University Libraries
University of Nevada, Las Vegas

Management, Entrepreneurship and Technology Faculty Publications

Management, Entrepreneurship and Technology

2008

Formative versus reflective measurement: Comment on Marakas, Johnson, and Clay (2007)

Andrew M. Hardin
University of Nevada, Las Vegas, andrew.hardin@unlv.edu

Jerry Cha-Jan Chang
University of Nevada, Las Vegas, jchang@unlv.nevada.edu

Mark A. Fuller
University of Massachusetts, Amherst

Follow this and additional works at: https://digitalscholarship.unlv.edu/met_fac_articles

Part of the Business Commons, Community-Based Research Commons, and the Computer Sciences Commons

Repository Citation

Hardin, A. M., Chang, J. C., Fuller, M. A. (2008). Formative versus reflective measurement: Comment on Marakas, Johnson, and Clay (2007). *Journal of the Association for Information Systems, 9*(9), 519-535.
https://digitalscholarship.unlv.edu/met_fac_articles/6

This Article is brought to you for free and open access by the Management, Entrepreneurship and Technology at Digital Scholarship@UNLV. It has been accepted for inclusion in Management, Entrepreneurship and Technology Faculty Publications by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.

This article originally appeared in the *Journal of the Association for Information Systems,* 9(9), 519-535

Used with permission from the Association for Information Systems, Atlanta, GA; 404-413-7444; www.aisnet.org. All rights reserved.

Journal of the Association for Information Systems
JAIS

IS Research Perspective

Formative vs. Reflective Measurement: Comment on Marakas, Johnson, and Clay (2007)*

Andrew M. Hardin
College of Business
University of Nevada, Las Vegas
andrew.hardin@unlv.edu

Jerry Cha-Jan Chang
College of Business
University of Nevada, Las Vegas
jchang@unlv.nevada.edu

Mark A. Fuller
College of Business
Washington State University
mark@wsu.edu

Abstract

In a recent issue of the Journal of the Association for Information Systems, Marakas, Johnson, and Clay (2007) presented an interesting and important discussion on formative versus reflective measurement, specifically related to the measurement of the computer self-efficacy (CSE) construct. However, we believe their recommendation to measure CSE constructs using formative indicators merits additional dialogue before being adopted by researchers. In the current study we discuss why the substantive theory underlying the CSE construct suggests that it is best measured using reflective indicators. We then provide empirical evidence demonstrating how the misspecification of existing CSE measures as formative can result in unstable estimates across varying endogenous variables and research contexts. Specifically, we demonstrate how formative indicator weights are dependent on the endogenous variable used to estimate them. Given that the strength of formative indicator weights is one metric used for determining indicator retention, and adding or dropping formative indicators can result in changes in the conceptual meaning of a construct, the use of formative measurement can result in the retention of different indicators and ultimately the measurement of different concepts across studies. As a result, the comparison of findings across studies over time becomes conceptually problematic and compromises our ability to replicate and extend research in a particular domain. We discuss not only the consequences of using formative versus reflective measures in CSE research but also the broader implications this choice has on research in other domains.

Keywords: Computer self-efficacy, formative measurement, reflective measurement, construct development, generalizability, reliability

* Rudy Hirschheim was the accepting senior editor. David Salisbury and two anonymous reviewers were the reviewers. This was submitted on December 20, 2007 and went through two revisions.

Volume 9, Issue 9, pp. 519-534, September 2008

Formative vs. Reflective Measurement: Comment on Marakas, Johnson, and Clay (2007)

1. Introduction

In a recent issue of the *Journal of the Association for Information Systems*, Marakas, Johnson, and Clay (2007) presented an interesting and important discussion on the measurement of computer self-efficacy (CSE). While many of their suggestions should be viewed as representing best practices for researchers utilizing the CSE construct, we believe their recommendation to measure CSE using formative indicators deserves additional discussion before being adopted by researchers. The need for additional dialogue on this issue extends beyond CSE research, as the decision to employ reflective or formative indicators is also an important consideration for researchers in other domains.

Despite recent endorsements of formative measurement (Diamantopoulos et al., 2008; Jarvis et al., 2003; Podsakoff et al., 2003), other researchers have begun to question the validity of formative measurement in general (Bagozzi, 2007; Wilcox et al., 2008). In fact, some researchers have suggested that whenever possible, reflective, rather than formative indicators should be used (Bagozzi, 2007; Howell et al., 2007b). This is because formative indicators' weights are dependent on the particular outcome variable used to estimate them. As a consequence, the meaning of formatively measured constructs can change substantially from study to study, potentially hindering scientific progress (Howell et al., 2007a). Thus, the use of formative measurement "can be a fatal flaw in theory testing" (Howell et al., 2007a p. 245).

In the case of Marakas et al. (2007), the potential problems with the specification of CSE as a formative construct are two-fold. First, CSE is a psychological concept. Because psychological concepts are underlying factors that give rise to observed scores, their indicators tend to be recognized as being reflective (Fornell and Bookstein, 1982). Second, while the use of formative indicators had the advantage of maximizing the variance explained in the outcome variable, the use of formative indicators to measure CSE was disadvantageous in that a different set of indicator weights would almost certainly have been produced if a *different* endogenous variable or sample had been used during the validation process. Based on the instrument development procedures employed, the change in indicator weights would have resulted in a different set of indicators being retained for the CSE latent constructs. Because the meaning of formative latent constructs relies on the indicators used (Diamantopoulos and Winklhofer, 2001), changing indicators changes the meaning of the CSE constructs. Such changes in meaning represent a significant threat to both construct and external validity in future studies (Shadish et al., 2002), as researchers cannot be certain as to the true meaning of the constructs being measured.

The implications of employing formative measurement extend beyond applicability to the CSE construct. In general, researchers should carefully consider the advantages and disadvantages of using formative or reflective indicators because the purposes of these methods differ (Bagozzi, 2007; Howell et al., 2007b). Specifically, formative indicators should be used when the researcher's desire is to explain abstract or unobserved variance at the latent construct level, while reflective indicators should be used when the desire is to account for variance among observable indicators (Diamantopoulos and Winklhofer, 2001). In the current study, we provide empirical examples illustrating how the purposes of these measurement methods differ, and how their misapplication may threaten both construct and external validity.

This paper is organized as follows. First, we review the results and recommendations of Marakas et al. (2007). Next, we discuss the properties and validation principles of formative measurement, and based on this discussion, we then discuss why formative measurement may not be suitable for CSE research. We follow this discussion with a series of empirical comparisons of formative and reflective CSE measures that illustrate how the conceptual meaning of these constructs can change as a result of changes in the endogenous variables, samples, and contexts. Finally, we discuss the implications of our research.

2. Marakas et al. Study Description

The purpose of the Marakas et al. (2007) study was to ascertain the properties of various CSE measures for both isolating the CSE construct and capturing variance in performance. The authors suggested that as a construct, CSE should be formative rather than reflective. This claim was substantiated with theoretical arguments and the visual examination of two existing CSE measures. As a first step in the measurement validation process, Marakas et al. (2007) generated formative items to measure Windows, word processing, database, and Internet CSE. Indicators for existing reflective measures of spreadsheet CSE (Johnson and Marakas, 2000) and general CSE were respecified as formative. PLS-Graph 3.0 was used to evaluate the contribution of the formative indicators to the respective constructs. Formative indicators were retained based upon the significance of their contribution to the target construct, as well as on the authors' collective belief that some indicators were instrumental to the construct. For the Windows, word processing, Internet, and general computer self-efficacy (GCSE) measures, only those indicators that significantly contributed to the construct (based on t-test results from bootstrapping) were retained. For the spreadsheet and database measures, only three out of the nine total items were retained based on t-tests, the other six items were retained based on the authors' beliefs that the indicators were instrumental to the constructs.

Marakas et al. (2007) concluded with specific recommendations on the development of CSE measures and their proper use in research studies. Among the contributions cited by the authors is the creation of a new set of formatively specified CSE measures.

3. Properties of Formative Measurement

Changes in formative indicators are suggested to cause changes in latent constructs (Nunnally and Bernstein, 1994). In other words, the direction of causality is from the indicators to the construct, and it is the collection of formative indicators that jointly determines the empirical and conceptual meaning of the latent construct (Jarvis et al., 2003). Formative indicators need not covary, and, in fact, can be mutually exclusive. Formative indicator weights are estimated such that they rely on other variable(s) or constructs(s) in the structural model (Bagozzi, 2007; Chin, 1998; Diamantopoulos and Winklhofer, 2001; Hair et al., 1998; Howell et al., 2007b). Thus "the meaning of the latent construct is as much a function of the dependent variable as it is a function of its indicators" (Heise, 1972). Formative measures are designed to capture the latent construct in its entirety, and as a natural consequence, dropping indicators is said to alter their conceptual meaning (Diamantopoulos and Winklhofer, 2001; Jarvis et al., 2003).

These theoretical properties are said to render formative measurement most useful for studies employing constructs that are conceived of as explanatory combinations of indicators that are determined by a combination of variables such as socioeconomic status,[1] population change (Diamantopoulos and Winklhofer, 2001; Fornell and Bookstein, 1982), or faculty performance (Nunnally and Bernstein, 1994). The concept of faculty performance provides an example of a construct suggested as suitable for formative measurement, since faculty performance is typically formed by the explanatory combination of dimensions related to scholarly productivity, teaching, and service to the university (Nunnally and Bernstein, 1994). Each of these dimensions uniquely contributes to faculty performance, and one would not necessarily expect all three factors to change together, for it is possible that a faculty member may excel in one area but be average or even deficient in the other two. Further, it is not faculty performance that causes teaching, but rather the other way around (similar to the argument made by Borsboom et al. (2003) when referencing the causal relationship between SES and salary).

While constructs that are conceived as explanatory combinations of indicators are best measured formatively, they are believed to be less appropriate for measuring psychological constructs such as

[1] Some researchers argue that formative measurement is problematic even for measures of SES. These researchers suggest that when an indicator such as education is the only significant contributor to a formatively measured construct, the construct is essentially a measure of education rather than of SES (Howell et al., 2007b).

attitude or personality, because such concepts "are typically viewed as underlying concepts that give rise to something that is observed" (Fornell and Bookstein, 1982, p. 442). Thus, when measuring psychological constructs, reflective indicators are recommended (Diamantopoulos and Winklhofer, 2001; Fornell and Bookstein, 1982).

4. Validation of Formative Measures

Formative measurement validation will also influence the choice of formative versus reflective measurement because establishing the conceptual definition of constructs is a key component of determining whether a construct should be measured formatively or reflectively (Diamantopoulos and Siguaw, 2006). It is theory that should drive the measurement development process (Law and Wong, 1999). Determining whether constructs should be measured formatively or reflectively early in the measurement development process is important because the indicators used to measure the construct may look different depending on the measurement method selected (Diamantopoulos and Siguaw, 2006). While some guidelines have been suggested for developing formative measures (Diamantopoulos and Winklhofer, 2001; Fornell and Bookstein, 1982; Jarvis et al., 2003; Loch et al., 2003; Petter et al., 2007), this topic is still recognized as an open empirical issue (Gefen and Straub, 2005). We discuss some of these guidelines below.

The development of formative indicators should begin with a formal literature review and the use of expert panels and techniques such as Q-sorting to ensure content validity (Petter et al., 2007). Merely reversing the direction of the path between the construct and its measures is inappropriate (Diamantopoulos and Siguaw, 2006). Following the development of a set of indicators that form a census of the concepts defining the construct (Jarvis et al., 2003), indicators may then be eliminated based upon their lack of contribution to the construct during model estimation (Fornell and Bookstein, 1982). However, care should be taken to ensure that content validity is maintained when nonsignificant indicators are eliminated (Diamantopoulos and Winklhofer, 2001), "because the consequences of dropping one of the indicators are potentially quite serious" (Jarvis et al. 2003, p. 202).

Because of the properties of formative measurement, procedures used to assess the validity and reliability of reflectively measured constructs (e.g., internal consistency and factor analysis) are not appropriate for constructs with formative indicators (Diamantopoulos and Winklhofer, 2001). Instead, recommendations for formative measurement validation include the establishment of convergent and discriminant validity through procedures such as the examination of the correlations between the individual indicators and an overall measure of the target latent construct. Valid indicators should be those more highly correlated with the overall measure than with other constructs in the model. Particular attention should also be paid to predictive or nomological validity (Bollen and Lennox, 1991; Jarvis et al., 2003).

5. Problems with Formative Measures of CSE

According to Chin (1998) the choice between measuring latent constructs with formative or reflective indicators should be based on the research objectives, the substantive theory for the latent construct, and the empirical conditions. Organized around these categories, in this section we discuss potential problems with Marakas et al.'s (2007) recommendation to measure CSE using formative indicators. We support our discussion using the properties and validation of formative measurement as discussed in the prior sections. We begin by discussing CSE in terms of the research objectives.

5.1. Research objectives

Research objectives (addressing the researcher's purpose for employing a particular model) must be considered when determining whether to employ formative or reflective measures (Fornell and Bookstein, 1982). For example, researchers must decide whether to account for observed variances (in which case reflective indicators should be used), or to account for unobserved variance at the abstract or construct level (in which case formative indicators should be used).

These differences have important implications for study results, as they lead to different outcomes. When researchers conduct studies designed to account for observed variances, the model can be estimated in a true measurement sense—that is, constructs can be evaluated for their measurement properties without consideration of the structural model. In addition, traditional validation procedures can be followed, and reflective indicators can be added or dropped without changing the meaning of constructs. Conversely, if the purpose is to explain variance at the abstract or unobserved level, measurement properties must be evaluated within the structural model. In the latter case, there are implications for the generalization of study results, because the estimation of the formative indicator weights is dependent on the nomological net that is employed, and changes in indicator contribution can result in changes in construct meaning. Thus, if researchers employ measures of the same constructs across studies, indicator strength will vary, changes in the meaning of constructs across studies will result, and findings will be difficult to compare (Bagozzi, 2007; Chin, 1998; Howell et al., 2007a, 2007b).

The model employed by Marakas et al. (2007) was designed to explain the mean variance in the reflective endogenous indicator(s) by the linear composite of formative indicators used to measure the CSE constructs. Thus, the model accounts for variance at the unobserved level, and as a result, the formative indicator weights were estimated such that they best predicted the endogenous construct that was employed (Howell et al., 2007b). Because of the estimation procedures used to evaluate this model (i.e., the minimization of the residual variance in the structural equation), a greater amount of variance was explained than would have been if reflective indicators had been used to measure the CSE constructs.

At first glance this appears to be a desirable situation because it accomplished one of the Marakas et al. (2007) study objectives. However, while additional variance in performance was explained through the use of formative indicators, the indicator weights are dependent on the endogenous variable (and sample) used to estimate them. If other endogenous constructs were specified or other samples utilized, different indicator weights would likely have been found significant, because they would have been the best combination of predictors for that endogenous construct or sample (as we will demonstrate in the next section). This would have resulted in the retention of a different set of formative indicators for measuring the respective CSE constructs. Thus, the objective to capture variance in performance represents a trade-off between the generalizability of the measure and the explanation of variance in the outcome variable. In other words, the formative CSE construct proposed by Marakas et al. lacks the stability typically demonstrated by valid reflective measures.

This concern extends to future studies employing the formative indicators recommended by Marakas et al. (2007), as the significance of the weights will vary because they are dependent on a different nomological net (Chin, 1998). The question then becomes: Should the affected formative indicators be retained or eliminated? Further, if different indicators *are* retained and used to measure the CSE latent construct, should the construct then be reconceptualized to ensure construct validity? If so, will this re-conceptualization, in turn, influence the external validity of the study (Shadish et al., 2002)?

5.2. Substantive theory

The substantive theory (which addresses the underlying conceptual properties of constructs), as well as the auxiliary measurement theory (which explains the nature of the relationships between constructs and their measures), should also be considered when deciding on formative versus reflective measurement (Howell et al., 2007b). For example, psychological constructs are best measured using reflective indicators, while constructs determined by an explanatory combination of variables are best measured using formative indicators (Bagozzi, 2007; Chin, 1998; Fornell and Bookstein, 1982; Howell et al., 2007b). Further, the auxiliary theory explaining the nature of the relationship between constructs and their measures should dictate the a priori development of indicators used to measure constructs.

In terms of the conceptual properties underlying CSE, much of Marakas et al.'s (2007) argument for formatively specifying CSE measures rests upon measurement principles; however, beyond such considerations, it is critical to discuss how theory underlying the CSE construct relates to the decision to employ formative versus reflective indicators. Bandura (1997, p. 3) defines self-efficacy as a “belief in one’s capabilities to organize and execute the courses of action required to produce given attainment". Self-efficacy is rooted in Social Cognitive Theory, where it is positioned as a psychological construct that forms the major basis for people’s actions and guides people’s lives (Bandura, 1997). Its influence on behavior is suggested to be through its reinforcement of an individual’s sense of personal agency, or the influence over deliberate human action. In other words, it is a person’s sense of self-efficacy that allows him to take control over his actions. People judge their capability to complete a given behavior based upon their belief in the ability to execute a specific course of action (i.e., their self-efficacy).

Self-efficacy is developed through enactive mastery (gained through prior experience and hands-on training), vicarious experience (gained primarily through the observation of others), verbal persuasion (gained through the encouragement or discouragement of referent others), and affective states (gained through factors such as anxiety, or physiological states such as muscle pain). People cognitively integrate these four sources of information to form their self-efficacy perceptions (Bandura, 1997). Thus, *self-efficacy represents a complex psychological process that is formed and then used to guide human action*. Such a description supports the notion that self-efficacy is an underlying factor that exists apart from any attempts to measure it, and further, that changes in the self-efficacy latent construct will precede changes in the indicators used to measure it. Therefore, the underlying theory supporting self-efficacy is consistent with reflective latent variable analysis (Borsboom et al., 2003).

Measuring Self-efficacy Constructs

As discussed above, self-efficacy is consistent with other psychological concepts suggested to be best measured reflectively. During a discussion on the development of self-efficacy measures, Bandura (2005) suggests that efficacy items should accurately *reflect* the construct, and then further recommends that self-efficacy items should be correlated and that their homogeneity should be established through factor analysis. Finally, Bandura states that internal consistency reliabilities should be computed using Cronbach’s alpha, and if the alpha coefficients are low, the affected items should be discarded (consistent with classical test theory and reflective measurement). We note that reflect, intercorrelation, homogeneity, and internal consistency reliabilities are all consistent with reflective rather than formative measurement.

Marakas et al. (1998, p. 127) define CSE as “an individual’s perception of efficacy in performing specific computer related tasks within the general computing domain.” Thus, CSE is a domain specific measure of self-efficacy that reflects a person’s belief in the ability to perform specific computer tasks. CSE is developed over time and interaction with computers, and consistent with self-efficacy theory, influences the effort put forth, persistence in the face of obstacles, resilience to adversity, and whether thought patterns are self-hindering or self-aiding. It is through such processes that CSE influences levels of accomplishment in the computer domain. CSE, like self-efficacy, is therefore a psychological process that exists independently of any attempt to measure it. Thus, due to its psychological origins, we believe the formative specification of CSE is inconsistent with its substantive theory.

The second consideration related to the substantive theory surrounding the CSE construct relates to the nature of the relationship between the CSE construct and its measures. Marakas et al. (2007) conducted a visual inspection of two existing, reflective measures of CSE. The indicators were examined in terms of the comparative list of properties of formative and reflective measures proposed by Diamantopoulos and Winklhofer (2001). Following this visual examination, Marakas et al. (2007) state that “both CSE and GCSE are formative indicators” (p. 21). While the list of properties proposed by Diamantopoulos and Winklhofer (2001) has been suggested as useful during the a priori development of indicators, and as a tool for identifying misspecified measures (Jarvis et al., 2003), we believe it is inappropriate to apply these properties during a post hoc visual review of existing measures for the purposes of respecifying the indicators as formative or reflective.[2] Specifically, while

[2] Note that although Jarvis et al. (2003) identify existing reflective measures that they believe should be measured formatively using the criteria proposed by Diamantopoulos and Winklhofer (2001), they do not, in fact, respecify the measures for the purposes of employing them in empirical studies

these two types of indicators may share common aspects of the construct, their specification is driven by measurement theory and, thus, should not be examined from any perspective other than their original intent (Howell et al., 2007b).

Further, recent studies have demonstrated how the instrument development process guided by a formative rather than a reflective perspective can result in a completely different set of indicators even when drawn from the same item pool (Diamantopoulos and Siguaw, 2006). In other words, the decision to specify indicators as formative or reflective should be made prior to their use, because the theoretical underpinnings of formative versus reflective measurement are incompatible (Diamantopoulos and Siguaw, 2006; Howell et al., 2007a). Further, the guidelines suggested by Diamantopoulos and Winklhofer (2001) for determining whether measures are formative or reflective are only guidelines for development. Thus, the visual adherence of previously specified reflective indicators to the formative measurement criteria applied by Marakas et al. (2007) cannot confirm them as either formative or reflective. In essence, the criteria are necessary but not sufficient for determining the specification of measurement indicators.

5.3. Empirical conditions

Empirical conditions (factors such as multicollinearity and sample size) should also be considered when deciding on formative vs. reflective measurement. For example, multicollinearity is said to be of particular importance in terms of indicator stability. Specifically, covariance among reflective indicators is expected, and multicollinearity is not problematic given that simple regressions are used to generate indicator loadings. For formative indicators, however, multicollinearity can adversely affect the stability of indicator coefficients, because the estimation process is based on multiple regression (Fornell and Bookstein, 1982).

In the case of Marakas et al. (2007), the indicators used to measure spreadsheet CSE and GCSE are from existing, reflective measures. The spreadsheet CSE measure has been used in several prior studies and has been demonstrated to exhibit acceptable reliability and validity (as it does in the data sets analyzed in this paper). Given that the existing reflective indicators have been shown to covary, multicollinearity could be a source of concern when evaluating the spreadsheet CSE construct as formative.

5.4. Summation

Returning to our position that the choice of modeling constructs with either formative or reflective indicators should be based on: a) the research objective, b) the substantive theory for the latent construct, and c) the empirical conditions (Chin, 1998), we believe that sufficient evidence exists for questioning the treatment of the CSE construct as formative. Given the recent concern with formative measurement (Bagozzi, 2007; Howell et al., 2007b), we believe that the prudent course of action is to avoid the use of formative CSE measures until sufficient dialogue on this issue has taken place.

Recent research on formative measurement has addressed the issue of misspecifying formative indicators as reflective (Bollen, 2007; Jarvis et al., 2003; Petter et al., 2007). In the current study, we have argued that Marakas et al. (2007) misspecified existing reflective CSE measures as formative based on the theoretical underpinnings of self-efficacy. This misspecification has long-term implications for the accumulation of knowledge in the CSE area. We address these implications by demonstrating how the stability of constructs over time (e.g., varying endogenous variables and research contexts) is affected when reflective measures are respecified as formative. Specifically, we evaluate two commonly used CSE constructs both formatively and reflectively to illustrate how the conceptual meaning of these constructs may change as a result of changes in the endogenous variables, samples, and contexts.

6. Methodology

To empirically test the measurement properties of the CSE construct, we evaluated two different CSE measures across different samples, assessment periods, and endogenous variables (i.e., computer anxiety, affect, and spreadsheet performance). In all, we used three different data sets to conduct four distinct analyses. The expected relationship between CSE and each of the endogenous constructs is consistent with efficacy theory (Bandura, 1997) and has been supported empirically (Compeau and Higgins, 1995; Johnson and Marakas, 2000). Data were collected from IS students enrolled in computer skills training courses administered at universities located on the east and west coasts of the United States We conducted our analyses using PLS-Graph 3.0.

Analysis 1 description

In Analysis 1, we assessed the relationship between CSE and computer anxiety (CA) three times during a six-week software training course (n = 164). Four independent models were evaluated. The first model specified the reflective spreadsheet CSE (SCSE) measure developed by Johnson and Marakas (2000) as a predictor of CA across three assessment periods. The same relationship was then evaluated with the spreadsheet CSE indicators specified as formative. The relationship between CSE and CA was then reevaluated using the software CSE (CH) measure developed by Compeau and Higgins (1995) (both reflectively and formatively) across the three assessment periods.

Analysis 2 description

In analysis 2, we evaluated the relationship between spreadsheet CSE and CA using different software training participants (n = 388). The indicators for the CSE latent construct were again specified both reflectively and formatively.

Analysis 3 description

In analysis 3, we evaluate the relationship between spreadsheet CSE and CA and Affect (Compeau and Higgins, 1995) using yet another group of software training participants (n = 224). The indicators for the spreadsheet CSE latent construct were once again specified both reflectively and formatively.

Analysis 4 description

Finally, we assessed the relationship between spreadsheet CSE and actual performance using the same training participants used in analysis 3. Actual performance was measured using a computer-administered hands-on exam.

7. Results and Discussion

Table 1 depicts the respective beta weights and variance explained for the four analyses. In all cases the formative measures explained a greater amount of variance in the respective endogenous variables than did the reflective measures. This is expected because the use of formative indicators minimizes the residual variance in the structural portion of the model (Chin, 1998; Diamantopoulos and Winklhofer, 2001; Fornell and Bookstein, 1982; Heise, 1972), thus resulting in greater explanatory power at the latent construct level (Diamantopoulos and Winklhofer, 2001). In fact, because we utilized the same set of indicators, the reflective specification can never explain more variance than the formative one (Diamantopoulos and Siguaw, 2006).

Table 1: Betas and variance explained

Analysis 1 (n = 164)												
	Time 1				Time 2				Time 3			
	Reflective		Formative		Reflective		Formative		Reflective		Formative	
Relationship	b	VAF	b	VAF	b	VAF	b	VAF	b	VAF	b	VAF
SCSE – CA	-.629	40%	-.655	43%	-0.535	29%	-.596	36%	-.524	28%	-.577	33%
CH – CA	-.590	36%	-.661	44%	-.510	26%	-.542	29%	-.516	27%	-.622	39%
	Analysis 2 (n = 338)				**Analysis 3 (n = 224)**				**Analysis 4 (n = 224)**			
SCSE – CA	-.358	13%	-.387	15%	-.399	16%	-.421	18%				
SCSE – Affect					.412	17%	.433	19%				
SCSE– Perf									.439	19%	.551	30%

Analysis 1 results

Table 2 depicts the Analysis 1 indicator loadings and weights for the respective CSE latent constructs when specified as predictors of computer anxiety. As can be seen, the reflective indicator loadings are generally consistent across the three assessments. This result is expected given our use of two previously validated reflective measures of CSE.

Table 2: Comparison of Reflective Loadings and Formative Weights for Spreadsheet Efficacy (SCSE) and Computer Anxiety (CA) for Analysis 1

	Time 1				Time 2				Time 3			
	Reflective		Formative		Reflective		Formative		Reflective		Formative	
Indicator	Loading	t	Weight	t	Loading	t	Weight	t	Loading	t	Weight	t
SCSE1	**0.721**	17.462	0.062	0.344	**0.820**	22.823	0.023	0.119	**0.899**	49.775	-0.244	0.949
SCSE2	0.658	11.197	0.239	1.575	**0.799**	26.207	0.197	0.959	**0.903**	58.761	**0.676**	**2.628**
SCSE3	**0.715**	14.345	0.286	1.895	**0.713**	10.351	**-0.372**	**2.173**	**0.795**	20.295	-0.337	1.778
SCSE4	**0.846**	23.348	-0.105	0.497	**0.913**	58.622	**0.784**	**2.309**	**0.913**	48.291	0.250	0.900
SCSE5	**0.780**	22.072	0.284	1.945	**0.799**	20.189	**0.463**	**2.364**	**0.868**	37.909	-0.060	0.234
SCSE6	**0.852**	30.130	0.243	1.244	**0.903**	52.268	-0.257	0.915	**0.925**	60.051	**0.790**	**2.313**
SCSE7	**0.847**	26.100	-0.104	0.451	**0.918**	55.295	-0.053	0.143	**0.944**	91.696	-0.046	0.123
SCSE8	**0.801**	29.561	0.161	0.888	**0.844**	25.954	-0.127	0.622	**0.839**	29.526	0.057	0.293
SCSE9	**0.729**	13.495	0.236	1.629	**0.811**	23.565	0.345	1.605	**0.873**	39.441	-0.146	0.459
CR	0.931				0.955				0.970			
AVE	0.600				0.702				0.784			
	Endogenous Indicator Loadings				Endogenous Indicator Loadings				Endogenous Indicator Loadings			
CA1	**0.910**	74.833	**0.909**	67.438	**0.902**	51.298	**0.889**	41.568	**0.908**	57.789	**0.908**	41.532
CA2	**0.751**	13.342	**0.753**	13.134	**0.787**	17.344	**0.802**	19.178	**0.845**	22.526	**0.847**	31.657
CA3	**0.859**	28.436	**0.859**	26.989	**0.860**	22.312	**0.871**	30.720	**0.861**	24.446	**0.860**	26.175
CA4	**0.837**	26.046	**0.835**	26.116	**0.890**	45.599	**0.883**	41.352	**0.881**	37.068	**0.881**	34.952
CR	0.906		0.906		0.919		0.920		0.928		0.928	
AVE	0.707		0.708		0.741		0.743		0.764		0.764	
CH1	**0.796**	21.361	**0.526**	**3.206**	**0.794**	22.147	**0.543**	**2.222**	**0.806**	16.983	-0.014	0.056
CH2	**0.730**	15.031	-0.106	0.625	**0.755**	17.083	-0.026	0.115	**0.791**	16.493	**0.442**	**2.839**
CH3	**0.771**	17.505	**0.442**	**2.334**	**0.864**	37.747	0.200	0.791	**0.854**	33.621	0.113	0.597
CH4	**0.760**	16.947	-0.012	0.068	**0.841**	32.559	0.246	0.871	**0.846**	25.907	-0.229	1.129
CH5	**0.764**	16.859	0.249	1.837	**0.832**	24.772	0.095	0.347	**0.842**	26.943	0.216	1.028
CH6	0.693	12.342	0.204	1.347	**0.844**	26.969	-0.133	0.525	**0.838**	18.478	**-0.885**	**3.581**
CH7	**0.772**	18.030	-0.152	0.975	**0.823**	31.990	0.059	0.264	**0.836**	30.653	0.051	0.232
CH8	**0.778**	15.826	-0.259	1.408	**0.820**	23.762	0.117	0.527	**0.875**	43.947	0.247	0.890
CH9	0.640	10.448	0.000	0.002	**0.729**	13.097	-0.088	0.432	**0.773**	17.542	**0.633**	**2.719**
CH10	**0.771**	20.987	0.285	1.503	**0.763**	19.475	0.137	0.687	**0.798**	26.033	**0.500**	**2.509**
CR	0.927				0.949				0.956			
AVE	0.561				0.652				0.683			
	Endogenous Indicator Loadings				Endogenous Indicator Loadings				Endogenous Indicator Loadings			
CA1	**0.887**	38.956	**0.886**	36.949	**0.856**	27.900	**0.852**	28.308	**0.895**	43.831	**0.897**	50.199
CA2	**0.785**	17.341	**0.783**	15.360	**0.849**	30.923	**0.856**	27.612	**0.872**	31.364	**0.870**	35.271
CA3	**0.877**	32.524	**0.879**	40.716	**0.899**	37.798	**0.900**	42.271	**0.871**	28.065	**0.874**	30.221
CA4	**0.811**	20.920	**0.812**	19.935	**0.845**	27.033	**0.840**	27.347	**0.860**	22.099	**0.858**	28.385
CR	0.906		0.906		0.921		0.920		0.929		0.929	
AVE	0.708		0.708		0.744		0.743		0.765		0.765	

For the formatively specified SCSE measure, the results are not consistent. No indicator weights are significant across all three assessments. For example, at time 1 no indicator weights significantly contribute to the SCSE construct. At time 2, items 3, 4, and 5 contribute significantly. At time 3, items 2 and 6 are significant contributors. Notably, none of the indicators significantly contributing to the CSE construct in the Marakas et al. (2007) study were significant contributors in any of our three assessments. Such results demonstrate the dependence of formative indicator weights on the endogenous variable (in this case even the same endogenous variable measured at different times within the same study) during the estimation process.

For the formatively specified software CSE measure (Compeau and Higgins, 1995) at time 1, indicators 1 and 3 contribute significantly; at time 2, item 1 is significant; and at time 3, indicators 2, 6, 9, and 10 are all significant contributors. Once again, these results demonstrate the reliance of formative indicator weights on the endogenous variable during different assessments within the same study. The software CSE indicators were not evaluated by Marakas et al. (2007), and thus no comparison can be made between their study and ours. However, if retention was based upon the significance of indicator weights, it would have been inappropriate to compare results across our assessments, because the conceptual meaning of the constructs would be different.

Analysis 2 results

Analysis 2 was designed to reevaluate the relationship between spreadsheet CSE and CA using a different sample (n = 388) than used for Analysis 1. Based upon the results in Table 3 (depicting the Analysis 2 indicator weights and loadings), it can be seen that all of the reflective indicators for the SCSE construct load above the standard metric of .707 necessary for the retention of reflective indicators (Hair et al., 1998). Once again, this is evidence of a properly validated SCSE measure. We can also see that formative indicators 2 and 4 significantly contribute to the SCSE latent construct. These indicators are different from those retained by Marakas et al. (2007), and further, are different from the indicators contributing significantly during the three assessments in Analysis 1. Using the significance of formative indicator weights as retention criteria, the end result would be to retain items different from those in Analysis 1.

Table 3: Comparison of Reflective Loadings and Formative Weights for Spreadsheet Efficacy and Computer Anxiety for Analysis 2

	Reflective		Formative	
Indicator	Loading	t	Weight	t
SCSE1	**0.835**	28.858	0.092	0.302
SCSE2	**0.869**	53.965	**0.473**	**2.007**
SCSE3	**0.769**	22.422	0.109	0.542
SCSE4	**0.856**	41.368	**0.444**	**2.330**
SCSE5	**0.821**	25.350	-0.103	0.442
SCSE6	**0.854**	39.378	-0.044	0.186
SCSE7	**0.842**	30.464	0.299	1.215
SCSE8	**0.846**	40.604	0.100	0.539
SCSE9	**0.829**	28.061	-0.293	1.527
CR	0.945			
AVE	0.699			
	Endogenous Indicator Loadings		Endogenous Indicator Loadings	
CA1	0.678	74.833	0.688	18.458
CA2	**0.725**	13.342	**0.712**	17.663
CA3	**0.870**	28.436	**0.875**	49.534
CA4	**0.909**	26.046	**0.905**	69.571
CR	0.876		0.876	
AVE	0.642		0.642	

This analysis demonstrates how the contribution of formative indicators can differ across studies, even when the same endogenous construct is being predicted. As a result, different indicators may be retained, and because dropping or adding formative indicators changes the meaning of latent constructs, comparing results across these two studies would be conceptually problematic.

Analysis 3 results

During Analysis 3 we evaluated the properties of the SCSE measure as a predictor of two different endogenous variables within a single study. The sample (n = 224) was different from that used in Analyses 1 or 2. Table 4 depicts the reflective loadings and formative weights for the SCSE measure as a predictor of CA and Affect. As was the case in Analysis 2, all reflective indicator loadings were above the .707 metric recommended. However, as a predictor of CA, no formative indicators contribute significantly to the construct. As a predictor of Affect, formative indicators 1 and 9 significantly contribute to the SCSE construct.

This analysis demonstrates yet another problem with the misspecification of formative indicators. Specifically, formative indicator contribution can differ across endogenous variables, even *within* the same study. Therefore, we may be observing a relationship between an endogenous variable and two completely different concepts, making comparisons of within- study relationships difficult, if not impossible, to justify.

Table 4: Comparison of Reflective Loadings and Formative Weights for Spreadsheet Efficacy and Computer Anxiety and Affect for Analysis 3

	Reflective		Formative			Reflective		Formative	
Indicator	Loading	t	Weight	t	Indicator	Loading	t	Weight	t
SCSE1	**0.820**	33.233	0.239	0.976	SCSE1	**0.826**	29.390	0.430	**2.206**
SCSE2	**0.793**	23.756	-0.163	0.740	SCSE2	**0.806**	25.076	0.218	0.884
SCSE3	**0.791**	28.635	0.121	0.408	SCSE3	**0.785**	24.414	-0.103	0.439
SCSE4	**0.852**	31.241	0.162	0.433	SCSE4	**0.845**	25.952	-0.032	0.096
SCSE5	**0.853**	30.146	-0.318	1.448	SCSE5	**0.855**	30.601	-0.052	0.264
SCSE6	**0.870**	48.802	0.509	1.713	SCSE6	**0.866**	36.440	-0.054	0.188
SCSE7	**0.840**	39.244	0.198	0.613	SCSE7	**0.832**	32.695	0.225	0.921
SCSE8	**0.799**	30.218	0.255	1.155	SCSE8	**0.801**	24.353	0.063	0.265
SCSE9	**0.836**	40.367	0.121	0.472	SCSE9	**0.839**	36.199	0.452	**2.089**
CR	0.952				CR	0.952			
AVE	0.687				AVE	0.687			
	Endogenous Indicator Loadings					Endogenous Indicator Loadings			
CA1	0.589	8.660	0.598	9.840	AFF1	**0.878**	54.037	**0.876**	52.783
CA2	**0.827**	32.186	**0.833**	31.238	AFF2	**0.899**	60.294	**0.903**	63.132
CA3	**0.852**	32.576	**0.853**	38.111	AFF3	**0.779**	19.378	**0.789**	19.372
CA4	**0.855**	37.799	**0.844**	33.227	AFF4	**0.759**	16.822	**0.754**	18.544
CR	0.866		0.866		AFF5	**0.727**	16.656	**0.719**	13.276
AVE	0.622		0.623		CR	0.905		0.905	
					AVE	0.658		0.658	

Analysis 4 results

Analysis 4 was designed to evaluate the properties of SCSE as a predictor of an objective measure of spreadsheet performance. Using the same sample as that used in Analysis 3, a time 2 measure of SCSE (administered immediately prior to the computer- delivered spreadsheet performance assessment) was specified as a predictor of performance both formatively and reflectively. Table 5 reveals that all reflective loadings are above the .707 metric. On the other hand, we can see that only formative indicators 3, 5, 6, and 8 significantly contribute to the SCSE construct. Note that the current analysis replicates Marakas et al. (2007) in that SCSE is predictive of performance. However, while Marakas et al. retained five of the original SCSE indicators, two of which significantly contributed to the construct (items 1 and 8), and three based on their judgments (items 2, 4, and 5); they did not retain items 3 and 6, which are significant in our analysis. Should we then retain the two indicators that are significant in our analysis *and* the five indicators retained by Marakas et al.? Or should we rely only on the five items, as they suggest? Taking this a step further, should future researchers rely on the original five items, or the now seven items, or should they start from scratch and reevaluate the formative indicator weights altogether?

This result demonstrates the problem with the formative SCSE measure proposed by Marakas et al. (2007), as well as formative measurement in general. The retention of formative indicators is dependent on the particular outcome variable used in the initial studies used to validate them. Given that a universe of outcome variables, settings, and samples could be utilized for this purpose, it is difficult to envision how researchers can confidently argue that the items they have retained are theoretically appropriate for capturing all the aspects of a given construct across all settings. While the alternative is to depend on the initiating researcher(s) belief that certain items are instrumental to the construct regardless of the significance of their contribution, it is not clear how important these indicators are for predicting outcomes.

Table 5: Comparison of Reflective Loadings and Formative Weights for Spreadsheet Efficacy and Performance for Analysis 4

	Reflective		Formative	
Indicator	Loading	t	Weight	t
SCSE1	**0.850**	29.383	0.011	0.050
SCSE2	**0.854**	35.553	0.068	0.366
SCSE3	**0.778**	22.768	-0.402	**2.013**
SCSE4	**0.902**	45.493	0.017	0.054
SCSE5	**0.898**	56.601	0.923	**3.523**
SCSE6	**0.898**	45.234	-0.532	**2.359**
SCSE7	**0.903**	57.617	0.211	0.628
SCSE8	**0.851**	31.352	0.693	**3.069**
SCSE9	**0.868**	29.177	-0.120	0.473
CR	0.965			
AVE	0.753			
	Endogenous Indicator Loadings		Endogenous Indicator Loadings	
Perf1	**1.000**	0.000	**1.000**	0.000
CR	N/A		N/A	
AVE	N/A		N/A	

8. Implications

The implications of the choice to use formative versus reflective measurement are significant. Researchers should fully understand the purpose of the respective measurement methods before employing them. Reflective indicators are invoked in an attempt to account for the observed variances or covariances (Fornell and Bookstein, 1982) and can be estimated in a true measurement sense (Howell et al., 2007b). As demonstrated by our results, this leads to relatively stable indicator loadings across variables and studies when reflective measures are properly developed and validated. Responses to reflective indicators change as a result of changes in the underlying construct (which exist apart from attempts to measure it), making reflective measurement appropriate for measuring psychological constructs such as attitude, personality, and in the current case, computer self-efficacy. Reflective items can be selected from the universe of items available for measuring a specific latent construct. Consistent with classical test theory, reflective items can also be dropped without altering the meaning of latent constructs. Thus, when indicators are dropped, both measurement and structural results can be generalized across studies (thus preserving external validity) and effect sizes can be used in meta-analyses.

In contrast, formative indicators are designed to minimize residuals in structural relationships (Fornell and Bookstein, 1982). Formative measures can thus be appropriately used in studies designed to maximize the explanation of unobserved variance at the latent construct level for a given outcome (and as a result minimize type II errors). However, because the estimation of formative indicator weights is dependent on other constructs, indicator retention is study-specific. As our analyses showed, different indicators were significant contributors to the same latent constructs across different assessments.

For example, during Analysis 1, at time 2 indicators 3, 4, and 5 were significant; while at time 3, items 2 and 6 were significant. If at time 2 we chose to retain only those formative indicators with significant weights, what of the indicators having significant weights at time 3? Further, if we retained both indicators with significant weights and indicators determined to be instrumental to the construct (as was done by Marakas et al.), how can we be sure that any eliminated items would not have significantly contributed to the latent construct in subsequent analyses? Alternatively, if we chose to evaluate the entire set of formative indicators for each assessment and retained both indicators that were significant, as well as those we felt were instrumental to the construct, we would be left with a different set of indicators across the respective assessment periods. Given that adding or dropping formative indicators changes the conceptual meaning of latent constructs, the end result would be the measurement of different concepts across the assessment periods. To maintain construct validity, we would then need to name the construct differently across the separate analyses, and in effect, would be evaluating the structural relationships between the endogenous variable and two different exogenous variables. As a consequence, study results cannot be compared, which, in turn, affects our ability to advance our understanding of CSE over time.

Another alternative in formative measurement is to keep one set of indicators regardless of whether or not their contribution to the construct is statistically significant. However, in such cases, how do we determine which initial set of indicators will best capture the latent construct? The choice to rely on a limited set of indicators that the researcher determines are instrumental poses an additional dilemma, because at least when retention is based upon the significance of formative indicator weights, researchers have at their disposal an empirical tool that can be used to evaluate the validity of formative indicators. If this metric is abandoned, and instead, researchers' subjective perceptions of indicator contribution are utilized, capturing the conceptual meaning of latent constructs in a consistent manner becomes even less likely. Further, if measures of the same construct are developed in parallel, and researchers' perceptions of indicator contribution differ, the result would be different sets of formative indicators for latent constructs purported to measure the same underlying concept. Regardless of whether retention is based upon the significance of indicator weights or researchers' perceptions regarding their contribution, the use of formative indicators remains problematic, and thus reflective indicators should be used until a consensus has been reached on this issue (Howell et al., 2007b).

Finally, although we have specifically addressed the elimination and retention of formative indicators throughout this paper because it relates to the instrument validation process used by Marakas et al. (2007), the problem with formative measurement extends beyond just this process. For example, even if the same set of indicators is retained across studies—whether based on the significance of the indicator weights or the researcher's belief that the indicators are instrumental to the construct—the strength of the individual indicator's contribution to the construct will vary as a function of its relationship with the associated constructs used to estimate it (Bagozzi, 2007; Chin, 1998; Howell et al., 2007a, 2007b). In turn, the relative contribution of the indicator serves as an indication of its importance to the overall latent construct (Chin, 1998). Given that formative indicators are purportedly measuring potentially exclusive concepts (Jarvis et al., 2003), the question then becomes: what is being measured? For example, if a formative indicator measuring a person's scholarly productivity contributes significantly to a measure of faculty performance, while indicators measuring teaching and service do not, is the latent formative construct a measure of performance, or only of scholarly contribution? Thus, even when researchers apply the same set of formative indicators across studies, when the associated constructs are distinct from those in which the formative measure was originally developed, the significance of their contribution will likely vary (Chin, 1998), making their interpretation both confusing and ambiguous (Howell et al., 2007a). In other words, the stability of the construct over time is compromised because the conceptual meaning of the construct will change as a result of changes in the indicator weights used to measure it.

9. Conclusion

This research explores the consequences of both the reflective and formative measurement of latent constructs, specifically examining this in the case of CSE. While the use of formative measurement maximizes unobserved variance at the latent construct level, and thus minimizes Type II errors, generalizability across studies is reduced. As suggested by Bagozzi (2007), "Formative measurement is limited in scope and typically ambiguous" (p. 235). Our analyses clearly demonstrate some of the challenges in using formative indicators.

In contrast, reflective indicators can be selected from a universe of items in a manner consistent with classical test theory. In reflective measurement, indicators can be added or dropped from measures based upon established reliability and validity metrics, without the alteration of conceptual meaning. As a result, properly validated reflective measures are relatively stable across assessments, allowing for confidence when comparing study results and, as we have pointed out, suitable for measuring constructs with psychological origins. We believe that researchers should carefully consider the theoretical and statistical implications of employing either measurement technique when examining latent constructs.

In the case of the CSE construct, we have not only demonstrated the statistical implications of misspecifying the construct as formative, but we have also argued that formative measurement is inconsistent with the substantive theory supporting self-efficacy. When constructs are conceived of as explanatory combinations of indicators forming constructs such as SES, population change, or marketing mix, the use of formative measures may be appropriate (Fornell and Bookstein, 1982). For constructs with psychological origins such as attitude, personality, and, as we have argued, self-efficacy, there seems to be little disagreement that indicators that “reflect” the underlying concept are most appropriate. We hope that our article serves to generate additional dialogue on the use of formative measurement by CSE researchers as well as researchers in other domains.

Acknowledgement

The authors would like to express their sincere appreciation for the expertise and developmental comments provided by the senior editor and three anonymous reviewers during the review process. We would also like to thank Professor Wynne Chin (University of Houston) and Professor Craig Parks (Washington State University) for providing subject matter expertise and early guidance on this topic.

References

Bagozzi, R. P. (2007). On the meaning of formative measurement and how it differs from reflective measurement: Comment on Howell, Breivik, and Wilcox. *Psychological Methods, 12*(2), 229-237.

Bandura, A. (1997). *Self-efficacy: The Exercise of Control*. New York, NY: W. H. Freeman/Times Books/ Henry Holt & Co.

Bandura, A. (2005). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Adolescence and education, Self-efficacy beliefs of adolescents.* (Vol. 4, pp. 1-39). Greenwich: Information Age Publishing.

Bollen, K. (2007). Interpretational confounding is due to misspecification, not to type
of indicator: Comment on Howell, Breivik, and Wilcox (2007). Psychological Methods, 12(2), 219-228.

Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Methods, 110*(2), 305-314.

Borsboom, D., Mellenbergh, G. J., & Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review, 110*(1), 203-219.

Chin, W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), *Modern Methods for Business Research* (pp. 295-336). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Compeau, D. R., & Higgins, C. A. (1995). Computer self-efficacy - Development of a measure and initial test. *MIS Quarterly, 19*(2), 189-211.

Diamantopoulos, A., Riefler, P., & Roth, K. P. (2008). Advancing Formative Measurement Models. *Journal of Business Research, In Press, Corrected Proof*.

Diamantopoulos, A., & Siguaw, J. A. (2006). Formative versus reflective indicators in organizational measure development: A comparison and empirical illustration. *British Journal of Management 17*, 263-282.

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators. *Journal of Marketing Research, 38*, 269-277.

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

Gefen, D., & Straub, D. (2005). A practical guide to factorial validity using PLS-Graph: Tutorial and annotated example. *Communications of the Association for Information Systems, 16*, 91-109.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate Data Analysis* (5th ed.). Upper Saddle River, NJ: Prentice-Hall.

Heise, D. R. (1972). Employing nominal variables, induced variables, and block variables in path analysis. *Sociological Methods and research, 1*, 147-173.
Howell, D. R., Breivik, E., & Wilcox, J. B. (2007a). Is formative measurement really measurement? Reply to Bollen (2007) and Bagozzi (2007). *Psychological Methods, 12*(2), 238-245.
Howell, D. R., Breivik, E., & Wilcox, J. B. (2007b). Reconsidering formative measurement *Psychological Methods, 12*(2), 205-218.
Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. (2003). A critical review of construct indicators and measurement model mis-specification in marketing and consumer research. *Journal of Consumer Research, 30*, 1999-1216.
Johnson, R. D., & Marakas, G. M. (2000). Research report: The role of behavioral modeling in computer skills acquisition: Toward refinement of the model. *Information Systems Research, 11*(4), 402-417.
Law, K. S., & Wong, C. (1999). Multidimensional constructs in structural equation analysis: An illustration using the job perception and job satisfaction constructs. *Journal of Management 25*(2), 143-160.
Loch, K. D., Straub, D., & Kamel, S. (2003). Diffusing the Internet in the Arab World: The role of social norms and technological culturation. *IEEE Transactions on Engineering Management, 50*(1), 45-63.
Marakas, G. M., Johnson, R. D., & Clay, P. F. (2007). The evolving nature of computer self-efficacy construct: An empirical investigation of measurement construction, validity, reliability and stability over time. *Journal of the Association for Information Systems, 8*(1), 16-46.
Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). New York: McGraw-Hill.
Petter, S., Straub, D., & Rai, A. (2007). Specifying formative constructs in information systems research. *MIS Quarterly, 31*(4), 623-656.
Podsakoff, P., MacKenzie, S. B., Podsakoff, N. P., & Lee, J. Y. (2003). The mismeasure of management and its implications for theoretical models using survey data. *The Leadership Quarterly, 14*, 615-665.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Design*. New York: Houghton Mifflin Company.
Wilcox, J. B., Howell, R. D., & Breivik, E. (2008). Questions about formative measurement. *Journal of Business Research, In Press, Corrected Proof*.

About the Authors

Andrew M. Hardin is an Assistant Professor of MIS at the University of Nevada Las Vegas. He received his Ph.D. in Information Systems from Washington State University. Dr. Hardin's research interests include virtual teamwork, self-efficacy, collective efficacy, technology-mediated learning, human-computer interaction, and the application of research methodologies in the IS discipline. Past research has appeared in journals such as *Journal of Management Information Systems*, *DATABASE, Group Decision and Negotiation,* and *Small Group Research*.

Jerry C. J. Chang is an Associate Professor of MIS in the College of Business, University of Nevada Las Vegas. He has a B.S. in Oceanography from National Ocean University (Taiwan), an M.S. in Computer Science from Central Michigan University, an MBA from Texas A&M University, an M.S. in MIS and Ph.D. in MIS from the University of Pittsburgh. His research interests include instrument development, IS performance measurement, self-efficacy, software piracy, use and impact of IT, culture aspects of IS, and management of IS. His works have appeared in *Management Information Systems Quarterly*, *Journal of Management Information Systems*, *Information & Management*, *Decision Support Systems*, *DATA BASE*, and *Communications of the ACM*.

Mark A. Fuller is the Chair of the Department of Information Systems at Washington State University, and holds the Philip L. Kays Distinguished Professorship in Information Systems. Professor Fuller received his Ph.D. in Management Information Systems from the University of Arizona. His research focuses on virtual teamwork, technology supported learning, and trust in technology-mediated environments, and has appeared in outlets such as *Information Systems Research*, *Management Information Systems Quarterly*, *Journal of Management Information Systems*, *Decision Sciences*, *Journal of the Association for Information Systems,* and *Decision Support Systems*. Professor Fuller has won multiple teaching awards, has published a textbook on Information Systems Project Management, and has taught graduate and undergraduate courses on a variety of topics, including global information systems and strategy, information systems project management, and collaborative technology.

Copyright © 2008, by the Association for Information Systems. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and full citation on the first page. Copyright for components of this work owned by others than the Association for Information Systems must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers for commercial use, or to redistribute to lists requires prior specific permission and/or fee. Request permission to publish from: AIS Administrative Office, P.O. Box 2712 Atlanta, GA, 30301-2712 Attn: Reprints, or via e-mail from ais@gsu.edu.

Journal of the Association for Information Systems

ISSN: 1536-9323

Editor
Kalle Lyytinen
Case Western Reserve University, USA

Senior Editors			
Robert Fichman	Boston College, USA	**Dennis Galletta**	University of Pittsburgh, USA
Varun Grover	Clemson University, USA	**Rudy Hirschheim**	Louisiana State University, USA
Robert Kauffman	University of Minnesota, USA	**Frank Land**	London School of Economics, UK
Jeffrey Parsons	Memorial University of Newfoundland, Canada	**Suzanne Rivard**	Ecole des Hautes Etudes Commerciales, Canada
Ananth Srinivasan	University of Auckland, New Zealand	**Bernard C.Y. Tan**	National University of Singapore, Singapore
Michael Wade	York University, Canada	**Ping Zhang**	Syracuse University, USA
Editorial Board			
Steve Alter	University of San Francisco, USA	**Kemal Altinkemer**	Purdue University, USA
Michael Barrett	University of Cambridge, UK	**Cynthia Beath**	University of Texas at Austin, USA
Michel Benaroch	University of Syracuse, USA	**Francois Bodart**	University of Namur, Belgium
Marie-Claude Boudreau	University of Georgia, USA	**Susan A. Brown**	University of Arizona, USA
Tung Bui	University of Hawaii, USA	**Andrew Burton-Jones**	University of British Columbia, Canada
Dave Chatterjee	University of Georgia, USA	**Patrick Y.K. Chau**	University of Hong Kong, China
Mike Chiasson	Lancaster University, UK	**Mary J. Culnan**	Bentley College, USA
Jan Damsgaard	Copenhagen Business School, Denmark	**Samer Faraj**	University of Maryland, College Park, USA
Chris Forman	Carnegie Mellon University, USA	**Ola Henfridsson**	Viktoria Institute & Halmstad University , Sweden
Hitotora Higashikuni	Tokyo University of Science, Japan	**Kai Lung Hui**	National University of Singapore, Singapore
Hemant Jain	University of Wisconsin-Milwaukee, USA	**Bill Kettinger**	University of South Carolina, USA
Rajiv Kohli	College of William and Mary, USA	**Mary Lacity**	University of Missouri-St. Louis, USA
Ho Geun Lee	Yonsei University, Korea	**Jae-Nam Lee**	Korea University
Kai H. Lim	City University of Hong Kong, Hong Kong	**Ji-Ye Mao**	Remnin University, China
Anne Massey	Indiana University, USA	**Emmanuel Monod**	Dauphine University, France
Michael Myers	University of Auckland, New Zealand	**Fiona Fui-Hoon Nah**	University of Nebraska-Lincoln, USA
Mike Newman	University of Manchester, UK	**Jonathan Palmer**	College of William and Mary, USA
Paul Palou	University of California, Riverside, USA	**Brian Pentland**	Michigan State University, USA
Yves Pigneur	HEC, Lausanne, Switzerland	**Jaana Porra**	University of Houston, USA
Sandeep Purao	Penn State University, USA	**T. S. Raghu**	Arizona State University, USA
Dewan Rajiv	University of Rochester, USA	**Balasubramaniam Ramesh**	Georgia State University, USA
Timo Saarinen	Helsinki School of Economics, Finland	**Susan Scott**	The London School of Economics and Political Science, UK
Ben Shao	Arizona State University,USA	**Olivia Sheng**	University of Utah, USA
Carsten Sorensen	The London School of Economics and Political Science, UK	**Katherine Stewart**	University of Maryland, USA
Mani Subramani	University of Minnesota, USA	**Burt Swanson**	University of California at Los Angeles, USA
Dov Te'eni	Tel Aviv University, Israel	**Jason Thatcher**	Clemson University, USA
Ron Thompson	Wake Forest University, USA	**Christian Wagner**	City University of Hong Kong, Hong Kong
Eric Walden	Texas Tech University, USA	**Eric Wang**	National Central University, Taiwan
Jonathan Wareham	ESADE, Spain	**Stephanie Watts**	Boston University, USA
Bruce Weber	London Business School, UK	**Tim Weitzel**	Bamberg University, Germany
Richard Welke	Georgia State University, USA	**George Westerman**	Massachusetts Institute of Technology, USA
Kevin Zhu	University of California at Irvine, USA	**Ilze Zigurs**	University of Nebraska at Omaha, USA
Administrator			
Eph McLean	AIS, Executive Director	Georgia State University, USA	
J. Peter Tinsley	Deputy Executive Director	Association for Information Systems, USA	
Reagan Ramsower	Publisher	Baylor University	