physical activity, self-esteem and academic achievement. This model provided a precise pathway by which health behaviours and other potential exogenous factors influence academic achievement of adolescents. Poor growth in children is known to be strongly associated with malnutrition but the impact of underlying factors such as biological factors, social, economic and cultural on malnutrition are complex and interrelated (Desai, 2000). Investigating the interrelationships of all these components to provide plausible explanation for the identified relationships can only be meaningfully done using SEM. In a study conducted by Cheah et al.,2010, SEM was used to demonstrate that environmental factors (household income, total expenditure, number of rooms in the house and socioeconomic status) had significant relationship with malnutrition while the relationship among wealth, obesity and health (Mazzocchi and Traill, 2008) using SEM showed that though a higher wealth is expected to determine lower weight and better health, but through a better diet and not extra exercise or lower calorie consumption.

Finally, in an occupational study with simultaneous sampling across exposure locations, SEM can be used for monitoring pollutant data. Actually, many occupational hygiene surveys are designed to simultaneously collect pollutant monitoring data from more than one location (Davis ME, 2012). Such multi-site study will better reflect the reality of work-related exposure in the area under investigation and leads to complex interrelationship between factors. To better analyse such data, the exposure model must account for the inherent complexities in such study design and must also be flexible to extrapolating exposures across an occupational cohort for dose–response modelling and risk assessment. One of the most powerful statistical techniques for handling such complexities is the SEM.

23.6 Empirical Application of SEM to Psychosocial Functioning and Quality of Life of Adolescents

23.6.1 Source of Data and Data Extraction

The current study is a secondary analysis of data from a State wide cross-sectional survey (among adolescents in Benue state, Nigeria) funded by Fogarty international through the Medical Education Partnership Initiative in Nigeria (MEPIN). The primary survey was a state-wide study involving a Local Government Area (LGA) from each senatorial district in Benue state; Oju, Vandekeya, Wannune and the state capital.

Data for the present analyses were extracted for a total of 983 students who participated in a State wide cross-sectional survey conducted among adolescents in Benue state, Nigeria. Specifically, in addition to the socio-demographic characteristics of respondents, data on the Strength and Difficulty Questionnaire and Quality of Life Questionnaire (Adapted WHO-QOL BREF) were extracted (for each respondent) from the database.

23.6.2 Data Management and Analysis

The IBM SPSS Statistics, version 20 and R Programming Software version 3.2.0 were employed for both descriptive and analytical techniques. Data screening and preliminary analyses, such as data cleaning, missing values/no-response, and systematic endorsement (e.g. endorsing the same response for the entire survey), the normality test and outliers test were performed so as to allow the results to be meaningfully interpreted. The screened dataset was used for Confirmatory Factor Analysis (CFA) and structural equation modelling (SEM) using Analysis of Moment Structure (AMOS) program version 21.

Specifically, CFA, a special case of SEM was performed to test the hypothesized factor structure of the two measures (SDQ and WHO-QOL BREF) used in the study and also to determine whether the hypothesized or the existing structure provides a better fit to the dataset. The SDQ was assessed using two competing models: Model A1 (the original theoretical 5-factor model of the SDQ) Model A1) and Model A2 (the hypothesized 3-components model of the SDQ). On the other hand, the WHO-QOL BREF was assessed using three competing models: Model B1 (the original theoretical 4-factor model of the WHO-QOL BREF), Model B2 (a hypothesized 2-factor model with no correlated error) and Model B3 (a hypothesized 2-factor model but with some correlated error terms).

The resulting models from CFA that adequately fit the dataset were then modelled together using SEM. The global fit to the data was assessed using Chi-square test setting level of significance alpha to 0.01, Relative χ^2 (χ^2/df) which adjust for sample size with an acceptable value ≤3.0, Root Mean Square Error of Approximation (RMSEA); value <0.05 was considered a good fit (Kline, 2011). Also, multiple fit indices including Goodness-of-fit Index (GFI), Adjusted Goodness-of-fit Index (AGFI), Comparative Fit Index (CFI), Incremental Fit Index (NFI) which test if the variables are uncorrelated, Tucker-Lewis Index (TLI), with a threshold values ≥0.90 were used to assess the model fit. The Akaike Information Criterion (AIC), the Consistent AIC (CAIC) which assigns greater penalty to model complexity and the Bayesian Information Criterion (BIC) were used for model comparison with smaller value indicating a better fit (Daire et al., 2008; Akpa et al., 2015).

23.7 Results

23.7.1 Participants' Characteristics

In Table 23.1, of the 983 adolescents in the study, more than half (53.8%) were males while 46.0% were females. Participants were mostly teenagers aged 13-17 years (76.0%) while only 9.5% were older adolescents, aged 18-19 years. Majority of the adolescents were from Monogamy family 65.3% while 31.4% were from Polygamy family. Most (72.3%) of them have their parents living together while 7.2% of them have parents living apart and 4.0% of them have parents who are divorced. Many of them have parents who have completed secondary school (36.4% of Farther and 26.3% of Mothers) and are farmers (32.9% of fathers and 30.4% of mothers)

23.7.2 Confirmatory Factor Analyses for the SDQ and the adapted WHO QOL BREF

In Table 23.2, the Chi-square goodness-of-fit test for the two models of the SDQ were significant. This is indicative of the large differences between the observed and expected covariance matrices. However, the Chi-square indicator is highly dependent on sample size, so relative Chi-square χ^2 $(\chi^2/df < 3 = 2.67)$ which adjusted for sample size shows that Model A2 better fits the dataset analysed. The fit indices (GFI=0.941, AGFI=0929, CFI=0.885, NFI=0.830, TLI=0.874, and the RMSEA=0.041) and the information criteria (AIC=833.25 and CAIC=1145.45) confirmed that Model A2 fits the data better than A1.

Also, in Table 2, the Chi-square goodness-of-fit test of the three models of the adapted WHO-QOL BREF were significant even with poor estimates for the fit indices (Table 4.32). This is indicative of the large differences between the observed and expected covariance matrices. However, the Chi-square indicator is highly dependent on sample size, so relative Chi-square χ^2 $(\chi^2/df < 3 = 2.98)$ which adjusted for the sample size shows that Model B3 best fits the analysed data. The fit indices (GFI=0.941, AGFI=0928, CFI=0.907, NFI=0.867, TLI=0.895, and the RMSEA=0.045) and the information criteria (AIC=773.20 and CAIC=1073.62) confirmed that Model B3 fits the data better than the other competing models.

23.7.3 Structural equation modelling of the best fitted factor structure of the SDQ and the adapted WHO-QOL BREF

The path diagram with the standardized path coefficients, as well as the coefficients between the latent variables for the hypothesized 3-factor model of the SDQ and the 2-factor model of the Adapted WHO-QOL BREF are presented in Figure 2. The Standardized error terms of each indicator variables were also reported in Table 3, because the solution was admissible. This suggests that the fitted structural model provide a good fit to the dataset.

All indicator variables were significantly related to their respective latent factors at 1% level of significance. Also, there is a negative association between SDQ and QoL, which implies that for every unit increase in the adolescents SDQ there is a corresponding 60% reduction in the quality of life. The R^2 values in the last column indicate the proportion of variation in the latent component as explained by each particular indicator. In the present report, approximately 5% of the variation in the latent components were explained by their indicator variables (Table 3).

In Table 2, the Chi-square goodness-of-fit test of the fitted SEM model was also significant, indicating that there is difference between the observed and expected covariance matrices. However, since the Chi-square test is highly dependent on sample size, so relative Chi-square χ^2 ($\chi^2/df < 3 = 2.91$) which adjusted for sample size shows that the model B best fits the dataset. The fit indices (GFI=0.882, AGFI=0.871, CFI=0.784, NFI=0.705, TLI=0.773, and the RMSEA=0.044) also confirmed that model B fits the dataset better than model A.

23.8 Discussion

This paper discussed the application of SEM in public health research. We also demonstrated how SEM could be used to analyse public health related survey data. Such documentation provides a systematic guide to the application of SEM and may be helpful to public health researchers with limited knowledge of statistical theories.

There are numerous opportunities and possibilities for the application of SEM in public health research but it must be acknowledged that the techniques of SEM require some fundamental understanding of the principles and the basic theories. For instance, though SEM is akin to regression model, among other things, the indicator variables of every latent construct must be well thought of a priori. Many previous guidelines on the application of SEM to research areas have emphasized these points and the critical point of study design and variable definition (Amorin et al, 2010, Tanya & Claudio, 2010 & Michel et al, 2010).

In the empirical study, the results of the Confirmatory Factor Analysis suggest that the adolescents are likely to be reporting their psychosocial functioning based on three separates but correlated underlying constructs. This finding is similar to the three-factor model tested by Cathal and Richard (20,12) among Irish adolescents but did not give a good fit when subjected to CFA. The Irish study was based on extensive literatures which had used the parent version of the SDQ. The difference in this result might be due to the construal bias as mention by Wayne and

Stephen (2004) implicated on parent's willingness to attribute desirable or undesirable qualities to a child. Also, adolescents are likely to report their perceived quality of life based on two separate but uncorrelated underlying components. Our finding differs from the results found among Indian (Shally Awasthi et al., 2010) and English (Sik-Yum Lee et al., 2005) adolescents. These studies used WHO-QOL BREF instrument with 24 items plus two items of overall QOL and general health on a 5-points likert scale. To be more specific, the study by Sik-Yum Lee et al., 2005 used a sample of the dataset that was used in the initial design of the adopted WHO-QOL BREF instrument. This is similar to model testing and not instrument validation as explained by Cathal and Richard (2012) and Jamie (1998). Also, the methodology used in the instrument validation in the Indian adolescents' population was reported to have some scientific flaws (Hengqing et al., 2010; Gadermann et al., 2012).

The psychosocial functioning was found to negatively impact on the QOL of life of apparently healthy Nigerian adolescents. Although previous studies exist on the psychosocial functioning and quality of life of Nigerian adolescents (Ayuk et al., 2013; David et al., 2004; Dejan et al., 2011; Akpa and Bamgboye, 2015), we were unable to find a very related study on the application of SEM with which to compare our findings. Notwithstanding, in order to ensure a very reliable results in the present study, a slightly different approach was employed using hierarchical model that is simply a second-order factor analysis model. This approach allows for the complete and simultaneous tests of the relationships between the domains of these two instruments (Ullman, 2006). The result of the analysis shows that poor psychosocial functioning of adolescents has a negative impact on their quality of life. This finding is corroborated by an earlier report by Akpa and Bamgboye (2015) where adolescents with deficient prosocial behaviors and those at the borderline psychosocial problems were more likely to have poor quality of life compared to those with normal traits.

Despite its numerous advantages, SEM has some criticisms and limitations requiring considerations before use in public health. For instance, SEM requires large samples size (Kline, 2011; Schumacker and Lomax, 2010) which may be cost intensive for public health studies. It may seem complex and difficult to use (Schumacker and Lomax, 2010); especially if the public health researcher has little knowledge of statistical modelling. Also, although many SEM softwares user friendly to users who are familiar with them, there are complex and demanding than other multivariate techniques (Hair et al., 1998). Such complexities may discourage a public health research from using SEM for their studies. The empirical data used in the present paper is from a cross-sectional study and does not permit assessment of any causal effect of the independent variables. Data was also for adolescents enrolled and attending schools which may hinder generalizing results on out of school adolescents.

23.9 Conclusion

Structural equation modeling can be applied and is applicable to any research areas in public health, provided the study is designed with the application of SEM in mind. The original model of the WHO-QOL BREF and the Strength and Difficulty Questionnaire published and validated in other countries does not fit the sample of adolescents considered in this study. Nigerian adolescents. The results of the SEM reveal that the psychosocial functioning (measured by the SDQ) is negatively associated with the quality of life of adolescents in the study sample.

Acknowledgment

The project described in this study was supported by the Medical Education Partnership Initiative in Nigeria (MEPIN) project funded by Fogarty International Centre, the Office of AIDS Research, and the National Human Genome Research Institute of the National Institute of Health, the Health Resources and Services Administration (HRSA) and the Office of the U.S. Global AIDS Coordinator under Award Number R24TW008878. OMA was awarded the grant for the study. The content is solely the responsibility of the authors and does not represent the views of the funding organizations.

Bibliography

Agnihotri K, Awasthi S, Chadra H, Singh U and Thakur S. (2010). Validation of WHO QOL-BREF Instrument in Indian Adolescents. *Indian Journal of Pediatrics*, 77(4): 381-386

Akpa OM, Bamgboye EA and Baiyewu O. (2015). The Adolescents' Psychosocial Functioning Inventory (APFI): scale development and initial validation using

Exploratory and Confirmatory Factor Analysis. *African Journal of Psychological Study of Social Issues,*18(1): 1-21.

Ayuk AC, Oguonu T, Ikefuna AN and Ibe BC (2013). Health-related quality of life in school-aged children with and without asthma in Enugu, South East Nigeria. *Nigeria Journal of Paediatrics, 40 (4): 364 –369*

Bayley N. (1993). The Bayley scales of infant development. 2nd Ed. New York: *Psychological Corporation.*

Boniface DR, Tefft ME. (1997). The application of structural equation modeling to the construction of an index for the measurement of health-related behaviours. *Statistician*. 46:505-14.

Caldwell B and Bradley RH. (1984) Home Observation for Measurement of the Environment (HOME): revised edition. Little Rock: *University of Arkansas Press*;

Cheah WL, Wan Abdul Manan WM and Zabidi-Hussin ZAMH. (2010). A structural equation model of the determinants of malnutrition among children in rural Kelantan, Malaysia. *Rural and Remote Health,* 10:1248.

Chern JY, Wan TTH and Begun JW. (2002). A structural equation modelling approach to examining the predictive power of determinants of individuals' health expenditures. *Journal of Medical Systems*, 26:323-36.

Cudeck R, du Toit S, and Sörbom D. (2001). Structural equation modelling: Present and future. *Chicago*: *SSI Scientific Software.*

Davis ME. (2012). Structural equation models in occupational health: an application to exposure modelling. *Occupational and Environmental Medicine*, 69(3):184-90.

Desai S. (2000). Maternal education and child health: a feminist dilemma. *Feminist studies*, 26(2): 425-448.

Gadermann AM, Martin G and Bruno DZ. (2012). Estimating ordinal reliability for likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation,* 17(3). Available online at: http://pareonline.net/getvn.asp?v=17&n=3

Hengqing T, Shudan L, Yang Y and Yichao P. (2010). Analysis of Quality of Life in Cancer Patients by Structural Equation Model. *Journal of Cancer Therapy, 1: 71-75*

Hooper D, Coughlan J and Mullen MR. (2008). Structural Equation Modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods, 6(1): 53-60*

Ivanovic DM, Perez HT, Olivares MG, Diaz NZ, Leyton BD and Ivanovic RM. (2004). Scholastic achievement: A multivariate analysis of nutritional, intellectual, socioeconomic, sociocultural, familial, and demographic variables in Chilean school-age children. *Nutrition*, 20: 878-889.

Jorëskog KG. (1966). Testing a simple structure hypothesis in factor analysis. *Psychometrika,* 32:165-78.

Kaplan D. (2000). Structural equation modelling: Foundations and extensions. Thousand Oaks, *CA: Sage.*

Kline RB. (2011). Principles and practice of structural equation modelling. 3rd Edition, *The Guilford Press:* New York.

Kristjansson AL, Sigfusdottir ID and Allegrante JP. (2010). Health behavior and academic achievement among adolescents: The relative contribution of dietary habits, physical, body mass index, and self-esteem. Health Education & Behaviour, 37(1): 51-64

Mazzocchi M and Traill WB. (2008). A structural model of wealth, obesity and health in the UK. 12th Congress of the European Association of Agricultural Economist - *EAAE* .

McCrory C and Layte R (2012). Testing competing models of the Difficulties Questionnaire's (SDQ's) factor structure for the parent-informant instrument. *Personality and Individual Differences, 52: 882-887.*

Peter C and Russell E. (2009). Structural Equation Modelling By Example, Applications in Educational, Sociological and Behavioural Research, *Cambridge University Press.*

Santos LM, Santos DN, Bastos ACS, Assis AMO, Prado MS and Barreto ML. (2008) Determinants of early cognitive development: hierarchical analysis of a longitudinal study. *Cad Saúde Pública,* 24:427-37.

Schumacker R and Lomax R (2010). A Beginners Guide to Structural Equation Modelling, 3rd Edition, *Lawrence Erlbaum*: New York.

Stafford M, Sacker A, Ellaway A, Cummiins S, Wiggins D, Mac-intyre S. (2008). Neighbourhood effects on health: a structural equation modelling approach. *Schmollers Jahrbuch,* 128:1-12.

Tabachnick B and Fidell L (2007). Using Multivariate Statistics, 5th Edition, *Allyn andBacon:* New York.

Ullman JB. (2006). Structural equation modelling: reviewing the basics and moving forward. *Journal of Personality Assessment, 87:35-50.*

Weston R and Gore PA Jr (2006). A brief guide to structural equation modelling. *Counselling Psychology, 35:719-51.*

Yanuar F, Ibrahim K and Aziz A. (2010). On the application of structural equation modeling for the construction of a health index. *Environmental Health and Preventive Medicine,* 15:285-291.

Table 23.1: Personal Characteristics of Respondents

Variable	Freq.	%
Current Age		
10-12years	143	14.5
13-17years	747	76.0
18-19years	93	9.5
Gender		
Male	529	53.8
Female	452	46.0
Not Reported	2	0.2
Family type		
Monogamy	642	65.3
Polygamy	309	31.4
Not Reported	32	3.3
Parental marriage		
Pat	711	72.3
Pad	39	4.0
Parents live apart	71	7.2
Single parent	140	14.2
Not Reported	22	2.2
Father's level of education		
None	105	10.7
Primary	108	11.0
Secondary	233	23.7
Tertiary	358	36.4
Others	150	15.3
Not Reported	29	3.0

Contd.

Variable	Freq.	%
Father's Occupation		
Farming	323	32.9
Trading	70	7.1
Civil servant	378	38.5
Employee of PO	69	7.0
Others	125	12.7
Not Reported	18	1.8
Mother's level of education		
None	142	14.4
Primary	188	19.1
Secondary	245	24.9
Tertiary	259	26.3
Others	117	11.9
Not Reported	32	3.3
Mother's Occupation		
Farming	299	30.4
Trading	278	28.3
Civil servant	214	21.8
Employee of PO	69	7.0
Others	96	9.8
Not Reported	27	2.7

Pat: Parents are together; Pad: Parents are divorced; PO: Private Organisation

Table 23.2: Fit Indices from Confirmatory Factor Analyses and SEM

Fit Indices	SDQ CFA Model		WHOQOL-BREF CFA Model			SEM models	
	Model A_1	Model A_2	Model B_1	Model B_2	Model B_3	Model A	Model B
χ^2	1210.58*	727.25*	1031.43*	798.90*	671.20*	2950.64*	3132.66*
df	265	272	246	229	225	1121	1075
CGF	4.57	2.67	4.19	3.45	2.98	3.11	2.91
RMR	0.035	0.023	0.024	0.023	0.021	0.085	0.034
GFI	0.89	0.941	0.91	0.929	0.941	0.875	0.882
AGFI	0.865	0.929	0.89	0.914	0.928	0.864	0.871
PGFI	0.725	0.787	0.726	0.771	0.768	0.803	0.807
NFI	0.717	0.83	0.801	0.842	0.867	0.679	0.705
RFI	0.679	0.812	0.777	0.825	0.85	0.665	0.691
IFI	0.764	0.886	0.841	0.882	0.907	0.757	0.785
TLI	0.73	0.874	0.821	0.869	0.895	0.745	0.773
CFI	0.762	0.885	0.84	0.881	0.907	0.756	0.784
PRATIO	0.883	0.907	0.891	0.905	0.889	0.956	0.953
PNFI	0.633	0.752	0.714	0.76	0.771	0.649	0.671
PCFI	0.673	0.803	0.749	0.797	0.807	0.723	0.747
NCP	945.58	455.25	785.43	567.9	446.2	2375.6	2057.66
RMSEA	0.06	0.041	0.057	0.05	0.045	0.046	0.044
AIC	**1330.58**	**833.25**	**1139.43**	**892.9**	**773.2**	**3701.6**	**3334.66**
BIC	**1624.02**	**1092.45**	**1403.52**	**1901.76**	**1022.62**	**4195.55**	**3828.61**
CAIC	**1684.02**	**1145.45**	**1457.52**	**1169.76**	**1073.62**	**4296.55**	**3929.61**

*x^2 =Chi-square statistics; df=degree of freedom; CGF=Chi-square Goodness-of-Fit; RMR=Root mean square residual; GFI=Goodness-of-Fit index; AGFI= Adjusted Goodness-of-Fit index; PGFI=Parsimony Goodness-of-Fit Index; NFI=Normed-fit index; RFI=Relative fit index; IFI=Incremental fit indices; TLI=Tucker-Lewis index; CFI=Comparative fit index; PRATIO=Parsimony ratio; PNFI=Parsimonious Normed-fit index; PCFI= Parsimonious Comparative fit index; NPC=Noncentrality parameter; RMSEA=; AIC=Akaike's information criterion; BCC=Browne-Cudeck criterion; BIC=Bayesian information criterion; CAIC=Consistent AIC *Significant at 1% level of significance*

Table 23.3: Estimates of the Model Parameters in the SEM

			Estimate	S.E.	P-value	R^2
QoL	<---	**SDQ**	**-0.600**	**0.080**	**<0.001**	**0.045**
Factor1	<---	SDQ	0.238	0.045	<0.001	0.896
Factor3	<---	SDQ	1.000			0.180
Factor2	<---	SDQ	-1.038	0.129	<0.001	0.448
FAC1	<---	QoL	1.000			0.778
FAC2	<---	QoL	0.227	0.126	0.070	0.886
SDQ25	<---	Factor3	1.000			0.744
SDQ21	<---	Factor3	1.338	0.116	<0.001	0.622
SDQ14	<---	Factor3	1.152	0.105	<0.001	0.694

SDQ11	<---	Factor3	0.964	0.100	<0.001	0.802
SDQ07	<---	Factor3	0.815	0.088	<0.001	0.828
SDQ20	<---	Factor2	1.000			0.669
SDQ17	<---	Factor2	0.974	0.079	<0.001	0.671
SDQ09	<---	Factor2	0.720	0.071	<0.001	0.818
SDQ04	<---	Factor2	0.995	0.079	<0.001	0.646
SDQ01	<---	Factor2	0.861	0.070	<0.001	0.686
SDQ23	<---	Factor1	1.000			0.907
SDQ19	<---	Factor1	1.566	0.201	<0.001	0.768
SDQ06	<---	Factor1	1.475	0.193	<0.001	0.793
SDQ15	<---	Factor1	1.588	0.200	<0.001	0.740
SDQ10	<---	Factor1	1.324	0.179	<0.001	0.828
SDQ02	<---	Factor1	1.728	0.217	<0.001	0.732
SDQ22	<---	Factor1	1.539	0.197	<0.001	0.761
SDQ18	<---	Factor1	1.749	0.219	<0.001	0.730
SDQ12	<---	Factor1	1.582	0.197	<0.001	0.718
SDQ05	<---	Factor1	1.228	0.174	<0.001	0.860
SDQ24	<---	Factor1	1.414	0.188	<0.001	0.812
SDQ16	<---	Factor1	1.694	0.214	<0.001	0.740
SDQ13	<---	Factor1	1.701	0.210	<0.001	0.700
SDQ08	<---	Factor1	1.504	0.195	<0.001	0.783
SDQ03	<---	Factor1	1.483	0.192	<0.001	0.780
QOL09	<---	FAC1	1.000			0.683
QOL12	<---	FAC1	0.736	0.063	<0.001	0.810
QOL15	<---	FAC1	0.787	0.063	<0.001	0.774
QOL18	<---	FAC1	1.063	0.074	<0.001	0.664
QOL22	<---	FAC1	1.193	0.077	<0.001	0.583
QOL02	<---	FAC1	0.565	0.058	<0.001	0.872
QOL10	<---	FAC1	0.714	0.062	<0.001	0.815
QOL20	<---	FAC1	0.963	0.072	<0.001	0.721
QOL24	<---	FAC1	1.220	0.079	<0.001	0.594
QOL03	<---	FAC1	0.714	0.067	<0.001	0.846
QOL14	<---	FAC1	0.972	0.070	<0.001	0.699
QOL16	<---	FAC1	0.559	0.067	<0.001	0.912
QOL06	<---	FAC1	0.813	0.069	<0.001	0.799
QOL07	<---	FAC1	1.105	0.078	<0.001	0.680
QOL08	<---	FAC1	1.110	0.076	<0.001	0.657
QOL11	<---	FAC1	0.947	0.071	<0.001	0.727
QOL17	<---	FAC1	0.894	0.069	<0.001	0.744
QOL19	<---	FAC1	0.806	0.065	<0.001	0.772
QOL21	<---	FAC1	1.112	0.077	<0.001	0.661
QOL23	<---	FAC1	1.113	0.077	<0.001	0.659
QOL01	<---	FAC2	1.000			0.985
QOL05	<---	FAC2	1.265	0.535	0.018	0.955

QOL13	<---	FAC2	0.976	1.976	0.065	0.531

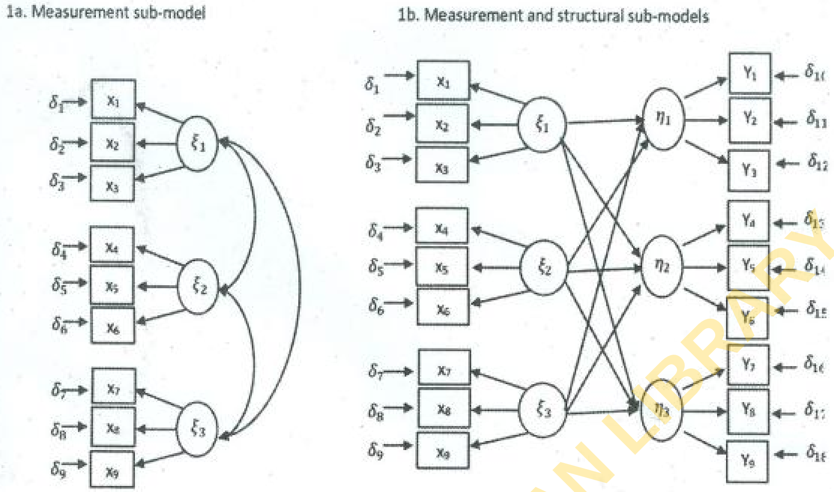

Figure 23.1: Measurement models and structural sub-models

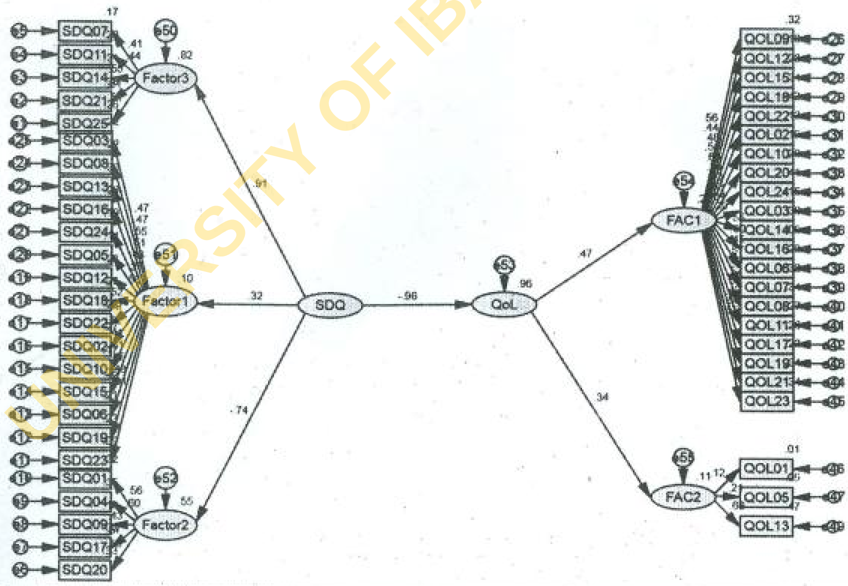

Figure 23.2: Model B: The best fitted structural equation model for the effect of SDQ on WHO-QOL BREF