POLAC INTERNATIONAL JOURNAL OF ECONS & MGT SCIENCE (PIJEMS)
DEPARTMENT OF ECONOMICS & MANAGEMENT SCIENCE NIGERIA
POLICE ACADEMY, WUDIL-KANO

APPLICATION OF PARTIAL LEAST SQUARES STRUCTURAL EQUATION MODELING (PLS-SEM) IN PSYCHOLOGICAL RESEARCH: MEASUREMENT AND STRUCTURAL MODEL ASSESSMENT USING SMARTPLS

Mohammed Shamsudeen T. Department of Psychology, Nasarawa State University, Keffi

Akeem A. Kenku, PhD Department of Psychology, Nasarawa State University, Keffi

Shafa A. Yunus, PhD Department of Psychology, Nasarawa State University, Keffi

Abstract

Partial Least Squares Structural Equation Modeling (PLS-SEM) has emerged as a powerful and flexible multivariate analytical technique increasingly adopted in psychological research to examine complex theoretical models involving latent constructs. Unlike covariance-based SEM (CB-SEM), PLS-SEM operates through a component-based estimation approach that accommodates small sample sizes, non-normal data distributions, and formative measurement specifications, conditions frequently encountered in psychological investigations. This methodological paper provides a comprehensive guide to applying PLS-SEM using SmartPLS software, covering the philosophical foundations, model specification, and sequential assessment procedures for both measurement and structural models. The study delineates the two-stage assessment protocol: first evaluating the outer measurement model through reliability (Cronbach's Alpha, Composite Reliability, rho_A) and validity (convergent validity via Average Variance Extracted, and discriminant validity via Fornell-Larcker criterion, cross-loadings, and HTMT ratio) indices; and second, assessing the inner structural model through path coefficients, coefficient of determination (R^2), effect sizes (f^2), predictive relevance (Q^2), and Standardized Root Mean Square Residual (SRMR). The paper also addresses advanced techniques, including mediation and moderation analysis, common method bias assessment, and reporting standards for publication. Practical recommendations are offered throughout to assist psychological researchers in making methodologically sound decisions within the PLS-SEM framework.

Keywords: Partial Least Squares, Structural Equation Modeling, Psychological, Research, SMARTPLS

1. Introduction

The landscape of quantitative methodology in psychological research has undergone a significant transformation over the past four decades, with structural equation modeling (SEM) emerging as a dominant analytical paradigm for testing theoretically grounded hypotheses involving multiple latent variables simultaneously (Hair et al., 2019). Within the broad SEM tradition, two major methodological streams have developed: covariance-based SEM (CB-SEM), represented by software packages such as LISREL, AMOS, and Mplus; and variance-based SEM, most prominently instantiated in Partial Least Squares SEM (PLS-SEM), implemented in software tools including SmartPLS, WarpPLS, and PLS-Graph (Ringle et al., 2015).

PLS-SEM was originally developed by Wold (1975) as a component-based alternative to full information maximum likelihood estimation methods, with early applications concentrated in econometrics and chemometrics. The technique gained increasing traction in business research during the 1980s and 1990s, particularly in information systems and marketing (Fornell & Bookstein, 1982; Chin, 1998), before being progressively adopted across the behavioral and social sciences, including psychology, education, health sciences, and organizational behavior (Hair et al., 2017a).

The appeal of PLS-SEM in psychological research stems from several distinctive methodological properties. First, PLS-SEM does not require multivariate normality assumptions, making it particularly suited to psychological data where Likert-scale distributions frequently exhibit skewness and kurtosis (Hair et al., 2011). Second, PLS-SEM operates efficiently with relatively small sample sizes, addressing the perennial challenge in psychological research of recruiting adequate participant numbers for complex model testing. Third, PLS-SEM accommodates both reflective and formative measurement model specifications, a critical advantage given the theoretical complexity of many psychological constructs that resist purely reflective operationalization (Coltman et al., 2008).

Despite these advantages, PLS-SEM has not been without controversy. Scholars have raised concerns about its tendency to overestimate path coefficients, its inability to confirm global model fit in the manner of CB-SEM, and the absence of universally accepted model fit indices in earlier iterations of the methodology (Rönkkö & Evermann, 2013). Recent developments in PLS-SEM methodology, particularly the incorporation of the Standardized Root Mean Square Residual (SRMR) as a model fit index and the development of consistent PLS (PLSc), have substantially addressed these limitations (Dijkstra & Henseler, 2015; Henseler et al., 2014).

SmartPLS, developed by Ringle et al. (2015), has become the most widely adopted software platform for PLS-SEM implementation, offering an intuitive graphical user interface, comprehensive algorithmic output, and an expanding suite of advanced analytical functions including mediation analysis, moderation analysis, multigroup analysis, and importance-performance map analysis (IPMA). The current version, SmartPLS 4, incorporates PLSc estimation, confirmatory composite analysis (CCA), and enhanced visualization tools that substantially reduce the technical barriers to PLS-SEM application for psychological researchers.

This methodological paper aims to provide psychological researchers with a systematic, step-by-step guide to applying PLS-SEM using SmartPLS. The paper is organized to mirror the sequential logic of a complete PLS-SEM analysis: from model conceptualization and data preparation, through measurement model evaluation, to structural model assessment and reporting. Throughout, the paper grounds its recommendations in established methodological literature, provides decision criteria with interpretive guidance, and illustrates key concepts with reference to a worked example drawn from the stress-coping research domain in clinical psychology.

2. Theoretical Foundations of PLS-SEM

2.1 The Nature of Latent Variables in Psychology

Psychological research is fundamentally concerned with constructs, theoretical entities that cannot be directly observed but are inferred from observable indicators (Borsboom et al., 2003). Constructs such as depression, self-efficacy, resilience, cognitive load, attachment security, and emotional regulation represent the core explanatory variables in psychological theory, yet none can be directly measured with the precision afforded by physical measurement instruments. This inherent measurement challenge necessitates the use of indicator variables, typically self-report items, behavioral observations, or physiological measures that serve as proxies for the underlying latent constructs of interest.

SEM addresses this measurement challenge by explicitly modeling the relationship between latent constructs and their observable indicators, thereby partitioning observed variance into systematic construct-related variance and measurement error. This capacity distinguishes SEM from path analysis (which treats all variables as directly measured) and from regression analysis (which does not model measurement error in predictors), rendering it the methodology of choice for theory testing in psychology when multiple latent constructs are involved (Kline, 2016).

2.2 The PLS-SEM Algorithm

PLS-SEM estimates model parameters through an iterative algorithm that alternates between two sets of weight estimations: outer weights connecting indicators to their respective constructs, and inner weights connecting constructs to each other according to the specified structural model. The algorithm proceeds as follows:

In the initialization stage, each construct receives an arbitrary starting value, typically defined as the unit-weighted sum of its indicators. The outer approximation stage then iterates between computing inner construct scores as weighted combinations of adjacent construct scores (using the inner weighting scheme), and computing outer construct scores as weighted combinations of indicator values. Outer weights are updated using ordinary least squares (OLS) regression of each indicator on the component score, or vice versa, depending on the mode specification (Mode A for reflective, Mode B for formative). This iteration continues until convergence, defined as negligible change (typically < 10^{-7}) in outer weights between successive iterations.

The resulting latent variable scores serve as proxies for the underlying constructs, upon which structural path coefficients are estimated using OLS regression. This sequential estimation approach contrasts sharply with the simultaneous, system-wide estimation employed by CB-SEM, with important implications for the statistical properties of resulting estimates (Henseler et al., 2009).

2.3 Reflective vs. Formative Measurement

A fundamental theoretical distinction in PLS-SEM concerns the directionality of the relationship between constructs and indicators. In the reflective measurement model, by far the most common specification in psychological research, indicators are conceptualized as effects or manifestations of the underlying construct, such that the construct causes the indicators. This implies that indicators sharing a common latent cause should be internally consistent and intercorrelated, should exhibit similar relationships with external variables (nomological equivalence), and should be interchangeable without altering the construct's meaning. Depression scales, anxiety inventories, and most psychological questionnaires employ reflective specifications.

Formative measurement models, by contrast, conceptualize indicators as causes of the construct; the indicators collectively define and constitute the construct rather than reflecting it. Formative constructs need not exhibit intercorrelation among indicators (indeed, multicollinearity among formative indicators is a concern rather than an expectation), and dropping an indicator changes the construct's conceptual domain. Composite constructs representing overall socioeconomic status, health behaviors, or life events are often better represented formatively. PLS-SEM accommodates both specifications, whereas CB-SEM handles formative models more difficultly due to identification constraints.

3. PLS-SEM vs. CB-SEM: A Comparative Overview

Understanding the appropriate conditions for PLS-SEM selection requires a systematic comparison with CB-SEM across multiple methodological dimensions. The following table summarizes key comparative characteristics:

Table 1. Comparative Summary of PLS-SEM and CB-SEM Characteristics

Estimation Goal	Prediction/exploration	Confirmation/theory testing
Distributional Assumptions	Distribution-free	Multivariate normality required
Sample Size	Minimum ~30–100 (10× rule)	Minimum ~200–400
Model Complexity	Handles complex models well	May face convergence issues
Measurement Specification	Reflective and formative	Primarily reflective

Latent Variable Scores	Explicitly calculated	Indeterminate
Global Model Fit	SRMR (approximate)	Multiple indices (CFI, RMSEA, etc.)
Parameter Consistency	Inconsistent (biased upward)	Consistent (unbiased)
Software	SmartPLS, WarpPLS, PLS-Graph	AMOS, LISREL, Mplus, lavaan
Common Application	Exploratory, predictive research	Confirmatory theory testing

3.1 When to Choose PLS-SEM

Hair et al. (2019) propose that PLS-SEM is preferable to CB-SEM under the following conditions: (1) the research objective is prediction or exploratory theory building rather than strict confirmatory theory testing; (2) the structural model is complex with many constructs and indicator variables; (3) the sample size is small or the data distribution deviates substantially from normality; (4) the research involves formative constructs; (5) the study employs single-item measures or constructs with small numbers of indicators; and (6) the research context values maximizing explained variance in key target constructs.

In contrast, CB-SEM is preferable when the research objective is strict confirmatory testing of an a priori theoretical model, when the constructs are purely reflective, when sample sizes are sufficient for maximum likelihood estimation, and when researchers require global fit indices with established benchmarks (e.g., CFI > .95, RMSEA < .06).

Many psychological research contexts meet the conditions that favor PLS-SEM. Studies in clinical psychology, health psychology, organizational psychology, educational psychology, and social psychology frequently involve complex multi-construct models, moderate sample sizes, non-normal data, and prediction-oriented research questions, all of which align with PLS-SEM's methodological strengths.

4. SmartPLS Software: An Overview

SmartPLS (www.smartpls.com) is a user-friendly, web-based, and desktop application developed by Ringle, Wende, and Becker (2015) at the University of Hamburg, currently maintained by SmartPLS GmbH. The software offers a graphical model editor in which researchers construct path diagrams by drawing constructs, connecting indicators, and specifying directional relationships using standard drag-and-drop operations. SmartPLS supports multiple estimation algorithms (PLS-SEM, PLSc, GSCA), bootstrapping for significance testing, blindfolding for predictive relevance assessment, and a comprehensive array of model quality indices.

4.1 Key Features of SmartPLS 4

SmartPLS 4 incorporates several important analytical capabilities relevant to psychological research. The Bootstrapping function generates empirical sampling distributions for all model parameters using nonparametric resampling, enabling confidence interval estimation and significance testing without distributional assumptions, a critical feature given the absence of closed-form standard error solutions for PLS-SEM estimates. The default recommendation is 10,000 bootstrap subsamples for publication-quality results (Hair et al., 2019).

The Blindfolding function implements Stone-Geisser's Q^2 statistic for assessing the model's predictive relevance, systematically omitting data points and assessing how well the model reconstructs omitted data. Confirmatory Composite Analysis (CCA) enables researchers to test the adequacy of composite measurement models against established criteria. The Importance-Performance Map Analysis (IPMA) extends structural model results to provide managerial or applied insights by mapping the importance (path coefficients) and performance (mean latent variable scores) of each predictor construct on a two-dimensional grid.

SmartPLS 4 also includes comprehensive multigroup analysis (MGA) capabilities, enabling researchers to test whether structural relationships differ significantly across subgroups defined by categorical variables such as gender, clinical diagnosis category, or treatment condition—a particularly valuable feature for psychological research examining individual differences or intervention effects.

5. Model Specification and Design

5.1 Conceptual Model Development

Rigorous PLS-SEM analysis begins with the articulation of a theoretically grounded conceptual model that specifies the hypothesized relationships among constructs of interest. In psychological research, this model emerges from a systematic review of relevant theoretical frameworks and prior empirical evidence, and should be presented as a formal path diagram in which constructs are represented as circles or ellipses, observed indicators as rectangles, and directional arrows represent hypothesized causal relationships.

Each hypothesized relationship (path) should be grounded in explicit theoretical rationale, and the directionality of each path should be specified a priori. PLS-SEM is a confirmatory technique in the sense that the model structure is fixed before estimation (unlike exploratory factor analysis), and post-hoc model modifications driven by empirical results without theoretical justification undermine the validity of inferential claims (Kline, 2016).

5.2 Construct Operationalization

Following model conceptualization, each construct must be operationalized through an appropriate set of indicator variables. For reflective constructs, best practices recommend a minimum of three indicators per construct to ensure model identification and adequate reliability estimation, though two indicators may be acceptable in special circumstances (Hair et al., 2017b). Indicators should be drawn from validated existing scales wherever possible, with preference for measures demonstrating established reliability and validity in populations similar to the target sample.

Formative constructs require indicator selection based on the theoretical comprehensiveness of content domain coverage rather than internal consistency. Researchers specifying formative constructs should: (1) conduct a thorough conceptual analysis of the construct's definitional domain to identify all relevant indicator dimensions; (2) assess potential multicollinearity among formative indicators using Variance Inflation Factor (VIF) analysis; and (3) assess indicator weights rather than loadings as indices of each indicator's relative contribution to the composite.

5.3 Data Requirements and Preparation

PLS-SEM imposes less demanding data requirements than CB-SEM, but several data quality considerations remain critical. Regarding sample size, Hair et al. (2017a) recommend the '10 times rule' as a minimum heuristic: the sample size should be at least ten times the maximum number of paths directed at any single construct in the model. However, this rule provides only a floor estimate; adequate statistical power for detecting small-to-medium effect sizes typically requires samples of 100–250 participants, and researchers are encouraged to conduct a priori power analysis using the G*Power software's linear multiple regression module, treating the largest number of predictors aimed at any endogenous construct as the parameter of interest.

Missing data should be examined and addressed before analysis. PLS-SEM is generally robust to missing data rates below 5%, but rates exceeding this threshold warrant systematic treatment using multiple imputation or full information maximum likelihood where appropriate. Researchers should examine data for univariate and multivariate outliers, as extreme values can distort parameter estimates even in distribution-free methods. Common data screening procedures include inspection of standardized residuals, Mahalanobis distance statistics, and Cook's influence measures. While PLS-SEM does not require multivariate normality, severely non-normal distributions (absolute skewness > 3, absolute kurtosis > 10) may affect the adequacy of bootstrap-based inference, particularly in small samples. Researchers should report distributional characteristics and consider data transformation where distributions are extreme.

5.4 Importing Data into SmartPLS

SmartPLS accepts data in comma-separated values (CSV) format, with cases in rows and variables in columns. Variable names in the header row should be concise and free of special characters. Indicator variables should be coded numerically; categorical variables used in multigroup analysis should be dummy-coded. Once data are imported, the model is specified by creating constructs in the graphical editor, assigning indicator variables to constructs by dragging from the indicator panel, and drawing structural paths between constructs. The measurement mode (reflective Mode A or formative Mode B) is specified for each construct individually, with reflective mode representing the default in most psychological applications.

6. Measurement Model Assessment (Outer Model)

The first stage of PLS-SEM evaluation focuses on the outer measurement model, assessing the quality of the construct operationalizations and the degree to which the selected indicators accurately and consistently capture their intended constructs. This stage is logically before structural model assessment because structural path coefficients are meaningful only if the constructs they connect are well-measured. Hair et al. (2019) recommend a systematic sequence of measurement model evaluation covering indicator reliability, internal consistency reliability, convergent validity, and discriminant validity.

6.1 Indicator (Item) Reliability

Indicator reliability in reflective models is assessed through outer loadings—the bivariate correlations between each indicator and its assigned construct score. Outer loadings reflect the proportion of indicator variance explained by the latent construct, analogous to factor loadings in exploratory factor analysis. The conventional threshold for acceptable outer loadings is .708 or above, corresponding to the indicator sharing at least 50% of its variance with the construct ($.708^2 \approx .50$), representing adequate indicator reliability (Hair et al., 2019).

Indicators with loadings between .40 and .708 may be retained if their inclusion substantially contributes to Average Variance Extracted (AVE) or Composite Reliability (CR), particularly in the case of newly developed scales or scales adapted across cultural contexts. Indicators with loadings below .40 should typically be eliminated, as they contribute minimally to construct definition and may introduce excessive measurement error. Elimination decisions should always be guided by both statistical criteria and theoretical considerations regarding the indicator's conceptual centrality to the construct.

Importantly, indicator elimination should proceed one indicator at a time, re-running the model after each elimination to assess the impact on all model quality criteria before deciding on further eliminations. Wholesale elimination of multiple low-loading indicators simultaneously can produce misleading results if the eliminated indicators share systematic variance.

6.2 Internal Consistency Reliability

Internal consistency reliability assesses the degree to which all indicators of a construct consistently measure the same underlying latent variable. Three reliability indices are commonly reported in PLS-SEM analyses:

6.2.1 Cronbach's Alpha (α)

Cronbach's Alpha (Cronbach, 1951) is the most familiar reliability index in psychological research, computed as the ratio of true score variance to total observed variance under the assumption of tau-equivalence (equal true score loadings across all indicators). In PLS-SEM contexts, Cronbach's Alpha tends to underestimate reliability because PLS-SEM outer weights are not constrained to equality. The acceptable threshold for Cronbach's Alpha is conventionally .70, with values above .90 suggesting potential item redundancy that may artificially inflate reliability estimates without improving construct validity.

6.2.2 Composite Reliability (CR / ρc)

Composite Reliability (Werts et al., 1974), also denoted ρc or rho_c, is generally preferred over Cronbach's Alpha in PLS-SEM because it weights indicators according to their actual reliability (outer loadings) rather than assuming equal contributions. CR is computed as the squared sum of outer loadings divided by the sum of squared outer loadings plus the sum of indicator error variances. The recommended threshold for adequate CR is .70 for exploratory research and .80 for confirmatory research contexts, with values above .95 potentially indicating problematic item redundancy (Hair et al., 2019).

6.2.3 Dijkstra-Henseler's rho_A (ρA)

Rho_A (Dijkstra & Henseler, 2015) was introduced as a reliability estimate that is both consistent and appropriate for PLS-SEM's mode A estimation. Unlike Cronbach's Alpha (which assumes tau-equivalence) and Composite Reliability (which assumes congeneric measurement), rho_A provides an exact reliability estimate under PLS-SEM's actual model assumptions. Rho_A values above .70 indicate acceptable reliability, and their values typically fall between Cronbach's Alpha and Composite Reliability. Hair et al. (2019) recommend reporting all three indices, as their convergence provides stronger evidence of reliability.

Table 2. Internal Consistency Reliability Thresholds for PLS-SEM

Cronbach's Alpha (α)	≥ .70	≥ .80	> .95
Composite Reliability (ρc)	≥ .70	≥ .80	> .95
rho_A (ρA)	≥ .70	≥ .80	> .95

6.3 Convergent Validity

Convergent validity establishes that the indicators assigned to a construct converge on a common latent variable, that is, they share substantial common variance. In PLS-SEM, convergent validity is primarily assessed through the Average Variance Extracted (AVE) index (Fornell & Larcker, 1981).

AVE is computed as the mean of the squared outer loadings for all indicators of a construct, representing the average proportion of indicator variance explained by the construct. The minimum acceptable AVE threshold is .50, indicating that the construct explains more variance in its indicators than the error variances do. An AVE below .50 implies that measurement error exceeds the variance captured by the construct, raising serious concerns about construct validity. Constructs with AVE between .40 and .50 may be retained in some contexts if Composite Reliability is sufficiently high (above .60), but this represents a compromise that should be acknowledged in research reporting.

AVE is mathematically equivalent to the mean of communalities across indicators, providing a construct-level summary of the degree to which indicators converge on their designated construct. Its calculation is:

$$AVE = (\Sigma\lambda_i^2) / n$$

where λ_i represents the outer loading of the i-th indicator and n is the number of indicators for the construct.

6.4 Discriminant Validity

Discriminant validity assesses the degree to which a construct is empirically distinct from other constructs in the model; that is, each construct captures a unique aspect of the psychological phenomenon under investigation that is not redundant with other constructs.

Three approaches to discriminant validity assessment are currently recommended in PLS-SEM literature:

6.4.1 Fornell-Larcker Criterion

The Fornell-Larcker criterion (Fornell & Larcker, 1981) establishes that the square root of each construct's AVE should exceed its correlations with all other constructs in the model. This criterion is assessed by comparing the AVE square roots on the diagonal of the inter-construct correlation matrix with the off-diagonal correlations. When the diagonal values uniformly exceed all off-diagonal values in the respective row and column, discriminant validity is supported.

While widely used historically, simulation research has demonstrated that the Fornell-Larcker criterion has limited sensitivity to discriminant validity violations, particularly when cross-loadings between constructs are moderate (Henseler et al., 2015). Researchers are therefore advised to supplement the Fornell-Larcker criterion with cross-loading analysis and the HTMT ratio.

6.4.2 Cross-Loading Analysis

Cross-loading analysis requires that each indicator's loading on its assigned construct exceeds its cross-loadings on all other constructs in the model. This ensures that each indicator is more strongly related to its intended construct than to any other construct, providing item-level evidence of discriminant validity. Cross-loadings are examined in the full loading and cross-loading matrix generated by SmartPLS. Any indicator with a cross-loading on another construct that approaches or exceeds its own construct loading represents a discriminant validity concern requiring examination.

6.4.3 Heterotrait-Monotrait Ratio (HTMT)

The Heterotrait-Monotrait (HTMT) ratio (Henseler et al., 2015) represents the current methodological gold standard for discriminant validity assessment in PLS-SEM. HTMT is defined as the ratio of the average heterotrait-heteromethod correlations (correlations between indicators of different constructs) to the average monotrait-heteromethod correlations (correlations among indicators of the same construct). Conceptually, HTMT estimates the true correlation between constructs in the population; values substantially below 1.0 indicate discriminant validity.

Two threshold criteria are proposed in the literature. A strict threshold of HTMT < .85 was originally recommended by Henseler et al. (2015), appropriate when constructs are expected to be theoretically distinct. A more lenient threshold of HTMT < .90 has been proposed for constructs that are conceptually related (Gold et al., 2001). SmartPLS generates bootstrap confidence intervals for HTMT, and discriminant validity is supported when the upper bound of the 95% confidence interval does not exceed the threshold value. If HTMT exceeds the threshold, researchers should consider whether the two constructs are genuinely distinct or may represent different facets of a single higher-order construct.

Table 3. Summary of Measurement Model Assessment Criteria

Indicator Reliability	Outer Loadings	≥ .708 (preferred) ≥ .40 (acceptable)	Proportion of indicator variance explained by construct
Convergent Validity	AVE	≥ .50	Mean communality of construct indicators
Internal Consistency	CR (ρc)	≥ .70–.80	Weighted reliability estimate
Internal Consistency	Cronbach's α	≥ .70	Unweighted reliability estimate
Discriminant Validity	Fornell-Larcker	√AVE > r with others	Supplementary check only

Discriminant Validity	HTMT	< .85 (strict) / < .90 (lenient)	Primary DV criterion
Discriminant Validity	Cross-Loadings	$\lambda_{ii} > \lambda_{ij}$ for all $j \neq i$	Item-level DV check

6.5 Model Fit in the Measurement Stage

SmartPLS computes the Standardized Root Mean Square Residual (SRMR) as a global fit index for the PLS-SEM model. SRMR quantifies the average discrepancy between the observed correlation matrix and the model-implied correlation matrix, with smaller values indicating a better fit. An SRMR value below .08 is generally considered indicative of acceptable model fit (Henseler et al., 2014), though values below .10 may be acceptable for complex models. The SRMR should be interpreted in conjunction with the Normed Fit Index (NFI), which compares the model's fit against a null (independence) model.

It is important to note that PLS-SEM model fit assessment remains an active area of methodological development, and the established benchmarks for CB-SEM fit indices (CFI > .95, RMSEA < .06) do not directly apply to PLS-SEM analyses. Researchers should exercise interpretive caution and explicitly acknowledge the limitations of PLS-SEM fit assessment relative to CB-SEM when reporting results.

7. Structural Model Assessment (Inner Model)

Having established adequate measurement model quality through the outer model assessment procedures described in Section 6, the researcher proceeds to evaluate the inner structural mode, the network of hypothesized relationships among the latent constructs. Structural model assessment in PLS-SEM proceeds through several sequential evaluations: variance inflation factors, path coefficients and their significance, coefficient of determination (R^2), effect sizes (f^2), predictive relevance (Q^2), and model fit.

7.1 Collinearity Assessment

Before examining path coefficients, researchers must assess collinearity among predictor constructs in the structural model. High collinearity among predictor constructs can distort path coefficient estimation, inflating standard errors and producing unstable estimates that are difficult to interpret meaningfully. The Variance Inflation Factor (VIF) is the standard collinearity diagnostic, computed for each predictor in each regression equation within the structural model. A VIF below 3.3 is recommended as the threshold for acceptable collinearity levels in PLS-SEM (Diamantopoulos & Siguaw, 2006; Hair et al., 2019), with values above 5 indicating serious collinearity concerns.

Where VIF values exceed the threshold, researchers have several remedial options: removing one of the highly collinear predictors based on theoretical grounds; combining collinear predictors into a higher-order construct; or acknowledging the collinearity limitation and interpreting individual path coefficients with caution.

7.2 Path Coefficients

Structural path coefficients (β) represent the hypothesized directional relationships between constructs, equivalent to standardized regression coefficients in ordinary least squares regression. Path coefficients range from -1 to +1, with values near ±1 indicating strong relationships and values near 0 indicating weak or negligible relationships. Conventionally, path coefficients below .10 are considered negligible, those between .10 and .20 are considered small, those between .20 and .30 are moderate, and those above .30 are substantial (Hair et al., 2019).

Because PLS-SEM does not produce closed-form standard error estimates, statistical significance of path coefficients is assessed through nonparametric bootstrapping. SmartPLS generates bootstrap-based t-statistics and confidence intervals for all path coefficients, using bias-corrected and accelerated (BCa) confidence intervals as the most accurate bootstrap interval type. A path coefficient is considered statistically significant at the conventional $\alpha = .05$ level if the 95% BCa confidence interval excludes zero, or equivalently if the bootstrap t-statistic exceeds the critical value of 1.96 for two-tailed tests (or 1.645 for one-tailed tests where directionality is theoretically specified).

Researchers are strongly encouraged to report bootstrap confidence intervals alongside t-statistics and p-values, as confidence intervals convey both statistical significance and effect precision (Cumming, 2014). Additionally, one-tailed tests should be employed only when the directional hypothesis is specified a priori and theoretically justified, as post-hoc adoption of one-tailed tests represents an analytic degree of freedom problem that inflates Type I error.

7.3 Coefficient of Determination (R^2)

The coefficient of determination (R^2) quantifies the proportion of variance in each endogenous (dependent) construct explained by all constructs that directly predict it in the structural model. R^2 values range from 0 to 1, with higher values indicating greater predictive accuracy. Chin (1998) proposed the following heuristic benchmarks for R^2 in behavioral research: values of .19 are considered weak, .33 are moderate, and .67 are substantial. Hair et al. (2011) suggest more pragmatic thresholds of .25 (weak), .50 (moderate), and .75 (substantial) depending on the research domain.

However, R^2 thresholds must be interpreted contextually. Psychological phenomena are inherently complex and multi-determined, and models explaining 20–30% of construct variance may represent meaningful theoretical contributions, particularly when predicting clinical outcomes from theoretically motivated predictors. Researchers should interpret R^2 in light of the complexity of the phenomenon, the number of predictors, and comparable findings in the relevant literature.

The adjusted R^2 (R^2adj) corrects for model complexity by penalizing for the number of predictor constructs relative to sample size, providing a less biased estimate of population-level predictive accuracy. R^2adj should be reported alongside R^2 in all PLS-SEM analyses to enable readers to assess model parsimony.

7.4 Effect Size (f^2)

Cohen's f^2 effect size index assesses the practical contribution of each predictor construct to the R^2 of the endogenous construct it predicts. f^2 is computed by comparing R^2 with and without the specific predictor of interest, standardized by the unexplained variance in the full model:

$$f^2 = (R^2\text{included} - R^2\text{excluded}) / (1 - R^2\text{included})$$

Cohen's (1988) benchmarks for f^2 interpretation are widely employed: $f^2 = .02$ represents a small effect, $f^2 = .15$ represents a medium effect, and $f^2 = .35$ represents a large effect. Effect sizes below .02 indicate that the predictor has a negligible practical impact on the endogenous construct, regardless of statistical significance. Reporting f^2 alongside path coefficients and R^2 enables readers to distinguish statistically significant but practically trivial relationships from those with meaningful predictive impact.

7.5 Predictive Relevance (Q^2)

Stone-Geisser's Q^2 statistic (Stone, 1974; Geisser, 1975) assesses the model's out-of-sample predictive relevance for each endogenous construct through a blindfolding procedure. Blindfolding systematically omits data points from the dataset (determined by an omission distance D, typically set to 7) and estimates parameters using the remaining data, then assesses how well the estimated model reproduces the omitted values. Q^2 is computed as:

$$Q^2 = 1 - (SSE / SSO)$$

where SSE is the sum of squared prediction errors, and SSO is the sum of squared observations (mean substitution). Q^2 values above 0 indicate that the model has predictive relevance for the corresponding endogenous construct, confirming that the model's structural specification generates predictions superior to naive baseline (mean substitution) predictions. Hair et al. (2019) propose the following interpretive benchmarks: $Q^2 > .00$ indicates small predictive relevance, $Q^2 > .25$ indicates medium predictive relevance, and $Q^2 > .50$ indicates large predictive relevance.

Table 4. Structural Model Assessment Criteria and Thresholds

VIF (Collinearity)	SmartPLS output	< 3.3	Acceptable predictor collinearity
Path Coefficient (β)	OLS regression	BCa CI excludes 0	Strength and significance of the relationship
R^2	Variance explained	.25 / .50 / .75 = Weak/Moderate/Substantial	Predictive accuracy of the model
Adjusted R^2	Complexity-adjusted R^2	Compare to R^2	Model parsimony
Effect Size (f^2)	R^2 change	.02 / .15 / .35 = Small/Medium/Large	Predictor-level practical significance
Predictive Relevance (Q^2)	Blindfolding	> .00 / .25 / .50	Out-of-sample prediction quality
SRMR	Residual matrix	< .08	Global model fit

8. Advanced Analytical Techniques

8.1 Mediation Analysis

Mediation analysis examines whether the relationship between a predictor construct (X) and an outcome construct (Y) is transmitted through an intervening mechanism construct (M)—a central analytical concern in psychological research, where theoretical models frequently posit mediating processes through which distal causes produce their effects. The Baron and Kenny (1986) causal steps approach historically dominated mediation analysis, but contemporary methodological consensus strongly favors the product-of-coefficients approach with bootstrapped confidence intervals (Hayes, 2018; Preacher & Hayes, 2008), which PLS-SEM naturally supports.

In PLS-SEM, mediation is assessed by examining the indirect effect of X on Y through M, computed as the product of path coefficients a (X → M) and b (M → Y): indirect effect = a × b. Bootstrap confidence intervals are generated for the indirect effect, with the absence of zero within the 95% BCa confidence interval indicating a statistically significant indirect effect. SmartPLS computes total effects (direct + indirect), specific indirect effects through each mediator, and total indirect effects automatically within its bootstrapping routine.

The characterization of mediation follows Baron and Kenny's (1986) original typology but is more precisely articulated by Zhao et al. (2010): complementary mediation occurs when both the direct path (c') and indirect path (a×b) are significant and point in the same direction; competitive mediation occurs when the direct and indirect effects are significant but opposing in direction; indirect-only mediation (full mediation) occurs when the indirect effect is significant but the direct effect is not; direct-only non-mediation occurs when the direct effect is significant but the indirect effect is not; and no-effect non-mediation occurs when neither effect is significant. Psychological researchers should report all relevant effect components (a, b, c, c', a×b) with bootstrap confidence intervals, and articulate the theoretical significance of the observed mediation pattern.

8.2 Moderation Analysis

Moderation analysis investigates whether the strength or direction of the relationship between a predictor and an outcome construct varies as a function of a third construct (the moderator). In psychological research, moderation effects are theoretically central to understanding boundary conditions of psychological processes, for example, whether the relationship between stress and burnout is stronger for individuals with lower emotional regulation capacity, or whether the effect of cognitive behavioral therapy on depression is moderated by patient motivation for change.

PLS-SEM implements moderation through the interaction term approach, in which the interaction between the predictor (X) and moderator (W) is specified as a new construct (X×W) that predicts the outcome (Y). SmartPLS supports three interaction term operationalization methods: (1) the product indicator approach, which creates all cross-products of X and W indicators; (2) the two-stage approach, which uses the latent variable scores from an initial PLS-SEM run as single indicators in a subsequent regression; and (3) the orthogonalizing approach, which residualizes cross-product indicators to remove collinearity. For most psychological research applications, the two-stage approach provides the best balance of statistical efficiency and interpretive clarity (Hair et al., 2019).

The significance of the moderation effect is assessed through the bootstrap confidence interval for the interaction path coefficient (β_XW). Where significant moderation is detected, the effect should be interpreted through simple slope analysis and plotted for three values of the moderator (typically 1SD, mean, +1SD) to facilitate visualization of the interaction pattern. Practical effect size for moderation is assessed through f^2 as described in Section 7.4.

8.3 Higher-Order Constructs

Higher-order construct (HOC) models, also termed hierarchical component models, specify constructs at multiple levels of abstraction, a theoretically important feature for psychological research involving multidimensional constructs such as personality (with its domain and facet levels), emotional intelligence, quality of life, or burnout. PLS-SEM supports Type I (reflective-reflective), Type II (reflective-formative), Type III (formative-reflective), and Type IV (formative-formative) higher-order specifications.

The repeated indicators approach (also called the hierarchical component model) is the most commonly employed method for HOC estimation in SmartPLS. In this approach, the indicators of all first-order constructs (dimensions) are assigned to both their respective first-order constructs and to the higher-order construct. This approach avoids construct score indeterminacy but increases model complexity and can affect discriminant validity assessment. Alternatively, the two-stage approach uses latent variable scores from first-order constructs as indicators of the higher-order construct in a subsequent analysis, providing cleaner conceptual separation between levels.

8.4 Multigroup Analysis (MGA)

Multigroup analysis (MGA) tests whether structural path coefficients differ significantly across distinct subgroups of participants, enabling researchers to examine group-level moderation hypotheses. In psychological research, MGA may be employed to test whether the predictors of depression severity differ between genders, age groups, cultural groups, or diagnostic categories. PLS-SEM supports both parametric MGA (based on bootstrap standard errors) and nonparametric MGA (Henseler's MGA) approaches.

MGA requires that measurement model equivalence (partial or full measurement invariance) be established before meaningful comparison of structural parameters across groups. Measurement invariance is assessed through the Confirmatory Composite Analysis (CCA) procedure in SmartPLS, analogous to confirmatory factor analysis-based measurement invariance testing in CB-SEM. At minimum, configural invariance (same model structure across groups) and metric invariance (equivalent outer loadings across groups) should be established before comparing structural path coefficients.

9. Discussion and Conclusion

9.1 Summary of Key Methodological Contributions

This paper has provided a comprehensive methodological guide to PLS-SEM application in psychological research, covering the full analytical workflow from model conceptualization through measurement and structural model assessment to advanced analytical extensions. The procedural framework presented integrates contemporary best practice recommendations from the methodological literature, offering psychological researchers a systematic, criterion-referenced approach to implementing and reporting PLS-SEM analyses.

Several methodological contributions merit emphasis. First, the paper positions the HTMT ratio as the primary discriminant validity criterion, superseding the historically dominant but methodologically limited Fornell-Larcker criterion. Second, the paper presents rho_A alongside Cronbach's Alpha and Composite Reliability as a reliability trinity that collectively provides more complete reliability evidence than any single index. Third, the paper demonstrates how the two-stage PLS-SEM approach addresses moderation analysis, and how bootstrap-based inference enables mediation analysis without distributional assumptions. Fourth, the worked example illustrates how all quantitative criteria and thresholds apply in a concrete psychological research context.

9.2 Limitations of PLS-SEM

PLS-SEM's methodological advantages are accompanied by genuine limitations that researchers must transparently acknowledge. Most importantly, PLS-SEM parameter estimates are statistically inconsistent; they do not converge to true population values as sample size increases, unlike maximum likelihood estimates in CB-SEM. While PLSc (consistent PLS) addresses this limitation for reflective models, it comes at the cost of requiring larger samples and imposing additional assumptions. Researchers using standard PLS-SEM should acknowledge the potential upward bias in parameter estimates and interpret effect magnitudes conservatively.

PLS-SEM's global model fit assessment capabilities remain more limited than those of CB-SEM, where well-established fit indices (CFI, RMSEA, SRMR) with validated benchmarks enable comprehensive model evaluation. The SRMR index available in SmartPLS provides a useful partial approximation, but its benchmarks are less firmly established, and its sensitivity to misspecification may differ from its CB-SEM counterpart. Researchers should supplement SRMR with examination of residual correlations and theoretically-motivated model comparison where possible.

Finally, PLS-SEM's prediction focus means it is less appropriate for strict confirmatory hypothesis testing contexts where the goal is to corroborate or refute a specific theoretically derived model. When the research question is genuinely confirmatory, and the data meet the requirements of CB-SEM, CB-SEM remains the preferred choice despite its greater complexity.

9.3 Future Directions

Several frontier developments in PLS-SEM methodology hold particular promise for psychological research. Bayesian PLS-SEM approaches combining the flexibility of PLS-SEM with Bayesian inference are beginning to appear in the literature, offering potential advantages for uncertainty quantification and integration of prior information. Machine learning extensions of PLS-SEM, particularly PLS-based tree methods for detecting complex interaction and nonlinear effects, are increasingly accessible through commercial and open-source software. The integration of PLS-SEM with network analysis approaches may offer new tools for examining complex, dynamic psychological systems beyond the confines of conventional path models.

The continuing methodological development of PLSc and its extensions offers the prospect of PLS-SEM analyses that combine the practical advantages of PLS-

SEM with the statistical consistency of CB-SEM, potentially reducing the tension between these methodological traditions. As SmartPLS and competing platforms continue to incorporate these methodological advances, psychological researchers are encouraged to remain current with the methodological literature and to update their analytical practice accordingly.

9.4 Conclusion

PLS-SEM represents a methodologically rigorous and practically powerful analytical tool for psychological research involving complex theoretical models with latent constructs. When properly implemented using SmartPLS and evaluated against contemporary methodological standards, PLS-SEM enables psychological researchers to examine theoretically grounded models with greater flexibility and under less restrictive data conditions than traditional CB-SEM approaches allow. This paper has provided the procedural and evaluative framework necessary for psychological researchers to apply PLS-SEM with confidence, rigor, and methodological transparency, qualities that are foundational to the cumulative advancement of psychological science.

References

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., Waltz, T., & Zettle, R. D. (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire–II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy*, *42*(4), 676–688. https://doi.org/10.1016/j.beth.2011.03.007

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*(2), 203–219. https://doi.org/10.1037/0033-295X.110.2.203

Carver, C. S. (1997). You want to measure coping, but your protocol's too long: Consider the Brief COPE. *International Journal of Behavioral Medicine*, *4*(1), 92–100. https://doi.org/10.1207/s15327558ijbm0401_6

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295–358). Lawrence Erlbaum Associates.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Coltman, T., Devinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business Research*, *61*(12), 1250–1262. https://doi.org/10.1016/j.jbusres.2008.01.013

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

Diamantopoulos, A., & Siguaw, J. A. (2006). Formative versus reflective indicators in organizational measure development: A comparison and empirical illustration. *British Journal of Management*, *17*(4), 263–282.

https://doi.org/10.1111/j.1467-8551.2006.00500.x

Dijkstra, T. K., & Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Quarterly, 39*(2), 297–316. https://doi.org/10.25300/MISQ/2015/39.2.02

Fornell, C., & Bookstein, F. L. (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *Journal of Marketing Research, 19*(4), 440–452. https://doi.org/10.2307/3151718

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.2307/3151312

Geisser, S. (1975). The predictive sample reuse method with applications. *Journal of the American Statistical Association, 70*(350), 320–328. https://doi.org/10.2307/2285815

Gold, A. H., Malhotra, A., & Segars, A. H. (2001). Knowledge management: An organizational capabilities perspective. *Journal of Management Information Systems*, *18*(1), 185–214. https://doi.org/10.1080/07421222.2001.11045669

Gray-Toft, P., & Anderson, J. G. (1981). The Nursing Stress Scale: Development of an instrument. *Journal of Behavioral Assessment*, *3*(1), 11–23. https://doi.org/10.1007/BF01321348

Hair, J. F., Henseler, J., Dijkstra, T. K., & Sarstedt, M. (2014). Common beliefs and reality about partial least squares: Comments on Rönkkö and Evermann (2013). *Organizational Research Methods, 17*(2), 182–209. https://doi.org/10.1177/1094428114526928

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). SAGE Publications.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed, a silver bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139–152. https://doi.org/10.2753/MTP1069-6679190202

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2017). *Advanced issues in partial least squares structural equation modeling*. SAGE Publications.

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). Guilford Press.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *Advances in International Marketing, 20*, 277–319. https://doi.org/10.1108/S1474-7979(2009)0000020014

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

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer.

Liang, H., Saraf, N., Hu, Q., & Xue, Y. (2007). Assimilation of enterprise systems: The effect of institutional pressures and the mediating role of top management. *MIS Quarterly, 31*(1), 59–87. https://doi.org/10.2307/25148781

Maslach, C., Jackson, S. E., & Leiter, M. P. (1996). *Maslach Burnout Inventory manual* (3rd ed.). Consulting Psychologists Press.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*(3), 879–891. https://doi.org/10.3758/BRM.40.3.879

Ringle, C. M., Wende, S., & Becker, J. M. (2015). *SmartPLS 3*. SmartPLS GmbH.

Rönkkö, M., & Evermann, J. (2013). A critical examination of common beliefs about partial least squares path modeling. *Organizational Research Methods, 16*(3), 425–448. https://doi.org/10.1177/1094428112474693

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B (Methodological), 36*(2), 111–133. https://doi.org/10.1111/j.2517-6161.1974.tb00994.x

Werts, C. E., Linn, R. L., & Jöreskog, K. G. (1974). Intraclass reliability estimates: Testing structural assumptions. *Educational and Psychological Measurement*, *34*(1), 25–33. https://doi.org/10.1177/001316447403400104

Wold, H. (1975). Path models with latent variables: The NIPALS approach. In H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon, & V. Cappecchi (Eds.), *Quantitative sociology: International perspectives on mathematical and statistical modeling* (pp. 307–357). Academic Press.

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research, 37*(2), 197–206. https://doi.org/10.1086/651257