Research Note

Re-Examining Perceived Ease of Use and Usefulness: A Confirmatory Factor Analysis

By: Albert H. Segars
Operations and Strategic Management Department
The Wallace E. Carroll School of Management
Boston College
Chestnut Hill, Massachusetts 02167 U.S.A.

Varun Grover
Department of Management Science
College of Business Administration
University of South Carolina
Columbia, South Carolina 29208 U.S.A.

Introduction

Recently, Adams, et al. (1992) presented the results of two studies designed to replicate previous work by Fred Davis (1989) regarding perceived usefulness, ease of use, and their influence on the usage of information technology. Specifically, these authors sought to: (1) evaluate the psychometric properties of the ease of use and usefulness scales, and (2) empirically examine the relationship between the constructs (or traits) of usefulness, ease of use, and reported levels of usage.

In study 1, a total 118 users from 10 different organizations were surveyed for their attitudes toward electronic mail (E-Mail) and voice mail (V-Mail) technologies. In essence, a heterogeneous group of users was asked to evaluate a largely similar set of technologies. In study 2, the scales were applied to a total of 73 student users of Lotus 1-2-3, Wordperfect, and Harvard Graphics; a homogenous set of users evaluated a heterogeneous set of technologies. To establish the measurement properties of the scales, the authors relied on two "classical" psychometric techniques—namely, a subset of Campbell and Fiske's (1959) Multi-Trait Multi-Method (MTMM) criteria and common factor analysis. In each study, the authors conclude that Davis' scales demonstrate properties of both reliability and validity. That is, the observed covariances in scale items seem to confirm the two-factor structure first postulated by Davis.

To measure the influence of usefulness and ease of use on reported levels of usage, Adams, et al. employed structural equation modeling (LISREL). Given the existence of a valid measurement model (i.e., the covariances in scale scores are sufficiently explained by the two constructs), this technique allows the researcher to analyze a set of latent constructs much like independent and dependent variables in regression analysis. In this particular instance, usage (the dependent construct) was regressed against the independent constructs of usefulness and ease of use. Unfortunately, in this part of the analysis, less-than-satisfactory model fits were observed. Further, inconsistencies within and across the two studies regarding the strength of causal influence of Davis' constructs on usage seem to suggest that these relationships may be more complex than previously thought. Citing these inconsistencies, the authors suggested that further analysis of these scales be undertaken to better establish their measurement properties, underlying structures, and stability over various technologies. The results of such efforts would provide important information to researchers seeking to statistically test relationships among these variables through structural equation modeling. In this research note, we undertake such an analysis.

Assessment of Construct Validity: Classical and Contemporary Methods

Within the social sciences, most theories and models are formulated in terms of unobservable or latent constructs. Fortunately, a sampling of indicators (alternatively known as items) from the population of all construct indicators can, in many instances, be used as an accurate measure of the unobservable phenomena (Cronbach, 1951). Just how accurately and consistently these indicators measure the construct of interest constitute the question of construct validity. Establishing acceptable levels of construct validity is critical, particularly when the measured constructs are to be further used for structural equation modeling (Anderson and Gerbing, 1988). Without unambiguous evidence of construct validity, structural equation estimates may become uninterpretable or counterintuitive—reflecting the confounding effects associated with poor measurement rather than the strength of relationships between variables measured.

Approaches to establishing construct validity can be broadly classified as "classical" and "contemporary" (Bagozzi, et al., 1991). *Classical* approaches include Campbell and Fiske's (1959) MTMM analysis, analysis of variance (ANOVA), and common factor analysis. *Contemporary* approaches include a variety of confirmatory factor models utilizing maximum likelihood estimation. Importantly, this classification of approaches is not meant to imply value; rather, it is meant to imply evolution and research intent. Specifically, as more is known about the theoretical and measurement properties of scales and their underlying constructs, methods for empirically evaluating these associations should evolve from the exploratory nature of classical techniques to the more exacting and confirmatory nature of contemporary techniques. As is discussed in subsequent sections, traditional approaches to construct validity do not always provide clear metrics for determining the quality of measurement. Hence, their use is more appropriate in exploring or discovering potential latent structures among indicators (Jöreskog and Sörbom, 1989). As these structures emerge and become theoretically grounded, contemporary techniques provide a means to statistically test theorized relationships against observed data. Within this lens of analysis, it is possible to more precisely evaluate the measurement properties of developed scales and subsequently the underlying theory that explains their structure.

Classical approaches to establishing construct validity

Within their study, Adams, et al. (1992) rely on two "classical" approaches for construct validation, namely, Campbell and Fiske's (1959) MTMM and common factor analysis. Utilizing the MTMM technique, convergent validity was evidenced by the high magnitude of correlations within trait and technology (i.e., inter-item correlations) across the two samples of E-Mail and V-Mail users. In turn, discriminant validity was evidenced by the low number of item correlations that were higher between trait and technology than within trait and technology. Based on these observations, it was concluded that the items seem to represent two factors and that the scales are relatively good in discriminating between very similar technologies. Although not a complete MTMM analysis (the methods were identical across samples), this approach is very typical of the technique (Campbell and Fiske, 1959).

MTMM has been widely used in the social sciences and contains a high degree of face validity. However, its use as a sole or confirmatory metric of measurement properties has a number of shortcomings (Bagozzi and Phillips, 1982; Bagozzi, et al., 1991). First, the procedure obscures the distinction between the validity of a concept and its measurement. Each observed correlation is a summary indicator of the effects of a number of factors, including spurious or extraneous elements. In other words, the individual correlations and their patterns reflect the relationships implied by latent sources, as well as the influence of both random and systematic error. Unless the errors are of small magnitude, it is not always possible to ascertain whether convergent and discriminant validity have been obtained. Also, use of the MTMM matrix does not provide criteria that are as conclusive as desired for assessing construct validation. By focusing on the number of times selected correlations are greater than certain others, the MTMM procedure neglects potentially important differences in magnitude between pairs of correlations. As noted by Bagozzi, et al. (1991), these magnitudes are functions of convergent and discriminant validity.

Another widely used classical validation technique is exploratory (also known as common) factor analysis (EFA) with varimax (orthogonal) rotation (Bagozzi and Phillips, 1982). EFA is useful in discovering potential latent sources of variation and covariation in observed measurements. Scales with good measurement properties should exhibit high factor loadings or "converge" on the latent factors of which they are indicators. Conversely, these same indicators should exhibit small loadings on factors that are measured by differing sets of indicators. Respectively, such results provide evidence of convergent and discriminant validity of scale items. Such a procedure was conducted by Adams, et al. (1992) and resulted in a two-factor structure with factor loadings that demonstrated strong evidence of convergent and discriminant validity.

Although this method represents a more rigorous assessment of measurement properties, it too has a number of significant shortcomings. As noted by Bagozzi and Phillips (1982), common factor analysis with orthogonal rotation assumes uncorrelated traits or factors, and its application to data exhibiting correlated factors can produce distorted factor loadings and incorrect conclusions regarding the number of factors.[1] In addition, exploratory factor models provide no explicit test statistics for ascertaining whether convergent and discriminant validity are achieved (also one of the more fundamental criticisms of MTMM). Further, the model does not permit one to easily partition variation that results from trait, method, and/or random error. Hence, these potentially important sources of indicator variation are lost in the factor solution. Finally, because the model searches for factors in an exploratory manner, each indicator is expressed as a function of all trait factors. Thus, the estimates obtained for factor loadings are not unique (i.e., the solution obtained is only one of an infinite number of possible solutions). Only by allowing a certain number of the indicators to be a function of a single trait factor can unique solutions be obtained. Importantly, unique solutions are hypothesis tested against data and are not a priori impositions such as those found in exploratory analysis (i.e., forcing a factor solution). Such factor models are the foundation for contemporary (structural equation modeling) approaches to construct validation.

[1] Common factor analysis may also be performed using oblique rotation. This factor solution assumes that factors are correlated. However, within the study by Adams, et al. (1982), orthogonal factor rotation is utilized.

Contemporary approaches to establishing construct validity

Over the last decade, use of structural equation modeling has been rapidly growing in psychology and the social sciences. The ability of these models to estimate and respecify multiple and interrelated dependence relationships as well as unobserved concepts make them quite useful for the complex research designs that characterize these scientific domains. Perhaps the most well-known and widely used model and accompanying software for generating these estimates is Jöreskog and Sörbom's (1989) LISREL. In its most general form, LISREL consists of two distinct parts: the measurement model (or confirmatory factor model) and the structural equation model. The measurement model specifies the relations of the observed measures (or indicators) to their posited underlying constructs. The structural equation model specifies the causal relationships of the constructs to one another as posited by underlying theoretical principles.

Although maximum likelihood estimates for both of these models can be generated simultaneously by the LISREL program, it is generally recommended that the measurement model first be assessed and "fixed" before estimation of the structural model (Anderson and Gerbing, 1988; Burt, 1976; MacCallum, 1986). A typical sequence in the assessment of the measurement model involves: (1) the development of an a priori model, representing the hypothesized pattern of relationships between observed and latent variables, (2) the fitting of the prespecified model to sample data, (3) the evaluation of the solution in terms of its parameter estimates and goodness of fit; and (4) the modification of the model to improve its parsimony and/or its fit to the data. The rationale for this approach is the avoidance of possible interaction between measurement and structural models. If indicators contain low levels of reliability or unmodeled multiple factor loadings, the potential exists for within-construct versus between-construct effects in estimation. These effects can be substantial and provide what Burt (1976) terms "interpretational confounding." Such effects may be in evidence within the study by Adams and his colleagues (1992). Rather than assessing the validity of the indicators through a confirmatory measurement model, the authors relied on the "classical" techniques outlined previously. Importantly, classical and contemporary techniques can yield differing conclusions regarding the measurement properties of indicators (Bagozzi, et al., 1991). Given an erroneous assumption regarding construct validity or factor structure, it is plausible that the inconsistent results (particularly among path coefficients in study 2) observed by Adams, et al. (1992) are attributable to unstable or improperly modeled measurement properties rather than influences implied in the structural equation model. Given this possibility and the well-developed theory surrounding these scales, it seems worthwhile to re-examine this data utilizing the confirmatory factor approach.

Testing the Measurement Model: Confirmatory Factor Analysis

Utilizing the notation of Jöreskog and Sörbom (1989), the confirmatory measurement model associated with Davis' (1989) scales can be expressed in matrix form as: $X = \Lambda_X \xi + \delta$ where X is a column vector of 10 indicators, ξ is a column vector of 2 constructs, δ is a column vector of 10 random errors, and Λ_X is a 10 x 2 matrix of coefficients relating each indicator to its posited construct. As illustrated in Figure 1, we postulate that the first six indicators (X_1-X_6) load only on the latent construct of usefulness (ξ_1), while the remaining four indicators load only on ease of use (ξ_2). We also postulate that some

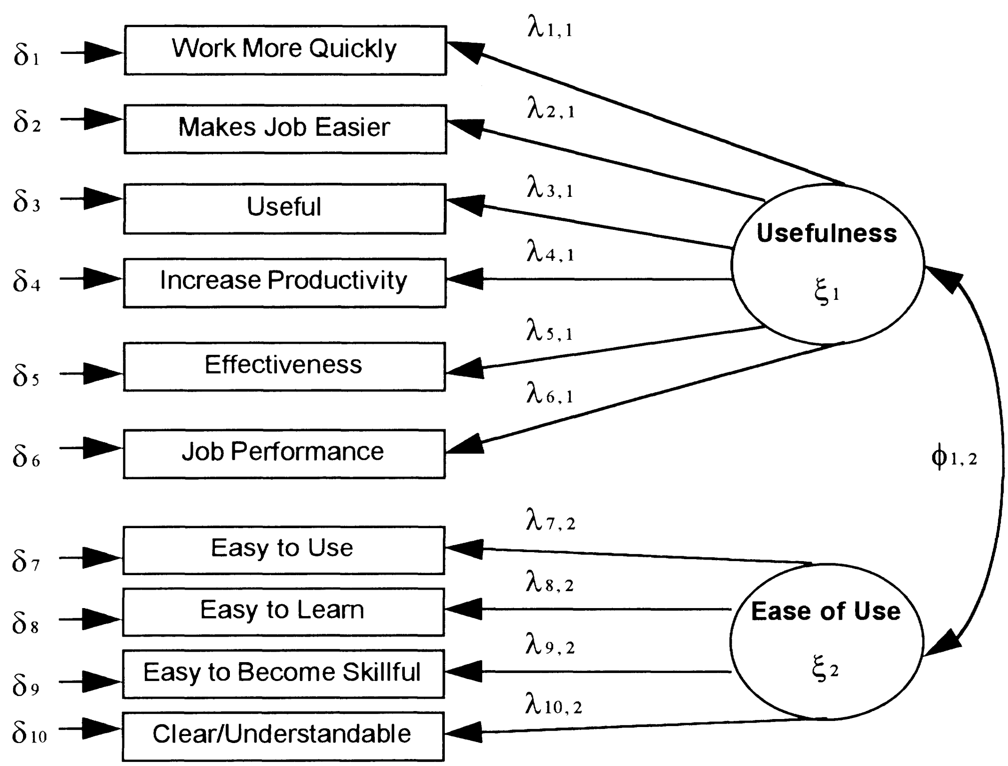

Figure 1. Two-Factor Confirmatory Measurement Model: Perceived Usefulness, Ease of Use

component of measurement error is inherent within the items through the vector of error terms (δ). With the general assumptions of uncorrelatedness between errors and constructs, multivariate normality, and an expected value of zero for δ and ξ, the 10 x 10 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 a 2 x 2 covariance matrix of latent constructs (ξ) and $\Theta\delta$ is a 10 x 10 diagonal matrix of measurement error variances. It is this set of matrices that are prespecified by the researcher and subsequently used by LISREL in generating maximum likelihood estimates.

The overall fit of a hypothesized model can be tested by using the maximum likelihood chi-square statistic provided in the LISREL output. 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 correlations between indicators, $\Sigma(\theta)$ is the implied correlation 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 correlations among indicators are well-modeled by the specified pattern matrices (Λ^X, Φ, $\Theta\delta$). Conversely, rejection of Ho implies poor model fit. Thus, in a general sense, smaller chi-square values are indicative of better-fitting models. The chi-square statistic is sensitive with respect to large sample sizes and models with large numbers of indicators. 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 should also be considered in assessing model adequacy (Bollen, 1989; Jöreskog and Sörbom, 1989).

Table 1 provides a summary of the model fit measures observed when the hypothesized two-factor structure illustrated in Figure 1 is applied to the correlation matrices observed by Adams, et al. (1992). In general, these measures suggest that the hypothesized model is a poor fit of the observed correlations in both E-Mail and V-Mail samples. Such conclusions are inconsistent with those obtained through the classical techniques employed by Adams, et al., and can perhaps be largely attributed to the methodological issues previously outlined.

Respecifying the Measurement Model

Given the statistical evidence of poor model fit, attention must now turn to respecification of the measurement model. This process is termed *specification search* (MacCallum, 1986) and is intended to detect and correct specification errors that represent a lack of correspondence between a proposed model and the true model characterizing the variables under study. Such analyses are not confirmatory, and a model arrived at through the specification search is not "confirmed" in any real sense. Rather, such analyses should be viewed as data-driven exploratory model fitting. Hence, resulting models cannot be statistically tested with any degree of validity, their goodness of fit and substantive meaning must be evaluated with caution, and their validity and replicability are open to question. To achieve any degree of substantive validation, models resulting from specification searches must be cross-validated.

In structuring the search for this analysis, the guidelines suggested by MacCallum (1986) and Anderson and Gerbing (1988) are followed.[2] In general, these observers recommend analysis of modification indices and standardized residuals as a starting point in determining specification error. Additionally, it is recommended that only one modification be made at a time during the search because a single change in a model can affect other parts of the solution. Finally, it is suggested that new parameters be added prior to deleting parameters. In other words, modifications should be made to improve the fit of the model prior to improving parsimony.

A useful diagnostic for initially locating the source of model misspecification is the patterning of

[2] These are only general guidelines and should not be construed as an optimal method in conducting specification searches. As of yet, such techniques have yet to emerge within the psychometric literature (MacCallum, 1986).

Table 1. Measures of Model Fit: Original Two-Factor Model

Fit Measure	Recommended Values	E-Mail Sample	V-Mail Sample
Chi-Square	$p \geq .05$	143.825 ($p = .001$)	70.768 ($p = .001$)
Chi-Square / df	≤ 3.0	4.230	2.081
Goodness of Fit	$\geq .90$	.807	.829
Adjusted Goodness of Fit	$\geq .80$	.688	.723
Fit Criterion	≤ 1.0	1.253	1.051
Bollen's Normed Index Rho	$\geq .90$	.823	.821
Root Mean Square Residual	≤ 1.0	.050	.053

residuals between the estimated (Σ) and observed (S) correlation matrices. The difference between the elements of this fitted matrix and the observed correlation matrix are *fitted residuals.* Given the potential for varying units of measurement between model indicators, standardized residuals are more commonly utilized and are easily obtained from the LISREL output.[3] Table 2 presents the pattern of standardized residuals observed in both E-Mail and V-Mail samples. The upper portion of the matrix contains residuals associated with the E-Mail sample; the lower contains those associated with the V-Mail sample. All values considered high (i.e., over 2.58) are highlighted in bold.

Given the need to validate a respecified model with an additional sample, only the top portion (E-Mail sample) of the matrix in Table 2 is utilized in deriving a modified factor structure. Examining this portion of Table 2, it is readily apparent that the two-factor model is not adequate in explaining the correlations between "Job Performance" and "Effectiveness."[4] In many instances, such a large residual between items is indicative of under-specification or need for an additional factor (Anderson and Gerbing, 1988; MacCallum, 1986). Based on the magnitude of this residual relative to the others, a three-factor solution was estimated, with "Job Performance" and "Effectiveness" indicators loading on an additional factor (termed "Effectiveness"). This re-estimated model resulted in a significantly improved fit over the two-factor model as evidenced by the reduction in chi-square (χ^2 difference = 25.45) and improvement in Akaike's Information Criterion[5] (AIC difference = 10.72). However, even with this incremental improvement in fit, the need for further refinement was evidenced by unfavorable measures of overall fit.

Analysis of standardized residuals associated with the three-factor solution revealed potential for further improvement in model fit through elimination of the indicators "Work Quickly" and "Understandable." Three related observations provide the rationale for this modification. First, high residuals between these two indicators and all other indicators suggest that their respective intercorrelations are not well-modeled. That is, the two indicators are not converging with others in explaining the latent sources of variation. Second, LISREL estimates of factor loadings between these indicators and their associated constructs suggest high levels of error variance (i.e., low reliabilities). Hence, these indicators are not capturing significant amounts of systematic variation in their respective constructs. (Jöreskog and Sörbom, 1989). Finally, the pattern and magnitude of original correlations among these and other indicators seem to further substantiate their elimination. Specifically, the indicator "Understandable" exhibits low intercorrelation with other measures of ease of use. Additional-

[3] A *standardized residual* is a fitted residual divided by its asymptotic standard deviation. Each standardized residual can be interpreted as a standard normal deviate and considered "large" if it exceeds 2.58 in absolute value.

[4] This extremely high residual was robust in that it also held true for the V-Mail sample (see Table 2).

[5] Akaike's (1987) Information Criterion (AIC) is based on information theory and is a comparative measure between models with differing numbers of constructs and/or indicators. The AIC will always be negative with values closer to zero, indicating better fit and greater parsimony (i.e., lack of "overfitting").

Table 2. Standardized Residuals: Hypothesized Two-Factor Model

	WkQuik	JobEasy	Useful	IncProd	Effectiv	Perform	EasyUse	EasyLn	Skillfl	Underst
WkQuik		**3.16**	−.91	−1.50	.107	−.45	.48	−1.82	**2.77**	−1.53
JobEasy	2.08		−.23	−1.03	−1.06	**−3.27**	**3.70**	−.01	2.37	1.86
Useful	−2.23	1.23		1.88	−.39	.30	.63	**−2.61**	.15	−1.30
IncProd	1.71	.434	−.96		−1.06	.33	.27	1.21	−.82	1.25
Effectiv	.76	−.82	−.96	.01		**4.8**	−1.32	−.87	−2.30	.25
Perform	−.99	−2.23	−.46	−.26	**3.50**		**−2.82**	.96	−2.24	1.27
EasyUse	−1.78	.02	2.44	−2.10	−1.06	.73		−2.15	.62	1.17
EasyLn	−1.89	−1.05	.93	.21	−1.86	.68	1.32		2.18	1.40
Skillfl	−.24	.31	.55	**2.69**	.44	.46	−1.64	.11		**−3.64**
Underst	1.46	.19	1.45	.54	−2.47	−1.19	.54	−.88	.33	

ly, "Works More Quickly" does not seem to converge strongly with other item measures of usefulness.

Fitting the reduced (eight indicator) three-factor model to the observed correlations results in significant improvement in incremental fit measures when compared to the full (10 indicator) three-factor model. Specifically, this model specification results in a significant reduction in chi-square over the earlier model (χ^2 difference = 84) as well as significant improvement in AIC (AIC difference = 34). Table 3 provides a summary of the overall fit measures observed for the reduced (eight indicator) model. Given the rather large sample size of E-Mail users (as mentioned earlier, the chi-square statistic is extremely sensitive to sample size), these measures provide evidence of relatively strong model fit.

To validate the respecified model, confirmatory factor analysis was applied to the observed correlations of the V-Mail sample. As shown in Table 3, all measures suggest a relatively strong fit of the hypothesized model to the observed correlations. Additionally, no standardized errors greater than 1.0 were observed. Hence, we can conclude statistically and with a certain degree of validity that the three-factor structure is a plausible representation of the observed correlation matrix.

Given the fitted model, its psychometric properties can now be assessed. Figure 2 is an illustration of the fitted model along with its estimated error variances, factor loadings, and correlations between constructs. The shared variance or reliability between each indicator and its respective underlying construct can be obtained by squaring the factor loading (called the squared multiple correlation). In general, evidence of convergent validity is achieved when the squared multiple correlations are greater than .50 and/or a significant t-value is observed for each indicator (Bollen, 1989; Jöreskog and Sörbom, 1989). In this analysis, all paths test highly significant ($t > |2.00|$) and all but one ("Easy To Become Skillful") are well above the .50 cutoff value. Hence, the three-factor structure seems to possess qualities of convergent validity.

Table 3. Measures of Model Fit: Reduced Three-Factor Model

Fit Measure	Recommended Values	E-Mail Sample	V-Mail Sample
Chi-Square	$p \geq .05$	58.141 (p = .001)	27.601 (p = .09)
Chi-Square / df	≤ 3.0	3.423	1.621
Goodness of Fit	$\geq .90$	.906	.903
Adjusted Goodness of Fit	$\geq .80$	.803	.807
Fit Criterion	≤ 1.0	.501	.415
Bollen's Normed Index Rho	$\geq .90$	.907	.907
Root Mean Square Residual	≤ 1.0	.031	.040

Note: Ease of Use

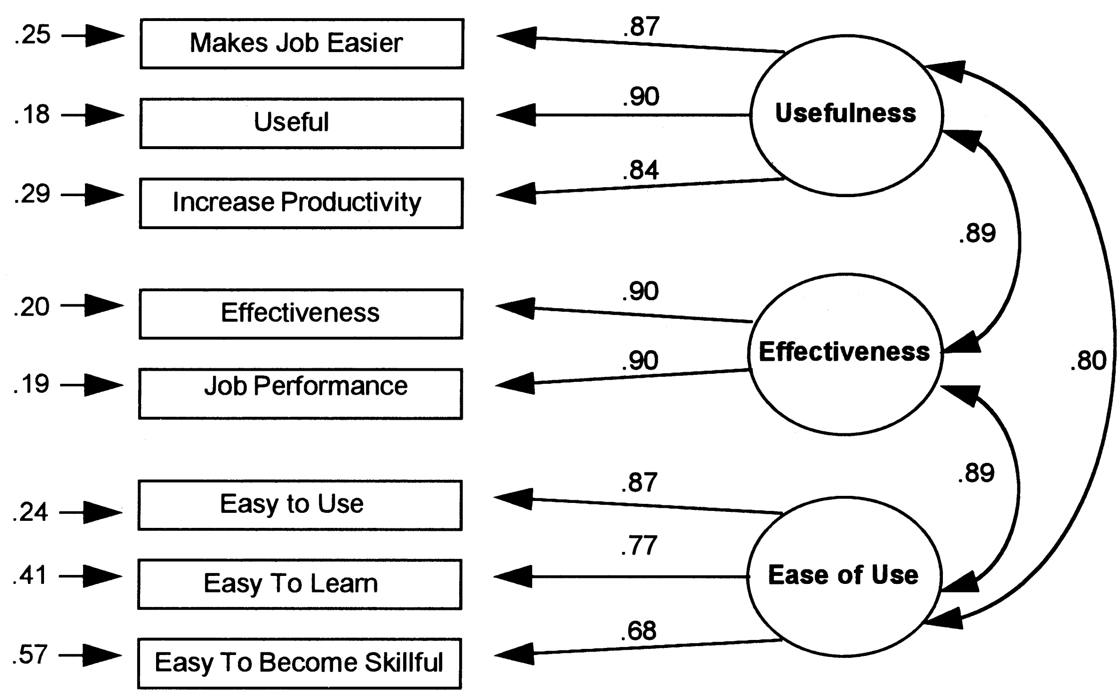

Figure 2. Reduced Three-Factor Model: V-Mail Sample[6]

Discriminant validity is assessed by fixing the correlation between various constructs at 1.0 then re-estimating the measurement model. Thus, three separate constrained models must be constructed and compared to the estimated unconstrained model. If the chi-square value for the constrained models is significantly higher than the unconstrained model illustrated in Figure 2, then evidence of construct unidimensionality is realized (Bagozzi and Phillips, 1982). This test is particularly important in this analysis because we have hypothesized the addition of a construct termed "effectiveness."

In each of these three comparisons, the chi-square values were highly significant ($\chi^2 (1) = 15.55$ $p<.001$; $\chi c^2 (1) = 10.23$ $p<.001$; $\chi^2 (1) = 6.53$ $p<.001$). Thus, we can conclude that each of the hypothesized factors is significantly different from the others. Alternatively stated, the three-factor structure exhibits properties of discriminant validity.

In summary, the confirmatory approach taken in this analysis has yielded three important findings. First, the correlations observed in study 1 of Adams, et al. (1992) do not appear well-modeled by the two-factor structure postulated by Davis (1989). Such a conclusion is in contrast to that which might have been reached through classical construct validation techniques. Second, a major source of the first result seems to be the existence of a third underlying construct termed "Effectiveness," which governs the pattern of correlations between "Job Performance" and "Effectiveness." Another potential source of model error is found in the indicators "Work More Quickly" and "Clear and Understandable." Patterns of correlations along with standardized residuals suggest elimination of these scales due to low reliability. Finally, a respecified eight indicator, three-factor structure seems well-suited to the underlying pattern of correlations. This structure was derived in an exploratory vein and then validated in a confirmatory analysis. The three-factor model exhibits sound psychometric properties and a certain degree of face validity. Importantly, any or all of these results can seriously confound the subsequent structural modeling of latent constructs. Particularly, in significant path coefficients or instability of the

[6] Factor loadings, error variances, and construct correlations were essentially the same in the E-Mail sample.

structural model across samples may be symptoms of an underlying measurement problem (Anderson and Gerbing, 1988). To eliminate these influences, measurement models must be rigorously assessed and, if necessary, respecified. Such analyses build upon the exploratory foundations of classical construct validation methodologies by providing a statistical and less ambiguous basis for measurement assessment and theory testing.

As this analysis has demonstrated, determining the structure of psychological constructs such as "ease of use" and "usefulness" is a complex activity. However, understanding how such concepts behave over varying sets of users and technologies is of critical importance in accurately explaining levels of usage. The findings reported in this paper and by Adams, et al. (1992) should serve to remind interested observers that no absolute measures for these constructs exist across varying technological and organizational contexts. Instead, it seems plausible that both task and user characteristics alter the nature and importance of perceptions that explain technology use. Such findings in no way diminish the value of Davis' (1989) original scales or the value of identifying measures that explain technology acceptance. Instead, they challenge the IS community to futher explore the nature and specific influences of factors that may alter the "user perception-usage" equation.

References

Akaike, H. "Factor Analysis and AIC," *Psychometrika* (52:3), September 1987, pp. 317-332.

Adams, D.A., Nelson, R.R., and Todd, P.A. "Perceived Usefulness, Ease of Use, and Usage of Information Technology: A Replication," *MIS Quarterly* (16:2), June 1992, pp. 227-247.

Anderson, J.C. and Gerbing, D.W. "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach," *Psychological Bulletin* (103:3), May 1988, pp. 411-423.

Bagozzi, R.P. and Phillips, L.W. "Representing and Testing Organizational Theories: A Holistic Construal," *Administrative Science Quarterly* (27:3), September 1982, pp.459-489.

Bagozzi, R.P. Yi, Y., and Phillips, L.W. "Assessing Construct Validity in Organizational Research," *Administrative Science Quarterly* (36:3), September 1991, pp. 421-458.

Bollen, K. A. *Structural Equations With Latent Variables*, Wiley, New York, NY, 1989.

Burt, R. S. "Interpretational Confounding of Unidimensional Variables in Structural Equation Modeling," *Sociological Methods and Research* (5:1), August 1976, pp. 3-51.

Campbell, D.T. and Fiske, D.W. "Convergent and Discriminant Validity by the Multitrait-Multimethod Matrix," *Psychological Bulletin* (56:2), March 1959, pp. 81-105.

Cronbach, L.J. "Coefficient Alpha and the Internal Structure of Tests," *Psychometrika* (16:3), September 1951, pp. 297-334.

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

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

MacCallum, R. "Specification Searches in Covariance Structure Modeling," *Psychological Bulletin* (100:1), July 1986, pp. 107-120.