12-31-1996

A Partial Least Squares Latent Variable Modeling Approach for Measuring Interaction Effects: Results from a Monte Carlo Simulation Study and Voice Mail Emotion/Adoption Study

Wynne Chin
University of Calgary

Barbara Marcolin
University of Calgary

Peter Newsted
University of Calgary

A PARTIAL LEAST SQUARES LATENT VARIABLE MODELING APPROACH FOR MEASURING INTERACTION EFFECTS: RESULTS FROM A MONTE CARLO SIMULATION STUDY AND VOICE MAIL EMOTION/ADOPTION STUDY

Wynne W. Chin
Barbara L. Marcolin
Peter R. Newsted
University of Calgary

Abstract

The ability to detect and accurately estimate the strength of interaction effects are critical issues that are fundamental to social science research in general and IS research in particular. Within the IS discipline, a large percentage of research has been devoted to examining the conditions and contexts under which relationships may vary, often under the general umbrella of contingency theory (McKeen, Guimaraes, and Wetherbe 1994; Weill and Olson 1989). In our survey of such studies where such moderating variables are explored, a majority fail to either detect and/or provide an estimate of the effect size. In cases where effects sizes are estimated, the numbers are generally small. These results have, in turn, led some to question the usefulness of contingency theory and the need to detect interaction effects (e.g., Weill and Olson 1989). This paper addresses this issue by providing a new latent variable modeling approach that can give more accurate estimates of such interaction effects by accounting for the measurement error in measures which attenuates the estimated relationships. The feasibility of this approach at recovering the true effects is demonstrated in two studies: a simulated data set where the underlying true effects are known and a Voice Mail adoption data set where the emotion of enjoyment is shown to have both a substantial direct and interaction effect on adoption intention.

1. INTRODUCTION

The goal of assessing the interaction effects between quantitative (i.e., continuous) variables represents a long tradition in IS research. Both empirical and theoretical models can easily be found going back several decades (e.g., Powers and Dickson 1973; Ginzberg 1979; Franz 1979). Yet, the ability to detect and accurately estimate such effects can be difficult. Recent articles bemoan the statistical difficulties inherent in the analysis of interactions in field research and observational studies (e.g., McClelland and Judd 1993). These problems include those resulting from measurement error and the low statistical power that can result from such error. In particular, under conditions of measurement error, traditional techniques such as analysis of variance (ANOVA) or moderated multiple regression (MMR) may not be able to detect such interaction effects. When detected, again due to measurement error, the estimate of the true effect may be biased downwards.

In general terms, an interaction effect involves a moderator variable which can be qualitative (e.g., gender, race, class) or quantitative (e.g., age, income). The moderator, in turn, affects the direction and/or strength of the relation between an independent or predictor variable and a dependent or criterion variable. Thus, moderator variables provide information as to the conditions in a which we would expect a relationship between two variables to exist.

Acknowledgments: This research was supported in part by a grant from the Social Sciences and Humanities Research Council of Canada, whose assistance is gratefully acknowledged. We also wish to thank Theodora Lo for her assistance in the literature review and simulation runs.

Consider the following example. Suppose we are studying the relationship depicted in the Technology Acceptance model (Davis 1989) which states that the stronger a person's belief in the usefulness of an IT for enhancing their job performance, the more they would intend to use it. Yet we also believe that the impact of this perception on IT usage intention is *negatively* moderated by the level of enjoyment the individual has during his or her use of the IT. In essence, when the usage experience is more enjoyable, the impact of perceived usefulness (at whatever level) on future intention to use is lower. Conversely, the less enjoyable one perceives the IT to be, the stronger the impact of one's perception of usefulness on intention to use. This phenomenon is based on a cognitive consistency argument where the underlying rule is that when IT usage is extremely enjoyable, instrumental issues such as perceived usefulness ought not to come into one's decision making criteria for future usage. Thus, all else being equal, if we had two different groups of people where the first group perceived the IT to be highly enjoyable and the second group the opposite, we would expect a low to possibly negative correlation between perceived usefulness and IT adoption intention for the first group and a high correlation for the second group. In this scenario, the dependent variable (Y) would be IT usage intention with the predictor (X) and moderator (Z) variables being perceived usefulness and enjoyment respectively.

Ideally, we would like to be able to detect such an interaction effect easily (should it exist) and more importantly provide an accurate estimate of the effect size of the interaction. The traditional approach for detection has been ANOVA, which generally provides an ability to detect the existence of an interaction effect. Yet, the effect size of an interaction as measured by additional variance explained is normally not provided and must be explicitly calculated from the sum of squares information. Alternatively, the use of MMR allows one to both detect and examine two estimates of the interaction effect. The first, identical to ANOVA, is the additional variance explained. This is assessed by the increase in R-square when the interaction term (i.e., X*Z) is included in a regression model with the main effects (i.e., X and Z). This is determined by subtracting the R-square for the two variable main effects model (Y regressed on X and Z) from the R-square for the three variable interaction model (Y regressed on X, Z, and X*Z). In addition to the change in R-square, the estimated beta for the interaction term provides additional information regarding the interaction effect. This estimate informs us as to how much a unit change in the moderator variable Z would change the regression relationship of Y on X.

Yet, both approaches assume that the measures used to predict the dependent variable are without error. In general, social science data are more than likely measured with error. This bias, in turn, can have substantial impact on the conclusions that IS researchers may make. For example, if the reliability of the predictor (X) and moderator (Z) variables is 0.70 and their correlation is 0, then the reliability of the product term (X*Z) representing the interaction effect would be 0.49. The implication is that, in the best of all situations, an estimate of the proportion of the dependent variable (Y) explained by the interaction term will be half of the true effect. In other words, suppose the real interaction effect has a R-square of 0.30. Yet, due to the measurement error, a researcher using MMR would end up with an estimate of 0.15! Furthermore, in addition to providing downward biased estimates, such unreliable measures will also result in a reduction in the statistical power to detect such effects.

Therefore, this paper addresses several issues and questions. The first primary question concerns the extent to which previous IS contingency research applied analytical techniques which assume infallible measures (i.e., ANOVA and multiple regression). In turn, what proportion of these studies were actually successful in detecting and estimating the interaction effect? If the majority of the studies suggest weak or negligible effects, can a meta-analytic type of conclusion be made questioning the value of such interaction studies and contingency theory in general? Alternatively, are the past results and conclusions possibly a byproduct of the analysis technique? A second primary concern addressed in this paper is to present a new latent variable modeling approach for better estimation and detection of the interaction effect between quantitative (i.e., continuous) predictor and moderator variables. To the best of our knowledge, the use of product indicators in a Partial Least Squares (PLS) analysis to model interaction effects has never been done. The performance ability of this approach is tested on simulated data where the underlying population parameters (i.e., true effects) are known. Finally, it is further tested on actual data, where the interaction effect size detected is stronger than those typically found in IS research.

2. LITERATURE REVIEW

Moderators have been used in the IS field for some time and are couched under terms such as multiplicative terms, moderators, contingency terms (sometimes under contingency theory), and interaction effects.

In our review of information systems literature back to 1980,[1] moderators were found to be present from the start. In 1980, Schonberger considered how information systems design approaches lead to good systems but only when the contingent factors are appropriately matched (i.e., information needs for MIS supporting functions and type of decision making). Recently, in 1994, McKeen, Guimaraes and Wetherbe summarized the contingency studies which existed on the relationship between user participation and satisfaction and found that two out of four moderating variables, task and system complexity, changed the outcome. However, these two moderators only had a small effect on satisfaction (beta=.03). The authors then concluded that these moderating variables were probably not the only two important ones and suggested five other promising moderators for future investigation. From their experience, the unmoderated relationship was not adequate for an investigation into user participation and satisfaction. Throughout the years, similar studies employing moderators can be found in information systems research.

Of the 8,110 potential articles over the fifteen year period for the journals we reviewed, seventy articles contained moderator variables (see Tables 1 and 2). Seven of the articles containing moderators had actually calculated path estimates, with the remaining sixty-three articles lacking sufficient information or analysis techniques to facilitate an estimate of the moderator effect size. Thus, while moderators are being examined in the information systems field, the dearth of detailed information may be a major reason for the slow progress and lack of cumulative knowledge. For instance, only three articles had standardized coefficients to enable comparison while the other articles published unstandardized coefficients without standard deviations (from which they could be standardized), split the sample, or failed to report the effect size.

Overall, moderating variables in IS studies tend to be evaluated in a quantitative sense, typically using categorical or interval scale variables. The majority of articles applied an analysis of variance technique for variables measured with categorical scales, while a small sub-group applied a regression/path analysis based technique for variables measured with interval Likert-type scales.

Of the articles identified with moderators (n=70), ANOVA was the predominate technique used for the analyses (51.4% including ANCOVA), with MANOVAs (14.3% including MANCOVA), multiple regression (10% including path analysis), and various sundry other analytic techniques (24.3% including theoretical moderators and meta-analyses) following in popularity.

The studies using analysis of variance approaches (i.e., ANOVA, MANOVA, MANCOVA, ANCOVA), however, did not provide an estimate of the effect size between the independent and dependent variables but only an indication that the relationship is significant or not. From these analyses it is not possible to determine the effect size for the significant results. The degree to which the phenomena exists in the population remains a mystery for 65.7% of the studies reviewed (see Table 1). A total of 321 moderators were investigated in the articles with 65 assessed as significant, but the majority of these (53) did not report effect size estimates.

Moreover, the regression and path analysis techniques, which do provide beta path coefficients, had few significant terms, small effect sizes and low power (see Table 2). Twelve of twenty-two moderating variables were found to be significant in these articles. The literature consistently reported moderators with a small effect size, beta averaging 0.10, suggesting that the moderating terms play only a small part for understanding information systems issues. Such a low number of significant moderators, overall, does not bode well for the impact these lines of investigation have had on the IS field. These moderating terms do not seem to be good contributors to the body of knowledge.

Statistical power also is rarely discussed in these studies. Given that power is a factor in assessing the ability of the tests to detect a significant result, it should be a pertinent subject within these studies. Only four articles were located which had discussed, or at least referred to, power. Of these, one indicated a 0.80 power level, which is the suggested level indicating the probability that the analytical tests will yield statistically significant results if the phenomena truly exist (Cohen 1977) while maintaining a Type I error level.

---

[1]All articles in *MIS Quarterly*, *Information Systems Research*, *Management Science*, *Journal of MIS*, *Decision Sciences*, *Communications of the ACM*, *Information & Management* and proceedings of ICIS and HICSS up to 1995 and as far back as the journal goes or 1980, whichever is more recent.

Table 1. Summary of Analysis of Variance Techniques

Sample Size	Significant Moderators*	Sample Size	Significant Moderators*	Sample Size	Significant Moderators*
7	3/3	132	2/6	48	2/43
441/387	2/22	359	4/4	1459	1/4
105	0/5	120	1/3	30	1/4
313,254	0/1	141	1/8	363	1/8
n/a	2/4	36	1/3	83	0/2
n/a	2/14	288	1/3	79	1/2
114	0/1	70	1/2	18	1/2
24	0/1	90	1/3	21	1/9
287/255	0/2	599	2/10	53	1/1
40	3/3	72	0/2	173	1/9
28	0/2	1454	1/4	80	0/4
256	1/2	43	1/3	39/27	1/2
48	1/8	768	1/3	504	0/1
36	2/3	16/26/86/58	1/5	48	0/3
49	0/1	112	2/4	36	1/1
28	2/4				

46 articles average sample size is 200 including articles with time series
51 samples average sample size is 149 excluding articles with time series
53/243 moderators had significant results
*n/m means n out of m moderators were found significant. For example, 2/43 means 2 out of 43 moderators tested were significant

Low power may have been a main culprit in the large number of non-significant results found in the review. Low precision (i.e., larger standard error) in the analytic calculations can be impacted by small sample sizes, random measurement error, and the shape of the data distributions. Sample sizes averaged 81 for articles using regression or path analysis, and 149 for articles using analysis of variance (see Tables 1 and 2). Given the fact that the greater the precision, the greater the probability of detecting a significant result, one wonders whether these studies reviewed had enough power to detect an effect if it existed. It is possible, given the average sample sizes, that many unknown interaction effects were left undetected.

Thus, moderators have been used to some extent in the IS field, but have generally been difficult to detect. Few studies have used analytic techniques that actually calculate the effect size; power levels tend to be low; and meta-analyses are either theoretical (Trice and Treacy 1986; Goodhue 1986) or report inconsistent results (Cavaye 1995). The inconsistent results in the meta-analyses are often blamed on different operationalizations of the constructs, uncontrolled research factors, inappropriate quantitative studies with little richness in the understanding of the influences, and poorly validated instruments (Cavaye 1995).

Regardless of the true effect of moderators, it can be concluded from the review that better reporting would drastically improve the body of knowledge being built within the information systems field. Sample sizes, effect sizes, power of statistical tests, and the usual significance levels (p=.05 or .01) are needed to provide consistency in the results.

Table 2. Summary of Regression and Path Analysis Articles

Sample Size	Power	Average Number of Indicators	Effect Size (beta)	Number of significant moderators*
36	no	n/a	n/a	0/1
30	no	5	unstandardized	3/4
30	no	n/a	split sample	2/4
249	n/a	3	unstandardized	1/3
151	no	8	.03, .026	2/4
47	no	1	.26	1/3
77/32	.80	1	unstandardized	3/3
average				
81		3.4		12/22

*n/m means n out of m moderators were found significant. For example, 3/4 means 3 out of 4 moderators tested were significant

Given the current state of the literature concerning IS moderating variables, in the remaining sections of this paper we try to make improvements. A new latent variable modelling approach for analyzing interaction effects is first offered, and the technique is then supported with evidence from a simulated data set and an IT adoption data set.

3. PARTIAL LEAST SQUARES PRODUCT INDICATOR APPROACH FOR MEASURING INTERACTION

To account for the deleterious effects of measurement error, a product indicator approach in conjunction with Partial Least Squares (PLS) is proposed. The predictor, moderator, and dependent variables are now viewed as latent variables (i.e., constructs) which cannot be measured directly. Instead, multiple indicators (i.e., measures) for these latent variables need to be obtained. Each indicator is modeled as being influenced by both the underlying latent variable and error. Product indicators reflecting the latent interaction variables are then created. Each set of indicators reflecting their underlying construct (i.e., latent variable) are then submitted to PLS for estimation resulting in a more accurate assessment of the underlying latent variables and their relationships.

The PLS procedure has been gaining interest and use among IS researchers in recent years (Compeau and Higgins 1995; Aubert, Rivard and Paltry 1994; Chin and Gopal 1995) because of its ability to model latent constructs under conditions of non-normality and small to medium sample sizes. Being a components-based structural equations modeling technique, PLS is similar to regression, but simultaneously models the structural paths (i.e., theoretical relationships among latent variables) and measurement paths (i.e., relationships between a latent variable and its indicators). Rather than assume equal weights for all indicators of a scale, the PLS algorithm allows each indicator to vary in how much it contributes to the composite score of the latent variable. Thus indicators with weaker relationships to related indicators and the latent construct are given lower weightings. In this sense, PLS is preferable to techniques such as regression which assume error free measurement (Lohmöller 1989; Wold 1982, 1985, 1989).[2]

[2]Additional information about the PLS method and how it compares to covariance based procedures is provided in Appendix A.

Under assumptions of error free measurement, product term regression analysis is generally recommended for continuous variables to detect interaction effects (Jaccard, Turrisi, and Wan 1990; Cohen and Cohen 1983). Yet, for variables that are measured with error (i.e., latent variables), regression produces biased and inconsistent coefficient estimates along with a loss of statistical power as the reliability of the measures decline (Busemeyer and Jones 1983; Aiken and West 1991). For example, Aiken and West (p. 164) show that while a sample of about 400 is needed to achieve a power of 0.80 to detect a small effect ($f^2 = 0.02$)[3] with perfect measures, sample requirements increase to 1,056 if the reliability of the measure is 0.80. In general, the power to detect an interaction effect at a given sample size is reduced by up to one half when reliabilities are 0.80 rather than 1.00 and up to two thirds when reliabilities decrease to 0.70. In addition, the estimates of the interaction effect sizes can also be significantly reduced as reliability decreases (Aiken and West, p. 161).

In addressing these problems, we apply a product indicator approach in conjunction with the PLS procedure to estimate the underlying interaction construct. Under the traditional regression based approach for modeling interaction effects, a single product term is used (x*z) to examine the influence that a moderator z would have on the relationship between the predictor x and the dependent variable of interest y. Yet, as discussed earlier, this approach is made under the assumption that each measure is error free. To account for the condition of measurement error, we can view the independent variables x, z, and x*z as indicators which reflect the true underlying continuous latent constructs of interest. The estimation of this true underlying construct, in turn, entails the use of multiple/parallel indicators.

First, multiple indicators of the underlying constructs (i.e., reflecting X, Z, and Y), need to be obtained. The predictor (X) and moderator (Z) indicators are then used to create product indicators. To do this, we need to determine whether the individual indicators for X and Z are theoretically parallel measures or if some are more important based on their measurement scale and/or variance. For the first scenario, we consider the indicators for each construct to be approximately equivalent and do not give any more credence to any one particular indicator in a set. Under this condition, we would standardize all indicators reflecting the predictor and moderator constructs to a mean of zero and variance of one (Jaccard, Turrisi, and Wan 1990; Aiken and West 1991). This approach would be reasonable for ordinal to interval level items such as Likert scaled attitude items. Alternatively, one would center the indicators to a mean of zero if it is felt the original metric of the items or the variances should be maintained.[4] For ratio level items, it is important to have all items transformed to the same metric in addition to being centered. Otherwise, the estimated latent variable score produced by PLS would be indeterminable.

Standardizing or centering the indicators helps avoid computational errors by lowering the correlation between the product indicators and their individual components (Smith and Sasaki 1979). Furthermore, without such a process, each product term would likely have a different location parameter leading to an inability for the PLS procedure to accurately estimate the underlying interaction construct. It also allows us an easier interpretation of the resulting regression beta for the predictor variable; that being the effect we would expect at the average value of the moderator variable (which is centered to zero). Standardizing is accomplished by calculating the mean and s.d. for each indicator. Then for each indicator score, the corresponding mean is subtracted and divided by the respective s.d. Centering refers only to subtracting the mean.

Using the standardized or centered indicators of the predictor variable (i.e., X) and the moderator variable (i.e., Z), product indicators are then developed by creating all possible products from the two sets of indicators. These product indicators are used to reflect the latent interaction variable. For example, if we have three measures reflecting the main predictor (X) and three measures for the moderator variable (Z), we would have nine measures for representing the interaction term (X*Z). Graphically, this is depicted in Figure 1. Since any indicator reflecting the predictor (X) or moderator (Z) is viewed as interchangeable with another from the same set, any product indicator (x_i*z_i) would represent an equivalent representation of the underlying latent interaction variable (X*Z).

The PLS procedure is then used to estimate the latent variables as an exact linear combination of its indicators with the goal of maximizing the explained variance for the indicators and latent variables. Following a series of ordinary least squares

---

[3] The effect size f^2 = [R^2 (interaction model) - R^2 (main effects)] / R^2 (interaction model).

[4] The results presented in this paper used standardized indicators. Similar results were obtained when we centered them.

analyses, PLS optimally weights the indicators such that a resulting latent variable estimate can be obtained.[5] The weights provide an exact linear combination of the indicators for forming the latent variable score which is not only maximally correlated with its own set of indicators (as in components analysis), but also correlated with other latent variables according to the structural (i.e. theoretical) model.

It should be noted that the underlying assumption of uncorrelated error terms among the indicators is not true (see Kenny and Judd [1984] for derivation). The error terms for the product indicators are partially correlated with the error terms associated with the indicators of the other exogenous constructs. Thus, the estimation procedure using PLS might be affected. The extent to which this is problematic will be assessed in the next section.

4. STUDY 1: MONTE CARLO SIMULATION

To test the efficacy of the Partial Least Squares– product indicator approach for detecting and estimating interaction effects where measurement error exists, we ran a Monte Carlo simulation study varying the sample size and number of indicators. Data were created using PRELIS 2.14 (Jöreskog and Sorbom 1993), to conform to an underlying population model where the standardized beta from X to Y is 0.3, the beta of Z on Y is 0.5, and the interaction effect is 0.3. Indicators for all primary constructs (i.e., latent variable) were modeled to have factor loadings of 0.70. Thus, we know what are to be the main and interaction effects. The goal here is to determine how good the PLS-product indicator approach is at detecting and recovering (i.e., estimating) the true effects under conditions of measurement error.

The two treatments considered were sample size and number of indicators per construct (i.e., latent variable). Sample sizes consisted of 20, 50, 100, 150, 200, and 500. Number of indicators per primary construct (i.e., X, Z, and Y) varied as follows: 1, 2, 4, 6, 8, 10, and 12. For each treatment cell, one hundred simulations were performed. Thus, for example, in the cell representing sample size 50 and three indicators per construct, we generated 100 data sets consisting of sample sizes of 50 where each case had three indicators for each of the three constructs X, Z, and Y respectively and nine product indicators for X*Z.

Table 3 provides the path and standard error results of our 6 by 7 study where PLS runs were made using PLS-Graph version 2.9 (Chin 1994). Our basis for how well the PLS-product indicator approach performed is multiple regression. The results of using multiple regression over varying sample sizes is provided in the column labeled single item per construct. When using PLS, this limiting case of one indicator per construct is identical to performing a multiple regression.

Keeping in mind that a perfect estimation procedure should result in 0.30 (x to y), 0.50 (z to y), and 0.30 (x*z to y), the results using multiple regression under conditions of measurement error consistently underestimated the true effects. The reason, as discussed earlier, is that regression does not explicitly take into account the attenuating effects of measurement error.[6] At a sample size of 500, for example, the estimation are at a half to one quarter of the real effects. Larger sample size, while providing smaller standard errors, did not result in closer approximations to the true effects.

The results do show that for any sample size, as the number of indicators increased, the estimation tended to move closer to the true effect. For example, at sample size 100, the estimation of the interaction effect approached the true effect of 0.30 at eight or more indicators. The estimation for the moderator z, while never reaching the true effect, increased until it reached an asymptotic limit at eight indicators.

---

[5]For standardized indicators, the default estimation procedure of unit variance/no location should yield similar results to the original scale/with locations option. For centered data, the original scale/with locations should be used.

[6]We should noted that what we conclude regarding effects size estimations using regression throughout this paper is equally or more so for ANOVA since the latter represents a special case of the former.

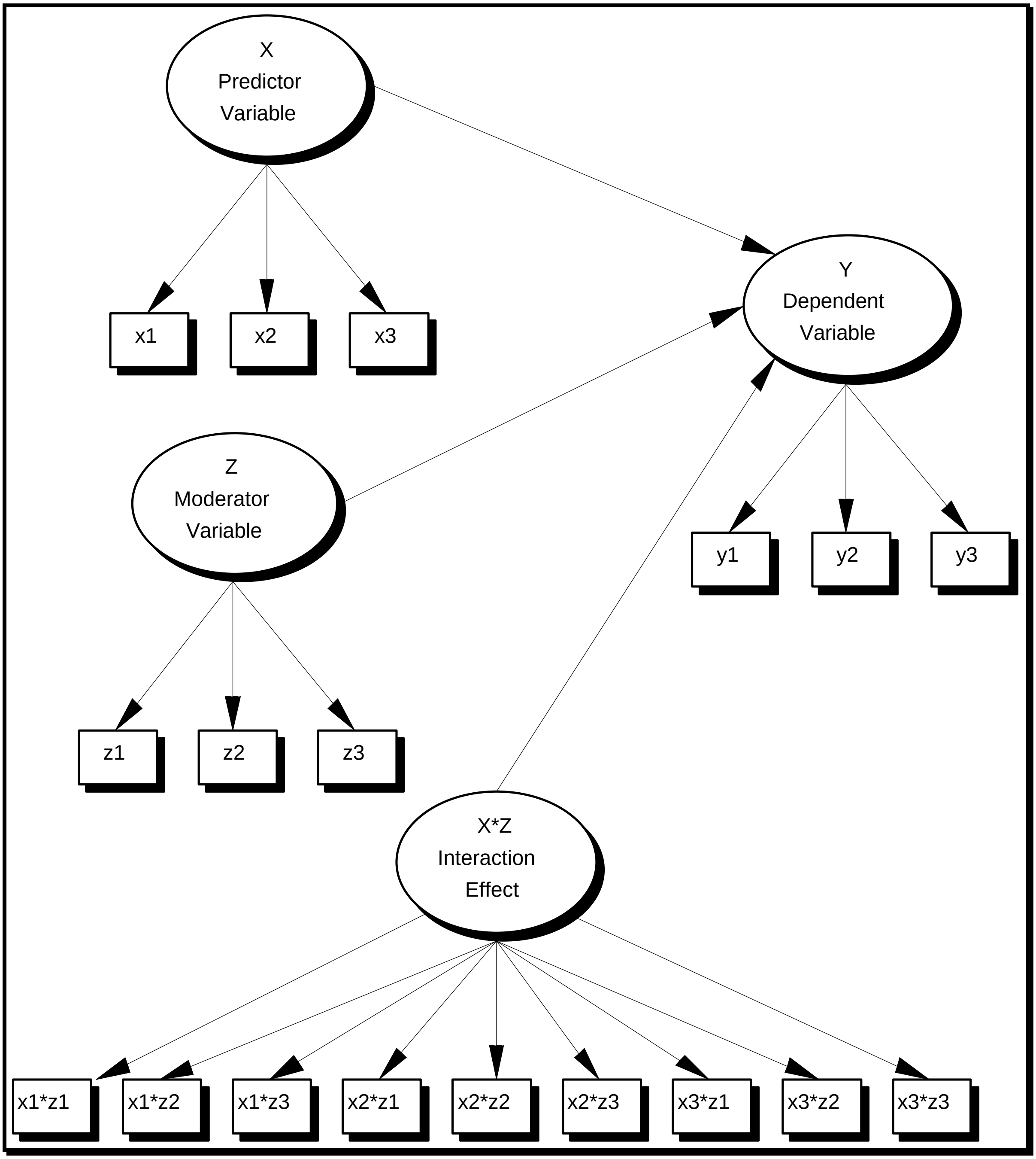

Figure 1. Model with Three Indicators per Main Constructs and Nine Product Indicators for the Interaction Construct

Table 3. Results of Monte Carlo Simulation (100 Runs per Cell)

	Indicators per construct						
Sample size	single item per construct	two per construct (4 for interaction)	four per construct (16 for interaction)	six per construct (36 for interaction)	eight per construct (64 for interaction)	ten per construct (100 for interaction)	twelve per construct (144 for interaction)
20	x -->y 0.1449 (0.2462) z --> y 0.2898 (0.2262) x*z--> y 0.1458 (0.2852)	x -->y 0.2177 (0.2340) z --> y 0.3113 (0.2296) x*z--> y 0.1609 (0.3358)	x -->y 0.2070 (0.2430) z --> y 0.3307 (0.2136) x*z--> y 0.2708 (0.3601)	x -->y 0.2352 (0.2160) z --> y 0.3168 (0.2037) x*z--> y 0.1897 (0.4169)	x -->y 0.2237 (0.2150) z --> y 0.3204 (0.2286) x*z--> y 0.1988 (0.4399)	x -->y 0.2391 (0.2332) z --> y 0.3176 (0.2232) x*z--> y 0.2788 (0.3886)	x -->y 0.2450 (0.2083) z --> y 0.3622 (0.1665) x*z--> y 0.3557 (0.3725)
50	x -->y 0.1008 (0.1442) z --> y 0.2409 (0.1276) x*z--> y 0.1133 (0.1604)	x -->y 0.2155 (0.1506) z --> y 0.3120 (0.1320) x*z--> y 0.1142 (0.2124)	x -->y 0.2279 (0.1434) z --> y 0.3914 (0.1013) x*z--> y 0.2795 (0.1873)	x -->y 0.2521 (0.1294) z --> y 0.3751 (0.1106) x*z--> y 0.2403 (0.2795)	x -->y 0.2593 (0.1523) z --> y 0.3941 (0.1011) x*z--> y 0.3066 (0.2183)	x -->y 0.2548 (0.1341) z --> y 0.3885 (0.1089) x*z--> y 0.3083 (0.2707)	x -->y 0.2694 (0.1090) z --> y 0.4160 (0.0999) x*z--> y 0.3615 (0.1848)
100	x -->y 0.1461 (0.0925) z --> y 0.2453 (0.0868) x*z--> y 0.1012 (0.0989)	x -->y 0.2129 (0.0828) z --> y 0.3070 (0.0821) x*z--> y 0.1614 (0.1276)	x -->y 0.2605 (0.0836) z --> y 0.3747 (0.0727) x*z--> y 0.2472 (0.1270)	x -->y 0.2674 (0.0901) z --> y 0.4090 (0.0870) x*z--> y 0.2669 (0.1301)	x -->y 0.2716 (0.0859) z --> y 0.4354 (0.0638) x*z--> y 0.3029 (0.0916)	x -->y 0.2716 (0.0789) z --> y 0.4354 (0.0724) x*z--> y 0.3029 (0.0805)	x -->y 0.2805 (0.0776) z --> y 0.4308 (0.0724) x*z--> y 0.3008 (0.1352)
150	x -->y 0.1459 (0.0779) z --> y 0.2474 (0.0713) x*z--> y 0.0953 (0.0843)	x -->y 0.2078 (0.0627) z --> y 0.3251 (0.0768) x*z--> y 0.1695 (0.0844)	x -->y 0.2478 (0.0623) z --> y 0.4007 (0.0540) x*z--> y 0.2427 (0.0778)	x -->y 0.2609 (0.0661) z --> y 0.4156 (0.0583) x*z--> y 0.2834 (0.0757)	x -->y 0.2704 (0.0600) z --> y 0.4424 (0.0598) x*z--> y 0.2805 (0.0916)	x -->y 0.2711 (0.0623) z --> y 0.4466 (0.0587) x*z--> y 0.3040 (0.0567)	x -->y 0.2862 (0.0563) z --> y 0.4530 (0.0514) x*z--> y 0.2921 (0.0840)
200	x -->y 0.1625 (0.0601) z --> y 0.2458 (0.0644) x*z--> y 0.0962 (0.0785)	x -->y 0.1919 (0.0670) z --> y 0.3289 (0.0611) x*z--> y 0.1769 (0.0674)	x -->y 0.2327 (0.0596) z --> y 0.3985 (0.0550) x*z--> y 0.2317 (0.0543)	x -->y 0.2559 (0.0509) z --> y 0.4184 (0.0550) x*z--> y 0.2730 (0.0528)	x -->y 0.2688 (0.0460) z --> y 0.4368 (0.0541) x*z--> y 0.2839 (0.0606)	x -->y 0.2687 (0.0542) z --> y 0.4533 (0.0527) x*z--> y 0.2843 (0.0573)	x -->y 0.2814 (0.0506) z --> y 0.4601 (0.0509) x*z--> y 0.3018 (0.0542)
500	x -->y 0.1425 (0.0450) z --> y 0.2421 (0.0463) x*z--> y 0.0965 (0.0436)	x -->y 0.1926 (0.0388) z --> y 0.3339 (0.0364) x*z--> y 0.1681 (0.0358)	x -->y 0.2426 (0.0385) z --> y 0.3960 (0.0353) x*z--> y 0.2275 (0.0419)	x -->y 0.2595 (0.0373) z --> y 0.4221 (0.0349) x*z--> y 0.2448 (0.0379)	x -->y 0.2731 (0.0350) z --> y 0.4396 (0.0373) x*z--> y 0.2637 (0.0377)	x -->y 0.2778 (0.0359) z --> y 0.4553 (0.0335) x*z--> y 0.2659 (0.0353)	x -->y 0.2773 (0.0381) z --> y 0.4574 (0.0316) x*z--> y 0.2761 (0.0375)

x-->y refers to the mean of 100 path estimates from predictor x to criterion y (standard errors within parenthesis).
z-->y refers to the mean of 100 path estimates from the moderator variable z to criterion y (standard errors within parenthesis).
x*z-->y refers to the mean of the 100 path estimates for the interaction effect of z on the path from x to y (standard errors within parenthesis)

Sample size, in general, did not influence the consistency of the estimation. Estimates tended to be quite similar for any column of Table 3. What sample size does influence instead is the ability to detect an effect which is reflected in the variation of the estimates.

Table 4. Tests of Significance for the Mean Estimates (100 Runs)

	Number of Inidcators						
Predictor: x							
	1	2	4	6	8	10	12
Sample Size							
20							
50				*	*	*	**
100		**	**	**	**	**	**
150	0	**	**	**	**	**	**
200	**	**	**	**	**	**	**
500	**	**	**	**	**	**	**
Moderator: z							
	1	2	4	6	8	10	12
Sample Size							
20							*
50	0	**	**	**	**	**	**
100	**	**	**	**	**	**	**
150	**	**	**	**	**	**	**
200	**	**	**	**	**	**	**
500	**	**	**	**	**	**	**
Interaction: x*z							
	1	2	4	6	8	10	12
Sample Size							
20							
50							*
100			*	*	**	**	*
150		*	**	**	**	**	**
200		**	**	**	**	**	**
500	0	**	**	**	**	**	**

* $p < .05$ (One-tailed t value: 1.66, df = 99)
** $p < .01$ (One-tailed t value: 2.36, df = 99)

In the case of multiple regression, for example, the ability to detect a beta effect of 0.30 (predictor X situation) did not occur until sample size 150 (see Table 4). But if we up the beta to a 0.50 effect, as in the case for moderator Z, significance occurred at a size of 50. If the effect size remained at 0.30, but the reliability drops in half (as in the case for the interaction effect where its reliability is the product of reliabilities of X and Z) only the sample of size of 500 was large enough. Until then, the standard errors relative to the estimated effect sizes were too large.

Table 4 also shows how increasing indicators will increase power. This is clearest for the interaction effect where a significant effect at the 0.01 level was obtained either at a larger sample size of 150 with four to six indicators or at a smaller size of 100 with eight indicators. For the more reliable predictor x (with similar effect size), significant detection at the 0.01 level was obtained for two indicators at a larger sample size 100 or twelve indicators at a smaller size of 50. When the effect size is larger (i.e., moderator z), a sample size of 100 is needed for detection using multiple regression or a smaller size of 50 with two or more indicators using PLS.

This simulation, therefore, partially demonstrates what Wold (1985, 1989) termed consistency at large for PLS estimation. This means that PLS estimation tends toward the true population parameter as the number of indicators and sample size increase.

We also see how measurement error, in the case of multiple regression, can attenuate the true effect and leave the IS researcher with the impression that an interaction effect is much smaller than it really is or possibly non-existent. The interaction estimate, in particular, would generally represent the worst among the estimations since its reliability is the product of the reliabilities of its predictor and moderator indicators.

Table 5 represents yet another side of the consistency at large influence in PLS estimation. While the loadings should ideally be estimated at the true value of 0.70 for main effects constructs and 0.49 for the interaction construct, it is a known fact that PLS tends to underestimate the structural paths connecting constructs to one another and overestimates the measurement paths connecting constructs to their indicators (Chin 1995). If we ignore the loadings for those situations in Table 4 that yielded non-significant estimations, we can see that the loadings tend to be inflated by more than ten percent in the two and four indicator situation. And as conjectured by Chin (1995, p. 319), it is not until we use ten to twelve indicators that an asymptotic limit is reached.

Other effect sizes and loadings should be tested in the future to get an even greater appraisal of how PLS estimates can vary. In the scenario we examined, it would suggest that a minimum sample size of 100 is needed in order to detect the interaction effect and six to eight indicators per construct to obtain structural path estimates within 10 percent of the true effects. This conclusion, in marked contrast to our earlier literature review that indicated the average sample size was 81 and number of indicators at 3.4, may provide a possible reason for the poor results of past IS studies.

5. STUDY 2: THE MODERATING EFFECT OF ENJOYMENT ON THE PERCEIVED USEFULNESS/IT ADOPTION INTENTION RELATIONSHIP

In this section, the PLS-product indicator approach is applied to detecting the interaction effect of enjoyment on the perceived usefulness/IT adoption intention relationship. Recall at the beginning of this paper, we had used the Technology Acceptance Model (Davis 1989) as the context for an example of an interaction effect. The same context will now be used as a test case. In the original model, perceived usefulness and perceived ease of use was modeled as having direct effects on adoption intention.

Davis, Bagozzi, and Warshaw (1992) note the difference between extrinsic and intrinsic sources of motivation to use computers in the workplace. They found that perceived usefulness, an extrinsic source of motivation, had a large significant effect on adoption intention as expected. However, its influence was complemented by enjoyment, an intrinsic source of motivation defined as the extent to which the activity of using the computer is perceived to be enjoyable in its own right, apart from any performance consequences. They also found that both variables mediated the influence of perceived ease of use on intention. Thus, perceived ease of use was not included in predicting intention for this analysis.

Table 5. Average Loadings for Each Construct[7]

sample size	X	Z	X*Z
two indicators			
20	0.758	0.809	0.570
50	0.793	0.825	0.617
100	0.827	0.848	0.635
150	0.841	0.859	0.656
200	0.853	0.862	0.703
500	0.857	0.861	0.729
four indicators			
20	0.672	0.715	0.416
50	0.729	0.764	0.536
100	0.767	0.782	0.564
150	0.773	0.782	0.581
200	0.781	0.781	0.585
500	0.783	0.785	0.612
six indicators			
20	0.698	0.726	0.390
50	0.716	0.744	0.466
100	0.743	0.757	0.512
150	0.752	0.757	0.547
200	0.754	0.756	0.555
500	0.756	0.756	0.571
eight indicators			
20	0.672	0.685	0.415
50	0.690	0.735	0.475
100	0.736	0.739	0.510
150	0.741	0.744	0.538
200	0.738	0.740	0.537
500	0.745	0.745	0.554
ten indicators			
20	0.656	0.704	0.402
50	0.699	0.726	0.456
100	0.714	0.737	0.513
150	0.725	0.731	0.510
200	0.731	0.734	0.524
500	0.735	0.736	0.537
twelve indicators			
20	0.667	0.693	0.395
50	0.694	0.723	0.473
100	0.703	0.726	0.484
150	0.728	0.728	0.520
200	0.730	0.732	0.535
500	0.728	0.730	0.530

[7]Average refers to taking the mean of the loadings (100 runs) for each indicator connected to a construct and then calculating the means. The number of indicators for the interaction construct is the square of the number of indicators for the predictor and moderate constructs.

Chin and Gopal similarly found both enjoyment and relative advantage (identical items to perceived usefulness) had a direct effect on Group Support System adoption intentions. In their article, they state "that there is also the possibility of interaction effects among the constructs that were not taken into account in this study. For example, Davis, Bagozzi and Warshaw indicate that a positive interaction may exist between enjoyment and usefulness. Because of its similarity to usefulness, the relative advantage construct used in this study may also have an interaction effect with enjoyment" (p. 58).

In order to test the possibility of such an interaction effect, the perceived usefulness and enjoyment items were used to examine the adoption intention of electronic mail (see Appendix B for construct definition and items used). The data were obtained from a single organization that had recently installed electronic mail.[8] A total of sixty questions relating to a recent introduction of electronic mail were presented. Of the 575 questionnaires distributed, 250 usable responses were analyzed representing 43.5 percent of the those surveyed. On average, the respondents had been using electronic mail for nine months, sent 2.53 messages per day (s.d. = 2.36) and received 4.79 messages per day (s.d. = 3.49). Respondents were on average 39 years old (s.d. = 9.28) and had worked for the company an average of eleven years (s.d. = 6.9). Of the respondents, 60 percent were were male. The respondents came from various levels in the organization, 13 percent were managers, 12 percent were engineers, 38 percent were technicians, and the remaining 37 percent were clerical workers.

The results (see Figure 2) give a standardized beta of 0.449 from usefulness to intention, 0.227 from enjoyment to intention, and an interaction effect of -0.209 with a total R-square of 0.50. Thus, these results imply that 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 (see Figure 3), as expected, resulted in slightly higher standardized beta and a smaller R-square of 0.465. The interaction effect, therefore, has an effect size *f* of 0.07 which is between a small and medium effect (Cohen and Cohen 1983) and represents about twice the average for past IS studies. More importantly, the beta estimates suggest the conditions under which enjoyment becomes a dominant factor equaling and potentially overshadowing perceived usefulness, that being for the group of people who perceive electronic mail to be extremely enjoyable.

To assess whether the interaction effect and main effects were significant, a bootstrap resampling procedure (Efron and Tibshirani 1993) was performed. The results of 200 resamples indicate that all paths, weights, and loadings (see Table 6) were significant at the 0.01 level.

We next need to assess how accurate the path estimates are to the "true" effect. As noted earlier, the estimates of the structural paths tend to be more accurate as the reliability of the estimated construct score increases. To assess the reliability of the latent variable estimated by PLS, the composite reliabilities as suggested by Werts, Linn, and Jöreskog (1974) were calculated and presented in Table 6. While Cronbach's alpha with its assumption of parallel measures represents a lower bound estimate of internal consistency, a better estimate can be gained using the composite reliability formula. The composite reliability ρ_c is defined as:

$$\rho_c = \frac{(\sum \lambda_i)^2 \mathrm{varF}}{(\sum \lambda_i)^2 \mathrm{varF} + \sum \Theta_{ii}}$$

where λ_i, F, and Θ_{ii}, are the factor loading, factor variance, and unique/error variance respectively. For our PLS analysis using unit variance scaling procedures, λ_i is given in Table 6, F is set at 1, and Θ_{ii} is 1 minus the square of λ_i. Using this formula, which does not assume a tau equivalency among the measures, typically provides more accurate estimates of the composite reliability. Overall, except for enjoyment with a three indicator composite reliability of 0.85, the composite reliabilities of the other constructs are very high (at or above 0.96). In the case of enjoyment, if we take the composite reliability as a bias correction factor, the 0.227 path between enjoyment and intention would increase slightly to 0.246.

---

[8]The authors wish to acknowledge Stephen Hayne for establishing contact with the organization and assisting in the delivery and receipt of the surveys.

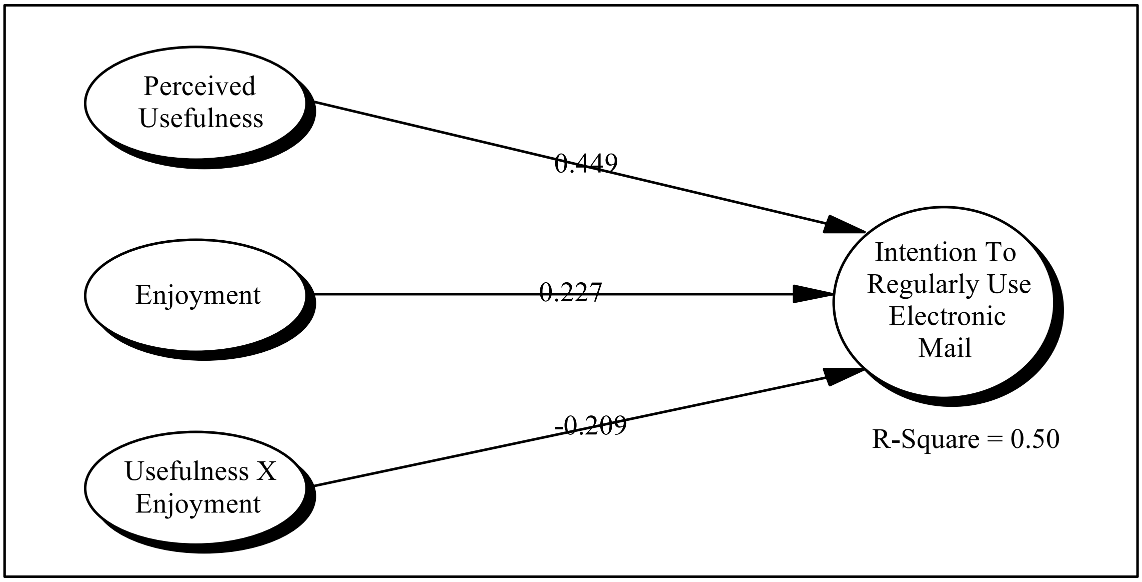

Figure 2. Results of the Interaction Model

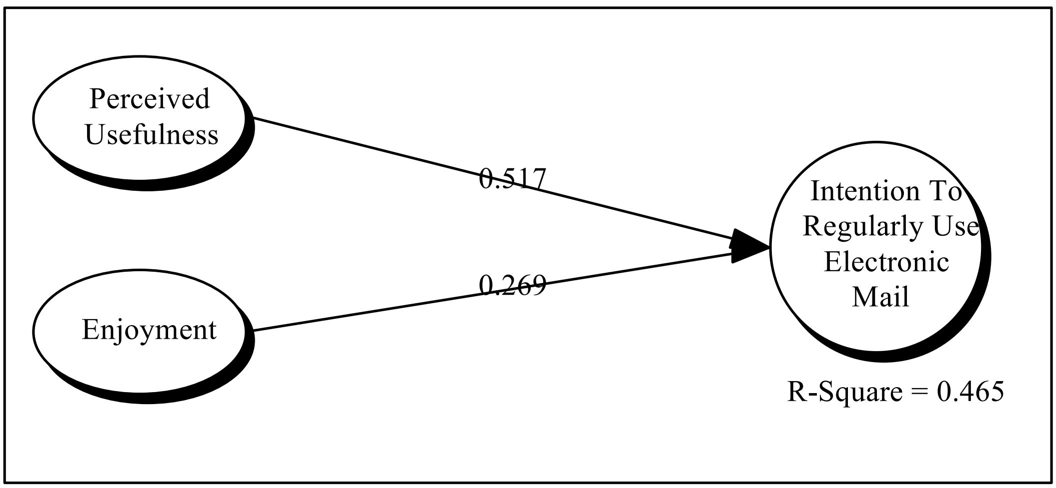

Figure 3. Results of the Main Effects Model

6. DISCUSSION AND CONCLUSION

This paper has provided a new approach toward the assessment of interaction effects among continuous variables. As Cronbach (1987, p. 417) has urged, "further investigation of statistical power in studies of interaction and invention of more sensitive research strategies are much to be desired." Using simulated data, we were able to provide an initial sense as to the efficiency and effectiveness of this approach. In general, assuming the true average loading is 0.70, sample sizes of approximately 100 are needed in order to detect the interaction effect with six to eight indicators per main effects constructs to yield reasonably consistent estimates. Under smaller sample sizes or number of indicators, the known bias in PLS for overestimating the measurement loading and underestimating the structural paths among constructs may occur. Yet, it is possible that with more reliable measures (e.g., true loadings of 0.80), the need for six to eight indicators per construct may be relaxed as well as the sample size.

Table 6. Weights and Loadings Results from PLS Analysis (Interaction Model) for Study 2

ITEM	WEIGHT	LOADING
Perceived Usefulness composite reliability = 0.975		
UFL1	0.175	0.890
UFL2	0.166	0.934
UFL3	0.179	0.951
UFL4	0.183	0.958
UFL5	0.177	0.944
UFL6	0.196	0.904
Enjoyment composite reliability = 0.853		
ENJ1	0.463	0.907
ENJ2	0.488	0.867
ENJ3	0.244	0.645
Intention composite reliability = 0.964		
INT1	0.329	0.945
INT2	0.362	0.949
INT3	0.364	0.949
Interaction composite reliability = 0.980		
UFL1ENJ1	0.079	0.890
UFL1ENJ2	0.070	0.848
UFL1ENJ3	0.063	0.831
UFL2ENJ1	0.068	0.878
UFL2ENJ2	0.065	0.828
UFL2ENJ3	0.050	0.802
UFL3ENJ1	0.070	0.906
UFL3ENJ2	0.064	0.842
UFL3ENJ3	0.050	0.813
UFL4ENJ1	0.071	0.903
UFL4ENJ2	0.069	0.882
UFL4ENJ3	0.050	0.836
UFL5ENJ1	0.072	0.908
UFL5ENJ2	0.067	0.873
UFL5ENJ3	0.052	0.829
UFL6ENJ1	0.079	0.878
UFL6ENJ2	0.074	0.843
UFL6ENJ3	0.053	0.789

Overall, more simulation studies need to be done under varying conditions to assess how PLS in general and the PLS–product indicator approach in particular performs. While the ability to detect and obtain consistent estimates will likely occur sooner with more reliable measures, it is not clear how the procedure will work for a nonhomogeneous set of indicators where the loadings vary as well as the degree of normality. Based on our understanding of the PLS algorithm, we would expect it to perform better than a strategy of simply summing the individual indicators to create a composite score and using multiple regression.

While the sample size suggested here may seem quite large, it is relatively small when contrasted to the sample size necessary for covariance based such as LISREL or EQS approached. As an extreme, consider the case of twelve indicators per construct. This represents, as a first approximation, 364 parameters to be estimated (180 factor loadings, 180 error terms, three structural paths, and one endogenous error term). While a power analysis or simulation study provides us with a more accurate means for estimating the necessary sample size, Bentler and Chou (1988, p. 172) gave a heuristic of five cases per parameter estimated under normal or elliptical conditions. Using this rule, we would need a sample size of 1,820. Even for the simpler two parameter per construct case, you would still need a sample size of 120. For non-normal conditions, Bentler and Chou recommend a minimum ratio of ten cases per parameter estimated.

Using an actual data set, we demonstrated the interaction effect between enjoyment and perceived usefulness on adoption intention. With the simulation study as a guide, it is likely that the structural effects are still being underestimated and the construct loadings are being overestimated since the number of indicators for the two constructs were three and six respectively. But in this case, using a 10 percent inflation rule, the true loadings are likely higher since the estimates are in the 0.90 region. The extent to which the negative moderating effect of enjoyment holds up in future studies is debatable, but the results do suggest the need for further research regarding the role of emotions in general and enjoyment in particular for predicting IT intentions and usage behaviors.

The approach presented here of creating nonlinear product indicators for PLS estimation can be extended to the other powers (e.g., X^2, X^3, or X^2*Z) as long as the indicators for the predictor and moderator constructs are viewed as reflective measures. Further, the indicators need not be interval level. We believe that ordinal or even dichotomous variables can be used as long as there is a large enough set to represent the underlying continuous latent variable.

Ideally, additional power constructs for the main components (i.e., X^2 and Z^2) should be included with the interaction construct in a stepwise regression fashion. Assuming equivalent composite reliabilities, having the product construct enter first relative to the other constructs would be indicative of an interaction effect. Without such a process, we cannot conclusively attribute the effect to the interaction term. Due to page limitations, we did not perform this with the Study 2 data since the goal of this paper was to demonstrate our new approach.

The current procedure cannot be used for formative indicators which are viewed as measures that create a latent construct (see Chin and Gopal for a good discussion regarding the formative/reflective distinction). One approach for formative indicators would be to follow a two step construct score procedure. The first step would entail using the formative indicators in conjunction with PLS to create underlying construct scores for the predictor and moderator variables. Step two would then consist of taking those single composite scores to create a single interaction term.

In summary, we hope this new approach will yield stronger results for future researchers interested in assessing interaction effects. Furthermore, it is hoped that issues such as appropriate sample size, multiple indicators, and power will be part of the standard information provided in future research, which should help the IS field build a more cumulative knowledge base than what we found in our literature review.

7. REFERENCES

Anderson, J. C., and Gerbing, D. W. "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach." *Psychological Bulletin*, Volume 103, 1988, pp. 411-423.

Aiken, L. S., and West, S. G. *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, California: Sage Publications, 1991.

Aubert, B. A.; Rivard, S.; and Patry, M. "Development of Measures to Assess Dimension of IS Operation Transactions." In J. I. DeGross, S. L. Huff, and M. C. Munro, Editors, *Proceedings of the Fifteenth International Conference on Information Systems*, Vancouver, British Columbia, 1994, pp. 13-26.

Bentler, P. M., and Chou, C.-P. "Practical Issues in Structural Modeling." In J. S. Long, Editor, *Common Problems/Proper Solutions: Avoiding Error in Quantitative Research*. Newbury Park: Sage Publications, 1988, pp. 161-192.

Busemeyer, J. R., and Jones, L. E. "Analysis of Multiplicative Combination Rules When the Causal Variables are Measure With Error." *Psychological Bulletin*, Volume 93, 1983, pp. 549-562.

Cavaye, A. L. M. "User Participation in System Development Revisited." *Information & Management*, Volume 29, 1995, pp. 311-323.

Chin, W. W. PLS-Graph Manual, unpublished, University of Calgary, 1994.

Chin, W. W. "PLS is to LISREL as Principal Components Analysis is to Common Factor Analysis." *Technology Studies*, Volume 2, 1995, pp. 315-319.

Chin, W. W., and Gopal, A. "Adoption Intention in GSS: Importance of Beliefs." *Data Base Advances*, Volume 26, 1995, pp. 42-64.

Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, Second Edition. New York: Academic Press, 1977, Chapter 1, "The Concept of Power Analysis."

Cohen, J., and Cohen, P. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, SEcond Edition. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1983.

Compeau, D. R., and Higgins, C. A. "Application of Social Cognitive Theory to Training for Computer Skills." *Information Systems Research*, Volume 6, 1995, pp. 118-143.

Cronbach, L. J. "Statistical Tests for Moderator Variables: Flaws in Analyses Recently Proposed." *Psychological Bulletin*, Volume 102, 1987, pp. 414-417.

Davis, F. D. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology." *MIS Quarterly,* Volume 13, 1989, pp. 319-340.

Davis, F. D.; Bagozzi, R. P.; and Warshaw, P. R. "Extrinsic and Intrinsic Motivation to Use Computers in the Workplace." *Journal of Applied Social Psychology*, Volume 22, 1992, pp. 1111-1132.

Efron, B., and Gong, G. "A Leisurely Look at the Bootstrap, the Jacknife, and Cross-Validation." *The American Statistician*, Volume 37, 1983, pp. 36-48.

Efron, B., and Tibshirani, R. J. *An Introduction to the Bootstrap* (Monographs on Statistics and Applied Probability, # 57). New York: Chapman and Hall, 1993.

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

Fornell, C.; Lorange, P.; and Roos, J. "The Cooperative Venture Formation Process: A Latent Variable Structural Modeling Approach." *Management Science*, Volume 36, 1990, pp. 1246-1255.

Fornell, C., and Yi, Y. "Assumption of the Two-Step Approach to Latent Variable Modeling." *Sociological Methods and Research*, Volume 20, 1992, pp. 291-320.

Franz, C. R. "Contingency Factors Affecting the User Involvement Role in the Design of Successful Information Systems." Unpublished Ph.D. dissertation, University of Nebraska, 1979.

Ginzberg, M. J. "A Study of the Implementation Process." *TIMS Studies in the Management Science*, Volume 13, 1979, pp. 85-102.

Goodhue, D. "IS Attitudes: Toward Theoretical and Definition Clarity." In L. Maggi, R. Zmud, and J. Wetherbe, Editors, *Proceedings of the International Conference on Information Systems*. San Diego, California, December 15-17, 1986, pp. 181-194.

Jaccard, J.; Turrisi, R.; and Wan, C. K. *Interaction Effects In Multiple Regression.* (Sage University Paper Series on Quantitative Applications in the Social Sciences, Series #. 07-072). Newbury Park, California: Sage Publications, 1990.

Jöreskog, K. G., and Sorbom, D. *PRELIS 2 User's Reference Guide*. Scientific Software International, 1993.

Jöreskog, K. G., and Wold, H. *Systems Under Indirect Observation – Causality Structure Prediction.* Amsterdam: North Holland Publishing Company, 1982.

Kenny, D. A., and Judd, C. M. "Estimating the Nonlinear and Interactive Effects of Latent Variables." *Psychological Bulletin*, Volume 96, 1984, pp. 201-210.

Lohmöller, J.-B. *LVPLS Program Manual: Latent Variables Path Analysis with Partial Least-Squares Estimation*. Köln: Zentralarchiv für empirische Sozialforschung, 1984.

Lohmöller, J.-B. *Latent Variable Path Modeling with Partial Least Squares*. Heidelberg: Physica-Verlag, 1989.

McKeen, J. D.; Guimaraes, T.; and Wetherbe, J. C. "The Relationship Between User Participation and User Satisfaction: An Investigation of Four Contingency Factors." *MIS Quarterly*, Volume 18, 1994, pp. 427-451.

McClelland, G. H., and Judd, C. M. "Statistical Difficulties Of Detecting Interactions And Moderator Effects." *Psychological Bulletin*, Volume 114, 1993, pp. 376-389.

Ping, J. R. A. "A Parsimonious Estimating Technique for Interaction and Quadratic Latent variables." *Journal of Marketing Research*, Volume 32, 1995, pp. 336-347.

Ping, J. R. A. "Latent Variable Interaction and Quadratic Effect Estimation: A Two-Step Technique Using Structural Equation Analysis." *Psychological Bulletin*, Volume 119, 1996, pp. 166-175.

Powers, R. F., and Dickson, G. W. "MIS Project Management: Myths, Opinions, and Reality." *California Management Review*, Volume 15, 1973, pp. 147-156.

Schonberger, R. J. "MIS Design: A Contingency Approach." *MIS Quarterly*, Volume 4, 1980, pp. 13-20.

Smith, K. W., and Sasaki, M. S. "Decreasing Multicollinearity: A Method for Models with Multiplicative Functions." *Sociological Methods and Research*, Volume 8, 1979, pp. 35-56.

Tabachnick, B. G., and Fidell, L. S. *Using Multivariate Statistics*, Second Edition. New York: Harper and Row, 1989.

Trice, A. W., and Treacy, M. E. "Utilization as a Dependent Variable in MIS Research." In L. Maggi, R. Zmud, and J. Wetherbe, Editors, *Proceedings of the International Conference on Information Systems*. San Diego, California. December 15-17, 1986, pp. 227-239.

Werts, C. E.; Linn, R. L.; and Jöreskog, K. G. "Intraclass Reliability Estimates: Testing Structural Assumptions." *Educational and Psychological Measurement*, Volume 34, 1974, pp. 25-33.

Weill, P.; and Olson, M. H. "An Assessment of the Contingency Theory of Management Information Systems." *Journal of Management Information Systems*, Volume 6, 1989, pp. 59-85.

Wold, H. "Soft Modeling – The Basic Design and Some Extensions." In K. Jöreskog and H. Wold, Editors, *Systems Under Indirect Observation II*. Amsterdam: North-Holland Press, 1982, pp. 1-53.

Wold, H. "Partial Least Squares." In S. Kotz and N. L. Johnson, Editors, *Encyclopedia of Statistical Sciences*. New York: Wiley, Volume 6, 1985, pp. 581-591.

Wold, H. "Introduction to the Second Generation of Multivariate Analysis." In H. Wold, Editor, *Theoretical Empiricism*. New York: Paragon House, 1989, pp. vii-xl.

APPENDIX A: PLS METHOD

Partial Least Squares (PLS) can be a powerful method of analysis because of the minimal demands on measurement scales, sample size, and residual distributions. Although PLS can be used for theory confirmation, it can also be used to suggest where relationships might or might not exist and to suggest propositions for later testing.

As an alternative to the more widely known covariance fitting approach (exemplified by software such as LISREL, EQS, COSAN, AMOS, and SEPATH), the component-based PLS avoids two serious problems: inadmissible solutions and factor indeterminacy (Fornell and Bookstein 1982). The philosophical distinction between these approaches is whether to use structural equation modeling for theory testing and development or for predictive applications (Anderson and Gerbing 1988). In situations where prior theory is strong and further testing and development is the goal, covariance based full-information estimation methods (i.e., Maximum Likelihood or Generalized Least Squares) are more appropriate. Yet, due to the indeterminacy of factor score estimations, there exists a loss of predictive accuracy. This, of course, is not of concern in theory testing where structural relationships (i.e., parameter estimation) among concepts is of prime concern.

For application and prediction, a PLS approach is often more suitable. Under this approach, it is assumed that all the measured variance is useful variance to be explained. Since the approach estimates the latent variables as exact linear combinations of the observed measures, it avoids the indeterminacy problem and provides an exact definition of component scores. Using an iterative estimation technique (Wold 1982), the PLS approach provides a general model which encompasses, among other techniques, canonical correlation, redundancy analysis, multiple regression, multivariate analysis of variance, and principle components.

As a consequence of using an iterative algorithm consisting of a series of ordinary least squares analyses, identification is not a problem for recursive (i.e., one way path) models nor does it presume any distributional form for measured variables. Furthermore, sample size can be smaller, with a standard rule of thumb suggesting that it be equal to the larger of the following: (1) ten times the scale with the largest number of formative (i.e., causal) indicators (note that scales for constructs designated with reflective indicators can be ignored), or (2) ten times the largest number of structural paths directed at a particular construct in the structural model. A weak rule of thumb, similar to the heuristic for multiple regression (Tabachnik and Fidell 1989, p. 129), would be to use a multiplier of five instead of ten for the preceding formulae. An extreme example is given by Wold (1989), who analyzed 27 variables using two latent constructs with a data set consisting of ten cases.

When estimating interaction effects, Ping (1996) noted that recent covariance based procedures "can produce specification tedium and errors or estimation difficulties in larger structural equations models." Ping (1995; 1996) has suggested an alternative two-step approach to overcome these limitations, however the ability to assess the reliability and validity of individual items using a two step approach has been questioned (Fornell and Yi 1992).

Second order factors can be approximated using various procedures. One of the easiest to implement is the approach of repeated indicators known as the hierarchical component model suggested by Wold (cf. Lohmöller 1989, pp. 130-133). In essence, a second order factor is directly measured by observed variables for all the first order factors. 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.

Finally, PLS is considered better suited for explaining complex relationships (Fornell, Lorange, and Roos 1990; Fornell and Bookstein 1982). As stated by Wold (1985, p. 589), "PLS comes to the fore in larger models, when the importance shifts from individual variables and parameters to packages of variables and aggregate parameters." Wold states later (p. 590), "In large, complex models with latent variables PLS is virtually without competition."

Nevertheless, being a limited information method, PLS parameter estimates are less than optimal regarding bias and consistency. The estimates will be asymptotically correct under the joint conditions of consistency (large sample size) and consistency at large (the number of indicators per latent variable becomes large). Otherwise, construct to loadings tend to be overestimated and structural paths among constructs underestimated. Furthermore, standard errors need to be estimated via resampling procedures such as jackknifing or bootstrapping (cf. Efron and Gong 1983). The significance of paths can also be determined by using the jackknife statistics resulting from a blindfolding resampling procedure (Lohmöller 1984, pp. 5-09 through 5-12). The blindfolding procedure omits a part of the data matrix for the construct being examined and then estimates the model parameters. This is done a number of times based on the blindfold omission distance. Results obtained from this resampling procedure include the jackknifed estimated means and standard deviations.

In this paper, all structural paths were significant at the 0.01 level using a bootstrapping sample of 200. Also, it should be noted that by using the PLS algorithm under a reflective mode for all constructs, we eliminate any concerns of collinearity within blocks of variables used to represent underlying constructs.

Overall, rather than being viewed as competitive models, the covariance fitting procedures (i.e., ML and GLS) and the variance-based PLS approach have been argued as complementary in nature. According to Jöreskog and Wold (1982, p 270): "ML is theory-oriented, and emphasizes the transition from exploratory to confirmatory analysis. PLS is primarily intended for causal-predictive analysis in situations of high complexity but low theoretical information."

APPENDIX B: CONSTRUCT DEFINITIONS AND ITEMS USED FOR STUDY 2

Perceived Usefulness

Construct Definition (Davis 1989): The degree to which a system is perceived to enhance one's performance.

UFL1 Using Electronic Mail in my job enables (would enable) me to accomplish tasks more quickly.
UFL2 Using Electronic Mail improves (would improve) my job performance.
UFL3 Using Electronic Mail in my job increases (would increase) my productivity.
UFL4 Using Electronic Mail enhances (would enhance) my effectiveness on the job.
UFL5 Using Electronic Mail makes it (would make it)easier to do my job.
UFL6 I find (would find) Electronic Mail useful in my job.

(Seven point scale using extremely likely, quite, slightly, neither, slightly, quite, extremely unlikely)
Future tense wording in parentheses were added to be meaningful for a percentage of respondents who have not used electronic mail.

Enjoyment

Construct Definition (Davis, Bagozzi, and Warshaw 1992): The extent to which the activity of using the computer is perceived to be enjoyable in its own right, apart from any performance consequences. (Electronic Mail was substituted for the behavior of using the computer).

ENJ1 I find (would find) using Electronic Mail to be enjoyable.
(Seven point scale using extremely likely, quite, slightly, neither, slightly, quite, extremely unlikely)

ENJ2 The actual process of using Electronic Mail is (would be)
(Seven point scale using extremely unpleasant, quite, slightly, neither, slightly, quite, extremely pleasant)

ENJ3 I have (would have) fun using Electronic Mail.
(Seven point scale using extremely likely, quite, slightly, neither, slightly, quite, extremely unlikely)

Intention To Regularly Use Electronic Mail

Construct Definition: A measure of the strength of one's intention to perform a behavior [behavioral intention]. In this case, the behavior is to use Electronic Mail regularly.

INT1 I presently intend to use Electronic Mail regularly:
(Seven point scale using extremely likely, quite, slightly, neither, slightly, quite, extremely unlikely)

INT2 My actual intention to use Electronic Mail regularly is:
(Seven point scale using extremely strong, quite, slightly, neither, slightly, quite, extremely weak)

INT3 Once again, to what extent do you at present intend to use Electronic Mail regularly:
(Eleven point scale from 0 to 10 with anchors Definite No, Definite Yes)