# Using WarpPLS in e-Collaboration Studies: Descriptive Statistics, Settings, and Key Analysis Results

*Ned Kock, Texas A&M International University, USA*

## ABSTRACT

*This is a follow-up on a previous article (Kock, 2010b) discussing the five main steps through which a nonlinear structural equation modeling analysis could be conducted with the software WarpPLS (warppls.com). Both this and the previous article use data from the same e-collaboration study as a basis for the discussion of important WarpPLS features. The focus of this article is on specific features related to saving and analyzing grouped descriptive statistics, viewing and changing analysis algorithm and resampling settings, and viewing and saving the various minor and major results of the analysis. Even though its focus is on an e-collaboration study, this article contributes to the broad literature on multivariate analysis methods, in addition to the more specific research literature on e-collaboration. The vast majority of relationships between variables, in investigations of both natural and behavioral phenomena, are nonlinear; usually taking the form of U and S curves. Structural equation modeling software tools, whether variance- or covariance-based, typically do not estimate coefficients of association based on nonlinear analysis algorithms. WarpPLS is an exception in this respect. Without taking nonlinearity into consideration, the results can be misleading; especially in complex and multi-factorial situations such as those stemming from e-collaboration in virtual teams.*

*Keywords:* *Multivariate Statistics, Nonlinear Analysis, Partial Least Squares, Structural Equation Modeling, Virtual Teams, WarpPLS*

## INTRODUCTION

The vast majority of the statistical methods used in the behavioral sciences, and many of those used in the natural sciences, can be seen as special cases of structural equation modeling (SEM). This applies to both univariate (a.k.a. bivariate) and multivariate statistical analysis methods (Hair et al., 1987). This can be demonstrated through a sequential logical inference process. In short, it can be shown that most of these methods are instances of multiple regression analysis, which is itself an instance of path analysis, which in turn is an instance of SEM.

Methods like ANOVA, ANCOVA, MANOVA and MANCOVA can be shown to be special cases of multiple regression analysis (Hair et al., 1987; Rencher, 1998). In multiple regression analysis, hypothesis testing is typically conducted through the calculation of coefficients of association between multiple independent

DOI: 10.4018/jec.2011040101

variables and one main dependent variable. These coefficients of association normally take the form of standardized partial regression coefficients (Rencher, 1998; Rosenthal & Rosnow, 1991). The corresponding P values are the probabilities that the relationships reflected in the coefficients are "real".

Path analysis is a method developed by Sewall Wright in the 1930s (Wolfle, 1999; Wright, 1934) and later "rediscovered" by statisticians and social scientists. Sewall Wright was an evolutionary biologist and animal breeder. He was also one of the founders of the field of population genetics. Population genetics unified Darwin's theory of evolution with Mendel's theory of genetics. Another co-founder of the field of population genetics was Ronald A. Fisher, who also has made many contributions to the field of statistics (Hair et al., 1987; Kock, 2009).

Any path analysis model can be decomposed into one or more multiple regression models (Gefen et al., 2000; Kline, 1998). Each of the multiple regression models can then be solved separately, and the solution combined into one main solution to the path analysis model. In this sense, multiple regression analysis can be seen as a special case of path analysis. Since SEM is essentially path analysis with latent variables (LVs), then path analysis can be seen as a special case of SEM (Maruyama, 1998). As a corollary, all of the methods discussed above can also be seen as special cases of the SEM.

In SEM, LV scores are calculated as weighted averages of their respective indicators. Usually there are two or more indicators for each LV, although that is not always the case. Once LV scores are calculated, the SEM solution problem is reduced to the solution of a path analysis model. That is achieved through the calculation of path coefficients and respective P values, as well as several other ancillary statistical coefficients. The path coefficients are standardized partial regression coefficients, which are mathematically identical to those obtained through multiple regression analyses.

The calculation of weights linking indicators to LVs is one of the key aspects that differentiate SEM approaches. Those approaches can be divided into two with main types: variance- and covariance-based (Chin et al., 2003; Gefen et al., 2000; Haenlein & Kaplan, 2004; Kline, 1998). One of the main advantages of variance-based SEM is that it employs robust statistics to calculate P values, and thus can be seen as a nonparametric equivalent to covariance-based SEM. That is, unlike covariance-based SEM, variance-based SEM typically yields robust results even in the presence of small samples and multivariate deviations from normality (Chin et al., 2003; Gefen et al., 2000). Variance-based SEM is often referred to as PLS-based SEM, where "PLS" stands for "partial least squares" or "projection to latent structures". The term "PLS-based SEM" is actually more commonly found in the literature than the term "variance-based SEM" (Chin et al., 2003; Haenlein & Kaplan, 2004).

The vast majority of relationships between variables, in investigations of both natural and behavioral phenomena, are nonlinear; usually taking the form of U and S curves. In spite of this, SEM software tools do not usually take nonlinear relationships between LVs into consideration in the calculation of path coefficients, respective P values, or other related statistical coefficients (e.g., R-squared coefficients). The SEM software WarpPLS (warppls.com), released as version 1.0 at the time of this writing, is an exception in this respect (Kock, 2010).

This article is a follow-up on a previous article (Kock, 2010b) discussing the five main steps through which a nonlinear structural equation modeling analysis could be conducted with the software WarpPLS. Both this and the previous article use the same e-collaboration study as a basis for the discussion of important WarpPLS features. Unlike in the previous article, the focus here is on specific features related to saving and analyzing grouped descriptive statistics, viewing and changing analysis algorithm and resampling settings, and viewing and saving the various minor and major results of the analysis.

## THE E-COLLABORATION STUDY

In the following sections, several screens are used to illustrate important features of WarpPLS. Those screens were generated based on an e-collaboration study if virtual teams. A total of 290 teams were studied. The teams were tasked with developing new products, goods or services, in a variety of organizations belonging to multiple industries and sectors (e.g., aerospace and banking; service and manufacturing; respectively). Data related to five LVs were collected as part of the study. The LVs are indicated here as "ECU", "ECUVar", "Proc", "Effi", and "Effe".

"ECU" refers to the extent to which electronic communication media were used by each team. "ECUVar" refers to the variety of different electronic communication media used by each team. "Proc" refers to the degree to which each team employed established project management techniques, referred to in the study as procedural structuring techniques. "Effi" refers to the efficiency of each team, in terms of task completion cost and time. "Effe" refers to the effectiveness of each team, in terms of the actual commercial success of the new goods or services that each team developed.

## SAVING AND ANALYZING GROUPED DESCRIPTIVE STATISTICS

Once steps 1 and 2 of an SEM analysis are completed through WarpPLS, you (the user) can then save and analyze grouped descriptive statistics. Through Step 1, you will open or create a project file to save your work. Through Step 2, you will read the raw data used in the SEM analysis.

When the "Save grouped descriptive statistics into a tab-delimited .txt file" option is selected, a data entry window is displayed (Figure 1). There you can choose a grouping variable, number of groups, and the variables to be grouped. This option is useful if one wants to conduct a comparison of means analysis using the software, where one variable (the grouping variable) is the predictor, and one or more variables are the criteria (the variables to be grouped). Arguably the comparison of means is the most common type of analysis used in the natural and behavioral sciences. One of the reasons for this is that this type of analysis is intuitively appealing and its results are easy to understand.

Figure 2 shows the grouped statistics data saved through the window shown in Figure 1. The tab-delimited .txt file was opened with a spreadsheet program, and contained the data on the left part of the figure.

That data on the left part of Figure 2 was organized as shown above the bar chart. The data are the means and standard deviations for each interval (or group). Next the bar chart was created using the spreadsheet program's charting feature. If a simple comparison of means analysis using this software had been conducted in which the grouping variable (in this case, an indicator called "ECU1") was the predictor, and the criterion was the indicator called "Effe1", those two variables would have been connected through a path in a simple path model with only one path. Assuming that the path coefficient was statistically significant, the bar chart displayed in Figure 2, or a similar bar chart, could be added to a report describing the analysis.

Some may think that it is an overkill to conduct a comparison of means analysis using an SEM software package such as this, but there are advantages in doing so. One of those advantages is that this software calculates P values using a nonparametric class of estimation techniques, namely resampling estimation techniques. (These are sometimes referred to as bootstrapping techniques, which may lead to confusion since bootstrapping is also the name of a type of resampling technique.) Nonparametric estimation techniques do not require the data to be normally distributed, which is a requirement of other comparison of means techniques (e.g., ANOVA).

Another advantage of conducting a comparison of means analysis using this software

*Figure 1. Save grouped descriptive statistics window*

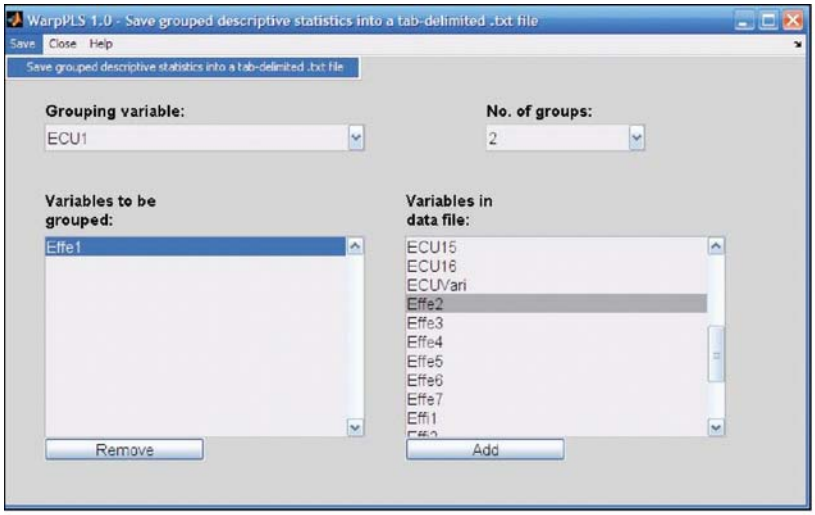
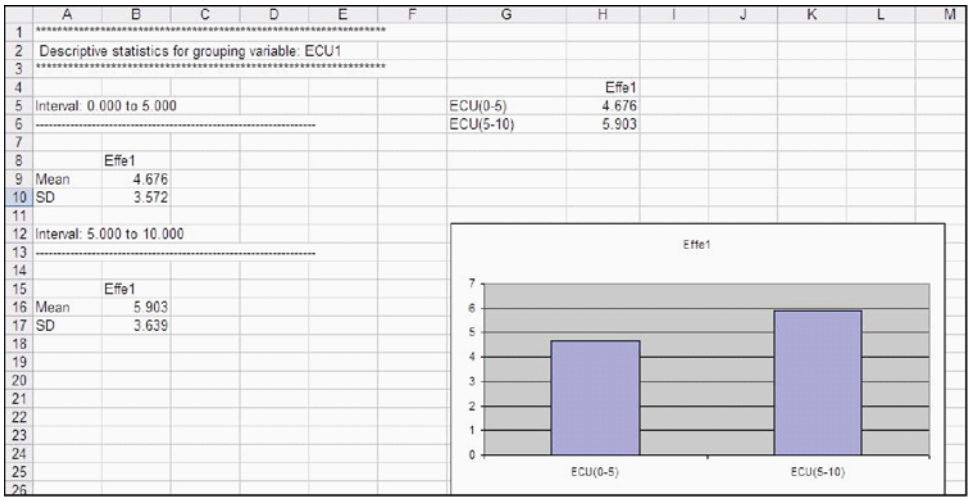



*Figure 2. Grouped descriptive statistics bar chart*

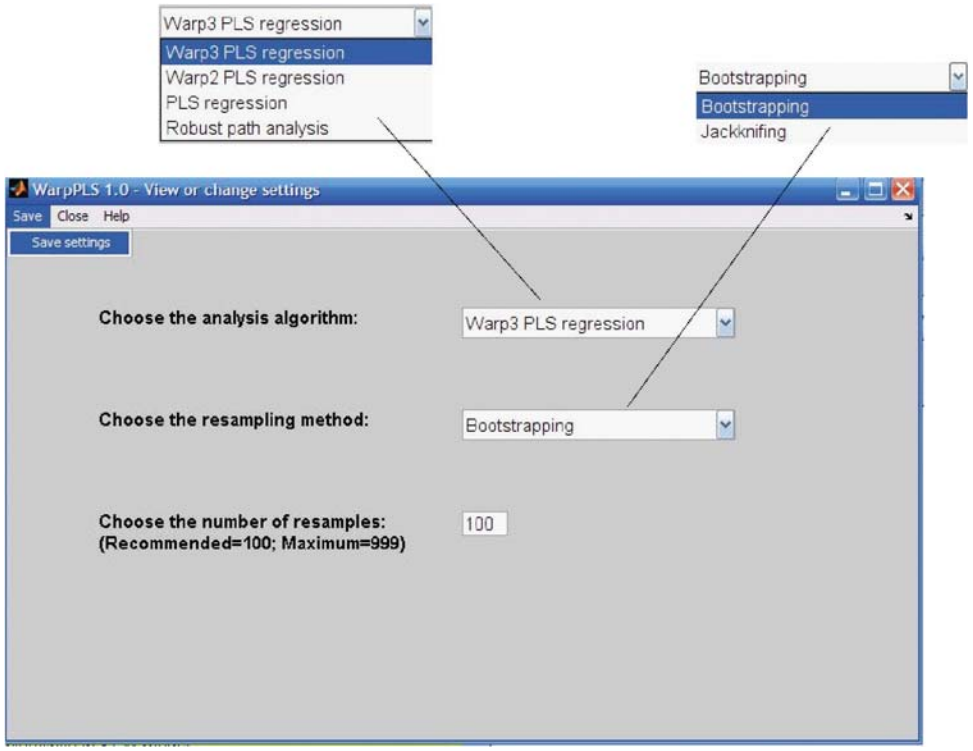



is that the analysis can be significantly more elaborate. For example, the analysis may include control variables (or covariates), which would make it equivalent to an ANCOVA test. Finally, the comparison of means analysis may include LVs, as either predictors or criteria. This is not usually possible with ANOVA or commonly used nonparametric comparison of means tests (e.g., the Mann-Whitney U test).

## VIEWING AND CHANGING ANALYSIS ALGORITHM AND RESAMPLING SETTINGS

The view or change settings window (Figure 3) allows you to select an algorithm for the SEM analysis, select a resampling method, and select the number of resamples used (this latter option is only useful if the resampling method selected

*Figure 3. View or change settings window*




was bootstrapping). The analysis algorithms available are Warp3 PLS Regression, Warp2 PLS Regression, PLS Regression, and Robust Path Analysis.

Many relationships in nature, including relationships involving behavioral variables, are nonlinear and follow a pattern known as U-curve (or inverted U-curve). In this pattern a variable may affect another in a way that leads to a maximum or minimum value, where the effect is either maximized or minimized, respectively. This type of relationship is also referred to as a J-curve pattern. This latter term is more commonly used in economics and the health sciences.

U curves also refer to sections of a complete U- or J-shaped curve. Therefore U curves can be used to model the majority of the usually seen functions in natural and behavioral studies. These are non-cyclical functions, such as logarithmic, exponential, and hyperbolic decay functions. When the relationships fit well with the forms of these common types of functions, S-curve approximations (which are mono-cyclical) will usually default to U curves.

The Warp2 PLS Regression algorithm tries to identify a U-curve relationship between LVs, and, if that relationship exists, the algorithm transforms (or "warps") the scores of the predictor LVs so as to better reflect the U-curve relationship in the estimated path coefficients in the model.

The Warp3 PLS Regression algorithm, the default algorithm used by the software, tries to identify a relationship defined by a function whose first derivative is a U-curve. This type of relationship follows a pattern that is more similar to an S-curve (or a somewhat distorted S-curve), and can be seen as a combination of two connected U-curves, one of which is inverted.

The PLS Regression algorithm does not perform any warping of relationships. It is

essentially a standard PLS regression algorithm (Wold et al., 2001), whereby indicators' weights, loadings and factor scores (a.k.a. LV scores) are calculated based on a least squares minimization sub-algorithm, after which path coefficients are estimated using a robust path analysis algorithm. A key criterion for the calculation of the weights, observed in virtually all PLS-based algorithms, is that the regression equation expressing the relationship between the indicators and the LV scores has an error term that equals zero. In other words, the LV scores are calculated as exact linear combinations of their indicators.

PLS regression (Wold et al., 2001) is the underlying weight calculation algorithm used in both Warp3 and Warp2 PLS Regression. The warping takes place during the estimation of path coefficients, and after the estimation of all weights and loadings in the model. It occurs only if the Warp3 or Warp2 algorithms are used. The weights and loadings of a model with LVs make up what is often referred to as outer model, whereas the path coefficients among LVs make up what is often called the inner model.

Finally, the Robust Path Analysis algorithm is a simplified algorithm in which LV scores are calculated by averaging all of the indicators associated with a LV; that is, in this algorithm weights are not estimated through PLS regression. This algorithm is called "Robust" Path Analysis, because, as with most robust statistics methods, the P values are calculated through resampling. If all LVs are measured with single indicators, the Robust Path Analysis and the PLS Regression algorithms will yield identical results.

One of two resampling methods may be selected: bootstrapping or jackknifing. Bootstrapping, the software's default, is a resampling algorithm that creates a number of resamples (a number that can be selected by the user), by a method known as "resampling with replacement". This means that each resample contains a random arrangement of the rows of the original dataset, where some rows may be repeated. (The commonly used analogy of a deck of cards being reshuffled, leading to many resample decks, is a good one. But it is not entirely correct because in bootstrapping the same card may appear more than once in each of the resample decks.)

Jacknifing, on the other hand, creates a number of resamples that equals the original sample size, and each resample has one row removed. That is, the sample size of each resample is the original sample size minus 1. Thus, the choice of number of resamples has no effect on jackknifing, and is only relevant in the context of bootstrapping.

The default number of resamples is 100, and it can be modified by entering a different number in the appropriate edit box. (Please note that we are talking about the number of resamples here, not the original data sample size.) Leaving the number of resamples for bootstrapping as 100 is recommended because it has been shown that higher numbers of resamples lead to negligible improvements in the reliability of P values; in fact, even setting the number of resamples at 50 is likely to lead to fairly reliable P value estimates (Efron et al., 2004). Conversely, increasing the number of resamples well beyond 100 leads to a higher computation load on the software, making the software look like it is having a hard time coming up with the results. In very complex models, a high number of resamples may make the software run very slowly.

Some researchers have suggested in the past that a large number of resamples can address problems with the data, such as the presence of outliers due to errors in data collection. This opinion is not shared by several researchers, including the original developer of the bootstrapping method, Bradley Efron (Efron et al., 2004).

Arguably jackknifing does a better job at addressing problems associated with the presence of outliers due to errors in data collection. Generally speaking, jackknifing tends

to generate more stable resample path coefficients (and thus more reliable P values) with small sample sizes (lower than 100), and with samples containing outliers. In these cases, outlier data points do not appear more than once in the set of resamples, which accounts for the better performance of jackknifing (Chiquoine & Hjalmarsson, 2009).

Bootstrapping tends to generate more stable resample path coefficients (and thus more reliable P values) with larger samples and with samples where the data points are evenly distributed on a scatter plot. The use of bootstrapping with small sample sizes (lower than 100) has been discouraged (Nevitt & Hancock, 2001).

Since the warping algorithms are also sensitive to the presence of outliers, in many cases it is a good idea to estimate P values with both bootstrapping and jackknifing, and use the P values associated with the most stable coefficients. An indication of instability is a high P value (i.e., statistically insignificant) associated with path coefficients that could be reasonably expected to have low P values. For example, with a sample size of 100, a path coefficient of .2 could be reasonably expected to yield a P value that is statistically significant at the .05 level. If that is not the case, there may be a stability problem. Another indication of instability is a marked difference between the P values estimated through bootstrapping and jackknifing.

P values can be easily estimated using both resampling methods, bootstrapping and jackknifing, by following this simple procedure. Run an SEM analysis of the desired model, using one of the resampling methods, and save the project. Then save the project again, this time with a different name, change the resampling method, and run the SEM analysis again. Then save the second project again. Each project file will now have results that refer to one of the two resampling methods. The P values can then be compared, and the most stable ones used in a research report on the SEM analysis.

## VIEWING AND SAVING THE VARIOUS RESULTS OF THE ANALYSIS

As soon as an SEM analysis is completed with WarpPLS, the software shows the results in graphical format on a window, which also contains a number of menu options that allow you to view and save more detailed results. The sections below refer to each of these various menu options.

### General Analysis Results

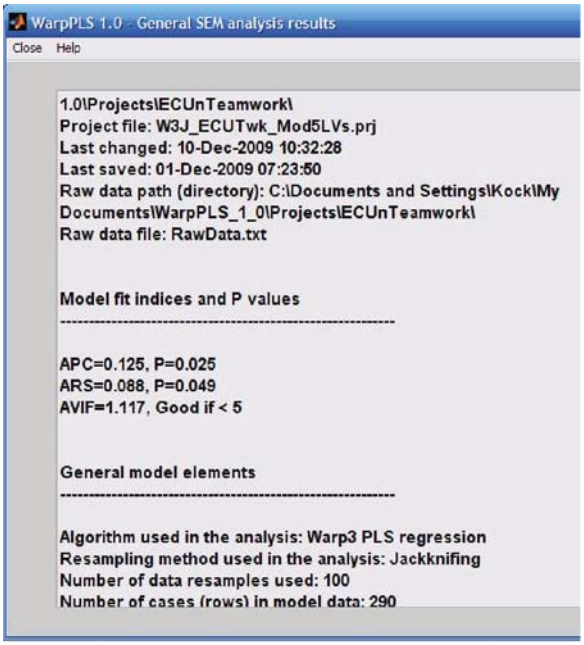
General SEM analysis results (Figure 4) include: project file details, such as the project file name and when the file was last saved; model fit indices, which are discussed in more detail below; and general model elements, such as the algorithm and resampling method used in the SEM analysis.

Under the project file details, both the raw data path and file are provided. Those are provided for completeness, because once the raw data is imported into a project file, it is no longer needed for the analysis. Once a raw data file is read, it can even be deleted without any effect on the project file, or the SEM analysis.

Three model fit indices are provided: average path coefficient (APC), average R-squared (ARS), and average variance inflation factor (VIF). For the APC and ARS, P values are also provided. These P values are calculated through a complex process that involves resampling estimations coupled with Bonferroni-like corrections. This is necessary since both fit indices are calculated as averages of other parameters.

The interpretation of the model fit indices depends on the goal of the SEM analysis. If the goal is to test hypotheses, where each arrow represents a hypothesis, then the model fit indices are of little importance. However, if the goal is to find out whether one model has a better fit with the original data than another, then the model fit indices are a useful set of measures related to model quality.

*Figure 4. General SEM analysis results window*

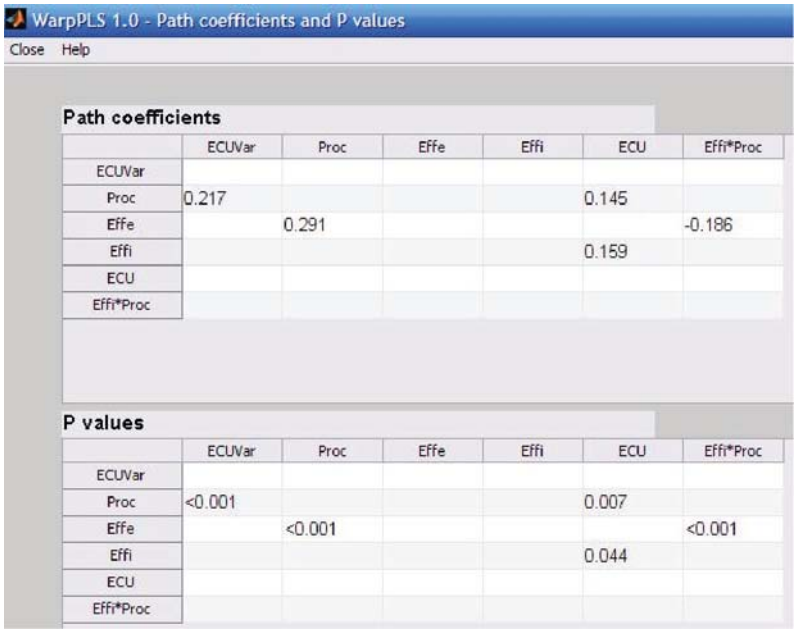



When assessing the model fit with the data, the following criteria are recommended. First, it is recommended that the P values for the APC and ARS be both lower than .05; that is, significant at the .05 level. Second, it is recommended that the AVIF be lower than 5. When comparing competing models, the ARS index should be given higher importance in terms of model fit than either AVIF or APC. Next in importance comes AVIF. APC comes third. One of the reasons for this is that the APC index may be low simply because there are many changes in path coefficient signs in the model. This is discouraged since it can lead to a phenomenon known as "suppression" (Kline, 1998), where competing path coefficients are distorted due to having different signs. However, different path coefficient signs may not have any significant distorting effect on competing path coefficients, which warrants placing APC as third in order of importance.

Typically the addition of new LVs into a model will increase the ARS, even if those LVs are weakly associated with the existing LVs in the model. However, that will generally lead to a decrease in APC, since the path coefficients associated with the new LVs will be low. Thus, the APC and ARS will counterbalance each other, and will only increase together if the LVs that are added to the model enhance the overall predictive and explanatory quality of the model. This assumes that all path coefficients have the same sign. Path coefficients can be made to all have the same sign (usually positive) by reversing variables as needed. For example, if a variable D (reflecting dullness) is negatively associated with P (performance), a new variable E (excitement) may be used instead of D. In this case, E is D reversed, and the association between E and P will be positive.

The AVIF index will increase if new LVs are added to the model in such a way as to add multicolinearity to the model, which may result from the inclusion of new LVs that overlap in meaning with existing LVs. It is generally undesirable to have different LVs in the same model that measure the same thing; those should be combined into one single LV. Thus, the

*Figure 5. Path coefficients and P values window*




AVIF brings in a new dimension that adds to a comprehensive assessment of a model's overall predictive and explanatory quality.

## Path Coefficients and P Values

Path coefficients and respective P values are shown together, as can be seen in Figure 5. Each path coefficient is displayed in one cell, where the column refers to the predictor LV and the row to the criterion. For example, let us consider the case in which the cell shows .145, and the column refers to the LV "ECU" and the row to the LV "Proc". This means that the path coefficient associated with the arrow that points from "ECU" to "Proc" is .145. Since the results refer to standardized variables, this means that a 1 standard deviation variation in "ECU" leads to a .145 standard deviation variation in "Proc".

The P values shown are calculated by resampling, and thus are specific to the resampling method and number of resamples selected by the user. As mentioned earlier, the choice of number of resamples is only meaningful for the bootstrapping method, and numbers higher than 100 add little to the reliability of the P value estimates.

One puzzling aspect of many publicly available PLS-based SEM software systems is that they do not provide P values, instead providing standard errors and T values, and leaving the users to figure out what the corresponding P values are. Often users have to resort to tables relating T to P values, or other software (e.g., Excel), to calculate P values based on T values.

This is puzzling because typically research reports will provide P values associated with path coefficients, which are more meaningful than T values for hypothesis testing purposes. This is due to the fact that P values reflect not only the strength of the relationship (which is already provided by the path coefficient itself) but also the power of the test, which increases with sample size. The larger the sample size, the lower a path coefficient has to be to yield a statistically significant P value.

## Indicator Loadings and Cross-Loadings

Indicator loadings and cross-loadings are provided in a table with each cell referring to an indicator-LV link (Figure 6). LV names are listed at the top of each column, and indicator names at the beginning of each row.

These indicator loadings and cross-loadings are from a pattern matrix, which is obtained after the transformation of a structure matrix through an oblique rotation. The structure matrix contains the Pearson correlations between indicators and LVs, which are not particularly meaningful prior to rotation in the context of measurement instrument validation. Because an oblique rotation is employed, in some cases loadings may be higher than 1 (Rencher, 1998), which should have no effect on their interpretation. The expectation is that loadings, which are shown within parentheses, will be high; and cross-loadings will be low.

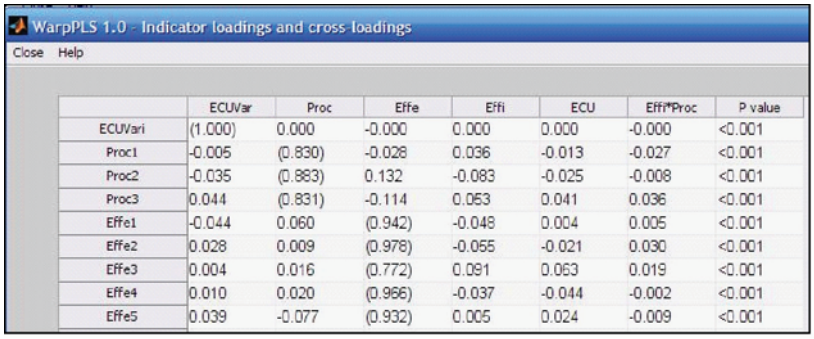
The main difference between oblique and orthogonal rotation methods is that the former assume that there are correlations, some of which may be strong, among LVs. Arguably oblique rotation methods are the most appropriate in PLS-based SEM analysis, because by definition LVs are expected to be correlated in SEM. Otherwise, no path coefficient would be significant. (Technically speaking, it is possible that a research study will hypothesize only neutral relationships between LVs, which could call for an orthogonal rotation. However, this is rarely, if ever, the case.)

P values are also provided, but only for reflective and moderating LVs. These P values are often referred to as validation parameters of a confirmatoty factor analysis, since they result from a test of a model where the relationships between indicators and LVs are defined beforehand. Conversely, in an exploratory factor analysis, relationships between indicators and LVs are not defined beforehand, but inferred based on the results yielded by a factor extraction algorithm; principal components analysis is one of the most popular of these algorithms.

For research reports, users will typically use the table of loadings and cross-loadings provided by this software when describing the convergent validity of their measurement instrument. A measurement instrument has good convergent validity if the question-statements (or other measures) associated with each LV are understood by the respondents in the same way as they were intended by the designers of the question-statements. In this respect, two criteria are recommended as the basis for concluding that a measurement model has acceptable convergent validity: that the P values associated with the loadings be lower than .05; and that the loadings be equal to or greater than .5 (Hair et al., 1987).

Indicators for which these criteria are not satisfied may be removed. This does not apply to formative LV indicators. If the offending indi-

*Figure 6. Indicator loadings and cross-loadings window*




| | ECUVar | Proc | Effe | Effi | ECU | Effi*Proc | P value |
|---|---|---|---|---|---|---|---|
| ECUVari | (1.000) | 0.000 | -0.000 | 0.000 | 0.000 | -0.000 | <0.001 |
| Proc1 | -0.005 | (0.830) | -0.028 | 0.036 | -0.013 | -0.027 | <0.001 |
| Proc2 | -0.035 | (0.883) | 0.132 | -0.083 | -0.025 | -0.008 | <0.001 |
| Proc3 | 0.044 | (0.831) | -0.114 | 0.053 | 0.041 | 0.036 | <0.001 |
| Effe1 | -0.044 | 0.060 | (0.942) | -0.048 | 0.004 | 0.005 | <0.001 |
| Effe2 | 0.028 | 0.009 | (0.978) | -0.055 | -0.021 | 0.030 | <0.001 |
| Effe3 | 0.004 | 0.016 | (0.772) | 0.091 | 0.063 | 0.019 | <0.001 |
| Effe4 | 0.010 | 0.020 | (0.966) | -0.037 | -0.044 | -0.002 | <0.001 |
| Effe5 | 0.039 | -0.077 | (0.932) | 0.005 | 0.024 | -0.009 | <0.001 |

cators are part of a moderating effect, then you should consider removing the moderating effect. Moderating effect LV names are displayed on the table as product LVs (e.g., Effi*Proc). Moderating effect indicator names are displayed on the table as product indicators (e.g., "Effi1*Proc1"). Low P values for moderating effects suggest possible multicolinearity problems. This is to be expected with moderating effects, since the corresponding product variables are likely to be correlated with at least their component LVs. Moreover, moderating effects add nonlinearity to models, which can in some cases compound multicolinearity problems. Because of these and other related issues, moderating effects should be used sparingly.

## Indicator Weights

Indicator weights are provided in a table, much in the same way as indicator loadings are (Figure 7). All cross-weights are zero, because of the way they are calculated through PLS regression. Each LV score is calculated as an exactly linear combination of its indicators, where the weights are multiple regression coefficients linking the indicators to the LV.

P values are provided for weights associated with formative LVs. These values can also be seen, together with those for loadings associated with reflective and moderating LVs, as the result of a confirmatory factor analysis. In research reports, users may want to report these P values as an indication that formative LV measurement items were properly constructed.

As in multiple regression analysis (Miller & Wichern, 1977; Mueller, 1996), it is recommended that weights with P values lower than .05 be considered valid items in a formative LV measurement item subset. Formative LV indicators whose weights do not satisfy this criterion may be considered for removal.

However, this criterion should not trump other criteria grounded on formative LV theory (Diamantopoulos, 1999; Diamantopoulos & Winklhofer, 2001; Diamantopoulos & Siguaw, 2006). Among other things, formative LVs are expected, often by design, to have many indicators (e.g., 15 or more). Yet, given the nature of multiple regression, indicator weights will normally go down as the number of indicators goes up , as long as those indicators are somewhat correlated. Respective P values will normally go up as well.

## Latent Variable Coefficients

Several estimates are provided for each LV that can be used in research reports for discussions

*Figure 7. Indicator weights window*


| | ECUVar | Proc | Effe | Effi | ECU | Effi*Proc | P value |
|---|---|---|---|---|---|---|---|
| Effi4 | 0.000 | 0.000 | 0.000 | (0.229) | 0.000 | 0.000 | |
| Effi5 | 0.000 | 0.000 | 0.000 | (0.224) | 0.000 | 0.000 | |
| ECU1 | 0.000 | 0.000 | 0.000 | 0.000 | (0.401) | 0.000 | <0.001 |
| ECU2 | 0.000 | 0.000 | 0.000 | 0.000 | (0.399) | 0.000 | <0.001 |
| ECU3 | 0.000 | 0.000 | 0.000 | 0.000 | (0.172) | 0.000 | 0.001 |
| ECU4 | 0.000 | 0.000 | 0.000 | 0.000 | (0.252) | 0.000 | <0.001 |
| ECU5 | 0.000 | 0.000 | 0.000 | 0.000 | (0.217) | 0.000 | <0.001 |
| Effi1*Proc1 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | (0.092) | |
| Effi1*Proc2 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | (0.094) | |
| Effi1*Proc3 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | (0.090) | |
| Effi2*Proc1 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | (0.091) | |
| Effi2*Proc2 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | (0.093) | |
| Effi2*Proc3 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | (0.086) | |
| Effi3*Proc1 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | (0.088) | |

on the measurement instrument's reliability and discriminant validity (Figure 8). R-squared coefficients are provided only for endogenous LVs, and reflect the percentage of explained variance for each of those LVs. Composite reliability and Cronbach alpha coefficients are provided for all LVs; these are measures of reliability. Average variances extracted (AVE) are also provided for all LVs, and are used in the assessment of discriminant validity.

The following criteria, one more conservative and the other two more relaxed, are suggested in the assessment of the reliability of a measurement instrument. These criteria apply only to reflective LV indicators. Reliability is a measure of the quality of a measurement instrument; the instrument itself is typically a set of question-statements. A measurement instrument has good reliability if the question-statements (or other measures) associated with each LV are understood in the same way by different respondents.

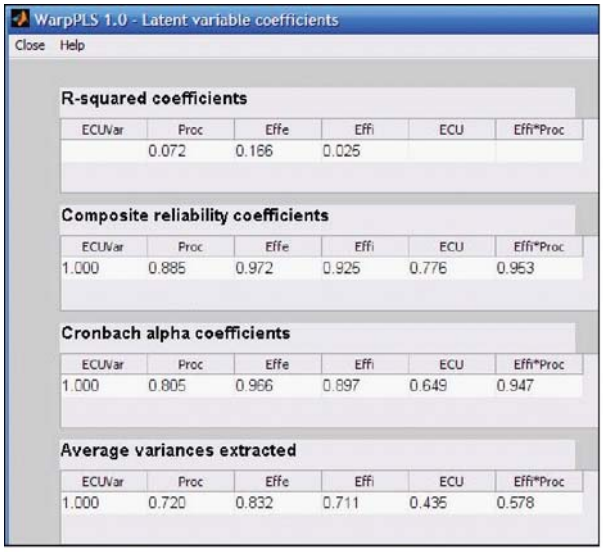
More conservatively, both the compositive reliability and the Cronbach alpha coefficients should be equal to or greater than .7 (Fornell & Larcker, 1981; Nunnaly, 1978; Nunnally & Bernstein, 1994). The more relaxed version of this criterion, which is widely used, is that one of the two coefficients should be equal to or greater than .7. This typically applies to the composite reliability coefficient, which is usually the higher of the two (Fornell & Larcker, 1981). An even more relaxed version sets this threshold at .6 (Nunnally & Bernstein, 1994). If a LV does not satisfy any of these criteria, the reason will often be one or a few indicators that load weakly on the LV. These indicators should be considered for removal.

Average variances extracted are normally used in conjunction with LV correlations in the assessment of a measurement instrument's discriminant validity. This is discussed below, together with the discussion of the table of correlations among LVs.

## Correlations Among Latent Variables

Among the results generated by this software are tables containing LV correlations, and the P values associated with those correlations (Figure 9). On the diagonal of the LV correlations table are the square roots of the average variances extracted for each LV. These results

*Figure 8. Latent variable coefficients window*



are used for the assessment of the measurement instrument's discriminant validity.

In most research reports, users will typically show the table of correlations among LVs, with the square roots of the average variances extracted on the diagonal, to demonstrate that their measurement instrument passes widely accepted criteria for discriminant validity assessment. A measurement instrument has good discriminant validity if the question-statements (or other measures) associated with each LV are not confused by the respondents to the questionnaire with the question-statements associated with other LVs, particularly in terms of the meaning of the question-statements.

The following criterion is recommended for discriminant validity assessment: for each LV, the square root of the average variance extracted should be higher than any of the correlations involving that LV (Fornell & Larcker, 1981). That is, the values on the diagonal should be higher than any of the values above or below them, in the same column. Or, the values on the diagonal should be higher than any of the values to their left or right, in the same row; which means the same as the previous statement, given the repeated values of the LV correlations table.

The above criterion applies to reflective and formative LVs, as well as product LVs representing moderating effects. If it is not satisfied, the culprit is usually an indicator that loads strongly on more than one LV. Also, the problem may involve more than one indicator. You should check the loadings and cross-loadings matrix to see if you can identify the offending indicator or indicators, and consider removing it.

Second to LVs involved in moderating effects, formative LVs are the most likely to lead to discriminant validity problems. This is one of the reasons why formative LVs are not used as often as reflective LVs in empirical research. In fact, it is wise to use formative variables sparingly in models that will serve as the basis for SEM analysis. Formative variables can in many cases be decomposed into reflective LVs, which themselves can then be added to the model. Often this provides a better understanding of the empirical phenomena under investigation, in addition to helping avoid discriminant validity problems.

*Figure 9. Correlations among latent variables window*

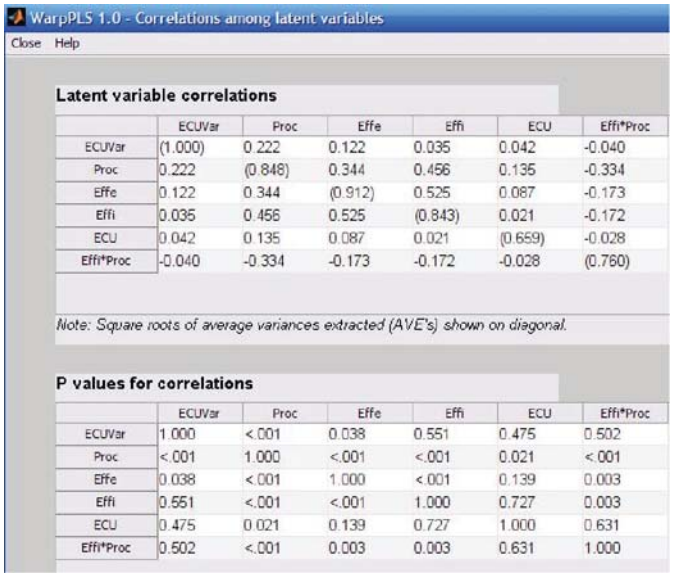
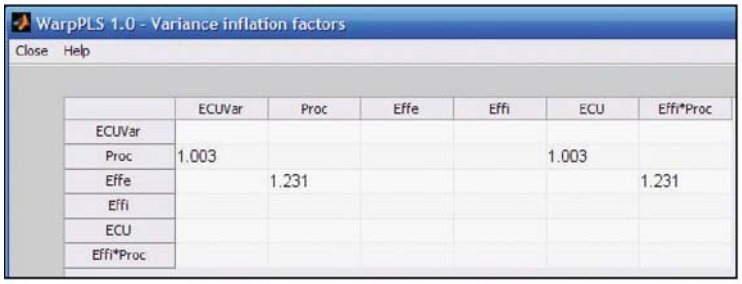


*Figure 10. Variance inflation factors window*




## Variance Inflation factors

Variance inflation factors are provided in table format (Figure 10) for each LV that has two or more predictors. Each variance inflation factor is associated with one predictor, and relates to the link between that predictor and its LV criterion. (Or criteria, when one predictor LV points at two or more different LVs in the model).

A variance inflation factor is a measure of the degree of multicolinearity among the LVs that are hypothesized to affect another LV. For example, let us assume that there is a block of LVs in a model, with three LVs: A, B, and C (predictors); pointing at one LV: D. In this case, variance inflation factors are calculated for A, B, and C, and are estimates of the multicolinearity among these predictor LVs.

Two criteria, one more conservative and one more relaxed, are recommended in connection with variance inflation factors. More conservatively, it is recommended that variance inflation factors be lower than 5; a more relaxed criterion is that they be lower than 10 (Hair et al., 1987; Kline, 1998). High variance inflation factors usually occur for pairs of predictor LVs, and suggest that the LVs measure the same thing; which calls for the removal of one of the LVs from the block, or the model.

## Correlations Among Indicators

The software allows users to view the correlations among all indicators in table format. Only the correlations for indicators included in the model are shown through the menu option "View correlations among indicators", available from the "View and save results" window. This option is useful for users who want to run a quick check on the correlations among indicators while they are trying to identify possible sources of multicolinearity.

The table of correlations among indicators used in the model is usually much larger, with many more columns and rows, than that of the correlations among LVs. For this reason, the P values for the correlations are not shown in the screen view option, but are saved in the related tab-delimited text file.

For correlations among all indicators, including those indicators not included in the model, use the menu option "Save general descriptive statistics into a tab-delimited .txt file". This menu option is available from the main software window, after Step 3 is completed (i.e., the data for the SEM analysis has been pre-processed). This option is generally more appropriate for users who want to include the correlations among indicators in their research reports, as part of a descriptive statistics table. This option also generates means and standard deviations for each of the indicators.

*Figure 11. Linear and nonlinear ("warped") relationships among latent variables window*

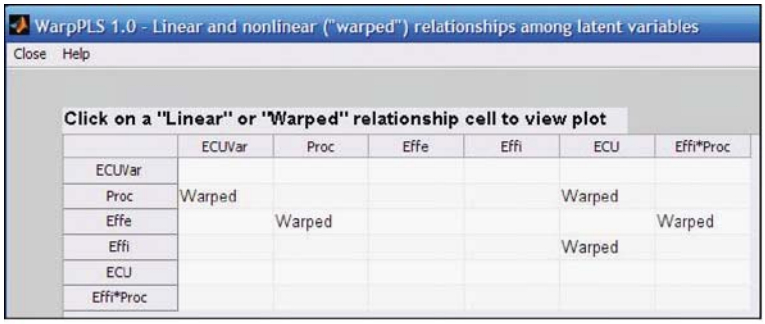
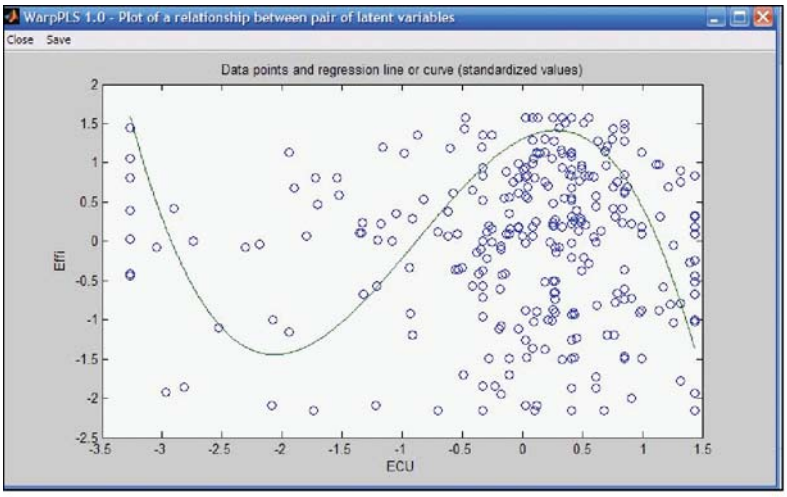



*Figure 12. Plot of a relationship between pair of latent variables*




Indicators that are not used in the model may simply be deleted prior to the inclusion in a research report.

## Linear and Nonlinear Relationships Among Latent Variables

The software shows a table with the types of relationships, warped or linear, between LVs that are linked in the model (Figure 11). The term "warped" is used for relationships that are clearly nonlinear, and the term "linear" for linear or quasi-linear relationships. Quasi-linear relationships are slightly nonlinear relationships, which look linear upon visual inspection on plots of the regression curves that best approximate the relationships.

Plots with the points as well as the regression curves that best approximate the relationships can be viewed by clicking on a cell containing a relationship type description. (These cells are the same as those that contain path coefficients, in the path coefficients table.) See Figure 12 for an example of one of these plots. In this example, the relationship takes the form of a distorted S-curve. The curve may also be seen as a combination of two U-curves, one of which (on the right) is inverted.

As mentioned earlier, the Warp2 PLS Regression algorithm tries to identify a U-curve relationship between LVs, and, if that relationship exists, the algorithm transforms (or "warps") the scores of the predictor LVs so as to better reflect the U-curve relationship in the estimated path coefficients in the model. The Warp3 PLS Regression algorithm, the default algorithm used by this software, tries to identify a relationship defined by a function whose first derivative is a U-curve. This type of relationship follows a pattern that is more similar to an S-curve (or a somewhat distorted S-curve), and can be seen as a combination of two connected U-curves, one of which is inverted.

Sometimes a Warp3 PLS Regression will lead to results that tell you that a relationship between two LVs has the form of a U-curve or a line, as opposed to an S-curve. Similarly, sometimes a Warp2 PLS Regression's results will tell you that a relationship has the form of a line. This is because the underlying algorithms find the type of relationship that best fits the distribution of points associated with a pair of LVs, and sometimes those types are not S-curves or U-curves.

The plots of relationships between pairs of LVs provide a much more nuanced view of how each pair of LVs is related. However, caution must be taken in the interpretation of these plots, especially when the distribution of data points is very uneven.

An extreme example would be a warped plot in which all of the data points would be concentrated on the right part of the plot, with only one data point on the far left part of the plot. That single data point, called an outlier, would influence the shape of the nonlinear relationship. In these cases, the researcher must decide whether the outlier is "good" data that should be allowed to shape the relationship, or is simply "bad" data resulting from a data collection error.

If the outlier is found to be "bad" data, it can be removed from the data set by a simple procedure. The user should save the standardized data into a text file, using the menu option "Save standardized pre-processed data into a tab-delimited .txt file". This menu option is available from the main software window, after Step 3 is completed (i.e., the data for the SEM analysis has been pre-processed). Next the user should open the file with spreadsheet software (e.g., Excel). The outlier should be easy to identify, on the dataset, and should be eliminated. Then the user should re-read this modified file as if it was the original data file, and run the SEM analysis steps again. This procedure may lead to a visible change in the shape of the nonlinear relationship, and significantly affect the results.

## REFERENCES

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research*, *14*(2), 189–218. doi:10.1287/isre.14.2.189.16018

Chiquoine, B., & Hjalmarsson, E. (2009). Jackknifing stock return predictions. *Journal of Empirical Finance*, *16*(5), 793–803. doi:10.1016/j.jempfin.2009.07.003

Diamantopoulos, A. (1999). Export performance measurement: Reflective versus formative indicators. *International Marketing Review*, *16*(6), 444–457. doi:10.1108/02651339910300422

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

Diamantopoulos, A., & Winklhofer, H. (2001). Index construction with formative indicators: An alternative scale development. *JMR, Journal of Marketing Research*, *37*(1), 269–177. doi:10.1509/jmkr.38.2.269.18845

Efron, B., Rogosa, D., & Tibshirani, R. (2004). Resampling methods of estimation . In Smelser, N. J., & Baltes, P. B. (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 13216–13220). New York, NY: Elsevier. doi:10.1016/B0-08-043076-7/00494-0

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

Gefen, D., Straub, D. W., & Boudreau, M.-C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the AIS*, *4*(7), 1–76.

Haenlein, M., & Kaplan, A. M. (2004). A beginner's guide to partial least squares analysis. *Understanding Statistics*, *3*(4), 283–297. doi:10.1207/s15328031us0304_4

Hair, J. F., Anderson, R. E., & Tatham, R. L. (1987). *Multivariate data analysis*. New York, NY: Macmillan.

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York, NY: The Guilford Press.

Kock, N. (2009). Information systems theorizing based on evolutionary psychology: An interdisciplinary review and theory integration framework. *Management Information Systems Quarterly*, *33*(2), 395–418.

Kock, N. (2010a). *WarpPLS 1.0 User Manual*. Laredo, TX: ScriptWarp Systems.

Kock, N. (2010b). Using WarpPLS in e-collaboration studies: An overview of five main analysis steps. *International Journal of e-Collaboration*, *6*(4), 1–13. doi:10.4018/jec.2010100101

Maruyama, G. M. (1998). *Basics of structural equation modeling*. Thousand Oaks, CA: Sage.

Miller, R. B., & Wichern, D. W. (1977). *Intermediate business statistics: Analysis of variance, regression and time series*. New York, NY: Holt, Rihehart and Winston.

Mueller, R. O. (1996). *Basic principles of structural equation modeling*. New York, NY: Springer.

Nevitt, J., & Hancock, G. R. (2001). Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. *Structural Equation Modeling*, *8*(3), 353–377. doi:10.1207/S15328007SEM0803_2

Nunnally, J. C. (1978). *Psychometric theory*. New York, NY: McGrawHill.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. New York, NY: McGrawHill.

Rencher, A. C. (1998). *Multivariate statistical inference and applications*. New York, NY: John Wiley & Sons.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis*. New York, NY: McGrawHill.

Wold, S., Trygg, J., Berglund, A., & Antti, H. (2001). Some recent developments in PLS modeling. *Chemometrics and Intelligent Laboratory Systems*, *58*(2), 131–150. doi:10.1016/S0169-7439(01)00156-3

Wolfle, L. M. (1999). Sewall Wright on the method of path coefficients: An annotated bibliography. *Structural Equation Modeling*, *6*(3), 280–291. doi:10.1080/10705519909540134

Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics*, *5*(3), 161–215. doi:10.1214/aoms/1177732676