SEM FOR SMALL SAMPLE AND NON-NORMAL DATA WORKSHOP

14-15 MAC 2011
UUMKL, KUALA LUMPUR

STRUCTURAL EQUATION MODELING (SEM)

Mokhtar Abdullah, UPNM

Organised by :
Center for Testing, Measurement and Appraisal (CeTMA)

STRUCTURAL EQUATION MODELING
(SEM)

Partial Least Squares

Mokhtar Abdullah, UPNM

Part I

- Overview of SEM.
- Estimation Methods for SEM
  - Covariance-based SEM (CBSEM) or LISREL
  - Partial Least Squares (PLS)
- Comparison between Covariance-based SEM and PLS

PART II

- PLS using software **SmartPLS**
- Goodness of Fits Statistics with PLS
  - Reliability & Validity measures
- SmartPLS Outputs

PART III

- Mediating & Moderating Effects in SEM with PLS

CONTENT

OVERVIEW OF SEM

- First generation regression models such as linear regression, Logit, ANOVA and MANOVA analyse only one level of linkage between independent and dependent variables at a time.
- Unlike first generation regression tools, SEM not only assesses causation of a set of dependent and independent constructs – but, in the same analysis, also evaluates the 'measurement model' – loadings of observed items (measurement) on their expected 'latent variables' (constructs).
- The combined analysis of the 'measurement' and the 'structural' models enables the measurement errors of the observed variables to be analysed simultaneously.

Why SEM?

FIRST GENERATION REGRESSION MODELS

$Y = \alpha + \beta X + \text{error}$ — **No error in X, only in Y** — **Simple Regression Model**

Y: Dependent X : Independent

$Y = \alpha 1 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_m X_m + \text{error1}$ — **Multiple Regression Model**

$$\left.\begin{aligned} Y1 &= \alpha 1 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_m X_m + \text{error1};\\ Y2 &= \alpha 2 + \gamma_1 X_1 + \gamma_2 X_2 + \ldots + \gamma_m X_m + \text{error2};\\ &\ \ .\\ Yp &= \alpha_p + \phi_1 X_1 + \phi_2 X_2 + \ldots + \phi_m X_m + \text{error2} \end{aligned}\right\} \text{Multivariate Regression}$$

SECOND GENERATION OF REGRESSION MODELS : SEM

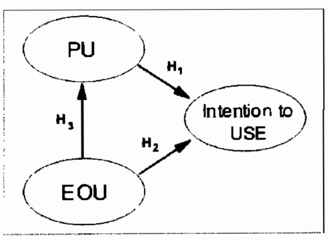
$Y = \alpha + \beta \xi + \text{error}$; $X = \xi + \delta$ — δ – Error in X, ξ – Latent Variable — **Simple SEM**

$$\left.\begin{aligned} Y1 &= \alpha 1 + \beta_1 \xi_1 + \beta_2 \xi_2 + \ldots + \beta_m \xi_m + \text{error1};\\ Y2 &= \alpha 2 + \gamma_1 \xi_1 + \gamma_2 \xi_2 + \ldots + \gamma_m \xi_m + \text{error2};\\ &\ \ .\\ Yp &= \alpha_p + \phi_1 \xi_1 + \phi_2 \xi_2 + \ldots + \phi_m \xi_m + \text{error2}\\ X1 &= \xi_1 + \delta 1;\ X2 = \xi_2 + \delta 2;\ \ldots;\ Xk = \xi_k + \delta k \end{aligned}\right\} \text{SEM}$$

Errors in X1, X2 ... Xk

EXAMPLE : Modeling Latent Variables/Constructs

Technology Acceptance Model (TAM)
(Fishbein & Ajzen, 1975; Davies, 1989)

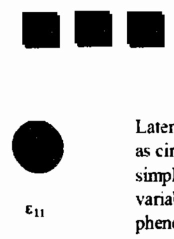


PU – Perceived Usefulness
EOU – ease of Use
IUSE – Intention to Use

PU, EOU, and IUSE – Latent Variables/Constructs, i.e, Variables that are NOT measurable directly – use Indicators / proxies

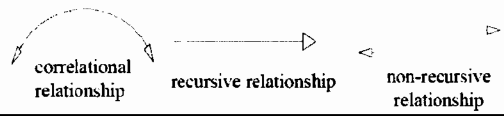
- Indicators (often called manifest variables or observed measures/variables)
- Latent variable (or construct, concept, factor)
- Path relationships (correlational, one-way paths, or two way paths).

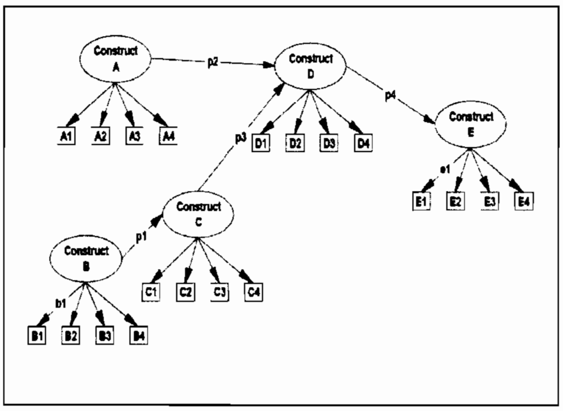
SEM with causal diagrams involve three primary components:



indicators are normally represented as squares. For questionnaire based research, each indicator would represent a particular question.

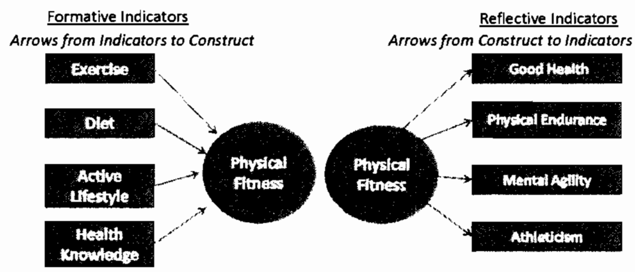
Latent variables are normally drawn as circles. In the case of error terms, for simplicity, the circle is left off. Latent variables are used to represent phenomena that cannot be measured directly. Examples would be beliefs, intention, motivation.





Reflective vs Formative (Indicators) Measurement Model

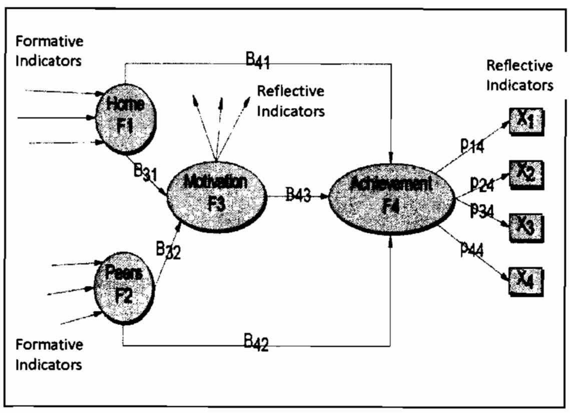
Note: Arrows in Opposite Directions



These indicators need not covary

These indicators should covary

EXAMPLE : Formative & Reflective Indicators



- Covariance-based Method (CBSEM or LISREL)
- Component- or PLS-based Method

Structural Equation Modeling (SEM) – Two Methods

CBSEM

Steps

1. **Define individual constructs:**
   - Define the constructs theoretically.
2. **Develop the overall measurement model:**
   - Develop indicators
3. **Design the study to produce the empirical results**
4. **Assess the measurement model validity** (using CFA)
5. **Specify the structural model** (loadings)
6. **Examine the structural model validity:**
   (like CFI, GFI, TLI, AGFI, etc.) and one badness of fit index (like RMR, RMSEA, SRMR, etc.) meet the predetermined criteria

Assumptions & Model Assessment

Comparison between PLS & CBSEM

Criterion	PLS	CBSEM
Objective	Prediction oriented	Parameter oriented
Approach	Variance based	Covariance based
Assumptions	Predictor Specification (non parametric)	Typically multivariate normal distribution and independent observations (parametric)
Parameter estimates	Consistent as indicators and sample size increase (i.e., consistency at large)	Consistent
Latent Variable scores	Explicitly estimated	Indeterminate

(ref. Chin & Newsted, 1999 In Rick Hoyle (Ed.). Statistical Strategies for Small Sample Research. Sage Publications, pp. 307-341)

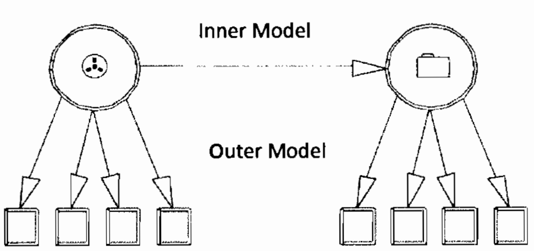
Comparison between PLS & CBSEM

Criterion	PLS	CBSEM
Epistemic relationship between a latent variable and its measures	Can be modeled in either formative or reflective mode	Typically only with reflective indicators
Implications	Optimal for prediction accuracy	Optimal for parameter accuracy
Model Complexity	Large complexity (e.g., 100 constructs and 1000 indicators)	Small to moderate complexity (e.g., less than 100 indicators)
Sample Size	Power analysis based on the portion of the model with the largest number of predictors. Minimal recommendations range from 30 to 100 cases.	Ideally based on power analysis of specific model - minimal recommendations range from 200 to 800.

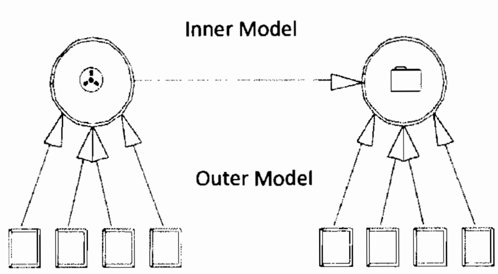
(ref. Chin & Newsted, 1999 In Rick Hoyle (Ed.). Statistical Strategies for Small Sample Research. Sage Publications, pp. 307-341)

Terminologies Used in CBSEM & PLS

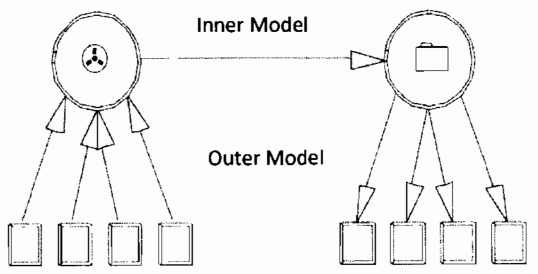
CBSEM	PLS
$X = \lambda\xi + \delta$	$X = \lambda\xi + e$
$y = \beta\xi + \varepsilon$	$y = \beta\xi + \varepsilon$
(Measurement Model)	(Outer Model)
$\eta = \pi\Pi + \varepsilon$	$\eta = \beta\xi + \varepsilon$
(Structural Model)	(Inner Model)



Interbattery factor analysis (mode A)



Canonical correlation analysis (mode B)



Redundancy Analysis (Mode C).

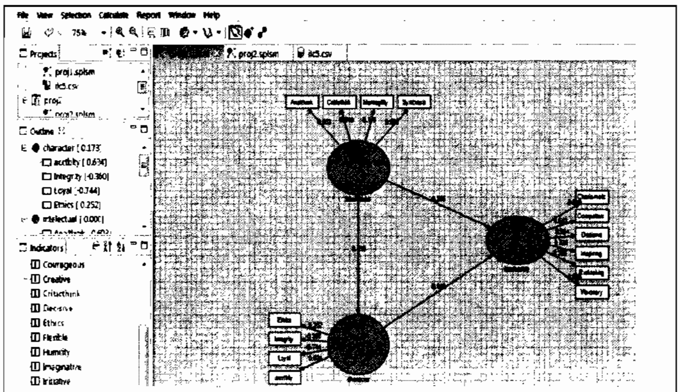
- AMOS
- LISREL
- EQS
- EZPath
- SEPATH
- CALIS
- MX
- RAMONA

CBSEM Softwares

- SmartPLS
- PLS-Graph
- PLS-PC 1.8 (DOS Version)
- XLSTAT-Pro
- PROC PLS (SAS)
- PLSgui (MATLAB)

PLS-Based SEM Softwares

PLS Using SmartPLS



Modeling Intellectual Leaders of Character

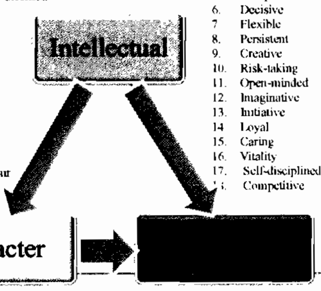
The Case of UPNM Student Cadets

THE FRAMEWORK

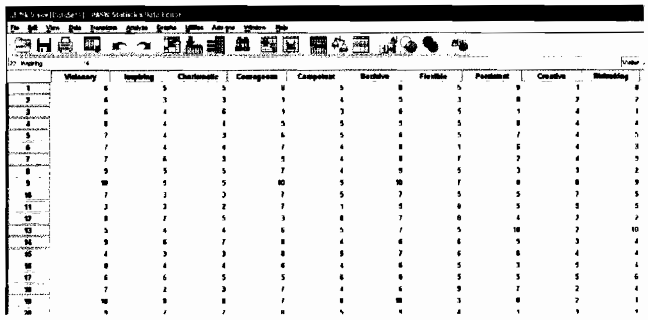
Intellectual Traits:
1. Strategic thinking (*analytical, critical, synthesis, mental agility*)
2. Innovative (*anticipative & responsive to uncertainties*) problem solving
3. Results-oriented

Leadership Traits:
1. Visionary
2. Inspiring
3. Charismatic (dominant, ambitious, self-confident, strong sense of purpose, role model, command respect)
4. Courageous
5. Competent
6. Decisive
7. Flexible
8. Persistent
9. Creative
10. Risk-taking
11. Open-minded
12. Imaginative
13. Initiative
14. Loyal
15. Caring
16. Vitality
17. Self-disciplined
18. Competitive

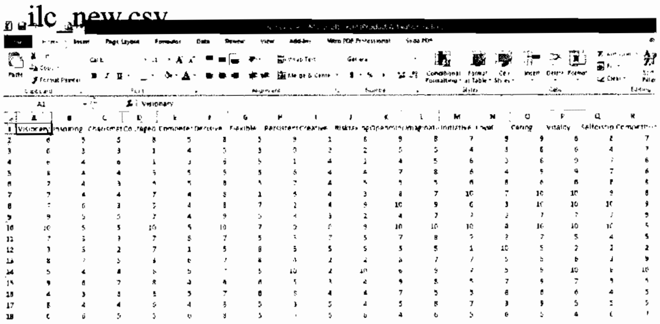
Character Traits:
1. Integrity (honesty)
2. Accountability
3. Moral Ethical Behaviour
4. Humility
5. Hardiness



Data File in SPSS : ilc5.sav



Data File in Excel (comma delimited): ilc_new.csv

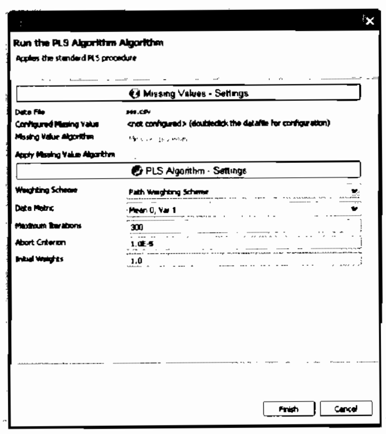


Data Input

- Select File, New, Create new project. When prompted, assign a project name (in the example, the project name is "Intellectual Leaders of Character") and check "Import indicator data," assuming the data are previously saved in a file.
- The filename may differ from the project name (in the example, the data is in a comma delimited Excel file, "ilc_new.csv", converted from SPSS file ilc5.sav).
- On the next dialog screen, browse to the data file and enter it.
- Click the Validate button to see if there are illegal cell entries, such as blanks, If there are illegal values one must quit, make the corrections, then restart.
- If there are missing values, check the box saying there are and ender the missing value code (ex., -99).
- Click Finish.

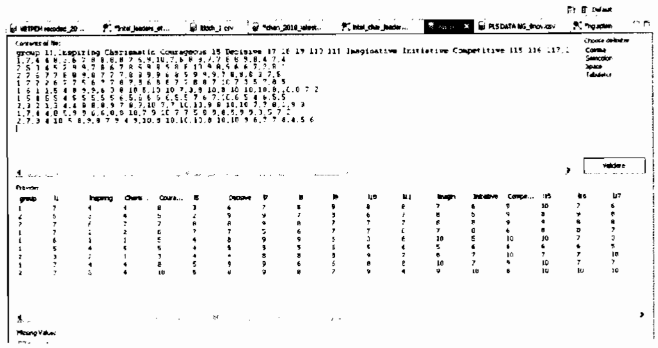
PLS Path Modeling Using SmartPLS

- Warning! SmartPLS works with standardized data. If one's data are not already standardized, it is essential when one of the Calculate menu choices is selected that the researcher specify "Mean 0, Var 1" as the data metric, causing SmartPLS to standardize the data. (See the figure below). If data are already standardized, specify "Original" at this step.

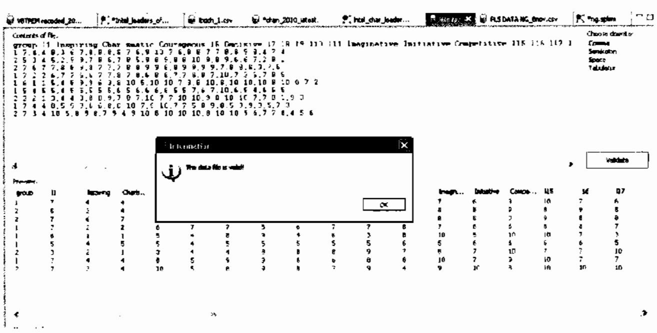


- The SmartPLS data screen shows the first several lines of raw data at the top and then in the window below, the converted data.
- Data are usually entered in comma-delimited text format (.csv). Data may also be delimited by tabs, spaces, or semi-colons.
- Click on the appropriate "Choose delimiter:" choice, which will cause data to appear in the Preview window below.
- Click 'Validate'

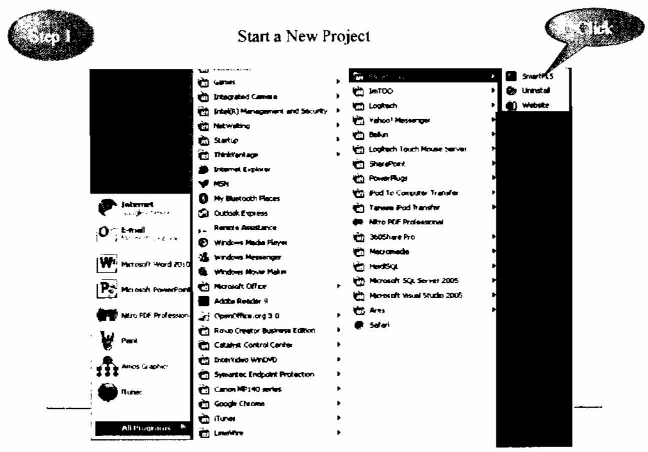
Display of Input Data – sss.csv

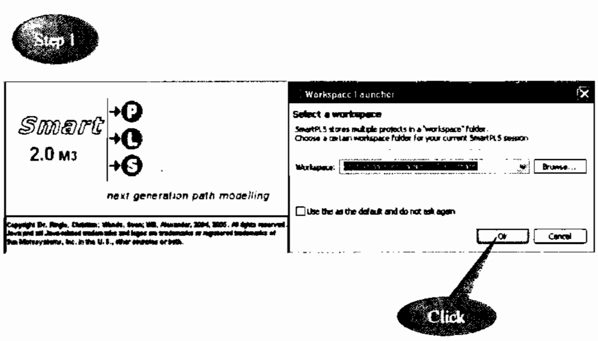


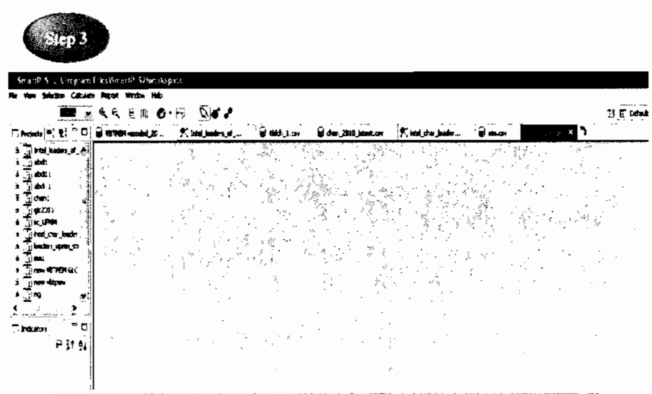
Checking the validity of Data File – sss.csv

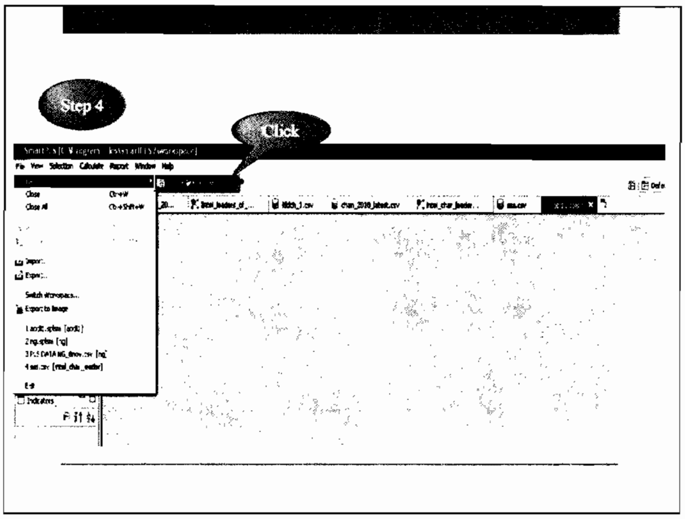


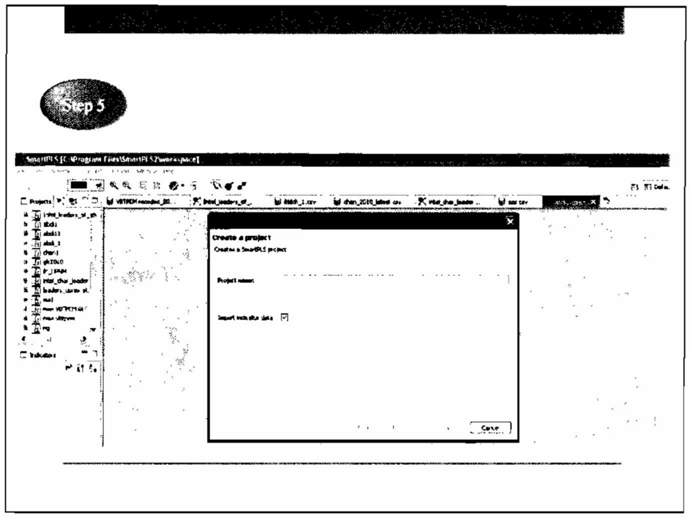
Start a New Project

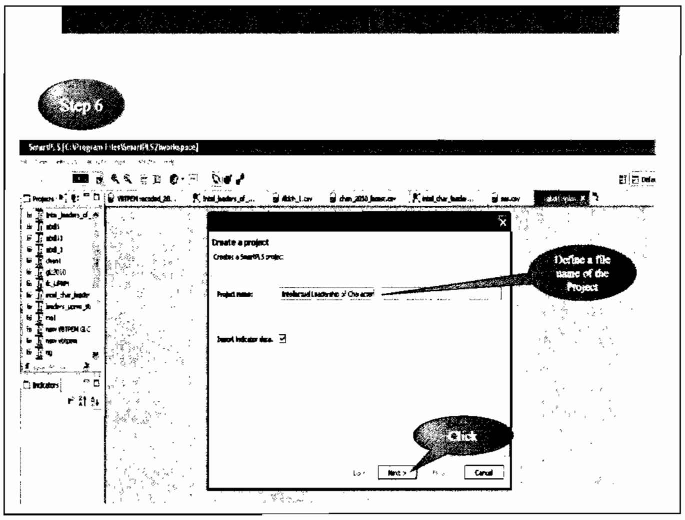


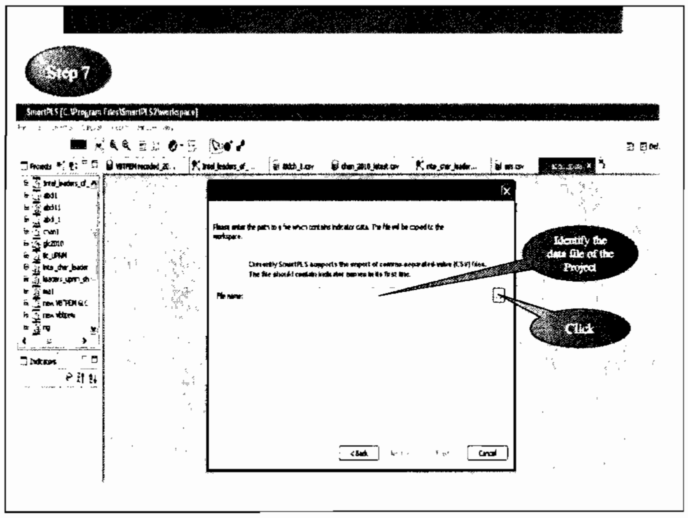


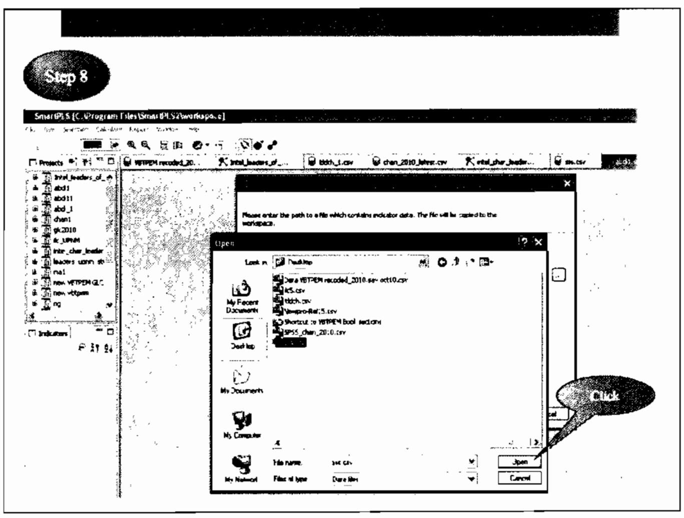


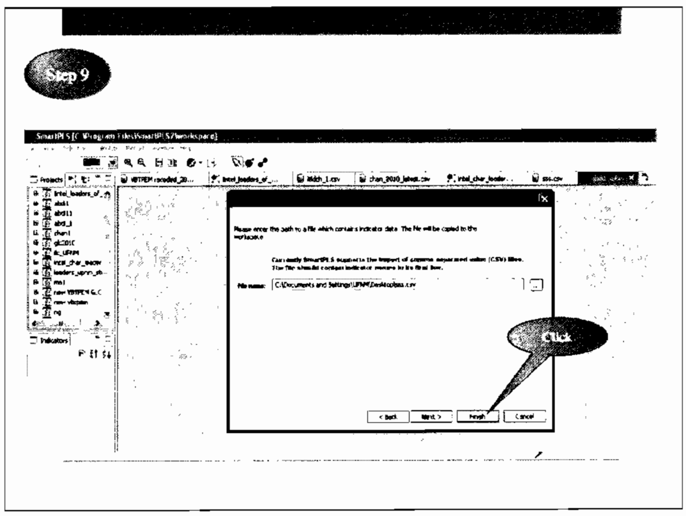

Step 4
Click

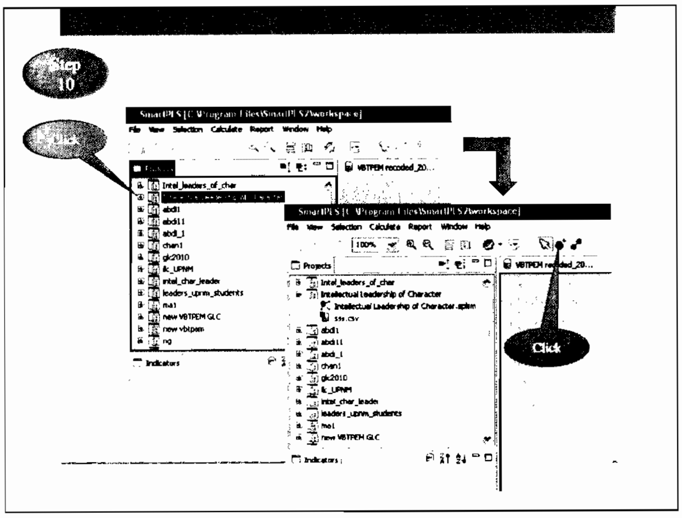

Step 5
Create a project

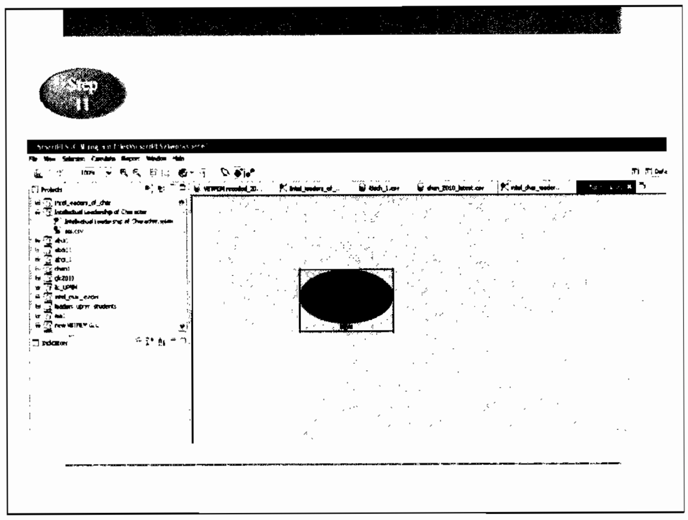

Step 6
Create a project
Define a file name of the Project
Click

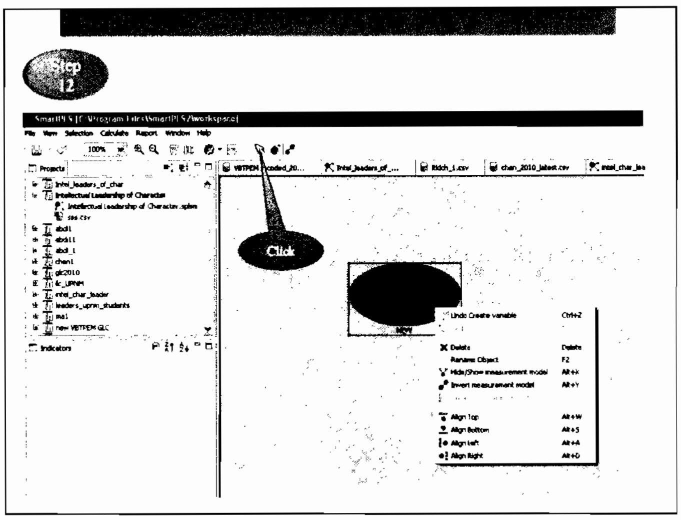

Step 7
Identify the data file of the Project
Click

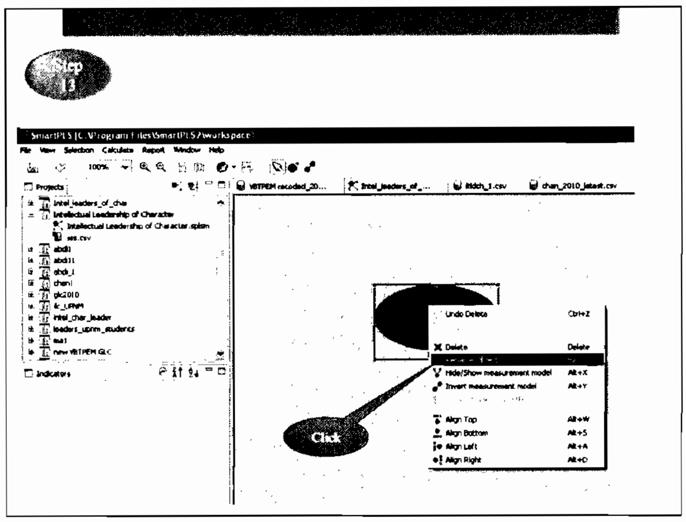

Step 8
Click

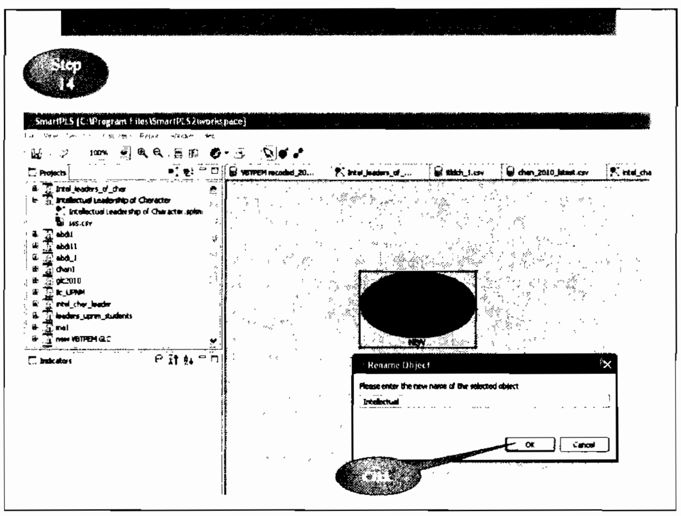

Step 9
Click

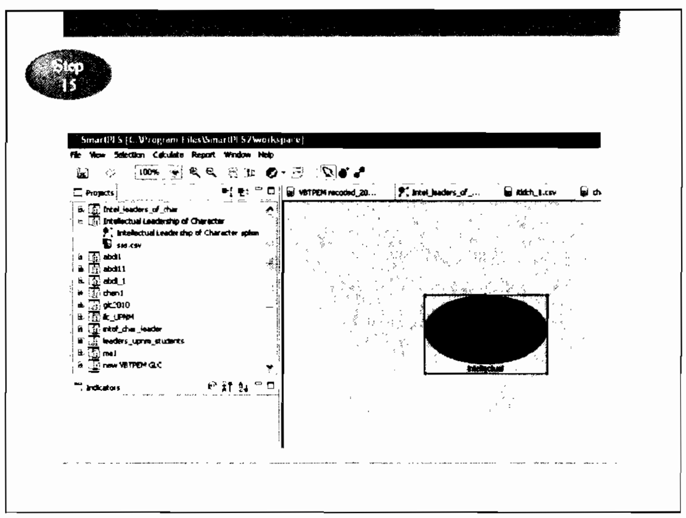

Step 10
Click
Click

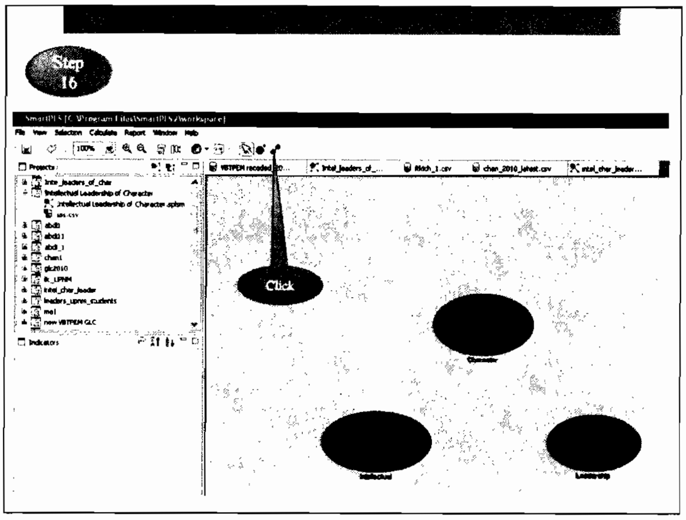

Step 11

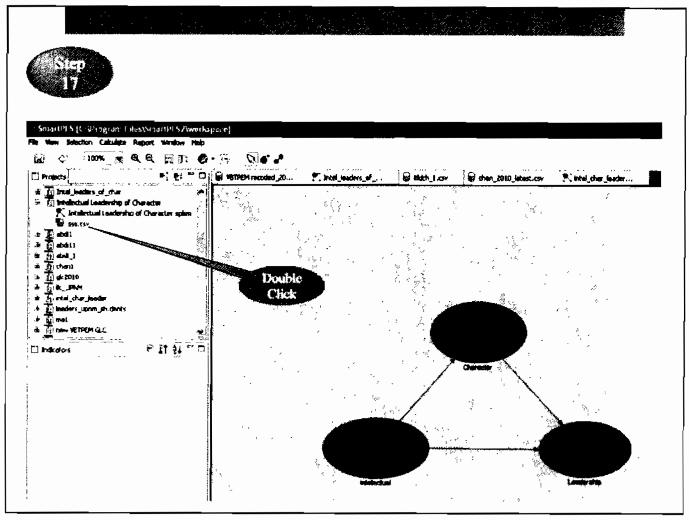

Step 12
Click

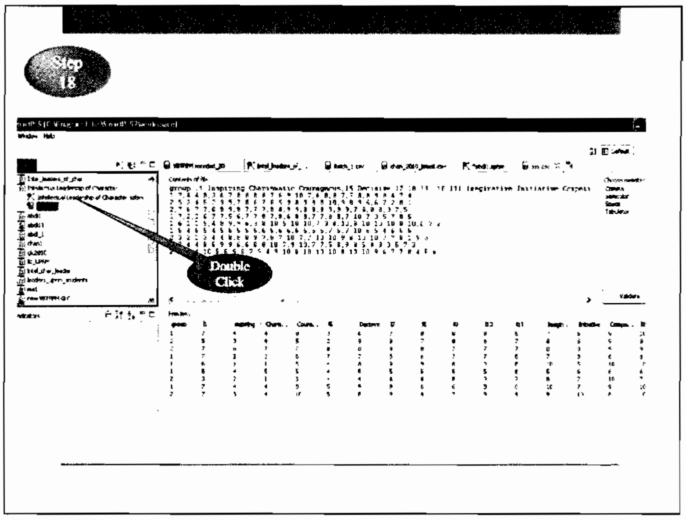

Step 13
Click

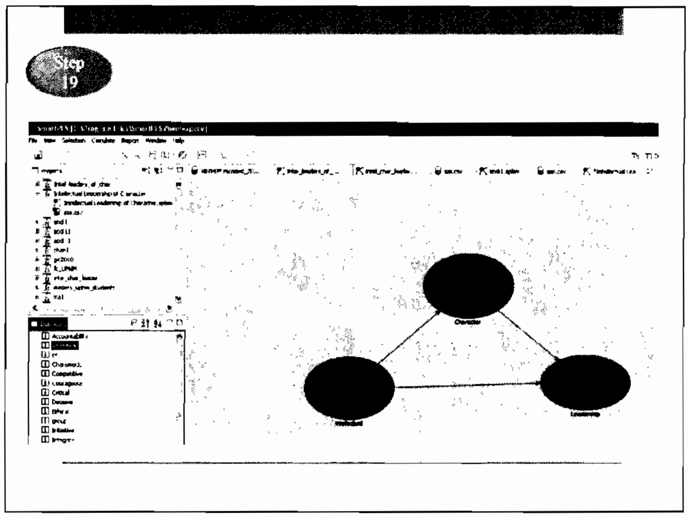

Step 14
Rename Object
OK
Cancel

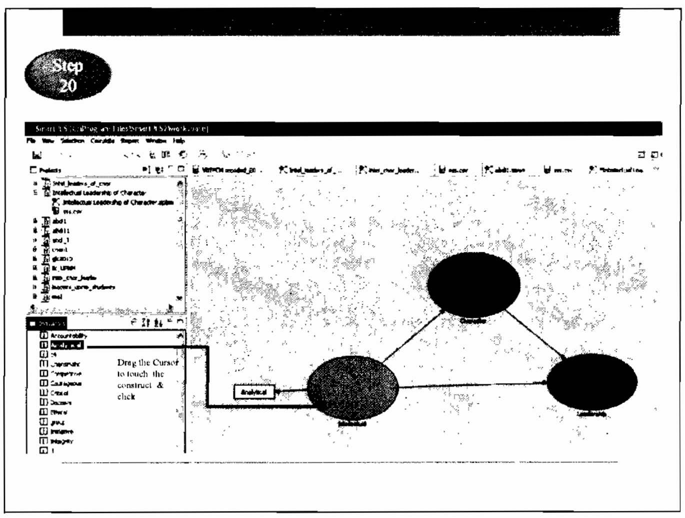

Step 15

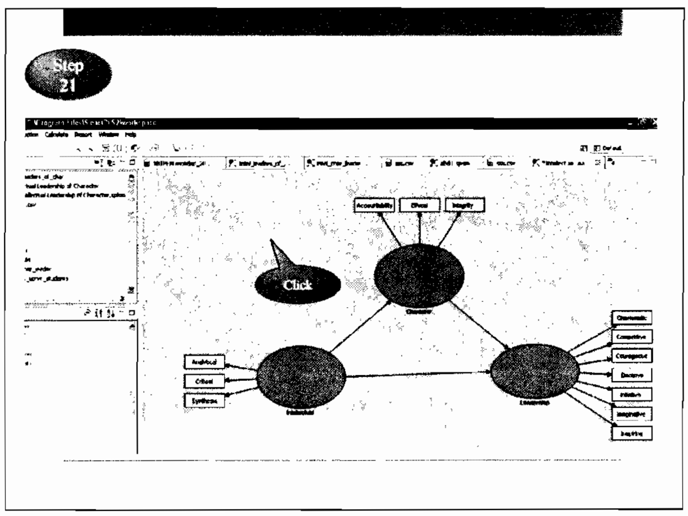

Step 16
Click

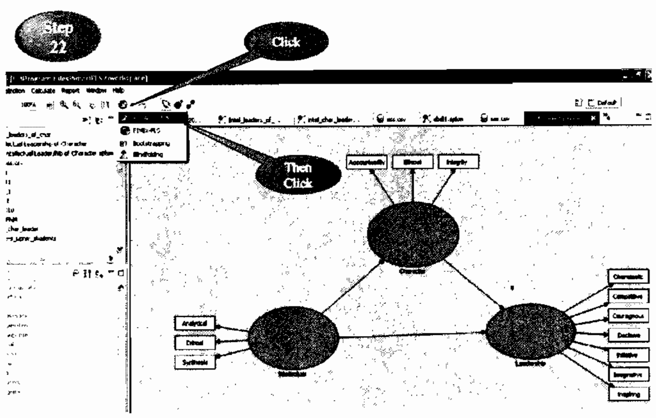

Step 17
Double Click

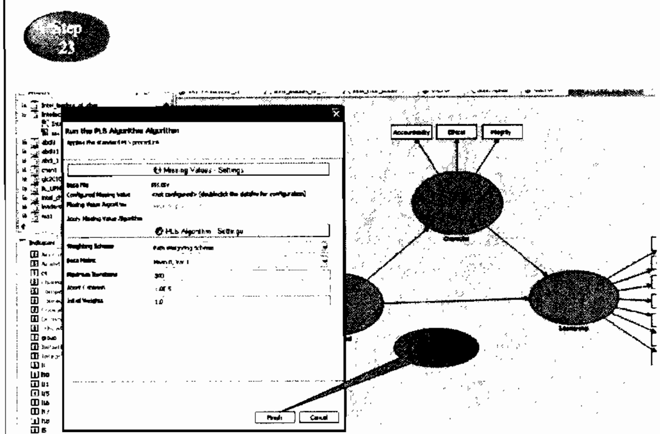

Step 18
Double Click

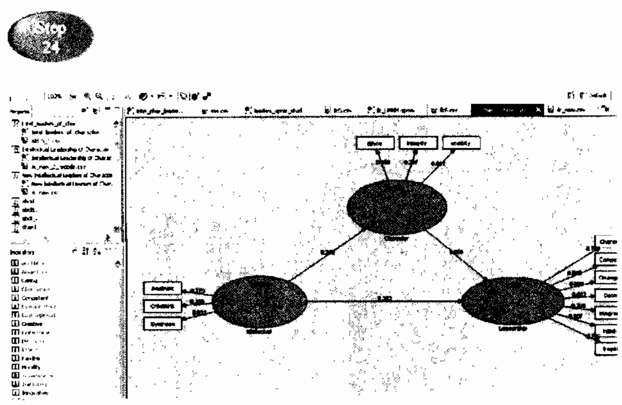

Step 19

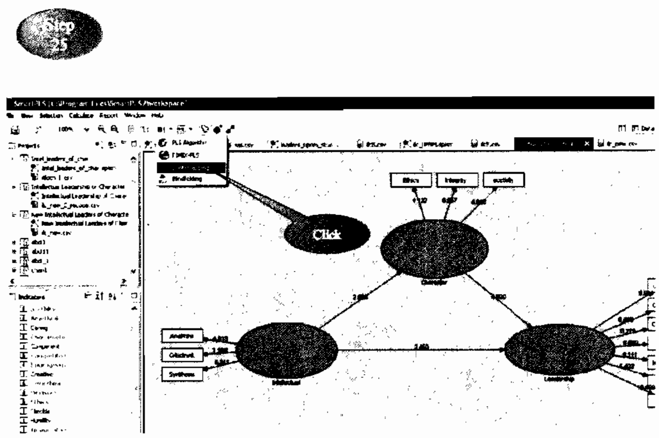

Step 20
Drag the Cursor to touch the construct & click


Step 21
Click









SmartPLS OUTPUTS

- Selecting Report, HTML (Print) Report
- *Model fit coefficients.* Immediately after the listing of the input data, the SmartPLS report displays various coefficients related to model fit ("model quality")
- *Cronbach's alpha.* By convention, alpha should be greater or equal to .80 for a good scale, .70 for an acceptable scale, and .60 for a scale for exploratory purposes

SmartPLS Output

Composite Reliability

- *Composite reliability* is a preferred alternative to Cronbach's alpha as a measure of reliability because Cronbach's alpha may over- or under-estimate scale reliability
- In an adequate model for exploratory purposes, composite reliabilities should be greater than .6 (Chin, 1998; Höck & Ringle, 2006: 15) and greater than .70 for an adequate model for confirmatory purposes

Average Variance Extracted (AVE)

- *AVE* is average variance extracted. It is reflects the average communality for each latent factor and is used to establish *convergent validity*. In an adequate model, AVE should be greater than .5 (Chin, 1998; Höck & Ringle, 2006: 15).
- *Communality*. Likewise, the communality measures the average percent of variance in the indicators for a row factor explained by that row factor and is sometimes interpreted as the reliability of row factor.

R-Square, Redundancy & Correlation

R-Square

- The cutoffs 0.67(substantial), 0.33(moderate) and 0.19(weak), respectively

Redundancy.

- The redundancy coefficient measures the percent of variance in the indicators for the dependent factor explained by the exogenous factors

Latent variable correlations. This shows the correlation coefficients for the factor scores for the three factors

Crossloadings, Outer Model Coefficients

- *Latent variable crossloadings*. The coefficients in this table show not only how much indicators are loaded on the expected factors but also how much each is cross-loaded on the other factors. An ideal model would have strong expected loadings and weak cross-loadings.
- *Measurement model coefficients (outer model coefficients)* are displayed in the "Outer Model (Weights or Loadings)" table. The "outer model" is the measurement model consisting of the indicators and their paths from their respective factors.

Inner Model Coefficients & Factor Scores

- *Structural model path coefficients (inner model coefficients)* for arrows linking the factors are shown in the "Total Effects" table and also displayed on the corresponding arrows in the graphical model. The "inner model" is the structural model consisting of the factors and the arrows connecting them.
- *Factor scores* in standardized form are displayed by default in the "Latent Variable Scores" table. The factor scores can also be analyzed to identify outlier cases (those with a greater absolute value than 1.96 are outliers at the .05 level, those greater than 2.58 are outliers at the .01 level, etc.).

Assumptions with PLS

- **Multicollinearity**. Since the PLS factors are orthogonal, multicollinearity is not a problem in PLS. This is a major reason why PLS models may be selected over OLS regression models or CBSEM structural equation modeling
- **Independence of observations is not required**
- **Distribution-free**. PLS is a distribution-free approach, unlike CBSEM using the usual maximum likelihood estimation method, which assumes multivariate *normality* (see Lohmoller, 1989: 31)

- **Appropriate sample size**. While PLS can be computed even for very small samples (ex., <20) or when cases are fewer than indicator variables, reliance on small samples can yield flawed results. Chin, 1998) recommend PLS users follow a similar "rule of 10" guideline as SEM users.
- Qureshi & Compeau (2009) found that PLS is better than CBSEM when data were normally distributed, with a small sample size and correlated exogenous variables.

Sample Size

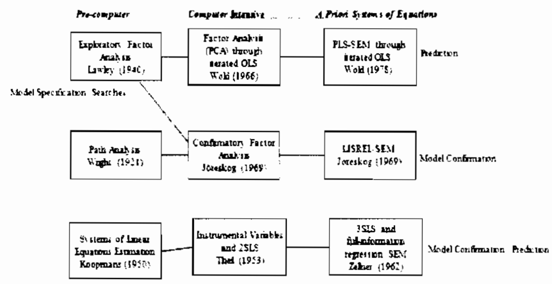
- **Bootstrap estimates of significance**. As the distribution of PLS is unknown, conventional significance testing is impossible. However, testing may be accomplished by bootstrap methods.

Bootstrapping

Chin (2000). Partial Least Squares For Researchers: An overview and presentation of recent advances using the PLS approach, http://disc-nt.cba.uh.edu/chin/indx.html

Superiority of CBSEM Over PLS

- CBSEM is superior to PLS on mathematical grounds. This point refers to the fact that CBSEM is a population based model for estimating loadings and structural path estimates. CBSEM is able to estimate the underlying population parameters.
- Because PLS is a limited information estimation procedure, an appropriate sample size tends to be much smaller than that needed for a full information procedure such as CBSEM.



History of SEM Statistical Methods and Their Precursors

Superiority of PLS Over CBSEM

Chin (2000). Partial Least Squares For Researchers: An overview and presentation of recent advances using the PLS approach, http://disc-nt.cba.uh.edu/chin/indx.html

- PLS is superior to CBSEM on practical grounds.
- PLS is computationally more efficient than CBSEM as a components analysis is faster than a Maximum Likelihood factor analysis. CBSEM estimation time increases dramatically as the number of indicators increase.
- The statement that PLS makes no distributional assumption relate to the asymptotic efficiency of the OLS estimator.

Superiority of PLS Over CBSEM

Chin (2000). Partial Least Squares For Researchers: An overview and presentation of recent advances using the PLS approach, http://disc-nt.cba.uh.edu/chin/indx.html

- PLS avoids two serious problems (Fornell and Bookstein, 1982):
  - *inadmissible solutions*
  - *factor indeterminacy*

  (Fornell, C., and Bookstein, F. (1982). "Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory," *Journal of Marketing Research*, 19, 440-452.)

- The measurement model in PLS is assessed in terms of item loadings and reliability coefficients (composite reliability), as well as the convergent and discriminant validity.
- Individual item **loadings greater than 0.7** are considered adequate (Fornell and Larcker, 1981). Interpreted like a Cronbach's alpha for internal consistency reliability estimate, a composite reliability of .70 or greater is considered acceptable (Fornell and Larcker, 1981).
- The **average variance extracted (AVE)** measures the variance captured by the indicators relative to measurement error, and it **should be greater than .50** to justify using a construct (Barclay, Thompson and Higgins, 1995). An AVE value of at least 0.5 indicates sufficient *convergent validity*, meaning that a latent variable is able to explain more than half of the variance of its indicators on average (e.g., Gotz, Liehr-Gobbers, & Krafft, 2009).

Goodness of Fit Criteria in PLS - Reliability

Composite Reliability

$$\rho_c = \frac{(\sum \lambda_i)^2 \operatorname{var} F}{(\sum \lambda_i)^2 \operatorname{var} F + \sum \Theta_{ii}}$$

where λ_i, F, and Θ_{ii}, are the factor loading, factor variance, and unique/error variance respectively. If F is set at 1, then Θ_{ii} is the **1-square of λ_i.**

Average Variance Extracted

$$AVE = \frac{\sum \lambda_i^2 \operatorname{var} F}{\sum \lambda_i^2 \operatorname{var} F + \sum \Theta_{ii}}$$

where λ_i, F, and Θ_{ii}, are the factor loading, factor variance, and unique/error variance respectively. If F is set at 1, then Θ_{ii} is the **1-square of λ_i.**

- The discriminant validity of the measures (the degree to which items differentiate among constructs or measure distinct concepts) was assessed by examining the correlations between the measures of potentially overlapping constructs.
- Items should **load more strongly on their own constructs** in the model, and the average variance shared between each construct and its measures should be greater than the variance shared between the construct and other constructs (Compeau, Higgins and Huff, 1999).

Goodness of Fit Criteria in PLS - Validity

- The structural model in PLS is assessed by examining the path coefficients (standardized betas).
- T statistics are also calculated to assess the significance of these path coefficients.
- R^2 is used as an indicator of the overall predictive strength of the model.

Goodness of Fit Criteria in PLS – Structural Model

ASSESSING FACTORIAL VALIDITY IN PLS

Convergent Validity: Two Steps

1. Examine the convergent validity of the scales
2. Generate the t-values with a bootstrap. Convergent validity is shown when the t-values of the Outer Model Loadings are above 1.96.

ASSESSING FACTORIAL VALIDITY IN PLS

DISCRIMINANT VALIDITY:

1. Examine item loadings to construct correlations.

2. Examine the ratio of the square root of the AVE of each construct to the correlations of this construct to all the other constructs.

The square root of the AVE of each construct needs to be much larger than any correlation between this construct and any other construct.

Partial Least Squares

EXAMPLE

Mediating & Moderating Effects

Article

Trustworthiness in mHealth Information Services: An Assessment of a Hierarchical Model with Mediating and Moderating Effects Using Partial Least Squares (PLS)

Akter, D'Ambra, and Ray (2010)

Purpose:
The model shows that trustworthiness is a 'second-order', reflective construct that has a significant direct and indirect impact on continuance intentions in the context of mHealth information services. It also confirms that consumer trust plays the key, mediating role between trustworthiness and continuance intentions, while trustworthiness does not have any moderating influence in the relationship between consumer trust and continuance intentions.

The second-order, hierarchical, reflective latent variable

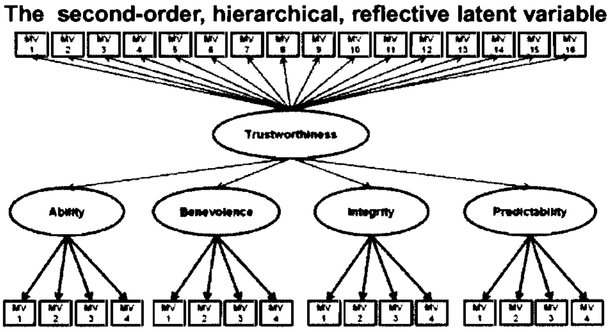


TABLE 4. Estimation of trustworthiness as a reflective, second-order, higher order hierarchical model using PLS.

First-order model	Second-order model
$y_i = \Lambda_y \cdot \eta_j + \varepsilon_i$	$\eta_j = \Gamma \cdot \xi_1 + \zeta_j$
y_i = *manifest variables (e.g., measures or indicators)*	η_j = *first order factors (e.g., ability)*
Λ_y = *loadings of first order latent variable*	Γ = *loadings of second order latent variable*
η_j = *first order latent variable (e.g., ability, benevolence, integrity and predictability)*	ξ_1 = *second order latent variable (e.g., trustworthiness)*
ε_i = *measurement error of manifest variables*	ζ_j = *measurement error of first order factors*

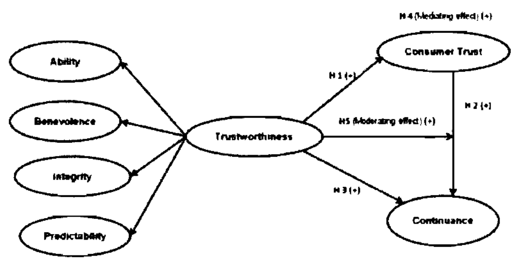


FIG. 3. Research model and hypotheses

Mediation

Mediation is defined as a situation when the predictor (trustworthiness), (1) has a significant influence on the mediator (consumer trust); (2) the mediator (consumer trust) has a significant influence on the criterion variable (continuance intentions), and, (3) the predictor (trustworthiness) has a significant influence on the criterion variable (continuance intentions) in the absence of the mediators' influence (consumer trust; Barron &, 1986).

Moderation

- Moderating variable is "a variable that affects the direction and/or strength of the relation between an independent or predictor variable and a dependent or criterion variable" (Barron & Kenny, 1986, p. 1174).

H1: Trustworthiness has a significant positive impact on consumer trust in mHealth information services.

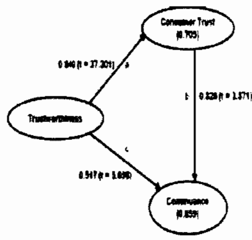
H2: Consumer trust has a significant positive impact on continuance intentions of mHealth information services.

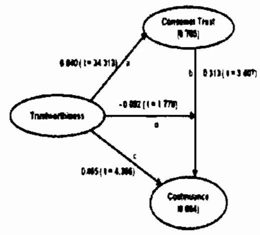
H3: Trustworthiness has a significant positive impact on continuance intentions of mHealth information services.

H4: Consumer trust mediates the relationship between trustworthiness and continuance intentions (***mediating effect***).

H5: Trustworthiness moderates the relationship between consumer trust and continuance intentions (***moderating effect***).

[redacted]





H1*: Trustworthiness ⟶ Consumer trust
H2*: Consumer Trust ⟶ Continuance
H3*: Trustworthiness ⟶ Continuance

H4** Mediating effect of consumer trust
H5***: Moderating effect (path d)

*significant at $p < 0.01$. **significant at $p < 0.05$. *** not significant.

[redacted]

- To establish the mediating effect, the indirect effect of a×b has to be significant.
- In this regard, we applied the z statistic (Sobel, 1982), which is significant at p<0.05. If the z value exceeds 1.96 (p<0.05), then we can accept H4, that is, there is an indirect effect of trustworthiness through consumer trust on continuance intentions. The z value is formally defined as follows:

The result supports the mediating effects of consumer trust (H4), which implies that it has an indirect influence on continuation intentions

$$z = \frac{a \times b}{\sqrt{b^2 \times s_a^2 + a^2 \times s_b^2 + s_a^2 \times s_b^2}}$$

$$z = \frac{0.840 \times 0.328}{\sqrt{(0.328)^2 \times (0.0239)^2 + (0.840)^2 \times (0.0950)^2 + (0.0239)^2 \times (0.0950)^2}} = 3.45 \quad > 1.96$$

- To estimate the size of the indirect effect, we used the variance accounted for (VAF) value, which represents 'the ratio of the indirect effect to the total effect'.

$$VAF = \frac{a \times b}{a \times b + c} = \frac{0.840 \times 0.328}{0.840 \times 0.328 + 0.517} = 0.348$$

Meaning: 34.8% of the total effect of trustworthiness on continuance intentions is explained by indirect effect (consumer trust).

Moderating Effect

- To test the possibility of the **moderating effect** of trustworthiness on the relationship between consumer trust and continuance intentions,
  - create an <u>interaction construct</u> **(consumer trust × trustworthiness)** to predict continuance intentions (Chin et al.; Henseler & Fassott, 2010).

- To test the moderating effect,
  - estimate the influence of predictor on criterion variable (b), the direct impact of the moderating variable on the criterion variable (c), and the influence of interaction variable on criterion variable (d).
- The significance of a moderator can be confirmed if the interaction effect (path d) is meaningful, independently of the size of the path coefficients b and c (Henseler & Fassott, 2010).
- In this example, the standardized path coefficient of −0.092 for the interaction construct (path d), which is not significant at p<0.05 (t =1.96).
- <u>Conclusion</u>: Moderation effect is not significant

Interaction Effect Size

- The effect size is calculated as follows:

$$f^2 = \frac{R_i^2 - R_m^2}{1 - R_i^2} = \frac{0.664 - 0.659}{1 - 0.664} = 0.015$$

(Here, i = *interaction model*, m = *main effect model*)

The results show that

- the size of the moderating effect is small (f^2 =0.02; Cohen 1988)
- the resulting beta changes are insignificant (β=−0.092, t =1.778).

<u>Conclusion:</u> Trustworthiness does not moderate the relationship between trust and continuance intentions, and we reject H5.

Global Fit Measure (GoF) for PLS

- GoF is defined as the geometric mean of the average communality and average R^2 (for endogenous constructs;Tenenhaus et al., 2005).

$$GoF = \sqrt{\overline{AVE} \times \overline{R}^2}$$

- The baseline values

  GoF small =0.1,

  GoF medium = 0.25

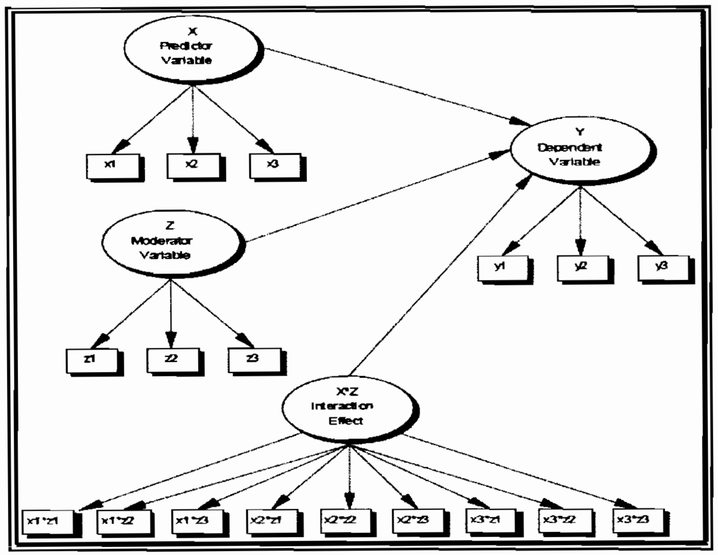
  GoF large =0.36
- GoF = 0.7803 > 0.36 , adequate support to validate the PLS model globally (Wetzels et al.).

Partial Least Squares

Interaction/Moderating Effects

Interaction/ Effects with reflective indicators

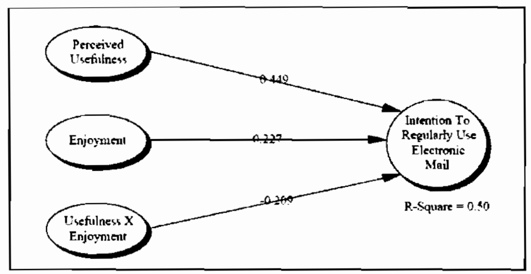
- Step 1: Standardize or center indicators for the main and moderating constructs.
- Step 2: Create all pair-wise product indicators where each indicator from the main construct is multiplied with each indicator from the moderating construct.
- Step 3: Use the new product indicators to reflect the interaction construct.



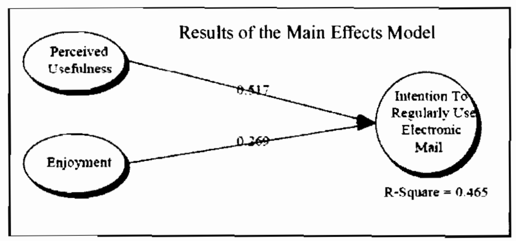
TECHNOLOGY ACCEPTANCE MODEL

Chin, Marcolin and Newsted

Results of the Interaction Effects Model



One standard deviation increase in enjoyment will not only impact intention by 0.227, but it would also decrease the impact of perceived usefulness to intention from 0.449 to 0.240.

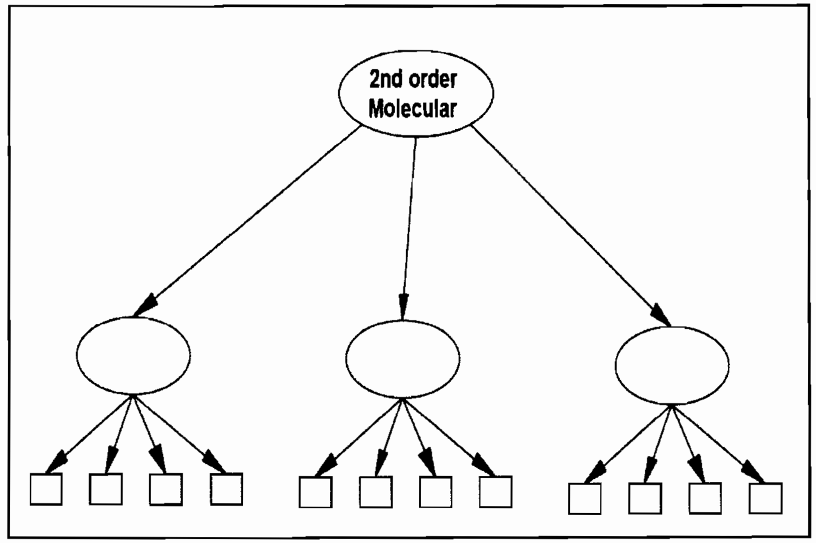


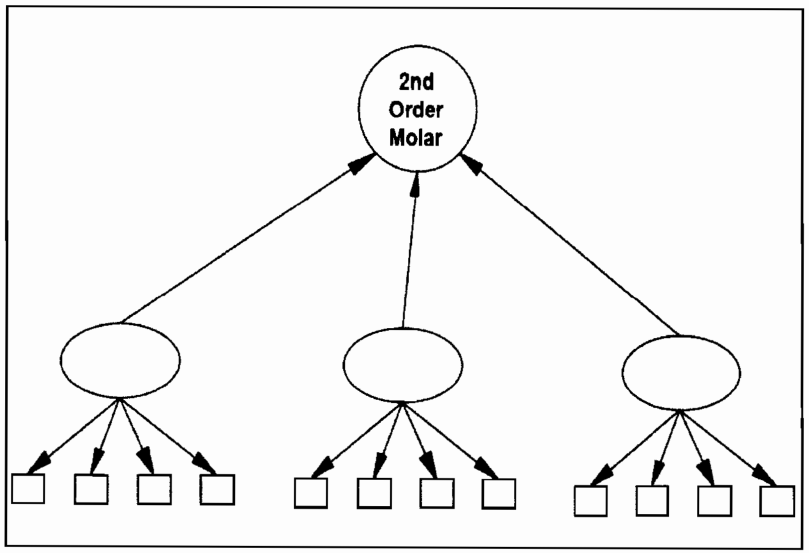
The main effects model, as expected, resulted in slightly higher standardized beta and a smaller R-square of 0.465.

Interaction Effects

Second Order Factors

- Second order factors can be approximated using various procedures.
- The method of repeated indicators known as the hierarchical component model suggested by Wold (cf. Lohmöller, 1989, pp. 130-133) is easiest to implement.
- Second order factor is directly measured by observed variables for all the first order factors that are measured with reflective indicators.
- While this approach repeats the number of manifest variables used, the model can be estimated by the standard PLS algorithm.
- This procedure works best with equal numbers of indicators for each construct.





Goodness of Fit Criteria

Goodness of Fit Criteria in PLS - Reliability

- The measurement model in PLS is assessed in terms of item loadings and reliability coefficients (composite reliability), as well as the convergent an discriminant validity.
- Individual item **loadings greater than 0.7** are considered adequate (Forne and Larcker, 1981). Interpreted like a Cronbach's alpha for internal consistency reliability estimate, a composite reliability of .70 or greater is considered acceptable (Fornell and Larcker, 1981).
- The **average variance extracted (AVE)** measures the variance captured by the indicators relative to measurement error, and it **should be greater than .50** to justify using a construct (Barclay, Thompson and Higgins, 1995).

Composite Reliability

$$\rho_c = \frac{(\sum \lambda_i)^2 \operatorname{var} F}{(\sum \lambda_i)^2 \operatorname{var} F + \sum \Theta_{ii}}$$

where λ_i, F, and Θ_{ii}, are the factor loading, factor variance, and unique/error variance respectively. If F is set at 1, then Θ_{ii} is the **1-square of λ_i**.

Average Variance Extracted

$$AVE = \frac{\sum \lambda_i^2 \operatorname{var} F}{\sum \lambda_i^2 \operatorname{var} F + \sum \Theta_{ii}}$$

where λ_i, F, and Θ_{ii}, are the factor loading, factor variance, and unique/error variance respectively. If F is set at 1, then Θ_{ii} is the **1-square of λ_i**.

Goodness of Fit Criteria in PLS - Validity

- The discriminant validity of the measures (the degree to which items differentiate among constructs or measure distinct concepts) was assessed by examining the correlations between the measures of potentially overlapping constructs.
- Items should **load more strongly on their own constructs** in the model, and the average variance shared between each construct and its measures should be greater than the variance shared between the construct and other constructs (Compeau, Higgins and Huff, 1999).

Goodness of Fit Criteria in PLS – Structural Model

- The structural model in PLS is assessed by examining the path coefficients (standardized betas).
- T statistics are also calculated to assess the significance of these path coefficients.
- R^2 is used as an indicator of the overall predictive strength of the model.

ASSESSING FACTORIAL VALIDITY IN PLS

Convergent Validity: Two Steps

1. Examine the convergent validity of the scales
2. Generate the t-values with a bootstrap. Convergent validity is shown when the t-values of the Outer Model Loadings are above 1.96.

ASSESSING FACTORIAL VALIDITY IN PLS

DISCRIMINANT VALIDITY:

1. Examine item loadings to construct correlations.

2. Examine the ratio of the square root of the AVE of each construct to the correlations of this construct to all the other constructs.

The square root of the AVE of each construct needs to be much larger than any correlation between this construct and any other construct.

Partial Least Squares

PLS Basic Reporting

TAM Framework

'An Empirical Study of the Roles of Affective Variables in User Adoption of Search Engines', Sun & Zhang (2004)

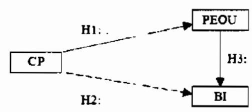
Problem Statement:
Computer playfulness (CP) may predict users' PEOU (Venkatesh, 2000). However, CP may not have a significant effect on behavioral intention. PEOU is one of the major factors that antecedes behavioral intention (e.g. Davis, 1989, Davis et al., 1989).

Therefore, we propose the following research hypotheses:

H1: Computer playfulness has a significant positive effect on perceived ease of use.

H2: Computer playfulness has a non-significant impact on behavioral intention.

H3: Perceived ease of use has a significant positive effect on behavioral intention.



The TAM Framework

Data Analysis

- Partial Least Square (version PLS-graph 03.00) was used to analyze the data. The measurement model in PLS is assessed in terms of item loadings and reliability coefficients (composite reliability), as well as the convergent and discriminant validity. Individual item loadings greater than 0.7 are considered adequate (Fornell and Larcker, 1981). Interpreted like a Cronbach's alpha for internal consistency reliability estimate, a composite reliability of .70 or greater is considered acceptable (Fornell and Larcker, 1981). The average variance extracted (AVE) measures the variance captured by the indicators relative to measurement error, and it should be greater than .50 to justify using a construct (Barclay, Thompson and Higgins, 1995). The discriminant validity of the measures (the degree to which items differentiate among constructs or measure distinct concepts) was assessed by examining the correlations between the measures of potentially overlapping constructs. Items should load more strongly on their own constructs in the model, and the average variance shared between each construct and its measures should be greater than the variance shared between the construct and other constructs (Compeau, Higgins and Huff, 1999).

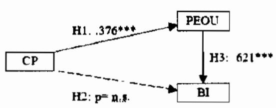
- The structural model in PLS is assessed by examining the path coefficients (standardized betas). T statistics are also calculated to assess the significance of these path coefficients. In addition, R^2 is used as an indicator of the overall predictive strength of the model.

Measurement Model

- The results show that the measures of the constructs examined are robust in terms of item loadings, their internal consistency reliability as indexed by composite reliabilities, and discriminant validity. Except for CPS1, all other item loadings were above the suggested 0.70 (Table 1). The internal reliabilities, assessed by composite reliability, were all greater than 0.70 (see Table 2). Therefore, no items were dropped.
- This allowed consistency with the measures used in prior studies. Table 2 also demonstrates satisfactory convergent and discriminant validity of the measures. Average variance extracted (AVE) for all constructs exceeded 0.50. As for the discriminant validity, Table 2 shows that all constructs were more strongly correlated with their own measures than with any of the other constructs. Therefore, discriminant validity was observed.

	CPS	PEOU	BI
CPS1	**0.67**	0.30	0.16
CPS2	**0.83**	0.22	0.06
CPS3	**0.81**	0.37	0.18
CPS4	**0.81**	0.25	0.06
CPS5	**0.75**	0.33	0.11
CPS6	**0.79**	0.22	0.09
CPS7	**0.76**	0.27	0.08
PEOU1	0.38	**0.92**	0.58
PEOU2	0.34	**0.92**	0.51
PEOU3	0.33	**0.91**	0.52
PEOU4	0.35	**0.95**	0.57
BI1	0.09	0.55	**0.95**
BI2	0.19	0.57	**0.95**

Table 1: Item Loadings



	CR	AVE	1	2	3
1. CP	.913	.601	**.775**		
2. PEOU	.960	.856	.376	**.925**	
3. BI	.946	.897	.149	.590	**.947**

CR: Composite Reliability; AVE: Average Variance Extracted.

Diagonal Elements are the square root of the variance shared between the constructs and their measurement (AVE). Off diagonal elements are the correlations among constructs. Diagonal elements should be larger than off-diagonal elements in order to obtain the discriminant validity.

Table 2: Reliability, Convergent and Discriminant Validity Coefficients

Structural Model

- The path coefficients from the PLS analysis are shown in Figure 1. Hypotheses 1, 2 and 3 were all supported. CP demonstrated a direct, statistically significant, and positive effect on PEOU (H1 p<.001). As hypothesized, CP did not have a direct impact on behavioral intention (H2 p= n.s.). Its impacts were fully mediated by users' perception of ease of use, which had a significant direct effect on BI (H3 p<.001).
- R^2 values can be used to evaluate the strength of the proposed model. In Model 1, 35.4% of variance in BI was explained by the model. In addition, 14.1% of variance in PEOU was explained by computer playfulness itself.