International Journal of Quantitative and Qualitative Research Methods, 13(1),12-31, 2025
Print ISSN 2056-3620 (Print)
Online ISSN 2056-3639(Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

Structural Equation Modeling: Principles, Assumptions, and Practical Application Using SPSS Amos

Soleman Abu-Bader
Professor, School of Social Work, Howard University
sabu-bader@howard.edu

Tiffanie Victoria Jones
Adjunct faculty, School of Social Work, Howard University
tiffanie.jones@howard.edu

doi: https://doi.org/10.37745/ijqqrm.13/vol13n11231 Published June 26, 2025

Citation: Abu-Bader S. and Jones TV (2025) Structural Equation Modeling: Principles, Assumptions, And Practical Application Using SPSS Amos, *International Journal of Quantitative and Qualitative Research Methods*, 13(1),12-31

Abstract: *Structural Equation Modeling (SEM) is a comprehensive multivariate statistical technique that permits the testing of complex theoretical models involving observed and latent variables. This article reviews the foundational principles and assumptions of SEM and outlines the practical steps for conducting SEM analysis using SPSS AMOS. As well, it presents implications to research and practice utilizing SEM, as well as directions for future research with the technique, with particular emphasis on social science fields. An example using a real data set is used to demonstrate how to construct a model, interpret path coefficients, evaluate model fit, and ensure that SEM assumptions are met. While the data used here are real, the emphasis remains on understanding the SEM framework and enhancing methodological rigor in research applications.*

Keywords: Structural equation modeling, SPSS AMOS, model fit, latent variables, model assumptions

INTRODUCTION

Structural Equation Modeling (SEM) has emerged as a cornerstone of empirical research across disciplines such as psychology, sociology, education, social work, and business. SEM integrates elements of multiple regression, path analysis, and factor analysis into a unified framework for testing complex theoretical models (Kline, 2016; Hoyle, 2012). It provides researchers with the tools to model latent constructs, test measurement validity, and evaluate hypothesized causal relationships simultaneously.

Unlike traditional multivariate methods, SEM accounts for measurement error and supports the estimation of both direct and indirect effects (Byrne, 2016). Given its theory-driven nature, careful model specification is essential to the integrity of SEM analyses. This paper reviews the core SEM assumptions and procedures, highlights model evaluation strategies, and demonstrates the practical implementation of SEM using SPSS AMOS.

Literature/Theoretical Underpinning

The origins of SEM date indirectly back to Charles Spearman's work on general intelligence, in which he designed the first factor analysis by creating a two-factor model to measure human beings' cognitive abilities

(Spearman, 1904). However, the works of Sewall Wright (1918) directly concern SEM, as he developed a path diagram to estimate the effect between variables – whether direct, indirect, or a total effect – and he showed the correlation between the variables and their links to the model parameters. The work began out of researchers' need to produce analyses for complex relationships, particularly, those that pertain to abstract or latent concepts, which cannot be directly measured, in contrast to manifest or observable variables.

Some examples of latent variables include mental health, life satisfaction, attitudes, burnout levels, and well-being. The importance of accurately measuring these concepts cannot be over-emphasized. Because these concepts are abstract in nature, they are not directly observable in the physical world (Dijkstra & Henseler, 2015). As such, they cannot be directly measured either. Handling an abstract concept or event as though it were concrete is a reification error (Levy, 2010). Therefore, to avoid this error in reasoning, it is vital to create instruments that can measure observable indicators for these types of variables. SEM is ideal for its capability to address latent variables and avoid error.

In addition to SEM, the analysis is known by various names including Jöreskog & Sörbom's LISREL (**LI**near **S**tructural **REL**ationships) or structural equations with latent variables, which accounts for measurement error in observed variables; covariance structure models, which estimates the relationships between variables; and latent variables models, a model that contains unobserved variables.

Purpose

SEM is a sophisticated, iterative method whose purpose is to develop a set of equations that analyze observable or manifest and latent variables to produce a causal path diagram to help explain complex social phenomena (Tarka, 2022). Kaplan (2000, p. 1) defines the technique in this way: "structural equation modeling can perhaps best be defined as a class of methodologies that seeks to represent hypotheses about the means, variances and covariances of observed data in terms of a smaller number of 'structural' parameters defined by a hypothesized underlying model."

SEM can also be understood more simply as a system of linear equations, a characterization that underlies the alternative term LISREL—**LI**near **S**tructural **REL**ations (Nachtigall et al., 2003). As an advanced multivariate technique, SEM is comprised of two interrelated components: the *measurement model* and the *structural model* (Kaplan, 2001). The measurement model shows the covariance between the different factors, identifying each factor and its corresponding indicators, while the structural model specifies the directions between the latent variables. Combined, these two models merge multiple analyses to develop the path diagram. SEM's strength lies in its ability to test entire models simultaneously rather than in isolated steps.

SEM proposes a system consisting of the varying relationships between multiple independent variables and multiple dependent variables, rather than a single bivariate relationship or even a single multivariate relationship that only involves either one independent variable or one dependent variable. These variables are those that are most subject to error, such as latent variables that rely on test construction due to their abstract nature. These are variables that are seen most often in social science research. SEMs are able to serve the following additional purposes:

- estimate relationships between these abstract concepts corrected for measurement error to get structural parameters;
- estimate the nature of measurement error in observed variables;
- answer research questions that assess the effect of mediators on an indirect relationship;
- allow multiple indicators of the same concept;
- assess reliability and validity of measures;
- allow new tests of fit of systems of equations;
- allow measurement errors to be correlated.

International Journal of Quantitative and Qualitative Research Methods, 13(1),12-31, 2025
Print ISSN 2056-3620 (Print)
Online ISSN 2056-3639(Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

Key Features of SEM

Key features of SEM include a classical measurement model, mediating factors, and graphical representation. A classical measurement model of SEM is $X=T+E$, in which "X" represents the observed indicator or measured variable; "T" represents the true score; and "E" represents error. It is assumed that because "X" is a latent variable that originally was not observable, development of a measurement that consists of observable indicators might include some errors. For example, perhaps, when responding to the measurement items, the study participant was sleepy or hungry; perhaps, the individual was distracted or needed to rush through the survey to attend to other business; perhaps, the respondent gave an answer that seemed socially desirable to the surveyor; perhaps, the questions on the survey were unclear; or perhaps, the individual simply made a mistake in responding. These types of errors can be classified as either systematic or random. In either case, the responses given might not be an accurate reflection of the person's true feelings regarding the measurement. The overarching idea is that the measurement is likely to include the participants' *true score* as well as some *error*. This is accounted for in the model.

Another key characteristic of SEM is its use of mediators to better specify more complex, causal theories. Mediating variables are useful for explaining the process of a relationship when an indirect relationship might exist between variables. In doing so, a causal chain explains how the independent variable impacts the dependent variable but through its influence on the mediating variable, as the mediator is the more proximal cause of the outcome.

SEM uses path diagrams to show the systems of relationships between sets of variables. Manifest variables are typically represented by squares or rectangles with the letter "x" and a number following, while latent variables are typically represented by circles or ellipses with the eta symbol inside followed by a number; error variance is usually represented using circles with the letter "e" and a number following. (Other types of variables may be included in the system as well, such as exogenous variables and endogenous variables). A covariance path is usually represented by a curved double-headed arrow to indicate a bi-directional correlation between the variables, and a single-headed, straight arrow implies a directional, causal relationship in the model. See figure 1 of a path diagram.

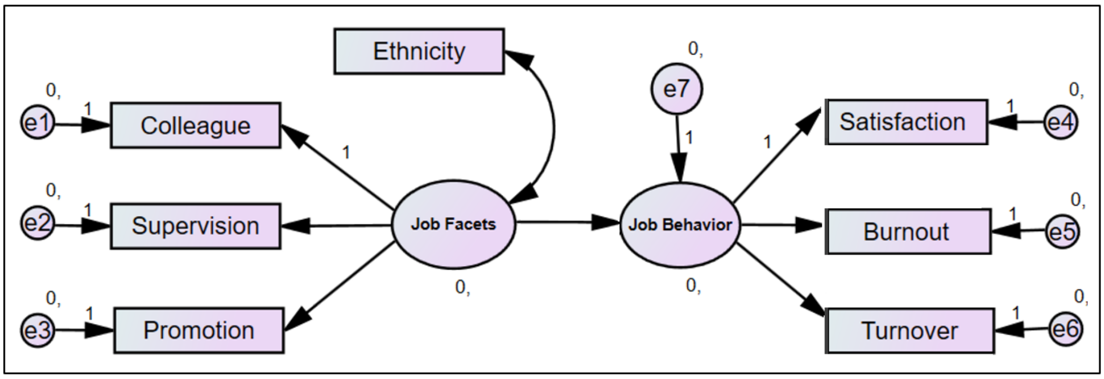

Figure 1. Path diagram

Types of SEM

Structural equation modeling (SEM) encompasses a range of analytical approaches that vary in complexity and purpose. The following outlines the primary types of SEM commonly used in empirical research.

1. **Confirmatory Factor Analysis (CFA):** Evaluates the extent to which observed variables accurately reflect their underlying latent constructs.
2. **Path Analysis:** Examines the directional relationships among observed variables to test hypothesized causal pathways.

3. **Full Structural Equation Modeling (SEM):** Integrates CFA and path analysis to assess complex relationships among both latent and observed variables within a unified framework.

Definitions

A clear understanding of SEM requires familiarity with its key terms. The following definitions outline these terms that underline the construction, interpretation, and evaluation of SEM models, particularly within the context of social science research. The following definitions provide a concise reference for the principal components and terminology associated with SEM:

1. *Observed or manifest variables*: Variables that can be directly measured or items on a scale that serve as observed indicators.
2. *Latent or abstract variables*: Variables which cannot be directly measured because they are not directly observable but instead, requires observable indicators.
3. *Indicator*: Observable items combined to indirectly measure latent variables.
4. *Endogenous variables*: They are dependent variables, or those influenced by other variables in the system.
5. *Exogenous variables*: They are completely independent of others but influence the outcomes in the model.
6. *Model*: It is a mathematical representation or statistical statement of observed data, depicting the relationships between variables.
7. *Measurement model*: It operationalizes the theoretical concepts through confirmatory factory analysis by delineating the relationships between manifest variables and their corresponding latent variables.
8. *Structural model*: It defines causal relationships between the latent variables through a path diagram that depicts the constructs' influence on each other.
9. *Path diagram*: It is a visual representation of causation, showing variables' hypothesized interconnectedness within a complex system.
10. *Path coefficients*: They are standardized regression coefficient (beta), indicating the direct effect an independent variable has on a dependent variable.
11. *Direct effects*: It is a causal relationship between two variables, measured directly through a straight-line or single path in the model, while holding all other variables constant.
12. *Indirect effects*: It is the pathway through a mediating variable that creates a relationship between the independent variable and the dependent variable that previously did not exist in the absence of the mediator.
13. *Errors*: They are variances that are not attributed to the variables itself but other factors.
14. *Model specification*: It is a definition of the hypothesized relationships by determining which variables will be included in the regression equation.

Research Questions and Hypothesis

Because SEM is composed of a variety of analyses, it is appropriate for addressing a number of research questions, including those that pertain to the relationships between latent variables as well as relationships that are formed through indirect effects, including mediation and moderation.

The overarching question being asked with an SEM is: "***Does the model produce an estimated population covariance matrix that is consistent with the sample (observed) covariance matrix?***" The hypothesis for this question is that there is a significant difference between the observed covariance matrix (obtained from the sample) and the implied covariance matrix (observed in the population). The researcher should seek ***not to reject the null hypothesis***, which would suggest that the observed covariance matrix and the implied covariance matrix are not statistically significantly different.

Several questions that might subsequently be asked to address this primary question include the following:

1. What are the causal relationships between multiple latent variables? Otherwise asked, how do latent variables influence other latent variables within a system?
2. A question pertaining to the relationship between latent variables can be the following: What is the correlation between the latent variables?
3. Mediation questions might take the following form: How does *Z* mediate the relationship between *X* and *Y*?
4. Moderation questions, on the other hand, might ask: Is *Y* strengthened when Z is combined with *X*?

Assumptions of SEM

Structural equation modeling relies on several key assumptions to ensure accurate and valid results. Violations can impact estimates and model fit, so it is important to assess these assumptions before interpreting findings. Below are the main assumptions and common methods to evaluate them:

1. Level of measurement: Independent and dependent variables can be categorical (nominal or ordinal) and continuous (interval or ratio) data (Abu-Bader, 2021).
2. Linearity: The endogenous (dependent) and exogenous (independent) variables should have a linear relationship, meaning that the data points form a straight line – whether increasing, decreasing, or remaining constant – and thus indicating that as one variable changes, the other variable changes proportionally. Linearity is evaluated by a scatterplot (Abu-Bader, 2021; West, Finch, & Curran, 1995).
3. Multicollinearity: SEM requires that multicollinearity is minimal and can be detected by using Variance Inflation Factor (VIF) and tolerance values. Multicollinearity occurs when exogenous variables have high intercorrelations, indicating that the items being used to operationalize the variables are too similar to distinguish between each variable's individual contribution. This assumption can be assessed through a correlation matrix, VIF, and tolerance values (Abu-Bader, 2010).
4. Univariate normality: Individual variables should be normally distributed. It can be evaluated by inspecting histograms, normality plots, and normality tests (Abu-Bader, 2021).
5. Multivariate normality: Multiple variables in a set should be normally distributed. Normal distribution for both univariate and multivariate normality should be assessed using histograms, q-q plots, skewness, kurtosis, and statistical tests, and data should be transformed if warranted. Normality can be tested using Mardia's coefficient (Mardia, 1970).
6. Independence of observations: Observations should be independent of each other, meaning that no data collected from one individual should impact the data collected from another individual (Preacher, Zyphur, & Zhang, 2010).
7. Missing data: Data should be complete, and where there is missing data, it should be random. Because missing data can impact the model's estimate, missing data can be addressed using the Full Information Maximum Likelihood (FIML) or Multiple Imputation (MI) techniques (Enders, 2010).
8. No Measurement Error in Exogenous Variables: Traditional regression assumes no error in predictors, but SEM allows modeling of measurement error, improving validity (Bollen, 1989).
9. Sample size: Because of the potential risk of type I and type II errors, SEM requires that the sample size be sufficiently large, with general recommendations ranging between 200-400 observations and 10-15 indicators per variable (Boomsma, 1982; Wolf et al., 2013). In general, the following parameters should be followed when determining sample size: 50 equals very poor; 100 equals poor; 200 equals fair; 300 equals good; 500 equals very good; and 1,000 equals excellent.
10. Model Identification: A model must be identified to be estimated. Each latent variable should be measured by at least three indicators (Brown, 2015).
11. Correct Model Specification: Models must be theoretically grounded. Misspecification can lead to biased estimates and poor fit (MacCallum et al., 1992).
12. Outliers: An outlier is an observation that significantly deviates – either much larger or much smaller – from other values in a dataset (Jenatabadi, 2015),

13. Non-spuriousness: The observed covariance between variables must indicate a causal relationship, such that no other variable influences the outcome (Bollen, 2013)
14. Sequence and Causality: A causal relationship between endogenous and exogenous variables must be established, such that the cause precedes the effect (Bollen, 2013).

METHODOLOGY

SEM consists of five logical steps, including model specification, model identification, parameter estimation, model fit evaluation, and model modification (Fan et al., 2016). Below, each step is described.

Model Specification: Model specification defines the hypothesized relationships by determining which relevant variables will be included in the regression equation. These hypotheses should be grounded in prior empirical research, based on a thorough literature review, and theoretical underpinnings. To conduct this step, the latent variables and their corresponding indicators should be identified. Structural relationships between the latent variables should be defined, with a path diagram representing the hypothesized relationships through directional arrows.

Model Identification: Next, model identification assesses if there is enough information to identify a unique solution for the model's parameters. The model can be under-identified, just-identified, or over-identified (Kenny, 2024). As reported by Fan (2016, p. 8), Kline (2010) states the following regarding model identification: "(1) 'the model degrees of freedom must be at least zero to assure the degrees of freedom (*df*) is greater than zero; (2) 'every latent variable (including the residual terms) must be assigned a scale, meaning that either the residual terms' (disturbance) path coefficient and one of the latent variable's factor loading should be fixed to 1 or that the variance of a latent variable should be fixed to 1'; and (3) 'every latent variable should have at least two indicators.'

While an over-identified model contains ample information to estimate the model parameters, a just-identified model reflects a perfect fit that will reflect the observed data and contains only enough information to estimate the model parameters. By contrast, an under-identified model refers to one that does not contain enough unique solutions for the parameter; in essence, there is a lack of information, and consequently, the model cannot be estimated.

Parameter Estimation: The third step consists of parameter estimation, estimating the unknown parameters, or betas, in the model which can be determined through maximum likelihood estimation (Fan et al., 2016). This requires determining the maximum likelihood of choosing parameters that represent data that we would observe in the population. Otherwise stated, multivariate normality must be assessed for continuous distributions (Nachtigall et al., 2003). One benefit of maximum likelihood estimation is that, as the sample size increases, it will yield an efficient and unbiased estimate (Kaplan, 2001). The model must be either over-identified or just-identified in order for model coefficients to be estimated.

Model Fit: Next, the model fit must be evaluated to determine the data's reproducibility or how well the model fits the data. This is achieved by determining the degree of closeness between the theory and reality, or the estimated covariance matrix and the observed covariance matrix, respectively, such that the data is reproducible (Bhale & Bedi, 2023; Kenny, 2024). Model fit allows for comparisons between models to determine which one most closely conforms to the data actually observed in the field. To determine if empirical data supports the model, there must be an acceptable fit between the assumed dependencies between the latent variables and the assumed relationships between the manifest and latent variables (Nachtigall et al., 2003). Because the latent variables, in and of themselves, are not measurable, observed indicators used to operationalize the independent variables are associated with the observed indicators of the dependent variables through the loadings of the measures on their corresponding latent factors. To determine if the model showing patterns of associations reasonably reflects the actual associations that we observe in our empirical data, we can use model fit tests.

Fit Indices: In SEM several tests can be used. One test that assesses exact fit is the standardized root mean squared residual (SRMR). Values less than .05 represent good fit, while values less than .08 represent an acceptable fit (Hu & Bentler, 1999; Pavlov, Maydeu-Olivares, & Shi, 2021). A traditional and notable test is the chi-square test of model fit (CMIN). The chi-square test will show a non-significant result if the model is a good fit ($p > .05$), small value, and small degrees of freedom, supporting the null hypothesis that the matrices are not statistically significantly different. By contrast, if the null hypothesis is rejected, then the model will indicate that the matrices are not equal and represent a poor fit (Bhale & Bedi, 2023). This analysis is sensitive to sample size and will likely reject the null hypothesis with large sample sizes (Hooper et al, 2008). While this test assesses exact model fit, the normed chi-square test adjusts χ^2 for model complexity. Less than 3.0 is considered acceptable, while less than 2.0 represents good fit (Schermelleh-Engel et al., 2003).

As an alternative to the chi-square test, Jöreskog and Sorbom's (1993) goodness of fit index (GFI) indicates how well theoretical model represents the actual data by calculating the proportion of variance in the observed covariance matrix that is accounted for by the estimated covariance matrix. The GFI summarizes the overall fit. With values ranging between 0 and 1, a GFI closer to 0 represents a poor fit, while a value closer to 1 represents a good fit. A value greater than .95 is optimal (Hooper et al., 2008). Similarly, the adjusted goodness of fit index (AGFI), which considers the degrees of freedom of a model in relation to the number of variables and prevents over-fitting, also suggests that values closer to 1 (>.90), represent a good-fitting model. Both of these indices are also sensitive to sample size – the AGFI increasing with larger sample sizes – and although their use has been discouraged (Sharma et al, 2005), others recommend continuing their use along with other indices, rather than as a standalone measure (Hooper, 2008).

Accommodating for large sample sizes and data complexity, the root mean square error of approximation (RMSEA) – one of the most widely used measures of model fit – is used to approximate a close fit in contrast to an exact fit (Kaplan, 2000). The following values represent the interpretation of fit: 0 = perfect fit, ≥.05 = good fit, .05 - .08 = fair fit, >.10 = poor fit.

The incremental fit index (IFI) places the researcher's model on a continuum of fit, from best fit (1) to worst possible fit (0) (Kenny, 2024). The simplest IFI – the normed fit index (NFI) – adjusts for the sample size and degrees of freedom and compares the null model to the baseline model (Bollen, 1989; Schmukle & Hardt, 2005). A value of one equals an ideal fit; >1 equals an overfit; <.9 suggests that improvement is needed. Hu and Bentler (1999) adjusted the standard of excellent fit to a value of <.95 but have also suggested that interpretations should consider model complexity and sample size, allowing for a value of .9 to be suitable. To prevent underestimation of fit that might be noted in small sample sizes for this index, the comparative fit index (CFI) can be used (Bentler, 1990). This index, which compares the theoretical model with data, should result in a value of ≥ .90 for a good fit, with a value of 1 equaling a perfect fit. Another IFI, but non-normed, is the Tucker-Lewis Index (TLI), which compares the fit of a given substantive model against a null model (Cai et al., 2021). This index has a penalty for adding parameters (Kenny, 2024). Because higher values suggest a better fit, as 1 equals ideal fit; >1 equals overfit, and <.9 equals poor fit.

Lastly, parsimony-adjusted fit indices (PRATIO) provide information about model fit relative to model complexity. Otherwise stated, when multiple explanations for fitting a wide range of observed data exist, the simplest model is usually the best fit; this conclusion harkens to Occam's Razor, which proposes that the simplest explanation is usually the correct one (Falk & Muthukrishna, 2021). While there are no universally accepted cutoff values for PNFI and PCFI, higher values indicate a better balance between model fit and parsimony (Mulaik et al., 1989). A chi-square test that also assesses parsimony is probability of close fit

(PCLOSE). Values above .05 – the null hypothesis – represents a close fit (MacCallum et al., 1996; Madhanagopal & Amrhein, 2019).

When determining which index to report, if all (or most) indices lead to the same conclusion, report one or more indices based on your preference or the journal's preference. If the results are inconsistent, re-examine and analyze the model. If inconsistency continues, report multiple indices. Table 1 presents widely used fit indices, their interpretations, recommended cutoff criteria, and authoritative sources, offering a comprehensive guide for evaluating SEM results in applied research.

Fit Index	Interpretation	Recommended Cutoff	Reference
Chi-Square (χ^2) - Chi-Square Test of Model Fit	Tests exact model fit	Non-significant ($p > .05$)	Kline (2016)
Chi-Square/df (Normed χ^2)	Adjusts χ^2 for model complexity	< 3.0 (acceptable); < 2.0 (good)	Schermelleh-Engel et al. (2003)
CFI (Comparative Fit Index)	Compares target model to null model	≥ .90 (acceptable); ≥ .95 (good)	Hu & Bentler (1999)
TLI (Tucker-Lewis Index)	Penalizes complex models	≥ .90 (acceptable); ≥ .95 (good)	Hu & Bentler (1999)
NFI (Normed Fit Index)	Compares model to null model	≥ .90 (acceptable)	Bentler & Bonett (1980)
SRMR - Standardized Root Mean Square Residual	Average standardized residuals	≤ .08 (acceptable); ≤ .05 (good)	Hu & Bentler (1999)
RMSEA - Root Mean Square Error of Approximation	Measures fit per degree of freedom	≤ .08 (acceptable); ≤ .06 (good)	Hu & Bentler (1999); Browne & Cudeck (1993)
PCLOSE - Probability of Close Fit	Tests if RMSEA ≤ .05 (close fit)	$p > .05$ indicates close fit	MacCallum et al. (1996)
GFI - Goodness-of-Fit Index	Proportion of variance explained	≥ .90 (acceptable)	Jöreskog & Sörbom (1984)
AGFI - Adjusted Goodness-of-Fit Index	GFI adjusted for model complexity	≥ .90 (acceptable)	Jöreskog & Sörbom (1984)
PNFI - Parsimony Normed Fit Index	NFI adjusted for parsimony	No strict cutoff; higher is better	Mulaik et al. (1989)
PCFI - Parsimony Comparative Fit Index	CFI adjusted for model parsimony	No strict cutoff; higher is better	Mulaik et al. (1989)

Table 1. SEMs Fit Indices

Model Modification: Finally, model modification requires adjustments to be made to the model to improve its fit. To adjust the model, paths can be added or removed. Additionally, model re-specification can be used to improve the fit. To maintain theoretical integrity, modifications should be based on theory. To conduct this step, there are two approaches that can be used: (1) releasing constraints, also considered as "forward search," consists of adding free parameters; and (2) imposing constraints, also considered as "backward search," consists of deleting free parameters.

RESULTS

This paper conducts path analysis using SPSS AMOS, a widely used software for SEM that provides a graphical interface to draw and estimate models. The program provides intuitive tools for drawing, estimating, and interpreting SEMs, making it accessible for researchers at various levels. The dataset, Job Satisfaction, was collected by the first author as part of research on job behavior among social service employees. It comprises responses from a random sample of 218 individuals employed in social services agencies, who completed a self-administered survey that included well-known and validated scales measuring various job behaviors and job-related facets, along with demographic and personal characteristics (Abu-Bader, 2021). The following steps outline the SEM procedure in AMOS (Arbukle, 2017):

Step 1: Specify the Model

- Use the SPSS AMOS draw tool (figure 2) to represent observed variables (rectangles) and latent variables (oval).

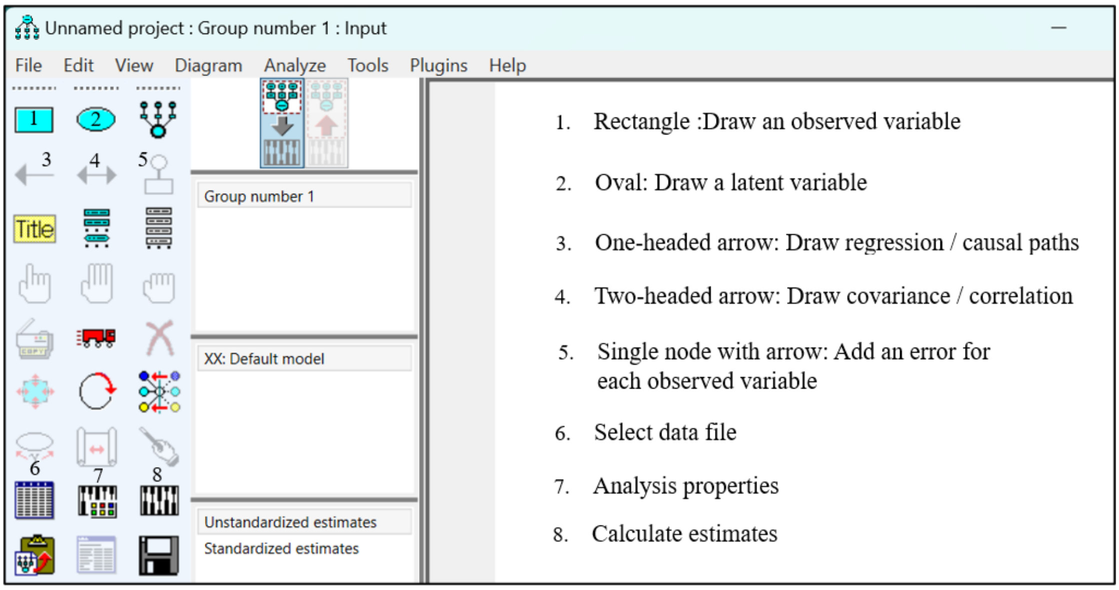

Figure 2. SPSS AMOS Main Screen

- Define causal paths using directional arrows and covariances using two-headed arrows (figure 3).
- Ensure that identification rules are met (e.g., one fixed loading per latent variable).

Step 2: Load the Dataset

- Import the dataset from SPSS (figure 2 icon 6).
- Assign variables by double-clicking on each box in the diagram.

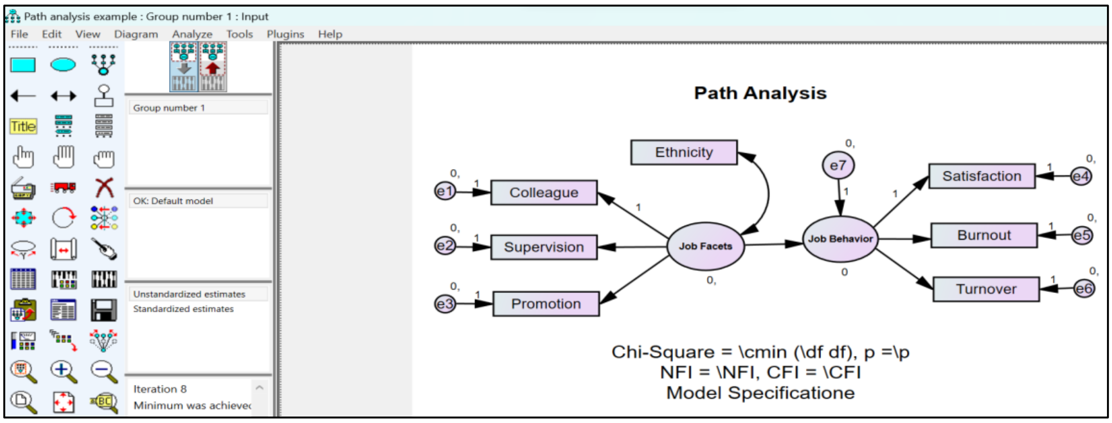

Figure 3. Path Diagram with Estimated Parameters – Input

Step 3: Estimate the Model

- Use Maximum Likelihood estimation or robust alternatives if data are non-normal.
- Click "Calculate Estimates" (figure 2 icon 8) to run the model.

Step 4: Assess Model Fit

AMOS provides multiple fit indices. The results of SEM path analysis are displayed in Table 2 (A-I) and figure 4. Here we will assess the model fit which can be evaluated using multiple fit indices provided by AMOS. These included the chi-square statistic (CMIN), the comparative fit index (CFI), the Tucker–Lewis index (TLI), the incremental fit index (IFI), the root mean square error of approximation (RMSEA), and the chi-square to degrees of freedom ratio (CMIN/DF).

As we discussed earlier, a non-significant chi-square value indicates acceptable model fit, although this test is known to be sensitive to sample size. Values greater than .95 for the CFI, TLI, and IFI are considered indicative of excellent fit (Hu & Bentler, 1999). An RMSEA value below .06 reflects good fit, and a CMIN/DF ratio less than 2 is typically viewed as desirable (Kline, 2016).

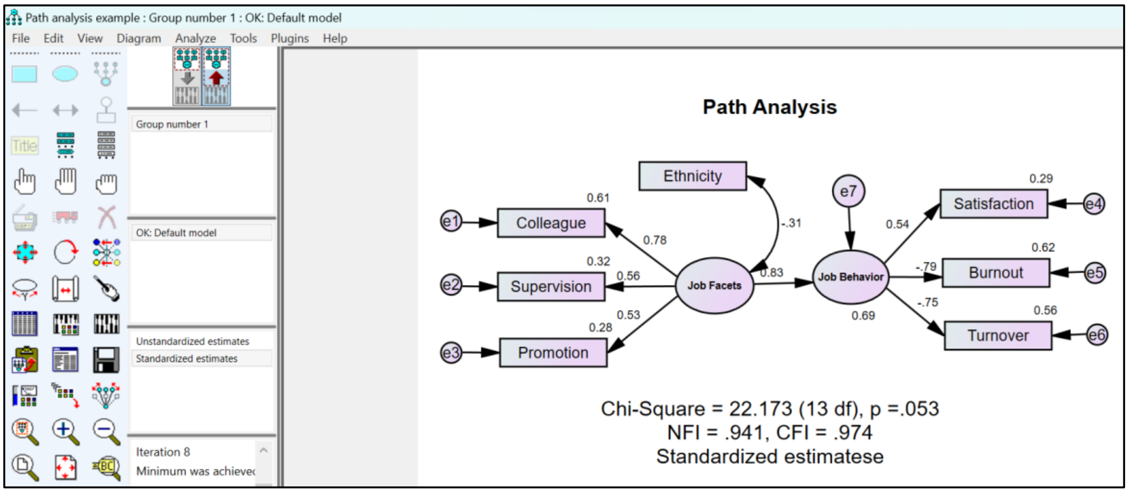

Figure 4. Path Diagram with Estimated Parameters – Output

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	22	22.173	13	.053	1.706
Saturated model	35	.000	0		
Independence model	7	377.433	28	.000	13.480

Table 2A. CMIN

The model fit indices for the default, saturated, and independence models are presented in Table 2A. The **default model** (the hypothesized model) included 22 estimated parameters (**NPAR**) and produced a chi-square (**CMIN**) value of 22.173 with 13 degrees of freedom, yielding a *p*-value of .053. This non-significant *p*-value indicates that the model's fit to the data is acceptable, as it suggests that the observed covariance matrix does not significantly differ from the model-implied matrix. The chi-square to degrees of freedom ratio (CMIN/DF) was 1.706, which is below the commonly recommended threshold of 2.0, further supporting a good model fit (Kline, 2016).

The **saturated model**, which perfectly reproduces the observed data by estimating all variances and covariances, had a chi-square of 0 and 0 degrees of freedom, as expected. The **independence model**, which assumes that all variables are uncorrelated, showed a poor fit with a chi-square of 377.433 (df = 28, $p < .001$)

and a very high CMIN/DF of 13.480. This stark contrast in fit between the independence model and the default model underscores the plausibility and improvement of the hypothesized model.
Overall, the fit indices suggest that the hypothesized (default) model provides an adequate to good fit to the data.

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.941	.873	.975	.943	.974
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Table 2B. Baseline Comparisons

Incremental fit indices for the default, saturated, and independence models are reported in Table 2B. The **default model** demonstrated strong fit based on several widely used indices. The Normed Fit Index (**NFI**) was .941, and the Relative Fit Index (**RFI**) was .873. While the **NFI** exceeds the commonly accepted threshold of .90, the **RFI** is slightly below the more conservative threshold of .90, indicating room for improvement in model parsimony (Bentler & Bonett, 1980).
The Incremental Fit Index (IFI = .975), Tucker–Lewis Index (TLI = .943), and Comparative Fit Index (CFI = .974) all exceeded the .95 benchmark, indicating excellent model fit (Hu & Bentler, 1999). These results support the conclusion that the hypothesized model provides a substantially better fit to the data compared to the null (independence) model.

As expected, the **saturated model**, which perfectly fits the data, yielded values of 1.000 across all indices. In contrast, the **independence model** produced values of 0.000 for all fit indices, reflecting its poor fit due to its assumption of no relationships among variables. Collectively, these results suggest that the default model demonstrates good to excellent fit according to conventional standards for incremental fit indices.

Model	PRATIO	PNFI	PCFI
Default model	.464	.437	.452
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

Table 2C. Parsimony-Adjusted Measures

As shown in Table 2C, the **default model** yielded a Parsimony Ratio (**PRATIO**) of .464, indicating that approximately 46.4% of the possible degrees of freedom were used by the model. The Parsimony Normed Fit Index (**PNFI**) and Parsimony Comparative Fit Index (**PCFI**) were .437 and .452, respectively. These results suggest that the model demonstrates a reasonable level of parsimony while maintaining good fit.

As expected, the **saturated model**, which has zero degrees of freedom and no model constraints, showed PRATIO, PNFI, and PCFI values of 0.000. In contrast, the **independence model**, which assumes no relationships among variables and uses all available degrees of freedom, yielded a PRATIO of 1.000, but its PNFI and PCFI remained at 0.000 due to poor model fit.
Overall, the parsimony indices support the adequacy of the default model by indicating an acceptable trade-off between model complexity and goodness of fit.

Model	NCP	LO 90	HI 90
Default model	9.173	.000	26.281
Saturated model	.000	.000	.000
Independence model	349.433	290.394	415.915

Table 2D. NCP

The non-centrality parameter (**NCP**) offers an estimate of model misfit, where lower values suggest better model fit. Table 2D presents the NCP and its 90% confidence interval for the default, saturated, and independence models. The **default model** yielded an NCP of 9.173, with a 90% confidence interval ranging from 0.000 to 26.281. The lower bound at zero and the relatively narrow upper bound indicate minimal model misfit, further supporting the adequacy of the hypothesized model.

As expected, the **saturated model**, which perfectly fits the data, had an NCP of 0.000 with both lower and upper bounds at zero, reflecting exact fit. In contrast, the **independence model** showed a very high NCP of 349.433, with a 90% confidence interval ranging from 290.394 to 415.915, indicating substantial misfit. These findings reinforce the conclusion that the default model fits the data substantially better than the independence model and introduces minimal misfit.

Model	FMIN	F0	LO 90	HI 90
Default model	.102	.042	.000	.121
Saturated model	.000	.000	.000	.000
Independence model	1.739	1.610	1.338	1.917

Table 2E. FMIN

Table 2E presents the minimum value of the discrepancy function (FMIN) and the estimated population discrepancy function (F0), along with its 90% confidence interval. The **default model** yielded an FMIN of .102 and an estimated F_0 of .042, with a 90% confidence interval ranging from 0.000 to .121. These values suggest that the model shows a low level of discrepancy between the observed and model-implied covariance matrices. The inclusion of zero in the lower bound of the confidence interval for F_0 indicates the possibility of a near-perfect fit in the population.

In contrast, the **saturated model**—which perfectly reproduces the sample data—had FMIN and F_0 values of 0.000 with a confidence interval also fixed at zero, as expected for a model with no degrees of freedom. The **independence model**, which assumes all variables are uncorrelated, exhibited a much higher FMIN of 1.739 and an F_0 of 1.610, with a 90% confidence interval from 1.338 to 1.917, indicating substantial misfit. Overall, the low F0 and narrow confidence interval for the default model support the conclusion that it adequately represents the underlying data structure with minimal discrepancy.

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.057	.000	.097	.348
Independence model	.240	.219	.262	.000

Table 2F. RMSEA

Table 2F displays the Root Mean Square Error of Approximation (RMSEA), its 90% confidence interval, and the *p*-value for close fit (PCLOSE) for the default and independence models. The **default model** produced an RMSEA value of .057, with a 90% confidence interval ranging from 0.000 to .097. This value falls just below the commonly accepted threshold of .06, indicating good approximate fit to the population data (Hu & Bentler, 1999). The confidence interval includes zero, which suggests that close fit cannot be ruled out. Furthermore, the

PCLOSE value of .348 exceeds the .05 threshold, indicating that the hypothesis of close fit (*i.e.*, RMSEA < .05) cannot be rejected.

In contrast, the **independence model** yielded an RMSEA of .240, with a 90% confidence interval from .219 to .262 and a PCLOSE value of .000, clearly indicating poor fit. These values reflect the model's severe misfit due to its assumption of uncorrelated variables. Overall, the RMSEA results support the conclusion that the default model provides an acceptable fit to the data, while the independence model performs poorly.

Model	AIC	BCC	BIC	CAIC
Default model	66.173	67.857		
Saturated model	70.000	72.679		
Independence model	391.433	391.968		

Table 2G. AIC

Table 2G reports several information criteria used to assess model fit while accounting for model complexity: Akaike Information Criterion (AIC), Browne–Cudeck Criterion (BCC), Bayesian Information Criterion (BIC), and Consistent AIC (CAIC). Lower values on these indices indicate better relative fit and parsimony.

The **default model** yielded an AIC of 66.173 and a BCC of 67.857, both of which are lower than those of the **saturated model** (AIC = 70.000; BCC = 72.679), suggesting that the default model achieves a better balance between model fit and complexity. The **independence model** showed much higher values (AIC = 391.433; BCC = 391.968), indicating poor fit.

While BIC and CAIC values were not reported for the default and saturated models in this table, the substantially higher AIC and BCC values for the independence model clearly reflect its inferior fit. The available information supports the conclusion that the default model outperforms both the saturated and independence models in terms of parsimony-adjusted fit**.**

Model	ECVI	LO 90	HI 90	MECVI
Default model	.305	.263	.384	.313
Saturated model	.323	.323	.323	.335
Independence model	1.804	1.532	2.110	1.806

Table 2H. ECVI

Table 2H presents the Expected Cross-Validation Index (ECVI), its 90% confidence interval, and the Modified ECVI (MECVI) for the default, saturated, and independence models. The **default model** had an ECVI of .305, with a 90% confidence interval ranging from .263 to .384, and a MECVI of .313. These relatively low values suggest that the model is likely to perform well when applied to other samples from the same population, indicating good generalizability.

The **saturated model**, which perfectly reproduces the sample data, showed a slightly higher ECVI of .323 and MECVI of .335. Although the saturated model achieves perfect fit, it may be overfitted and less parsimonious. Notably, the ECVI of the default model is slightly lower than that of the saturated model, which suggests that the default model achieves a better balance between fit and complexity.

In contrast, the **independence model** yielded a substantially higher ECVI of 1.804 (90% CI: 1.532 to 2.110) and a MECVI of 1.806, indicating poor fit and weak generalizability. Overall, the ECVI and MECVI values further support the adequacy and relative parsimony of the default model compared to the saturated and independence models.

Model	HOELTER .05	HOELTER .01
Default model	219	271
Independence model	24	28

Table 2I. HOELTER

Table 2I presents Hoelter's Critical N values at the .05 and .01 significance levels, which estimate the minimum sample size required for the model to be considered a good fit based on the chi-square test. For the **default model**, the CN was 219 at the .05 level and 271 at the .01 level. Both values exceed the commonly cited threshold of 200, suggesting that the model has an acceptable level of fit and that the current sample size is adequate for detecting meaningful model–data discrepancies (Hoelter, 1983).

In contrast, the **independence model** yielded substantially lower CN values—24 at the .05 level and 28 at the .01 level—indicating extremely poor fit. These values reflect the model's failure to adequately represent the observed data structure. Taken together, the Hoelter CN values support the conclusion that the **default model fits the data well and is supported by a sufficient sample size**, whereas the independence model is inadequate.

Step 5: Interpret Parameters

- Examine regression weights, standardized estimates, and R-squared values.
- Significant paths validate hypothesized relationships.

Tables 3 (A – B) display the results of the regression weights and squared multiple correlation coefficients.

			Estimate	S.E.	C.R.	P	Label
Job Behavior	<---	Job Facets	.596	.099	6.050	***	par_6
Satisfaction	<---	Job Behavior	1.000				
Turnover	<---	Job Behavior	-.723	.102	-7.072	***	par_1
Colleague	<---	Job Facets	1.000				
Supervision	<---	Job Facets	1.251	.178	7.030	***	par_2
Promotion	<---	Job Facets	.595	.089	6.658	***	par_3
Burnout	<---	Job Behavior	-1.044	.145	-7.187	***	par_5

Table 3A. Regression Weights: (Default model)

Table 3A presents the unstandardized regression weights, standard errors (S.E.), critical ratios (C.R.), and *p*-values for the hypothesized structural relationships in the model. All estimated paths were statistically significant at $p < .001$, as indicated by the "***" notation.

The latent construct **Job Behavior** was significantly predicted by **Job Facets** (Estimate = 0.596, S.E. = 0.099, C.R. = 6.050, $p < .001$), supporting the hypothesized positive relationship. In turn, **Job Behavior** significantly predicted several outcomes. It had a strong negative effect on **Turnover** (Estimate = –0.723, S.E. = 0.102, C.R. = –7.072, $p < .001$) and on **Burnout** (Estimate = –1.044, S.E. = 0.145, C.R. = –7.187, $p < .001$), while positively predicting **Satisfaction**, which was fixed at 1.000 for model identification.
Within the **Job Facets** construct, **Colleague** was used as a reference indicator (fixed at 1.000). Other indicators loaded significantly: **Supervision** (Estimate = 1.251, S.E. = 0.178, C.R. = 7.030, $p < .001$) and **Promotion** (Estimate = 0.595, S.E. = 0.089, C.R. = 6.658, $p < .001$), indicating they are valid and reliable indicators of the latent construct. Overall, the path estimates support the hypothesized relationships in the model, with all paths showing strong statistical significance and meaningful effect sizes.

			Estimate
Job Behavior	<---	Job Facets	.831
Satisfaction	<---	Job Behavior	.540
Turnover	<---	Job Behavior	-.747
Colleague	<---	Job Facets	.783
Supervision	<---	Job Facets	.563
Promotion	<---	Job Facets	.526
Burnout	<---	Job Behavior	-.790

Table 3B. Standardized Regression Weights: (Default model)

Table 3B reports the standardized regression weights for the structural and measurement paths in the model. All reported paths indicate the strength and direction of the relationships between latent constructs and their observed indicators, as well as between latent variables.

The latent variable **Job Facets** had a strong positive effect on **Job Behavior** ($\beta = .831$), indicating that higher levels of perceived job facets (e.g., colleague relations, supervision, and promotion opportunities) are strongly associated with more positive job behavior. **Job Behavior**, in turn, positively predicted **Satisfaction** ($\beta = .540$) and negatively predicted both **Turnover** ($\beta = -.747$) and **Burnout** ($\beta = -.790$). These findings suggest that greater engagement in positive job behaviors is associated with increased job satisfaction and reduced intentions to leave and emotional exhaustion.

Regarding the measurement model, the standardized factor loadings of the indicators for **Job Facets** were all substantial: **Colleague** ($\beta = .783$), **Supervision** ($\beta = .563$), and **Promotion** ($\beta = .526$). These values indicate that all three observed variables are reliable indicators of the underlying construct, with **Colleague** being the most strongly associated. Overall, the standardized estimates provide strong empirical support for the proposed model, confirming both the strength and directionality of the hypothesized relationships.
Example Case: Job Behavior Model A demonstration model tests the impact of Job Facets on Job Behavior and various outcomes. The model includes two latent variables (Job Facets and Job Behavior) and several observed outcomes (e.g., Satisfaction, Turnover).

Key results:

- **Model Fit:** $\chi2(13) = 22.173$, $p = .053$; CFI = .974; RMSEA = .057
- **Path Coefficients:** Job Behavior ← Job Facets: $\beta = .831$; Burnout ← Job Behavior: $\beta = -.790$.

The hypothesized structural equation model demonstrated acceptable fit to the data, $\chi^2(13) = 22.173$, $p = .053$, CFI = .974, RMSEA = .057, indicating that the model adequately represents the observed relationships. The structural path from **Job Facets** to **Job Behavior** was strong and positive ($\beta = .831$), suggesting that favorable perceptions of job characteristics are associated with more constructive job-related behaviors. In turn, **Job Behavior** negatively predicted **Burnout** ($\beta = -.790$), indicating that higher levels of positive job behavior are associated with lower levels of employee burnout.

DISCUSSION

In assessing the model fit, the structural equation model (SEM) demonstrated acceptable to excellent fit across multiple indices. The chi-square test of model fit was non-significant, $\chi^2(13) = 22.173$, $p = .053$, indicating that the model's implied covariance structure did not significantly differ from the observed data. The chi-square/degrees of freedom ratio (CMIN/DF = 1.71) also fell below the recommended threshold of 2.0, suggesting good model fit. Incremental fit indices further supported the adequacy of the model. The

Comparative Fit Index (CFI = .974), Incremental Fit Index (IFI = .975), and Tucker–Lewis Index (TLI = .943) all exceeded or approached the .95 benchmark for excellent fit (Hu & Bentler, 1999). The RMSEA value was .057 with a 90% confidence interval of [.000, .097], indicating good approximate fit, and the PCLOSE value (.348) suggested that the hypothesis of close fit (RMSEA < .05) could not be rejected. Parsimony-adjusted indices also supported model adequacy, with PNFI = .437 and PCFI = .452, reflecting a balance between model complexity and goodness of fit. The model demonstrated lower values on AIC (66.173) and BCC (67.857) compared to the saturated and independence models, indicating superior parsimony-adjusted fit. Additionally, the Expected Cross-Validation Index (ECVI = .305) and Modified ECVI (MECVI = .313) were lower than those of the saturated and independence models, suggesting better generalizability. Hoelter's Critical N values (CN = 219 at $p = .05$; CN = 271 at $p = .01$) exceeded the threshold of 200, indicating adequate sample size to support model stability.

Regarding the path estimates, all hypothesized structural paths were statistically significant (all *p*s < .001). Standardized estimates revealed a strong positive relationship between **Job Facets** and **Job Behavior** (β = .831), suggesting that more favorable job characteristics (e.g., quality of supervision, promotion opportunities, and collegial support) are associated with increased constructive job behavior.
In turn, **Job Behavior** positively predicted **Satisfaction** (β = .540) and negatively predicted both **Turnover** (β = –.747) and **Burnout** (β = –.790), indicating that employees who engage in more positive job behaviors are more satisfied and less likely to experience emotional exhaustion or consider leaving their jobs. The results of the SEM path analysis are displayed in tables 4 and 5.

Predictor	Outcome	Estimate (β)	SE	CR	p
Job Facets	Job Behavior	.831	.099	6.050	< .001
Job Behavior	Satisfaction	.540	—	—	—
Job Behavior	Turnover	-.747	.102	-7.072	< .001
Job Facets	Colleague	.783	—	—	—
Job Facets	Supervision	.563	.178	7.030	< .001
Job Facets	Promotion	.526	.089	6.658	< .001
Job Behavior	Burnout	-.790	.145	-7.187	< .001

Table 4. Standardized Regression Weights (N = 218)

Fit Index	Value	Recommended Cutoff
χ^2 (13)	22.173	p = .053 (ns)
χ^2/df	1.71	< 3.00
CFI	.974	> .95
TLI	.943	> .90
RMSEA	.057	< .06
AIC	66.17	(for model comparison)
Hoelter (0.05)	219	> 200

Table 5. Model Fit Indices

Implications to Research and Practice

SEM offers numerous implications to research and practice. Contrasting, traditional research methods, which focused on testing one bivariate or multivariate relationship using a single analysis, SEM integrates several disciplines into one model fitting framework. These include confirmatory factor analysis from education (Abrahim, et al., 2019; Ampofo & Aidoo, 2022); measurement error theory from psychology (DeShon, 1998); regression analysis from statistics (Musil, Jones, & Warner, 1998); path analysis from nutrition and epidemiology (Harris, et al., 2022; Christ, Lee, Lam, & Zheng, 2014); and simultaneous equations from econometrics (Bentler, 1983). Other special cases of SEM include multiple and multivariate regression; analysis of variance and covariance (ANOVA / ANCOVA); recursive & nonrecursive models; and classical test theory. The example model tested in this paper, which represents research from the social science field, demonstrates SEM's ability to test complex relationships while controlling for measurement error.

Research methodologies that are well-suited to investigate complex relationships are especially important to social science fields, where client groups often present with multi-layered complexities due to intersecting social dynamics that might compound problems. Such complexities inform theory development, helping to explain the multi-faceted nature of the relationship between observed constructs. Beyond that, SEM allows for the inclusion of unobservable constructs, without which, deeper explanations might be limited or omit critical contributions underlying complex problems. SEM's integration of theory into statistical analysis addresses the limitations of traditional research methods – which were unable to build a theory – while simultaneously analyzing multiple causal relationships consisting of unobservable variables through longitudinal patterns (Cao, 2023). Ultimately, theoretical models developed from advanced methodologies that account for measurement error, such as SEM, can underpin important interventions that improve clients' well-being.

CONCLUSION

Many techniques have been used to analyze relationships, but few of them excel above SEM. An interdependent technique that evaluates multiple hypotheses simultaneously, SEM has evolved from a two-factor model (Spearman, 1904) to a multi-factor system, complete with both latent and observable variables that create a path diagram useful for explaining complex relationships. This technique is especially valuable for social science research, which notably investigates problems layered with multi-dimensional complexities. Good model fit and significant paths underscore the theoretical validity of the hypothesized relationships.

Though the technique is lauded for its abilities to evaluate such complexities, one important limitation of SEM is the possibility of overfitting or erroneous conclusions if assumptions are not properly tested or theoretically grounded. Researchers must remain vigilant regarding these considerations. If misapplied, SEM's power can lead to overfitting or erroneous conclusions (MacCallum et al., 1992). Additionally, researchers must contend with issues pertaining to sample size requirements and causation when using SEM.

In this paper, we developed a path analysis using SPSS AMOS, with results demonstrating acceptable to excellent fit across multiple indices. This paper sought to provide insight into the foundational principles and assumptions of SEM and its application using SPSS AMOS, while adding to the body of knowledge of advanced methodological techniques useful for social science research.

Future Research

Future research is ripe for use of SEM, especially in fields such as those in the social sciences. As previously mentioned, use of SEM within the context of multi-layered complexities seen in such fields allow for the examination of endless latent constructs, furthering the development of tailored theoretical models and interventions for a larger variety of intersecting group dynamics. This feature of SEM is one of the many benefits of the technique, and there are others presented throughout the literature – its correction for unreliability of measurement error; its ability to test and compare model fitness; and its ability to specify causal chains for complex sets of relationships. However, the disadvantages of SEM have also been noted. First, the technique is not well-suited for small samples (Ramlall, 2016). Large sample size is a requirement for SEM, as the accuracy of SEM decreases with smaller sample sizes. To this, Wolf et al. (2013) suggests that statistical power analysis is not the sole requirement for achieving an adequate sample size; they also recommend addressing considerations pertaining to bias and missing data (Tomarken & Waller, 2005) when achieving adequate sample size. Future research can target methods that allow for utilizing SEM with smaller sample sizes without compromising statistical power or model fit.

Additionally, despite SEM's ability to merge several analytical methods together, the method does not inherently allow for causal conclusions to be drawn, although, under specific assumptions, the technique can be conducted to assume causality (Nachtigall et al., 2003). When causation is sought, SEM achieves this goal, moving beyond mere correlation – a term often incorrectly used interchangeably with causation – to allow for the creation of accurate theoretical models that offer a deeper and more nuanced explanation of the connection between variables through path analyses or longitudinal structural causal models (Cao, 2023; Madhanagopal & Amrhein, 2019). Given the importance of the development of complex theoretical models for social science research, an area of future research for this consideration can focus on methods that allow SEM to draw causal conclusions without current assumption restraints. Lastly, SEM does not currently allow for the conclusive determination of model validity. In a general sense, validity ensures accuracy. While theoretical models often undergo rigorous testing that account for errors, the inability for SEM to ensure model validity presents an area for future research, in which SEMs would produce theoretical models that are accurate, better improving the underpinning of theoretically-based interventions.

REFERENCES

Abrahim, S., Mir, B.A., Suhara, H. et al. (2019) Structural equation modeling and confirmatory factor analysis of social media use and education. International Journal of Education Technology in Higher Education, 16(32). https://doi.org/10.1186/s41239-019-0157-y

Abu-Bader, S. (2021) Using Statistical Methods in Social Science Research with a Complete SPSS Guide, (3rd ed.). Oxford University Press.

Abu-Bader, S. (2010) Advanced and Multivariate Statistical Methods for Social Science Research. Oxford University Press.

Ampofo, R. T., & Aidoo, E. N. (2022) Structural equation modelling of COVID-19 knowledge and attitude as determinants of preventive practices among university students in Ghana. Scientific African, 16, e01182. https://doi.org/10.1016/j.sciaf.2022.e01182

Arbuckle, J. L. (2017) IBM SPSS Amos 25. IBM.

Bentler, P. M. (1983) Simultaneous equation systems as moment structure models : With an introduction to latent variable models. Journal of Econometrics, 22(1-2),13-42.

Bentler P. M. (1990) Comparative fit indexes in structural models. Psychological Bulletin, 107(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Bentler, P. M., & Bonett, D. G. (1980) Significance tests and goodness of fit in the analysis of covariance structures. Psychological Bulletin, 88(3), 588–606. Bhale, U. & Bedi, H. (2023). SEM model fit indices meaning and acceptance literature support. Retrieved from SSRN http://dx.doi.org/10.2139/ssrn.4659352

Bollen, K. A. (1989) Structural equations with latent variables. Wiley.

Boomsma, A. (1982) The robustness of LISREL against small sample sizes. In K. G. Jöreskog & H. Wold (Eds.), Systems under indirect observation (Vol. 1, pp. 149–173).

Brown, T. A. (2015) Confirmatory factor analysis for applied research (2nd ed.). Guilford Press.

Browne, M. W., & Cudeck, R. (1993) Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), Testing structural equation models (pp. 136–162). Sage.

Byrne, B. M. (2016) Structural equation modeling with AMOS: Basic concepts, applications, and programming (3rd ed.). Routledge.

Cai, L., Chung, S. W., & Lee, T. (2023) Incremental model fit assessment in the case of categorical data: Tucker–Lewis index for item response theory modeling. Prevention Science, 24(3), 455–466. https://doi.org/10.1007/s11121-021-01253-4

Cao, X. (2023) The application of structural equation model in psychological research. CNS Spectrums, 28(S1), S17–S19. https://doi.org/10.1017/S1092852923000858

Christ, S. L., Lee, D. J., Lam, B. L., & Zheng, D. D. (2014) Structural equation modeling: a framework for ocular and other medical sciences research. Ophthalmic epidemiology, 21(1), 1–13. https://doi.org/10.3109/09286586.2013.867508

DeShon, R.P. (1998) A cautionary note on measurement error corrections in structural equation models. Psychological Methods, 3(4), 412–423. https://doi.org/10.1037/1082-989X.3.4.412

Dijkstra, T. K., & Henseler, J. (2015) Consistent partial least squares path modeling. MIS Quarterly, 39(2), 297–316. https://www.jstor.org/stable/26628355

Enders, C. K. (2010) Applied missing data analysis. Guilford Press.

Falk, C. F., & Muthukrishna, M. (2023) Parsimony in model selection: Tools for assessing fit propensity. Psychological Methods, 28(1), 123–136. https://doi.org/10.1037/met0000422

Fan, Y., Chen, J., Shirkey, G. et al. (2016) Applications of structural equation modeling (SEM) in ecological studies: An updated review. Ecological Processes, 5(19). https://doi.org/10.1186/s13717-016-0063-3

Harris, J. E., & Gleason, P. M. (2022) Application of Path Analysis and Structural Equation Modeling in Nutrition and Dietetics. Journal of the Academy of Nutrition and Dietetics, 122(11), 2023–2035. https://doi.org/10.1016/j.jand.2022.07.007

Hooper, D., Coughlan, J., & Mullen, M. R. (2008) Structural equation modelling: Guidelines for determining model fit. The Electronic Journal of Business Research Methods, 6, 53-60.

Hoyle, R. H. (2012) Handbook of structural equation modeling. Guilford Press. Kline, R. B. (2016). Principles and practice of structural equation modeling (4th ed.). Guilford Press.

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

Jenatabadi, H.S. (2015) A Critical Story About Sample Size, Outliers, and Normality Criteria in Structural Equation Modelling. http://doi.org/10.2139/ssrn.2599163

Jöreskog, K. G. & Sörbom, D. (1984) LISREL-VI user's guide (3rd ed.). Scientific Software.

Jöreskog, K. G., & Sörbom, D. (1993) LISREL 8: Structural equation modeling with the SIMPLIS command language. Scientific Software International; Lawrence Erlbaum Associates, Inc.

Kaplan, D.W. (2000) Structural Equation Modeling: Foundations and Extensions. Sage Publications, Inc.

Kenny, D. (2024) Measuring model fit. Retrieved from https://davidakenny.net/cm/fit.htm

Kline, R. B. (2016) Principles and practice of structural equation modeling (4th ed.). Guilford Press.

Levy, D. A. (2010) Tools of critical thinking: Metathoughts for psychology (2nd ed.). Waveland Press.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996) Power analysis and determination of sample size for covariance structure modeling. Psychological Methods, 1(2), 130–149.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992) Model modifications in covariance structure analysis: The problem of capitalization on chance. Psychological Bulletin, 111(3), 490–504.

Madhanagopal, B. & Amrhein, J. (2019) Analyzing structural causal models using the CALIS procedure. Proceedings of the SAS Global Forum 2019 Conference. 3765-2019.

Mardia, K. V. (1970) Measures of multivariate skewness and kurtosis with applications. Biometrika, 57(3), 519–530.

Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., & Stilwell, C. D. (1989) Evaluation of goodness-of-fit indices for structural equation models. Psychological Bulletin, 105(3), 430–445.

Musil, C.M., Warner, C.B., Yobas, P.K., & Jones, S.L. (2002) A comparison of imputation techniques for handling missing data. Western Journal of Nursing Research. 24(7), 815-829. http://doi.org/10.1177/019394502762477004

Nachtigall, C., Kroehne, U., Funke, F., Steyer, R., & Schiller, F.V. (2003) (Why) should we use SEM? Pros and cons of structural equation modeling. Methods of Psychological Research Online, 8(2), 1-22

Pavlov, G., Maydeu-Olivares, A., & Shi, D. (2021) Using the standardized root mean squared residual (SRMR) to assess exact fit in structural equation models. Educational and Psychological Measurement, 81(1), 110-130. https://doi.org/10.1177/0013164420926231

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010) A general multilevel SEM framework for assessing multilevel mediation. Psychological Methods, 15(3), 209–233.

Ramlall, I. (2016) Applied structural equation modelling for researchers and practitioners: Using R and Stata for behavioural research. Emerald Group Publishing Limited.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003) Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. Methods of Psychological Research Online, 8(2), 23–74.

Schmuckle, S. C., & Hardt, J. (2005) A cautionary note on incremental fit indices reported by Lisrel. Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 1(2), 81–85. https://doi.org/10.1027/1614-1881.1.2.81

Sewall, W. (1918) On the nature of size factors. Genetics, 3(4), 367-374, https://doi.org/10.1093/genetics/3.4.367

Sharma, S., Mukherjee, S., Kumar, A., & Dillon, W.R. (2005) A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models. Journal of Business Research, 58(1), 935-43.

Spearman, C. (1904) 'General intelligence,' objectively determined and measured. The American Journal of Psychology, 15(2), 201–293. https://doi.org/10.2307/1412107

Tarka, P. (2022) Correction to: An overview of structural equation modeling: Its beginnings, historical development, usefulness and controversies in the social sciences: Quality & Quantity: International Journal of Methodology, 56(2), 885–887. https://doi.org/10.1007/s11135-022-01326-8

Tomarken, A. J., & Waller, N. G. (2005) Structural equation modeling: Strengths, limitations, and misconceptions. Annual Review of Clinical Psychology, 1, 31-65. https://doi.org/10.1146/annurev.clinpsy.1.102803.144239

West, S. G., Finch, J. F., & Curran, P. J. (1995) Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and applications (pp. 56–75). Sage.

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013) Sample size requirements for structural equation models. Educational and Psychological Measurement, 73(6), 913–934.