ICIS 1994 Proceedings

International Conference on Information Systems (ICIS)

12-31-1994

Scale Development in Information Systems Research: A Paradigm Incorporating Unidimensionality and Its Assessment

Albert Segars
Boston College

Michel Globensky
Vice President, Administration and Security, Equifax Canada

SCALE DEVELOPMENT IN INFORMATION SYSTEMS RESEARCH: A PARADIGM INCORPORATING UNIDIMENSIONALITY AND ITS ASSESSMENT

Albert H. Segars
Wallace E. Carroll School of Management
Boston College

ABSTRACT

Because of their value in assessing many aspects of information systems (IS) productivity, the development and psychometric evaluation of scales which measure unobservable (latent) phenomenon continues to be an issue of high interest among researchers in the IS community. Typically, the measurement properties of developed scales are evaluated through traditional techniques such as item-to-total correlations, coefficient alpha, and exploratory factor analysis. While potentially useful in exploratory situations, these metrics provide little assessment of scale unidimensionality. Scales which are unidimensional measure a single trait. This property is a basic assumption of measurement theory and is absolutely essential for accurate (unconfounded) measurement of variable interrelationships. In this paper, a paradigm for developing unidimensional scales is presented. Drawing from well developed techniques within marketing research, education, and psychology, this paradigm incorporates the use of confirmatory factor analysis (CFA) as a means of assessing measurement properties. Importantly, CFA provides a stricter interpretation of unidimensionality than traditional methods and in many instances will lead to different conclusions regarding scale acceptability.

1. INTRODUCTION

Increasingly, relationships of interest within the information systems (IS) community contain variables which cannot be directly observed and therefore cannot be measured with complete accuracy. Some of these variables (or constructs) include individual traits such as perceived ease of use and usefulness of information technology (Davis 1989), perceived equity of technological resource allocation (Joshi 1989), user involvement in systems design (Ives and Olson 1984), and user satisfaction with information and information systems (Bailey and Pearson 1983; Ives, Olson, and Baroudi 1983; Doll and Torkzadeh 1988). Other variables deal with organizational traits such as strategic planning (Raghunathan and King 1988), competitive advantage from IT (Sethi and King 1991), and organizational factors impacted by IT (Mamood and Soon 1991). Although diverse, a common theme among these and many other scales developed in IS research is their purpose: the measurement of value-added through investment in, or use of, information technologies. Therefore, it is not surprising that issues surrounding the assessment of measurement properties remains a major concern of researchers in the field (Segars and Grover 1993; Sethi and King 1991; Straub 1989). Such properties reveal the accuracy (validity) and consistency (reliability) of construct measurement (Churchill 1979). Sound measurement properties can yield valuable insight into the structure and interrelationships among complex variables. Conversely, poor measurement properties can lead to erroneous conclusions regarding such phenomena.

It is generally recommended that measurement of latent variables be accomplished through use of multi-item scales (Churchill 1979). Typically, respondents are administered two or more items that are intended to be alternative indicators of the same underlying construct. A composite score, which is calculated as the unweighted sum of item scores, is then used as an estimate of the corresponding construct. It then becomes necessary to establish the goodness of measurement inherent in the scale. Often, metrics such as coefficient alpha and item-to-total correlations are used to measure the amount of explained variance captured by the scale. Exploratory factor analysis (EFA) furthers the assessment by providing a means of assessing convergent and discriminant validity. Together, these and other metrics of reliability and validity provide empirical evidence regarding the measurement value of the developed scale.

An essential, but often overlooked, property of measurement value is unidimensionality. This property states that a *single* construct underlies a set of scale items. Importantly, the computation of a composite score is meaningful only if each of the items is acceptably unidimensional (Gerbing and Anderson 1988; Hattie 1985). Unfortunately, none of the traditional techniques typically utilized for establishing measurement value directly tests unidimensionality according to its formal definition. However, through confirmatory factor analysis (CFA) a stricter (and therefore more accurate) assessment of unidimensionality can be gained. Further, because CFA and traditional techniques are based on different criterion, opposite conclusions regarding the acceptability of the scale may be drawn depending on the choice of method (Gerbing and Anderson 1988). It is the purpose of this paper to present a paradigm which incorporates the formal assessment of unidimensionality through CFA. Drawn from similar frameworks within marketing research (Gerbing and Anderson 1988), psychology (Bentler 1986), and education (Jöreskog and Sörbom 1989), this paradigm is described and illustrated as a means of improving the quality of construct development within IS research.

2. UNIDIMENSIONAL MEASUREMENT

The mathematical definition of the traditional common factor model is based on the concept of unidimensional measurement in which a set of indicators share only a single underlying factor, ξ (Gerbing and Anderson 1988; McDonald 1981). Assuming linearity, the equation for a single indicator is given as

$$x_i = \lambda_i \xi + \delta_i$$

where x_i is the i^{th} indicator from a set of unidimensional indicators, λ_i is the corresponding factor loading, and δ_i is the corresponding error term assumed to be uncorrelated with any factors or other residuals (Jöreskog and Sörbom 1989). In most research settings, relationships among several distinct constructs are of interest. Therefore, several sets of postulated unidimensional measures will be present. The relationship of these measures to their respective constructs is represented formally by a multiple-indicator measurement model in which each estimated construct is defined by at least two indicators and each indicator is intended to be an estimator of only one construct (Gerbing and Anderson 1988; Anderson 1987).

Two product rules, each of which represents a necessary condition for unidimensionality, follow from this mathematical definition. The first is the product rule for internal consistency and is formulated as

$$\rho_{ij} = \rho_{i\xi}\rho_{j\xi}$$

In essence, this rule states that the correlation of two indicators i and j, which are alternative measures of a single underlying construct ξ, is equal to the product of their respective factor loadings. The second product rule concerns external consistency and is formulated as

$$\rho_{ip} = \rho_{i\xi}\rho_{\xi\xi^*}\rho_{p\xi^*}$$

This rule states that the correlation of two indicators i and p, which are measures of two different constructs ξ and ξ^*, is equal to the product of indicator i's loading on factor ξ, indicator p's loading on factor ξ^*, and the correlation between factor ξ and ξ^*. When ξ is the same as ξ^* (i.e., the correlation between the two equals 1.0), the two product rules are equivalent. In other words, internal consistency represents a special case of external consistency. It is interesting to note that the product rule for external consistency implies that items from other scales provide a means to assess the unidimensionality of items in a given scale. Given this definition, unidimensionality can now be formulated in more quantifiable terms. Succinctly stated, items which are truly unidimensional will exhibit a parallel correlational pattern (within sampling error) with other measures in the set and with items that are not in the set (i.e., measures of different constructs). It is this criteria upon which assessment of unidimensionality can be made.

As noted earlier, the development and evaluation of measurement scales has traditionally relied upon analyses such as coefficient alpha, item-to-total correlations, and exploratory factor analysis. The use of confirmatory factor models for such purposes is a relatively new phenomena within IS research (Segars and Grover 1993). Yet, the value of this technique in establishing sound measurement properties has long been acknowledged within the areas of marketing research, education, and psychology (Gerbing and Anderson 1988; Bentler 1986). Importantly, only CFA directly tests unidimensionality as formally defined in the equations developed earlier. In the following sections, a formal definition of the CFA model as contained within Jöreskog and Sörbom's larger LISREL model is developed. Based on this definition, a formal paradigm is then presented which incorporates assessment of unidimensionality.

2.1 The Confirmatory Measurement Model

Utilizing the notation of Jöreskog and Sörbom, the confirmatory measurement model can be expressed in matrix form as

$$X = \Lambda_x \xi + \delta$$

where X is a column vector of p indicators, ξ is a $k < p$ vector of latent constructs, δ is a vector of unique scores (random errors), and Λ_x is a p x k matrix of pattern coefficients relating each indicator to its posited underlying construct. Applying this expression to a two factor model with five indicators per factor results in the following set of simultaneous equations:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \end{bmatrix} = \begin{bmatrix} \lambda_{1.1} & 0 \\ \lambda_{2.1} & 0 \\ \lambda_{3.1} & 0 \\ \lambda_{4.1} & 0 \\ \lambda_{5.1} & 0 \\ 0 & \lambda_{6.2} \\ 0 & \lambda_{7.2} \\ 0 & \lambda_{8.2} \\ 0 & \lambda_{9.2} \\ 0 & \lambda_{10.2} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \\ \delta_7 \\ \delta_8 \\ \delta_9 \\ \delta_{10} \end{bmatrix}$$

As shown, it is postulated that the first five indicators (x_1-x_5) load on the first latent construct (ξ_1) while the remaining five indicators load on the second (ξ_2). Error in measurement is modeled through the vector of error terms (δ). Alternatively stated, each indicator is posited as a function of its correlation with a *single* latent construct and error due to random and measure specific effects.

With the general assumptions of uncorrelatedness between errors and constructs and an expected value of zero for δ and ξ, the p x p variance covariance matrix for the indicators x denoted as Σ can be expressed as:

$$\Sigma = \Lambda_x \Phi \Lambda_{x'} + \Theta_\delta$$

where Φ is the k x k covariance matrix of latent constructs (ξ) and Θ_δ is a p x p diagonal matrix of measurement error variances.

It is this set of pattern matrices that are pre-specified by the researcher based on theoretical expectation. Subsequently, the matrices are used by LISREL in maximum likelihood estimation. Thus, in the present example, Λ_x is the 10 x 2 matrix of pattern loadings specified above, Φ is a 2 x 2 symmetric matrix of construct covariances, and Θ_δ is a 10 x 10 diagonal matrix of indicator error terms.

The overall fit of a hypothesized model can be tested by using the maximum likelihood chi-square statistic provided in the LISREL output. Importantly, this measure of fit is a function of external and internal consistency (i.e., the difference between observed covariances and those implied by the model's estimated parameters). Assuming that the model is valid and that the observed measures follow a multivariate normal distribution, the statistic is asymptotically distributed as a chi-square variable with its associated degrees of freedom. Formally, the null and alternative hypotheses of the confirmatory model are specified as

$$\text{Ho: } \Sigma = \Sigma(\theta)$$
$$\text{Ha: } \Sigma = \Sigma\alpha$$

where Σ is the population matrix estimated by the observed covariances between indicators, $\Sigma(\theta)$ is the implied covariance matrix that would result from the pattern matrices specified by the researcher, and $\Sigma\alpha$ is any positive definite matrix. Retainment of Ho implies that the observed covariances among indicators are well modeled by the specified pattern matrices (Λ_x, Φ, Θ_δ). Conversely, rejection of Ho implies poor model fit. Importantly, the chi-square statistic is sensitive with respect to large sample sizes and models with large numbers of indicators (Bearden, Sharma, and Teel 1982). In these instances, even trivial discrepancies between a model and data can result in significant chi-square values. Therefore, other measures of model fit such as adjusted chi-square, goodness of fit indices, and mean square residual must also be considered in assessing model adequacy (Jöreskog and Sörbom 1989).

Assuming an adequate model fit, a number of important psychometric analyses can be performed with the LISREL measurement model. Convergent validity can be assessed by examining the ratio of factor loadings to their respective standard errors. In general, if these t values are greater than $|2.00|$, then they are considered significant at the .05 level. The reliability of individual items may also be assessed by squaring their respective standardized loadings. In general, if these estimates of reliability are greater than 0.50, then the item explains more variance than is explained by the error term. Also using the estimated loadings, overall construct reliability and average variance explained may be calculated as a further assessment of convergent validity (Fornell and Larcker 1981).

When two models are nested such that one model can be created from another model by imposing additional restraints, the difference between the two chi-square values is also distributed as a chi-square with degrees of freedom equal to the difference between the degrees of freedom between the two models. Comparison of such models provides a direct test of discriminant validity. For example, the discriminant validity of the two factor model can be assessed by fixing the correlation parameter Φ (2,1) between the two factors at 1.0 and then performing a chi-square difference test on the values obtained for the constrained and unconstrained models. A significantly lower chi-square value for the model in which the trait correlations are not constrained to unity would indicate that the traits are not perfectly correlated and that discriminant validity is achieved (Anderson 1987; Bagozzi and Phillips 1982).

If the prescribed model proves to be an ill fit for the observed correlation matrix, LISREL still provides a number of useful diagnostics. Modification indices, which are themselves chi-square variates, specify the incremental improvement in model fit given the estimation of an additional parameter in Λ_x. For example, a high modification index for element 1,2 of Λ_x suggests that x_1 may share a significant amount of variance with the alternative construct (i.e., the indicator is not unidimensional). Therefore, to reduce the overall chi-square of model fit by the amount of the modification index (i.e., improve model fit) a path can be estimated between this respective indicator and construct. Another useful model diagnostic is the matrix of standardized residuals reported by LISREL. These residuals, which are the standardized differences between the observed and estimated covariance matrices, can be extremely useful in identifying indicator covariances which are not well explained by the hypothesized model. Taken together, these two diagnostics can identify indicators which may be complex (i.e., load significantly on multiple factors) or contain abnormal amounts of random variation. Importantly, existence of either of these conditions may confound the results of subsequent structural modeling and therefore the validity of reported findings (Burt 1976).

2.2 A Paradigm for Scale Development that Incorporates Unidimensionality

Figure 1 illustrates a paradigm for scale development which incorporates CFA and, subsequently, assessment of unidimensionality. This paradigm is an amalgamation of similar paradigms developed in the areas of marketing research (Churchill 1979; Peter 1979; Gerbing and Anderson 1988), education (Jöreskog and Sörbom 1989), and psychology (Bentler 1986). While the presented technique is neither new nor novel, its discussion and illustrated use within the context of IS research is sorely needed, particularly in light of the growing sophistication of IS constructs and research designs for exploring their relationships.

As illustrated, the process begins with a pre-determination of items which are congeneric measures of a given construct. Items are said to be congeneric if they measure the same trait except for errors of measurement (Jöreskog 1971). Subsequently, any pair of such items will have linearly related true scores. Identification of congeneric items should be accomplished through the structured scale development approach suggested by Churchill and/or through Q-Sorting techniques (Moore and Benbasat 1991). Importantly, strong theoretical underpinnings are required for proper model development and testing. Therefore, it is critical that much effort be given to theoretically deriving and refining items at this initial stage. Upon development of scale items, the confirmatory factor model is estimated. Statistical packages such as LISREL (contained in SPSS), EQS, or CALIS (contained in SAS) may be utilized for this purpose. Although each is different with respect to modeling statements and syntax, all are consistent in their use of maximum likelihood estimation (MLE) for the derivation of parameter estimates.

Next, a formal assessment of convergent validity and unidimensionality can be performed. As noted earlier, examination of fit indices, indicator loadings, t-values, and modification indices provides the researcher with specific evidence regarding these important measurement characteristics. Depending on these measures, model fit may be improved by eliminating items with low reliabilities, respecifying indicators to load on more than one factor, or adding additional factors. In the development of unidimensional scales, it is critical that indicators with multiple factor loadings (i.e., high modification indices) be identified and eliminated. In other words, should model fit be established with some indicators loading on two or more factors, these indicators should be eliminated in order to reduce potential confounds in meaning and effect. This reduced model must then be reestimated to assess measurement properties.

Upon identification of an acceptable model, factor reliability and discriminant validity can be assessed. Composite factor reliability is defined by the following formula (Fornell and Larcker 1981; Jöreskog 1971):

$$\frac{\left(\sum_i \lambda i\right)^2}{\left(\sum_i \lambda i\right)^2 + \left(\sum_i (1 - \lambda i^2)\right)}$$

where i represents the i^{th} factor loading of indicator x on factor ξ. In essence, this composite measure assesses whether the specified indicators are sufficient in their

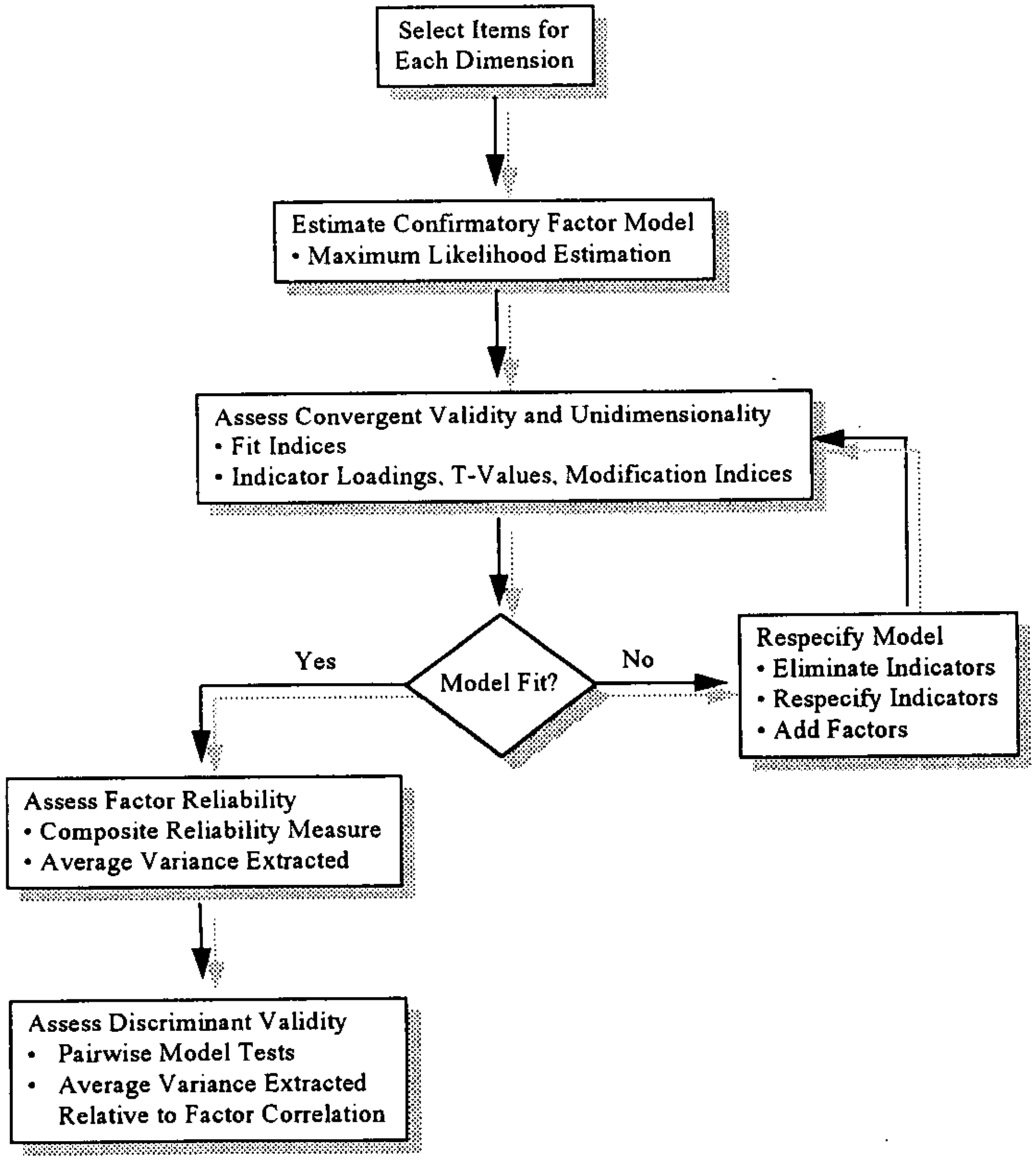

Figure 1. A Paradigm Incorporating Assessment of Unidimensionality

representation of their respective constructs. Although no exact criteria exists for factor reliability, a value of 0.70 is often cited as a lower threshold (Jöreskog 1971).

As noted by Fornell and Larcker, neither item estimates of reliability nor a composite measure indicate the amount of variance that is captured by the construct in relation to the amount of variance due to measurement error. These authors suggest a measure termed average variance extracted (AVE) as a means of acquiring this information. AVE is calculated as

$$\frac{\left(\sum_i \lambda i^2\right)}{\left(\sum_i \lambda i^2\right) + \left(\sum_i (1 - \lambda^2)\right)}$$

If AVE is less than 0.50, then the variance due to measurement error is larger than the variance captured by the respective construct. Subsequently, the validity of the individual indicators, as well as the construct, is suspect.

As noted earlier, discriminant validity is assessed by estimating an alternative model where the correlation between constructs is constrained to 1.0. The difference in chi-square values between this restricted and freely estimated models provides statistical evidence of discriminant validity. As further evidence of discriminant validity, the AVE of each construct can be compared with the estimated correlation between constructs. To fully satisfy the requirements for discriminant validity, AVE for each construct should be greater than the squared correlation between constructs. Such results suggest that the items share more common variance with their respective constructs than any variance the construct shares with other constructs (Fornell and Larcker 1981).

Importantly, the presented paradigm should not be interpreted as a "treasure hunt" in which numerous configurations are tested until a good fitting model emerges. Rather, the technique is designed to identify and eliminate indicators which, as a result of their factor complexity or excessive error, may confound the estimation and interpretation of structural paths between latent variables (Anderson and Gerbing 1988). Radical respecification of measurement models clearly calls for reassessment of construct domain and indicator content. Further, excessive respecification can only be meaningfully validated through retesting on an independent sample. Finally, it should be noted that other methods are available for assessing unidimensionality which are not dependent on factor analytic models. In particular, item response analysis has been proposed as a theoretically sound means of assessing measurement value. Such measures overcome the inherent limitations of traditional factor models (linearity, indeterminacy) but are themselves limited by a lack of rationale and formal empirical testing (Hattie 1985).

3. APPLYING THE PARADIGM IN PRACTICE

To illustrate the use of the presented paradigm as well as highlight some of the methodological caveats of traditional scale development techniques, a formal analysis of two well known constructs within the IS field was undertaken. The first of these constructs, *technological infusion*, measures the importance of information technologies in the realization of strategic objectives (McFarlan 1984; Sullivan 1985). Typically framed within the context of McFarlan's strategic grid, this concept has been important in uncovering optimal methods for strategic IS planning (Sullivan 1985) as well as differences in acceptance and usefulness of planning (Raghunathan and Raghunathan 1990). It has also provided a theoretical framework for identifying organizational and competitive characteristics which explain varying levels of technological dependence (Sullivan 1985; McFarlan 1984). The second construct, *technological diffusion*, measures the extent to which technology has been dispersed or distributed throughout the organization (Allen and Boynton 1991; Sullivan 1985). This technological characteristic has also been associated with differences in optimal IS planning and management approaches (Boynton, Jacobs, and Zmud 1992; Donovan 1988; Sullivan 1985). Together, these constructs have provided a lens through which organizations can assess their technological disposition, and subsequently, structure appropriate management and planning policies.

In Table 1, the items which are posited to be congeneric measures of each respective construct are listed. The items associated with diffusion are taken from studies by Grover (1993) and Donovan (1988). Items associated with infusion are taken from a recent study by Neumann, Ahituv, and Zviran (1992). These items were part of a larger survey mailed to 250 IS executives throughout the United States. These executives were identified through *The Directory of Top Computer Executives* (1993); a quarterly listing which is continuously updated and has provided the sampling frame for a number of IS surveys (Premkumar and King 1992). Each executive was asked to rate on a scale of 1 to 7 his or her agreement with each statement. A total of 153 usable surveys were returned for an effective response rate of 61.2%. This rate is significantly higher than that usually obtained from IS studies and can be largely attributed to the use of monetary and informational incentives which accompanied the instrument. To detect a potential effect of non-response bias, patterns of "early" and "late" respondents were compared along dimensions of industry, size, and executive rank (Schiltz 1988). In each instance, contingency tables revealed no significant differences among these groups, suggesting no bias due to non-response.

3.1 Item to Total Correlations

The use of item-to-total correlations has long been advocated in the assessment of scale unidimensionality. As noted by Nunnally (1978),

> Items within a measure are useful only to the extent that they share a common core — the attribute which is to be measured...the items that correlate most highly with total scores are the best items for a general-purpose test.

Although such measures account for the internal consistency of scale items, they fail to account for external consistency. In other words, by not accounting for the relations of the posited indicators with those of different factors, an item-to-total analysis may fail to discriminate

Table 1. Variable Operationalizations

Diffusion	Items (Strongly Agree - Strongly Disagree)
x_1	Our divisions/SBUs are dependent on centralized hardware.
x_2	Our major databases are in one geographic region.
x_3	Most major data processing is centralized in one location.
x_4	Most major hardware is centralized in one geographic location.
x_5	Storage and processing technologies are widely-distributed throughout the firm.
Infusion	**Items (Strongly Agree - Strongly Disagree)**
x_6	IS under development will enable our firm to deliver new products/services.
x_7	Present IS greatly aid our firm in increasing profitability.
x_8	Present IS are critical for effective competition within our market.
x_9	IS under development will greatly aid our firm in increasing profitability.
x_{10}	IS under development are critical for effective competition within our market.

Table 2. Correlation Matrix of Indicators

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
x_1	1.0									
x_2	0.76	1.0								
x_3	0.69	0.76	1.0							
x_4	0.68	0.71	0.82	1.0						
x_5	0.66	0.65	0.65	0.61	1.0					
x_6	0.23	0.28	0.29	0.22	0.29	1.0				
x_7	0.16	0.13	0.14	0.11	0.28	0.58	1.0			
x_8	0.11	0.15	0.15	0.13	0.29	0.60	0.63	1.0		
x_9	0.12	0.12	0.14	0.12	0.25	0.52	0.64	0.65	1.0	
x_{10}	0.18	0.12	0.18	0.14	0.29	0.55	0.62	0.63	0.63	1.0

between sets of indicators that represent different, though correlated, factors (Gerbing and Anderson 1988). For example, an indicator, x, will correlate more highly with its own factor, ξ, than with another factor, ξ^*, even if the factors themselves are highly correlated. However, in such instances, it is also likely that the correlation between x and ξ^* will be substantial. As a specific example of this phenomena, consider falsely combining the ten items of the diffusion and infusion scales whose correlations are presented in Table 2. In this sample, the correlations of each item with the ten item total score range from a low of 0.67 to a high of 0.73. Correcting for the overlap in item and total score, these measures range from 0.62 to 0.69. In essence, item-to-total correlations in the mid to high sixties can be obtained from two sets of five items whose average between group correlation is only 0.18. Such results seem to suggest that item-to-total correlations do not establish unidimensionality, rather, it is an assumed property of the data.

3.2 Coefficient Alpha and Reliability

Coefficient alpha (Cronbach 1951) is perhaps the most widely used metric for gauging the reliability of scale items (Peter 1979). Reliability is classically defined as the ratio of true score variance to observed score variance. Accordingly, the less error inherent within the scale, the more likely the measure will yield consistent results across observations and research settings. Often, coefficient alpha is misinterpreted as an index of unidimensionality. Consistent with this misinterpretation, researchers often develop scales by selecting items which maximize reliability (Gerbing and Anderson 1988). This practice is encouraged by computer packages such as SPSS which contain specialized routines for just such purposes.

The distinction between unidimensionality and reliability lies in the mathematical definition of each concept. As shown earlier, the dimensionality of a scale can be evaluated by examining the patterning of indicator correlations. In contrast, reliability is a function of the number of items that define the scale and the respective reliabilities of those items. Specifically, alpha is an application of the Spearman-Brown formula used in the computation of split-half reliability. The only difference is that instead of two splits, the number of splits for coefficient alpha is equal to the number of items in the scale. Formally, alpha is defined as

$$\alpha = \frac{p(r)}{1 + (p - 1)r}$$

where p is equal to the number of scale items and r is equal to the average off-diagonal correlation. As shown, regardless of scale unidimensionality, reliability tends to increase as the average off-diagonal item correlation increases and/or the number of scale items increases. Applying this formula to the correlations of Table 2, alpha for a combined ten item scale is a rather high 0.86. This result is especially surprising given an average within item correlation of 0.70 (diffusion) and 0.61 (infusion) relative to the mean between item correlation of 0.18. As this example shows, unidimensionality is assumed, not proven, in the computation of coefficient alpha.

3.3 Exploratory Factor Analysis

Exploratory factor analysis (EFA) is a useful technique for reducing a large set of indicators into a more manageable subset. A typical use of EFA in the development of scales is to factor an overall set of items and then construct scales based on the result of factor loadings. Items which load high on a particular factor but low on other factors are combined to form composite construct measures. This has been a particularly popular methodology within the realm of IS research (Davis 1989; Sethi and King 1992; Joshi 1989). Although useful in research settings where the factor structure of scale items is unknown, EFA does not provide an explicit test of unidimensionality. Specifically, factors in an exploratory analysis do not correspond directly to the constructs represented by each set of indicators because each factor is defined as a weighted sum of all observed variables (Segars and Grover 1993; Gerbing and Anderson 1988). In essence, the factor solution obtained from EFA is one of an infinite number of possible solutions. Only by prespecifying indicators to be a function of their respective factors can unique models be obtained. Unlike the factor loadings associated with EFA, unique solutions are hypotheses tested against data and can be statistically analyzed (Bagozzi and Phillips 1982).

In contrast to the findings of the item-to-total and alpha techniques, a factor analysis of the correlation matrix contained in Table 2 seems to imply a two-factor structure underlying the scale items. The first three eigenvalues extracted from the correlation matrix are 4.55, 2.71, and 0.51 respectively. Only two eigenvalues are greater than 1.0 and analysis of the associated screen plot (Cattell 1978) suggests a steep gradient after the second factor. Therefore, two principle axis factors were extracted by iterating for communalities. The oblique (promax) rotation of the resulting two factors also suggested a clean two-factor solution. As shown in Table 3, factor loadings of the diffusion indicators range from a low of 0.77 to a high of 0.89. Loadings of the infusion items range from 0.74 to 0.84. Based on these results, it seems safe to conclude that the ten items are adequately explained by two factors. Additionally, it is tempting to conclude that both scales are unidimensional. However, as can be detected in the correlation matrix of Table 2, x_5 and x_6 seem to violate the product rule of external consistency. Unfortunately, it is not possible to formally test this suspicion within the framework of the EFA model.

3.4 Confirmatory Factor Analysis

Utilizing the paradigm illustrated in Figure 1, we now formally assess the convergent validity, unidimensionality, and discriminant validity of the diffusion and infusion measures. To begin, the hypothesized factor structure of the observed correlations must be developed. Figure 2 is a path diagram of this expected structure.

As shown, each indicator is posited to be a function of the correlation (λ_i) with its respective factor (ξ_i) plus unique variance (δ_i). In essence, if properties of both external and internal validity hold, the observed correlations of Table 2 will approximate (within sampling error) the correlation matrix implied by the model of Figure 2. The maximum likelihood solution obtained with LISREL 7 suggests that the hypothesized model is not a good representation of the observed correlation matrix. The reported chi-square value, χ^2 (34) = 55.36, is highly significant (p=0.012), suggesting that the observed and implied correlation matrices are substantially different. While the goodness of fit index is a moderately strong 0.93, as is the adjusted goodness of fit index of 0.89, examination of modification indices and the residual matrix revealed substantial room for model improvement. Specifically, modification indices suggested that chi-square could be reduced by 5.80 through estimation of a path between ξ_1 and x_6; and by 10.06 through estimation of a path between ξ_2 and x_5. Additionally, high residual values between these items and other indicators suggested a need for model respecification.

Table 3. Rotated (Promax) Factor Matrix

Variable	Factor 1	Factor 2
x_1	0.86	0.08
x_2	0.89	0.07
x_3	0.89	0.09
x_4	0.87	0.05
x_5	0.77	0.26
x_6	0.22	0.74
x_7	0.07	0.83
x_8	0.07	0.84
x_9	0.05	0.83
x_{10}	0.09	0.82
Estimated Correlation Between Factors = 0.24		

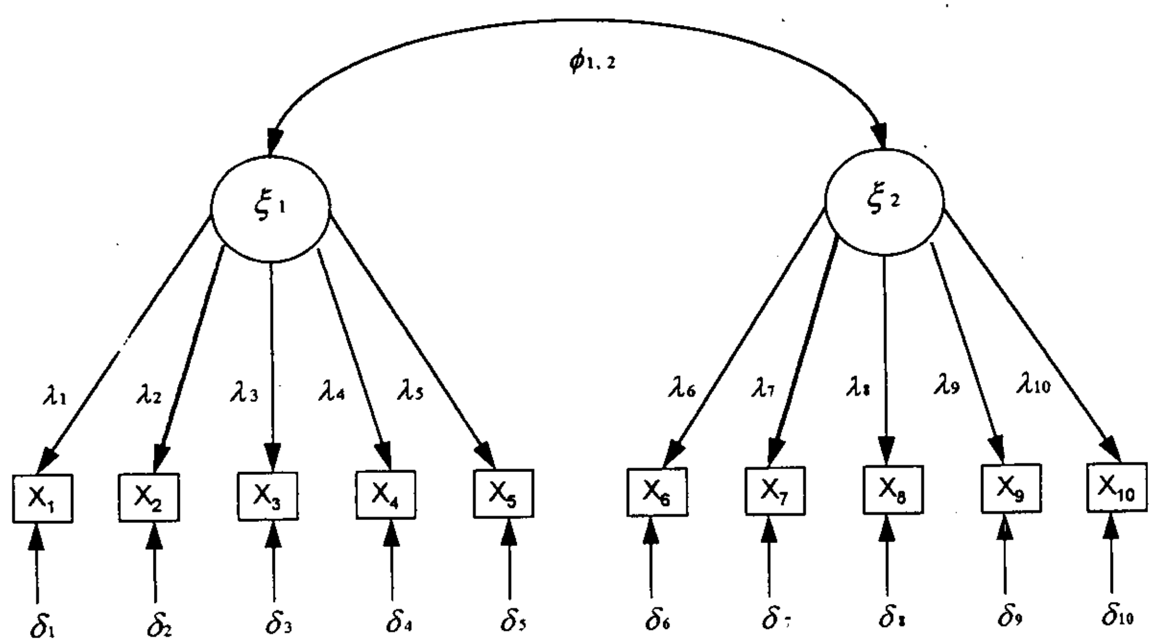

Figure 2. Hypothesized Two Factor — Ten Indicator Model

Given the apparent lack of unidimensionality in the items x_5 and x_6, these measures were eliminated from further analysis and a reduced eight item model was estimated. The reported chi-square value for this reduced model, χ^2 (19) = 25.29, is not significant (p=0.15), implying adequate modeling of the observed correlations. The goodness of fit index is a strong 0.96 as is the adjusted goodness of fit index of 0.92. Examination of modification indices and residuals suggested that no further improvement in fit could be realized through respecification. Observed factors loadings and the estimated correlation between factors are shown in Figure 3. As illustrated, all indicator loadings are well above the recommended cutoff value of 0.70. In essence, it can be concluded from this analysis that a two factor model with indicators x_1 through x_4 loading on the factor diffusion and indicators x_7 through x_{10} loading on the factor infusion exhibits properties of both convergent validity and unidimensionality.

Given evidence of unidimensional scales, factor reliabilities and discriminant validity can now be assessed. Using the composite factor reliability formula presented earlier, the reliability of the factor diffusion is 0.91. The composite score for infusion is 0.87. Both values are far above the cutoff of 0.70 suggesting that the underlying items are sufficiently representative of their respective constructs. The average variance extracted by the constructs diffusion and infusion is 0.73 and 0.63 respectively. These values are well above the suggested value of 0.50 implying that the variance captured by the constructs is significantly greater than that attributable to error. To establish discriminant validity, a model which constrained the correlation between the factors to unity was estimated. The observed chi-square value of this restricted model was a rather large 308.96 (df=20). The chi-square difference of the restricted and freely estimated models is 308.96 - 25.29 = 283.67 (df=1, p<.0001). This highly significant difference suggests that the constructs are distinct and that their underlying scales exhibit the property of discriminant validity. Additionally, AVE measures for both constructs are much larger than the square of the correlation between them, providing further evidence of discriminant validity.

In sum, this practical application has highlighted some common misunderstandings regarding traditional scale development techniques. Further, it has demonstrated the utility and structure of the CFA paradigm in the development of scales which exhibit sound properties of reliability and validity. Importantly, unidimensionality is a necessary condition for both reliability and validity. Yet, assessment of internal and external consistency of scale indicators is rarely undertaken in practice. As demonstrated, many traditional techniques do not test for unidimensionality, rather, it is an assumed property of the items. Further, these results show that it is possible to draw different conclusions regarding scale properties depending on the technique chosen. Such findings by no means render traditional techniques obsolete. Rather, they should serve to guide proper use of methodology in the development of measurement scales.

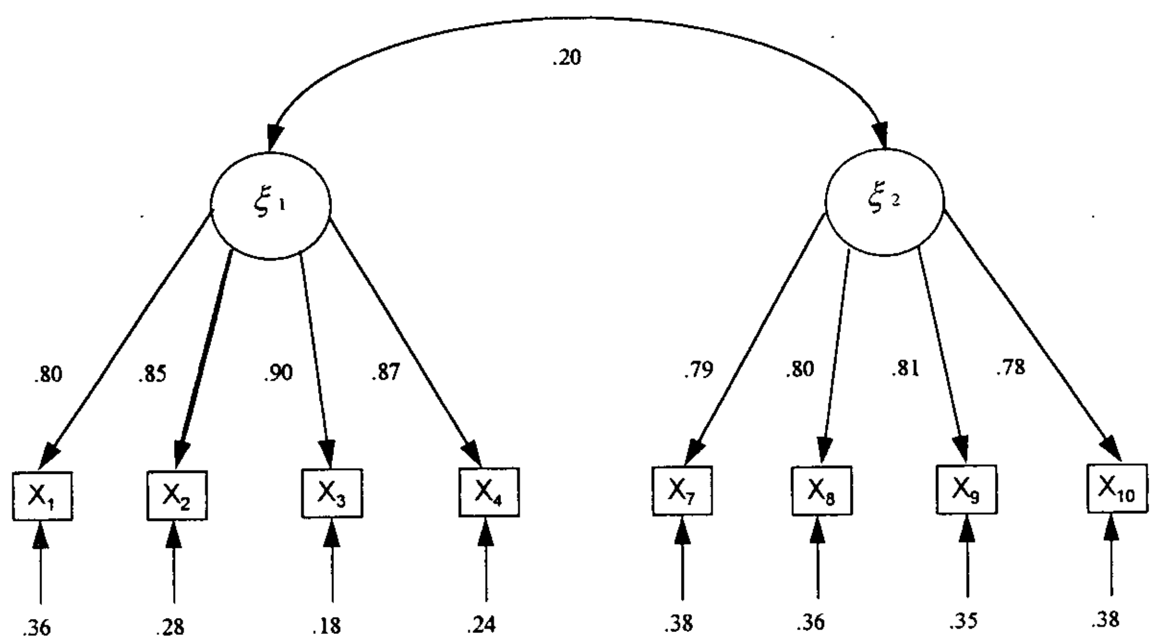

Figure 3. Reduced Two Factor — Eight Indicator Model

4. CONCLUSIONS

As noted by DeLone and McLean (1992), if information systems research is to make a contribution to the world of practice, well-defined outcome measures are essential. Such measures are useful in answering the questions of *how, when, where,* and *to what extent* information technologies influence organizational life and contribute to the firm's financial well-being. Unlike variables such as marketing expenditures or research and development spending, it is difficult to accurately quantify in a financial or productivity sense the value of IT to the organization. In essence, the same is true with many concepts related to IT practice. Traits such as involvement, participation, satisfaction, and equity are complex in nature. They are an amalgamation of perceptions, beliefs, and attitudes about the state of IS development and use within the organization. Accurate measurement of such complex variables requires careful analysis and is the building block for generating valid relationships among a system of variables. The paradigm developed here extends techniques which have typically been utilized in IS research for such purposes. By providing a stricter test of unidimensionality, this paradigm improves the quality of resulting scales, thereby strengthening the validity of variable interrelationships.

5. REFERENCES

Allen, B. R., and Boynton, A. C. "Information Architecture: In Search of Efficient Flexibility." *MIS Quarterly,* Volume 15, Number 4, December 1991, pp. 435-445.

Anderson, J. C. "An Approach For Confirmatory Measurement and Structural Equation Modeling of Organizational Properties." *Management Science,* Volume 33, Number 4, April 1987, pp. 525-542.

Anderson, J. C., and Gerbing, D. W. "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach." *Psychological Bulletin,* Volume 103, Number 3, 1988, pp. 453-60.

Bagozzi, R. P., and Phillips L. W. "Representing and Testing Organizational Theories: A Holistic Construal." *Administrative Science Quarterly,* Volume 27, September, 1982, pp. 459-489.

Bailey, J. E., and Pearson, S. W. "Development of a Tool for Measuring and Analyzing Computer User Satisfaction." *Management Science,* Volume 29, Number 5, May 1983, pp. 530-545.

Bearden, W. O.; Sharma, S.; and Teel, J. E. "Sample Size Effects on Chi-Square and Other Statistics Used in Evaluating Causal Models." *Journal of Marketing Research,* Volume 19, 1982, pp. 425-430.

Bentler, P. M. "Structural Modeling and Psychometrika: An Historical Perspective on Growth and Achievements." *Psychometrika,* Volume 51, Number 1, March 1986, pp. 35-51.

Boynton, A. C.; Jacobs, G. C.; and Zmud, R. W. "Whose Responsibility is IT Management?" *Sloan Management Review,* Volume 33, 1992, pp. 32-38.

Burt, R. S. "Interpretational Confounding of Unidimensional Variables in Structural Equation Modeling." *Sociological Methods and Research,* Volume 5, 1976, pp. 3-51.

Cattell, R. B. *The Scientific Use of Factor Analysis in Behavioral and Life Sciences.* New York: Plenum Press, 1978.

Churchill, G. A., Jr. "A Paradigm for Developing Better Measures of Marketing Constructs." *Journal of Marketing Research,* Volume 16, February 1979, pp. 64-73.

Cronbach, L. J. "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika,* Volume 16, 1951, pp. 297-334.

Davis, F. D. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology." *MIS Quarterly,* Volume 13, Number 3, September 1989, pp. 319-340.

Delone, W. H., and McLean E. R. "Information Systems Success: The Quest for the Dependent Variable." *Information Systems Research,* Volume 3, Number 1, March 1992, pp. 60-95.

Directory of Top Computer Executives. Applied Research Inc., Spring 1993.

Doll, W. J., and Torkzadeh, G. "The Measurement of End User Computing Satisfaction." *MIS Quarterly,* Volume 12, Number 2, June 1988, pp. 259-274.

Donovan, J. J. "Beyond Chief Information Officer to Network Manager." *Harvard Business Review,* Volume 66, 1988, pp. 124-140.

Fornell, C., and Larcker, D. F. "Evaluating Structural Equation Models With Unobservable Variables and Measurement Error." *Journal of Marketing Research,* Volume 18, 1981, pp. 39-50.

Gerbing, D. W., and Anderson, J. C. "An Updated Paradigm For Scale Development Incorporating Unidimensionality and Its Assessment." *Journal of Marketing Research*, Volume 25, May 1988, pp. 186-192.

Grover, V. "An Empirically Derived Model for the Adoption of Customer-Based Interorganizational Systems." *Decision Sciences*, Volume 24, Number 3, May/June 1993, pp. 603-640.

Hattie, J. "Methodology Review: Assessing Unidimensionality of Tests and Items." *Applied Psychological Measurement*, Volume 9, June 1985, pp. 139-164.

Ives, B., and Olson, M. "User Involvement and MIS Success: A Review of Research." *Management Science*, Volume 30, Number 5, May 1984, pp. 586-603.

Ives, B.; Olson, M.; and Baroudi, J. J. "The Measurement of User Information Satisfaction." *Communications of the ACM*, Volume 26, Number 10, October 1983, pp. 785-793.

Jöreskog, K. G. "Statistical Analysis of Sets of Congeneric Tests." *Psychometrika*, Volume 36, Number 2, June 1971, pp. 109-133.

Jöreskog, K. G., and Sörbom, D. *LISREL 7 User's Reference Guide*. Chicago: Scientific Software Inc., 1989.

Joshi, K. "The Measurement of Fairness or Equity Perceptions of Management Information Systems Users." *MIS Quarterly*, Volume 13, Number 3, October 1989, pp. 343-358.

Mamood, M. A., and Soon, S. K. "A Comprehensive Model for Measuring the Potential Impact of Information Technology on Organizational Strategic Variables." *Decision Sciences*, Volume 22, Number 4, 1991, pp. 869-897.

McDonald, R. P. "The Dimensionality of Tests and Items." *British Journal of Mathematical and Statistical Psychology*, Volume 34, May 1981, pp. 100-117.

McFarlan, F. W. "Information Technology Changes The Way You Compete." *Harvard Business Review*, Volume 62, 1984, pp. 98-103.

Moore, G. C., and Benbasat, I. "Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation." *Information Systems Research*, Volume 2, Number 3, September 1991, pp. 192-222.

Neumann, S.; Ahituv, N.; and Zviran, M. "A Measure For Determining The Strategic Relevance of IS to the Organization." *Information and Management*, Volume 22, 1992, pp. 281-299.

Nunnally, J. C. *Psychometric Theory*, Second Edition. New York: McGraw-Hill, 1978.

Peter, J. P. "Reliability: A Review of Psychometric Basics and Recent Marketing Practices." *Journal of Marketing Research*, Volume 16, 1979, pp. 6-17.

Premkumar G., and King, W. R. "An Empirical Assessment of Information Systems Planning and the Role of Information Systems in Organizations." *Journal of Management Information Systems*, Volume 9, Number 2, Fall 1992, pp. 99-125.

Raghunathan, T. S., and King, W. R. "The Impact of Information Systems Planning on the Organization." *OMEGA*, Volume 16, Number 2, 1988, pp. 85-93.

Raghunathan, B., and Raghunathan, T. S. "Planning Implications of the Information Systems Strategic Grid: An Empirical Investigation." *Decision Sciences*, Volume 21, 1990, pp. 287-300.

Schiltz, M. E. "Professional Standards for Survey Research." *Research in Higher Education*, Volume 28, 1988, pp. 67-75.

Segars, A. H., and Grover, V. "Re-Examining Perceived Ease of Use and Usefulness: A Confirmatory Factor Analysis." *MIS Quarterly*, Volume 17, Number 4, December 1993, pp. 517-525.

Sethi, V., and King, W. R. "Construct Measurement in Information Systems Research: An Illustration in Strategic Systems." *Decision Sciences*, Volume 22, 1991, pp. 455-472.

Straub, D. W. "Validating Instruments in MIS Research." *MIS Quarterly*, Volume 13, Number 2, June 1989, pp. 147-169.

Sullivan, C. H. "Systems Planning In The Information Age." *Sloan Management Review*, Volume 26, Number 2, Winter 1985, pp. 3-12.