Holistic Procedures for Contemporary Formative Construct Validation using PLS: A Comprehensive Example

ABSTRACT

Sound advice from knowledgeable methodologists synthesized from the literature, has yielded herein a complete and readily useable set of guidelines for best-practice contemporary formative construct validation using partial least squares (PLS). Procedures are logically organized around indicator level and construct level validation. Indicator level validation is concerned that each indicator contributes to the formative construct by carrying the intended meaning. Tests advocated are in attention to: (i) potential multicollinearity, (ii) formative indicator weights, (iii) formative indicator loadings, and (iv) formative indicator criterion validity. These procedures are advocated at all formative levels of the focal construct, thus guidelines presented are sufficient whether validating formative constructs that are multidimensional (multiple levels) or unidimensional (single level). Construct level validation includes tests of: (i) nomological validity, and (ii) external validity. Assessing nomological validity involves evaluating the extent to which the formative construct behaves as expected within its network of hypotheses. Testing the external validity of a formative model entails assessing the extent to which the formative indicators in combination capture the full domain of the construct.

The prescriptions advocated are instantiated through reference to a recently completed study employing the full procedures described. The referent study conveniently included items whose statistical indications were mixed, inviting illuminating discussion around the basis on which key decisions are made. Guidance compiled herein and demonstrated through example, will be of value to both novice and more experienced researchers working with or considering formative phenomenon. The combination of procedures outlined makes clear the necessary data and tests required in order to facilitate strong formative construct validity testing.

Key words:

Information Systems Evaluation, IS-Impact, Formative Construct Validation, PLS, SEM

DOI : 02.IJIS.2015.1.2

Introduction

Existing for decades, formative constructs have only relatively recently attracted the central attention of the IS research community, both through their use in research and in methodological writings. While validation techniques for formative constructs have been evolving, (Wilcox, Howell, & Breivik, 2008) argue that formative constructs have been understudied, and that methodological guidance on how to develop and empirically estimate such constructs is insufficient. (Diamantopoulos, Riefler, & Roth, 2008) suggest a main barrier to formative construct use, is the lack of support for formative constructs in popular covariance-based structural equation modeling (SEM) software packages such as LISREL and AMOS, and the concomitant difficulty identifying and estimating formative constructs with these tools.

At the same time there is increasing consciousness of formative phenomenon, and though formative constructs have their detractors, a substantial proportion of the research community nonetheless yet believes formative constructs serve an important purpose. Further, explicit and implicit formative construct use is more widespread than apparent. Petter, Straub and Rai (2007) suggest there is a significant threat of misspecifying and validating constructs as reflective that, on closer scrutiny, are, in fact, formative. In their review of marketing literature, Jarvis, MacKenzie and Podsakoff (2003) found that about one-third of constructs have been misspecified as reflective instead of formative. Misspecification of constructs as formative or reflective results in measurement error, which impacts the structural model, thereby increasing the potential for type I and type II errors (Petter et al., 2007).

Component-based SEM offers some salvation, more readily accommodating formative constructs. Though the argument for component-based rather than covariance-based SEM for formative model testing is not as clear as was once generally held (Hair, Ringle, & Sarstedt, 2011), a majority of proponents tend to use Partial Least Squares (PLS) in formative model testing.

Nonetheless, as a community, our understanding of appropriate methods of formative construct validation using PLS have evolved somewhat piecemeal. While conceptual papers on formative model assessment exist (e.g. Diamantopoulos & Siguaw, 2006; Diamantopoulos & Winklhofer, 2001; Götz, Liehr-Gobbers, & Krafft, 2010; Henseler, Ringle, & Sinkovics, 2009; Petter et al., 2007; Urbach & Ahlemann, 2010), (Jarvis et al., 2003; Petter et al., 2007) observe, that guidelines on how to interpret formative model results are scarce. Kim, Shin and Grover (2010) argue that there is yet a lack of consensus on what comprise appropriate tools and techniques to validate formative models. Having as a community learned much from recent methodological writings, lessons have nonetheless been scattered, often with little in the way of tangible examples. This paper represents a synthesis of past and recent sage advice from knowledgeable methodologists, yielding a complete and readily useable set of prescriptions for best practice contemporary, quantitative formative construct validation using PLS. The prescriptions are then instantiated through reference to recently published research employing the full procedures described.

The procedures and referent study reported in this paper, proceed from the point in the construct development lifecycle at which the construct has been operationalized. The focus herein is on quantitative validation employing main study data[1], subsequent to construct conceptualisation, specification and operationalisation, which tend in the main to be qualitative[2] - It is assumed these prior stages have been conducted well. We include early discussion comparing formative and reflective constructs, for novices looking ahead to what formative validation entails. The subsequent synthesized procedures may appeal to both novice and experienced researchers, as may their detailed instantiation.

[1] Ignoring any pilot quantitative testing undertaken as part of Operationalisation.
[2] Though proper execution of these earlier stages is crucial, they are outside the scope of this presentation.

DOI : 02.IJIS.2015.1.2

This paper uses the IS-Impact measurement model (Gable, Sedera, & Chan, 2008) as the referent study, to empirically illustrate full procedures to be used when validating formative constructs. The referent study reports validation of a 2nd-order hierarchical formative construct, formative on both levels, thereby exercising all procedures required to validate a formative construct having any number of formative levels[3]. That is, this paper includes a 'complete' set of prescriptions (i.e. guidelines) that should be used when validating formative constructs whether they are multidimensional (multiple levels) or unidimensional (single level).

The paper is structured around 6 main sections: (1) Reflective versus formative constructs; (2) A digest on structural equation modeling; (3) A digest on PLS path modeling; (4) Formative construct validation using PLS path modeling, as synthesized from methodological writings; (5) The referent study, instantiating the full procedures described; and (6) Conclusions.

Reflective versus Formative constructs

Reflective constructs have observed measures that are affected by an underlying latent, unobservable construct, while formative constructs are a composite of multiple measures (MacCallum & Browne, 1993; Petter et al., 2007) (See Figure 1).

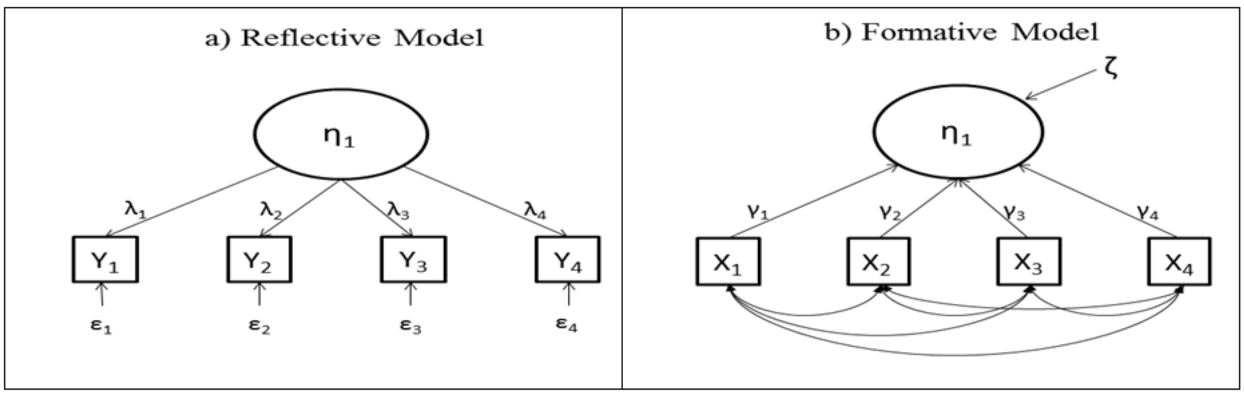


Figure 1: Specifying reflective and formative models.

With reflective specification (Figure 1.a), measures (i.e. indicators) are referred to as *'effects'* indicators (Bollen & Lennox, 1991; MacCallum & Browne, 1993), as the co-variation among indicators is explained by variation in an underlying common latent variable (MacKenzie, Podsakoff, & Jarvis, 2005). Therefore, causality in models of this type is from the latent variable to the indicators (Diamantopoulos et al., 2008; Jarvis et al., 2003; MacCallum & Browne, 1993). In reference to Bollen (1984), Diamantopoulos et al. (2008) characterise reflective models in two ways, namely: (1) a change in the latent variable causes variation in all measures simultaneously, and (2) all indicators in a reflective measurement model must be positively inter-correlated. Bollen and Lennox (1991) demonstrate the relationship between reflective indicators and their latent variable as:

$$Yi = \beta i1\ X1 + \varepsilon i \qquad \textbf{Equation (1)}$$

Where: Yi = the i^{th} indicator

βi1= coefficient representing effect of latent variable on indicator

X1= latent variable (i.e. the reflective construct)

εi = measurement error for indicator i

[3] In constructs entailing mixed formative and reflective levels, procedures described herein pertain to the formative levels and reflective procedures pertain to the reflective levels. Reflective construct validation is well understood and documented and not a focus of this discussion.

DOI : 02.IJIS.2015.1.2

Each indicator of a reflective model is represented by its own equation (Bollen & Lennox, 1991; Diamantopoulos et al., 2008; Petter et al., 2007). Furthermore, internal consistency (i.e. reliability) is essential for reflective latent variables. Hence, Cronbach's alpha (α) (Cronbach, 1951), or other reliability measures, are used to ensure the reliability of indicators in a reflective model (Petter et al., 2007). However, if an individual reflective indicator proves to be unreliable, it can be removed to improve construct validity without affecting the content validity of that construct (Petter et al., 2007).

In the second form of specification, the formative mode (Figure 1.b), formative indicators *"are not used to account for observed variances in the outer model, but rather to minimize residuals in the structural relationship"* (Petter et al., 2007, p: 626). In fact, formative indicators determine the latent variable, which receives its meaning from the former (Diamantopoulos et al., 2008). That is, formative indicators *'cause'* the construct (Petter et al., 2007), as distinct from reflective indicators which are referred to as effects indicators (Bollen & Lennox, 1991; MacCallum & Browne, 1993).

Diamantopoulos and Winklhofer (2001) point out four distinct characteristics of a formative model. These include: (1) formative indicators characterise a set of distinct causes which are not interchangeable, as each indicator captures a specific aspect of the construct's domain; (2) there are no specific expectations about patterns or magnitude of intercorrelations between the indicators; (3) formative indicators have no individual measurement error terms (i.e. they are assumed to be error-free in a conventional sense); and (4) while reflective measurement models with more than two indicators are identified and can be estimated, a formative measurement model, in isolation, is 'under-identified' and cannot be estimated. Additionally, while internal consistency (reliability) is essential for reflective latent variables, it is unimportant in formative models, as formative indicators are assessing different facets of the construct (Petter, et al., 2007). Also, strong correlation, which is desired among reflective indicators, is a problem for indicators in a formative model (Jarvis et al., 2003; Petter et al., 2007), as it can result in multicollinearity and is evidence of potential redundancy (overlap among composite items) and lack of mutual exclusivity; note that formative items might correlate positively or negatively or lack any correlation (Diamantopoulos et al., 2008; Diamantopoulos & Winklhofer, 2001; Jarvis et al., 2003; MacKenzie et al., 2005; Petter et al., 2007). Bollen and Lennox (1991) demonstrate the formative construct as:

$$\mathbf{Y = \beta 1\ X1 + \ldots\ \beta n\ Xn + \zeta} \qquad \textbf{Equation (2)}$$

Where: Y = the formative construct being estimated
Bi = beta weight for indicators
Xi = indicator scores/observations
ζ = a disturbance term

One main difference between reflective and formative indicators is the extent to which an indicator contributes to the construct under investigation (Jarvis et al., 2003; Petter et al., 2007). Hence, formative indicators are assigned beta weights (Petter et al., 2007) as shown in Equation (2). Consequently, *"dropping a measure from a formative-indicator model may omit a unique part of the conceptual domain and change the meaning of the variable, because the construct is a composite of all the indicators"* (MacKenzie et al., 2005, p: 712). That is, removing a nonsignificant formative indicator will remove the beta weight associated with it, no matter how large or small it might be (Petter et al., 2007, p: 627). Table 1 summarises the main differences between reflective and formative model types.

DOI : 02.IJIS.2015.1.2

Table 1: Comparison of Formative and Reflective Models (Jarvis et al. (2003, p: 203)

Criteria	Formative	Reflective
1. Direction of causality from construct to measure implied by the conceptual definition	*Direction of causality is from items to construct.*	*Direction of causality is from construct to items.*
a. Are the indicators (items) (a) defining characteristics or (b) manifestations of the construct?	Indicators are defining characteristics of the construct.	Indicators are manifestations of the construct.
b. Would changes in the indicators/items cause changes in the construct or not?	Changes in the indicators should cause changes in the construct.	Changes in the indicator should not cause changes in the construct.
c. Would changes in the construct cause changes in the indicators?	Changes in the construct do not cause changes in the indicators.	Changes in the construct do cause changes in the indicators.
2. Interchangeability of the indicators/items	*Indicators need not be interchangeable.*	*Indicators should be interchangeable.*
a. Should the indicators have the same or similar content? Do the indicators share a common theme?	Indicators need not have the same or similar content/indicators need not share a common theme.	Indicators should have the same or similar content/indicators should share a common theme.
b. Would dropping one of the indicators alter the conceptual domain of the construct?	Dropping an indicator may alter the conceptual domain of the construct.	Dropping an indicator should not alter the conceptual domain of the construct.
3. Covariation among the indicators	*Not necessary for indicators to covary with each other*	*Indicators are expected to covary with each other.*
a. Should a change in one of the indicators be associated with changes in the other indicators?	Not necessarily	Yes
4. Nomological net of the construct indicators	*Nomological net of the indicators may differ.*	*Nomological net of the indicators should not differ.*
a. Are the indicators/items expected to have the same antecedents and consequences?	Indicators are not required to have the same antecedents and consequences.	Indicators are required to have the same antecedents and consequences.

Structural Equation Modeling (SEM) in a Nutshell

The purpose of many research projects (in general, including IS research) is to identify important variables pertaining to the study context and goals, and analyze causal relationships between these variables. *"Structural equation modeling (SEM) has become a quasi-standard... when it comes to analyzing the cause–effect relations between latent constructs"* (Hair et al. 2011, p: 139). SEM is a statistical technique, used to test and estimate such relationships based on qualitative/ theoretical assumptions and statistical data (Urbach & Ahlemann, 2010). SEM techniques such as LISREL and Partial Least Squares (PLS), are considered second generation multivariate analysis techniques (Bagozzi & Phillips, 1982); *"in contrast to first-generation techniques, such as factor analysis, discriminant analysis, or multiple regression, SEM allows the researcher to simultaneously consider relationships among multiple independent and dependent constructs"* (Urbach and Ahlemann, 2010, p: 9) and is considered state-of-the art for high quality statistical analysis in research (e.g. Gefen, Rigdon, & Straub, 2011; Gefen, Straub, & Boudreau, 2000).

SEM techniques provide fuller information about the extent to which the research model is supported by the data than in regression techniques (Gefen et al. 2000, p: 6). In contrast to its first generation predecessors, SEM techniques allow complicated variable relationships to be expressed through hierarchical or non-hierarchical, recursive or non-recursive structural

DOI : 02.IJIS.2015.1.2

equations represented in a diagrammatic structure (see Gefen et al. (2000, p. 21)), to present a more complete picture of the entire model (Bullock, Harlow, & Mulaik, 1994; Gefen et al., 2000, p: 4). SEM not only assesses the 'structural model' – the assumed causation among a set of dependent and independent constructs – but, in the same analysis, also evaluates the 'measurement model' – loadings of observed indicators (measures) on their expected latent variables (constructs) (e.g. Esposito Vinzi, Chin, Henseler, & Wang, 2010; Gefen et al., 2011). The combined analysis of the measurement and the structural models enables: measurement errors of the observed variables to be analyzed as an integral part of the model, and factor analysis to be combined in one operation with the hypothesis testing (Gefen et al., 2000, p: 5). Thus, SEM techniques have the potential to improve the rigorous analysis of the proposed research model (Gefen et al., 2011), and, very often, is a better methodological assessment tool (e.g. Bullock et al., 1994; Gefen et al., 2011).

Overall, SEM has many advantages compared to traditional statistical analysis techniques (Garson (1998) and is the *"a priori method of choice"* (Gefan et al 2011, p: iv), when analyzing path diagrams that involve latent variables with multiple indicators. SEM has become progressively more popular in IS research for purposes such as testing linkages between different constructs, and instrument validation (e.g. Chin, 1995, 1998b; Gefen et al., 2011; Gefen et al., 2000). Since its first appearance in 1990 in the major IS journals, SEM usage has grown increasingly (Gefen et al., 2000; Urbach & Ahlemann, 2010). Gefen et al. (2000) report that 18% of empirical research articles published in three major IS journals (during 1994-1997); MIS Quarterly (MISQ), Information & Management (I&M), and Information Systems Research (IRS), made use of SEM techniques such as PLS, LISREL, EQS, and AMOS, with PLS and LISREL being the two most common techniques.

Overview of PLS Path Modeling

There are two primary types of SEM analysis; (i) Covariance-based SEM (CBSEM), as implemented in, for example in LISREL, AMOS, and EQS; and ii) Component-based partial-least-squares (PLS) SEM, as implemented in, for example, SmartPLS and PLS-Graph. PLS and CBSEM techniques differ in terms of objectives, assumptions, parameter estimates, latent-variable scores, implications, epistemic relationship between a latent variable and its measures, model complexity, and requisite sample size (Chin & Newsted, 1999). Urbach & Ahlemann (2010, p: 13) provide a useful summary comparison of the two approaches. The CBSEM method attempts to calculate model parameters that will minimize the difference between the calculated and observed covariance matrices, yielding goodness of fit indices as a result of the magnitude of these differences (Andreev, Heart, Maoz, & Pliskin, 2009). The PLS approach, on the other hand, attempts to estimate all model parameters in such a way that the result should be a minimized residual variance of all dependent variables, Latent variables (LVs), and indicators (i.e. maximize the explained variance) (Chin, 1998b; Gefen et al., 2000). Put differently, the main objective of the PLS approach is to best predict a latent variable (LV), instead of assessing the model fit with the data (which is the main goal of the CBSEM approach) (e.g. Andreev et al., 2009; Gefen et al., 2000). "*The philosophical distinction between CB‑SEM and PLS‑SEM is straightforward. If the research objective is theory testing and confirmation, then the appropriate method is CB‑SEM. In contrast, if the research objective is prediction and theory development, then the appropriate method is PLS‑SEM*" (Hair et al., 2011, p. 140). The ultimate choice between the two approaches should be determined by the relevant objectives of the study. Hair et al (2011, p. 144) present a set of guidelines for deciding whether to use CB-SEM or PLS-SEM, based on five decision criteria, also claiming that "*one method is not superior to the other in general"* (Hair et al, 2011, p. 149). While prediction and theory development are the main reasons to use PLS- SEM, it can also be used for confirmatory theory testing (Hair et al., 2011), though the potentially more sample-specific results do raise issues around generalisability.

DOI : 02.IJIS.2015.1.2

Some scholars seem to view the PLS method as 'less rigorous' (and therefore less suitable) for confirming relationships between latent variables. Hair et al. (2011) state that one likely reason for such less favourable views reported, is individual authors' previous experience and familiarity with CB-SEM and its software, hence their preference (and defence) of this technique over the PLS method, which may be less familiar. Tenenhaus (2008) discusses some of the PLS-SEM weaknesses. These include: (1) PLS-SEM software suffering from lack of widespread accessibility due to historical limited diffusion of the PLS software as compared with CBSEM software, (2) PLS more commonly used for exploratory research, and (3) unlike CBSEM, PLS does not enable testing of equality constraints on path coefficients, or imposing specific values on different model paths.

More recent PLS-SEM studies show that the method is developing rapidly and has demonstrated numerous (technical and algorithmic) advancements to address prior known limitations; "*recent methodological developments in PLS-SEM are radical*" (Hair et al 2011, p 148). PLS-SEM has been used by an increasing number of researchers from different disciplines (Chin, 2010b; Esposito Vinzi et al., 2010; Henseler et al., 2009) including: Information Systems (e.g. Dibbern, Goles, Hirschheim, & Jayatilaka, 2004), strategic management (e.g. Hulland, 1999), marketing (e.g. Duarte & Raposo, 2010; Reinartz, Krafft, & Hoyer, 2004), organisational behavior (e.g. Higgins, Duxbury, & Irving, 1992), and consumer behaviour (e.g. Fornell & Robinson, 1983). "…*negative perceptions of PLS-SEM are unfortunate and short-sighted. When properly applied, the method has many benefits not offered by CB-SEM... if CB-SEM assumptions cannot be met, or the research objective is prediction rather than confirmation of structural relationships, then variance-based PLS-SEM is the preferred method*" (Hair et al., 2011, p. 139).

PLS use in the Information Systems discipline is growing (Urbach & Ahlemann (2010, p: 7). Urbach and Ahlemann (2010, p: 8), based on an archival analysis of two of the most prestigious international IS journals, namely Information Systems Research (ISR) and Management Information Systems Quarterly (MISQ), state *"the numbers indicate that in the empirical studies published in the two journals investigated, PLS has been used even more frequently than the covariance-based approaches"*. These results strongly support the findings of Goodhue, Lewis, and Thompson (2006, p: 2) who observe that *"PLS has been wholeheartedly accepted as an important statistical method in the MIS field"*. (Goodhue, Lewis, & Thompson, 2012), based on a recent archival analysis of three top MIS journals (namely; Information Systems Research [ISR], Journal of Management Information Systems [JMIS]), and MIS Quarterly [MISQ]), found that 49% of the path analysis papers published from 2006-2010 used PLS. Ringle et al., (2011, p. iv-v), in their recent review of PLS-SEM use reported in MISQ (from 1992 through 2011), reveal that "*there were 65 studies containing 109 structural equation model estimations deploying the PLS-SEM technique*". The proliferation of PLS has been accredited to various strengths of PLS. Urbach and Ahlemann (2010, p: 9), based on their archival analysis of the IS discipline's application of PLS, summarize the researchers' arguments for choosing PLS as follows:

- PLS-SEM can work efficiently with a much wider range of sample sizes, it makes fewer demands regarding sample size than other methods (e.g. Agarwal & Karahanna, 2000; Ahuja & Thatcher, 2005; Bearoch, Lichtenstein, & Robinson, 2006).
- PLS-SEM does not require normally-distributed input data (e.g. Ahuja & Thatcher, 2005; Malhotra, Gosain, & El Sawy, 2007; Pavlou, Liang, & Xue, 2007).
- PLS-SEM can be applied to complex structural equation models with a large number of constructs (e.g. Bassellier & Benbasat, 2004; Burton-Jones & Straub, 2006; Wixom & Todd, 2005).
- PLS-SEM is able to handle both reflective and formative constructs (e.g. Choudhury & Karahanna, 2008; Liang, Saraf, Hu, & Xue, 2007; Limayem, Hirt, & Cheung, 2007).

DOI : 02.IJIS.2015.1.2

- PLS-SEM is better suited for theory development than for theory testing[4] (e.g. Chwelos, Benbasat, & Dexter, 2001; Kanawattanachai & Yoo, 2007; Komiak & Benbasat, 2006)).
- PLS-SEM is especially useful for prediction (Au, Ngai, & Cheng, 2008; Moores & Chang, 2006; Rai, Patnayakuni, & Seth, 2006).

Recent literature has critiqued several of these claims, cautioning against overestimating PLS capabilities (e.g. Gefen et al., 2011; Goodhue et al., 2006; Marcoulides & Saunders, 2006). There is "*apparent misuse of perceived leniencies such as assumptions about minimum sample size in partial least squares (PLS)"* (Gefan et al., 2011, p. iii). (Ringle, Sarstedt, & Straub, 2012), critique how some researchers misuse PLS-SEM's accommodation of single-item constructs. While they acknowledge that there are times the researcher may have no choice but to use single item constructs, they caution against this saying that a "*small number of items for construct measurement (in the extreme, the use of a single item) works against PLS-SEM's tendency to bias estimates*" (p. vii) . Goodhue et al., (2012), specifically contest claims that PLS has advantages over other techniques when analyzing small sample sizes or data with non-normal distributions, arguing that "*When used with small sample sizes, PLS, like the other techniques, suffers from increased standard deviations, decreased statistical power and reduced accuracy"* (p:1).

However, when applied appropriately, PLS is a valid and useful technique; "*PLS is an adequate choice if the research problem meets certain characteristics and the technique is properly used*" (Urbach and Ahlemann, 2010, p: 5). Chin (1998b) and Chin and Newsted (1999) argue that PLS can be an adequate alternative to CB-SEM. Hair et al. (2011) describe the method's strengths and limitations, guiding researchers on the *appropriate* application of PLS-SEM; "*Using "good" measures and data, both approaches practically yield the same results*" (Hair et al., 2011, p: 140). They highlight that "*researchers must always be aware of the differences in interpretation of the results, particularly as they relate to the constructs' measurement properties*" (p. 140). Gefen et al., (2011), provide detailed guidelines on when to choose PLS (over Covariance-based SEM) and what to report in PLS studies. They argue that; "*PLS shines forth in exploratory research and shares the modest distributional and sample size requirements of ordinary least squares linear regression*..." (p: v); it is a tool "*for situations that are 'data-rich but theory-primitive'* " and it is suited for studies that are *"measuring a construct with formative scales"* (p. vi). PLS-SEM path modelling, when applied appropriately, is indeed very practical and useful. Hair et al 2011 refer to PLS- SEM as a *" 'silver bullet' for estimating causal models in many theoretical models and empirical data situations"* (Hair et al., 2011, p: 139 & 148). Ringle et al., (2012, p: vii), in their critique of the use of PLS-SEM, also support this view that "*PLS-SEM can indeed be a "silver bullet" in certain research situations*", especially when models are relatively complex and representative sets of data are rather small, but "*PLS-SEM is not immune to threats from data inadequacies and researchers should make every effort to provide support for its statistical power in the research setting at hand*".

We join with the growing number of researchers (e.g. Hair et al., 2011; Ringle et al., 2012) who strenuously acknowledge the virtues of PLS and who endorse its use in particular situations as described preceding. We particularly espouse its merits when testing formative models, to which we now turn more specifically.

Overview of formative model assessment in PLS-SEM

Whether or not formative constructs can be empirically and statistically validated is contended (Diamantopoulos et al., 2008, p: 13). Some researchers state that no quantitative quality tests are applicable for measuring the appropriateness of formative indices. Other researchers note that the appropriateness and applicability of statistical procedures is limited because the choice of formative indicators determines the conceptual meaning of the construct. For instance,

[4] However, PLS-SEM may also represent a reasonable methodological alternative for theory testing and extension (Henseler et al., 2009).

DOI : 02.IJIS.2015.1.2

Rossiter (2002, p: 315) questions the need for any validity assessment of formative indicators, claiming *"all that is needed is a set of distinct components as decided by expert judgment"*.

However, we and other researchers (e.g. Diamantopoulos et al., 2008; Edwards & Bagozzi, 2000; Götz et al., 2010; Henseler et al., 2009; Jarvis et al., 2003; MacKenzie et al., 2005; Petter et al., 2007; Rabaa'i & Gable, 2012) do not share this view, and stress the need to assess the validity of formative models. As discussed previously, formative models reverse (as compared to reflective) the direction of causality in as far as the indicators form or compose the construct (Götz et al., 2010, p: 697). Therefore, the validation of formative models requires different procedures and techniques than those applied with reflective models (e.g. Götz et al., 2010; Henseler et al., 2009; Petter et al., 2007; Urbach & Ahlemann, 2010). That is, traditional validity assessments do not apply to formative models (e.g. Albers, 2010; Ali, Tate, Rabaa'i, & Zhang, 2012; Diamantopoulos, 1999, 2006; Diamantopoulos & Siguaw, 2006; Götz et al., 2010; Rabaa'i & Gable, 2012).

Diamantopoulos (2006, p: 11) states, with respect to formative models, that "*reliability becomes an irrelevant criterion for assessing measurement quality"*. It is the assumption of error-free measures that makes the question of indicator reliability irrelevant (Henseler et al., 2009). Unlike reflective indicators, the error term in a formative structure has no measurement error but rather a disturbance term, which represents the remainder of the construct domain unexplained by the presented indicators (Andreev et al., 2009, p: 5).

While reliability becomes an irrelevant criterion for assessing formative models (e.g. Bollen, 1984; Bollen, 1989; Diamantopoulos, 2006; Diamantopoulos & Siguaw, 2006), the examination of validity becomes essential (Diamantopoulos, 2006; Diamantopoulos & Siguaw, 2006; Götz et al., 2010; Henseler et al., 2009; Nils Urbach & Ahlemann, 2010). Accordingly, the literature (e.g. Andreev et al., 2009; Götz et al., 2010; Henseler et al., 2009; Petter et al., 2007; Urbach & Ahlemann, 2010) suggests that the assessment of formative measurement models should entail: (1) assessment of the content validity of the formative construct, (2) assessment at the indicators level, and (3) assessment at the construct level.

Content Validity

Petter et al. (2007) suggest that attention to content validity asessment should be mandatory when evaluating formative models. In a formative measurement model, content validity should be ensured when the model is specified (i.e. before the data is collected) (Götz et al., 2010, p: 697), as misspecification of the indicators could lead to forming a latent construct inappropriate for the content domain being explored and to biased estimation results (Andreev et al., 2009, p: 5). Content validity for formative models is concerned with whether the indicators capture the entire scope of the construct as described by the construct's domain (Andreev et al., 2009; Diamantopoulos & Winklhofer, 2001; Straub, Boudreau, & Gefen, 2004). Hence, it is critical to identify a broad set of indicators that covers all aspects of the formative model (Andreev et al., 2009; Diamantopoulos, 2006). The literature (e.g. Straub et al., 2004) proposes a number of techniques to ensure content validity of a construct, such as: conducting a thorough literature review related to the construct domain, and the use of qualitative methods like panel discussions, expert interviews and Q-sorting. Content validity should be largely addressed earlier, in conceptualisation, specification and operationalisation, using these kinds of qualitative procedures. While the procedures and referent study reported in this paper, proceed from the point in the construct development lifecycle at which the construct has been operationalized, it is important to understand the notion of content validity.

Assessment of Formative Indicators

Assessing formative models at the indicators level is in attention to the concern that each indicator indeed contributes to the formative construct by carrying the intended meaning (Henseler et al., 2009, p: 301). Various statistical tests can be performed to evaluate whether an indicator should be included in the formative construct or not (e.g. Cassel, Hackl, & Westlund,

DOI : 02.IJIS.2015.1.2

2000; Diamantopoulos et al., 2008; Götz et al., 2010; Grewal, Cote, & Baumgartner, 2004; Henseler et al., 2009; Petter et al., 2007; Urbach & Ahlemann, 2010). These include assessing:

- the degree of multicollinearity,
- formative indicator weights,
- formative indicator loadings, and
- formative indicator criterion validity.

Multicollinearity Testing

Assessing the degree of multicollinearity among formative indicators is important in formative model validation, as high multicollinearity could mean that a formative indicator's information is redundant (Henseler et al., 2009). That is, the existence of multicollinearity may suggest that specification of the formative indicators was not performed successfully since formative indicators should represent distinct characteristics of the content domain and high covariance might mean that formative indicators explain the same aspect of the domain (Andreev et al., 2009, p: 6).

In order to check for multicollinearity, researchers should calculate the variance inflation factor (VIF) or the tolerance values (e.g. Götz et al., 2010; Henseler et al., 2009; Petter et al., 2007; Urbach & Ahlemann, 2010). The VIF indicates how much of an indicator's variance is explained by the other formative indicators of the same construct (Urbach & Ahlemann, 2010). The VIF is calculated as the inverse of the tolerance value (J., Black, Babin, Anderson, & Tathum, 2006). Tolerance is 1- r^2, where r^2 is the multiple r of a given indicator, regressed on all other indicators of the same construct. A rule of thumb from econometrics states that VIFs greater than 10 (more than 90% of the variance in the item is explained by the other items) reveal a critical level of multicollinearity (e.g. Diamantopoulos & Siguaw, 2006; Gefen et al., 2000; Gefen et al., 2011; Götz et al., 2010; Henseler et al., 2009; Petter et al., 2007).

Indicator Weights

In PLS, the significance of formative indicator weights can be determined by means of bootstrapping (e.g. Chin, 1998b; Davison & Hinkley, 2003; Götz et al., 2010; Henseler et al., 2009; Tenenhaus, Vinzi, Chatelin, & Lauro, 2005; Urbach & Ahlemann, 2010). Formative indicator weights must not be interpreted as factor loadings (Götz et al., 2010, p: 698), but should be assessed and compared to determine their relative contribution to the formative construct (Henseler et al., 2009; Sambamurthy & Chin, 1994). Formative indicator weights explain the amount of variance in the formative construct that is explained by the indicator. Hence, a high indicator weight suggests that the indicator is making a substantive contribution to the formative construct (Diamantopoulos, 2006). However, formative indicator weights are often smaller than the loadings of reflective indicators (Götz et al., 2010). A significance level of at least .050 suggests that an indicator is relevant for the construction of the formative construct and, thus, demonstrates a sufficient level of validity (e.g. Urbach & Ahlemann, 2010, p: 20). It is also recommended that the path coefficients (between formative indicators and their respective construct) should be greater than .100 (Andreev et al., 2009; Jahner, Leimeister, Knebel, & Krcmar, 2008) or .200 (Chin, 1998).

While, reflective indicators with small loadings are frequently omitted from reflective models (e.g. Götz et al., 2010, p: 698), *"indicator elimination – by whatever means – should not be divorced from conceptual consideration when a formative measurement model is involved"* (Diamantopoulos & Winklhofer, 2001, p: 273). Diamontopolous and Winklhofer (2001) suggest that if any of the formative indicators are non-significant, it may be appropriate to remove them (one at a time) until all paths are significant and a good fit is obtained (Petter et al., 2007). However, when removing indicators, it is important to ensure that the construct is still measuring the entire domain and content validity is preserved (Diamantopoulos & Winklhofer, 2001; Jarvis et al., 2003; Petter et al., 2007).

DOI : 02.IJIS.2015.1.2

Indicator Loadings

Cenfetelli and Bassellier (2009, p: 695) argue that: *"formative indicators essentially "compete" with one another to be explanatory of their targeted construct. In this competition to explain variance, only a limited number of indicators will likely be significant while the others will be nonsignificant"*. As such, formative models with a relatively large number of indicators will generally have several low indicator weights (Cenfetelli & Bassellier, 2009). Put differently, a greater number of formative indicators will result in a greater possibility that some of the indicator weights will be low in magnitude and statistically non-significant (Cenfetelli & Bassellier, 2009, p: 694). A non-significant weight for a formative indicator may lead one to conclude that an indicator has no relationship with the formative construct it is intended to measure, hence, permitting its exclusion from the model. However, as mentioned previously, MacKenzie et al. (2005, p: 712) state that: *"dropping a measure from a formative-indicator model may omit a unique part of the conceptual domain and change the meaning of the variable, because the construct is a composite of all the indicators"*.

Cenfetelli and Bassellier (2009) argue that as important as formative indicator weights are for determining their *'relative'* contribution to their assigned construct, formative indicator weights are not the only determinant for retaining or omitting the indicator from a formative model. They suggest (2009, p: 697) that: *"…it is also possible to evaluate the 'absolute' importance of an indicator to its construct. This is provided by the loading of the indicator and so its bivariate correlation with the formatively measured construct"*. Cenfetelli and Bassellier (2009) supported their argument by reflecting on what Nunnally and Bernstein (1994) refer to as *"validity"*, the zero-order correlation between a predictor and a criterion. Cenfetelli and Bassellier (2009, p: 697) conclude that: *"Just as formative indicator weights are analogous to the beta weights of a multiple regression; formative indicator loadings are analogous to this zero-order correlation. In some cases, indicators may have a low or even nonsignificant weight, and therefore a low or nonsignificant relative contribution to the construct. However, an indicator with a low or nonsignificant weight may still have an important absolute contribution if the indicator is assessed independently from the other indicators"*. Hence, relying on only formative indicator weights may lead to misinterpretation of formative indicator analysis results.

Indicator Criterion Validity

For assessing formative indicator criterion validity, Diamantopoulos and Winklhofer (2001, p: 272) propose that researchers correlate formative items with a *"global item that summarizes the essence of the construct"*. That is, estimating formative indicators' correlations with an external variable (i.e. external to the formative construct) (Diamantopoulos et al., 2008). Assuming that the overall measure is a valid criterion, the relationship between a formative indicator and the overall measure implies indicator validity (Diamantopoulos et al., 2008, p: 13; Diamantopoulos & Winklhofer, 2001, p: 272; MacKenzie et al., 2005). In this case, only those indicators that are significantly correlated with the construct of interest should be included in the formative construct (Diamantopoulos et al., 2008; Diamantopoulos & Winklhofer, 2001).

Assessing the Formative Construct

Construct validity refers to the wider, out of the construct, validation of its measures (Straub et al., 2004). For instance, construct validity is concerned with whether or not indicators of the construct indeed measure what they intend to from the perspective of relationships between constructs, and between constructs and their relative indicators (Andreev et al., 2009, p: 6). Construct validity for formative models can be assessed in terms of (1) nomological validity (e.g. Andreev, et al., 2009; Henseler et al., 2009; Urbach & Ahlemann, 2010) and (2) external validity (e.g. Götz et al., 2010; Henseler et al., 2009; Reinartz et al., 2004)[5].

[5] While formative models are not expected to demonstrate convergent validity (e.g. Götz et al., 2010; Henseler et al., 2009), Loch, Straub, and Kamel (2003) suggest the use of a modified multi-trait-multi-method (MTMM) matrix analysis (Campbell & Fiske, 1959) for testing convergent validity for a formative model. However, MacKenzie et al. (2005) express concerns regarding the relevance

DOI : 02.IJIS.2015.1.2

Nomological Validity

A nomological network includes a (i) theoretical framework of research objects, (ii) an empirical framework of how these objects will be measured, and (iii) specification of the relationships between these two frameworks (Campbell & Fiske, 1959). Assessing nomological validity involves evaluating the extent to which the formative construct behaves as expected within a net of hypotheses (Diamantopoulos & Winklhofer, 2001; Henseler et al., 2009; Urbach & Ahlemann, 2010). Accordingly, those relationships between the formative construct and other of the structural model constructs, which have been sufficiently referred to in prior literature, should be strong and significant (Andreev et al., 2009; Diamantopoulos & Winklhofer, 2001; Henseler et al., 2009; Straub et al., 2004; Urbach & Ahlemann, 2010). That is, testing the nomological validity of a formative construct involves (Andreev et al., 2009, p: 8):

- First, linking the focal construct with its hypothesised antecedents and consequence constructs, and
- Second, evidencing nomological validity where the hypothesised linkages (structural paths) between the constructs are found to be significantly greater than zero and their signs are in the expected causality direction.

External Validity

Testing the external validity[6] of a formative model entails assessing the extent to which the formative indicators in combination capture the full domain of the construct (Andreev et al., 2009; Chin, 1998; Götz et al., 2010; Henseler et al., 2009; Jahner et al., 2008; Reinartz et al., 2004). In assessing the external validity of a formative model, one should be concerned about the construct's error-term (v), which represents the part of the construct that is not captured by any formative indicator (Götz et al., 2010; Henseler et al., 2009). External validity can be assessed by means of regressing the formative construct on a reflective indicator of the same construct (Henseler et al., 2009), as it is often possible to operationalise a construct formatively as well as reflectively (Götz et al., 2010; Reinartz et al., 2004). In this case, reflective indicators can be used to measure the error terms (Götz et al., 2010, p: 699). That is, the operationalization of a formative construct by means of reflective indicators allows the measurement error to be determined (Chin, 1998a).

A Multiple Indicators and Multiple Causes (MIMIC) model (Hauser & Goldberger, 1971; Jöreskog & Goldberger, 1975) should be applied for the model identification procedure (Andreev et al., 2009; Götz et al., 2010), where both formative and at least two reflective indicators measure one construct (Diamantopoulos et al., 2008; Diamantopoulos & Winklhofer, 2001). Figure 2 represents an example of a MIMIC model.

of assessing convergent validity for models with formative indicators. As a result, most formative model assessment literature eliminates convergent validity from the validly assessment. Conversely, we argue that the inter-indicator condition is in fact problematic in the context of formative models, as (1) formative indicators may be positively or negatively correlated, or uncorrelated at all (e.g. Albers, 2010; Bollen, 1989; Bollen & Lennox, 1991), and (2) excessive correlation among formative indicators suggests collinearity, and excessive collinearity suggests that redundant indicators may be included in the model (e.g. Petter et al., 2007).

[6] Not be confused with the sometimes usage of 'external validity' to refer to 'generalisability'

DOI : 02.IJIS.2015.1.2

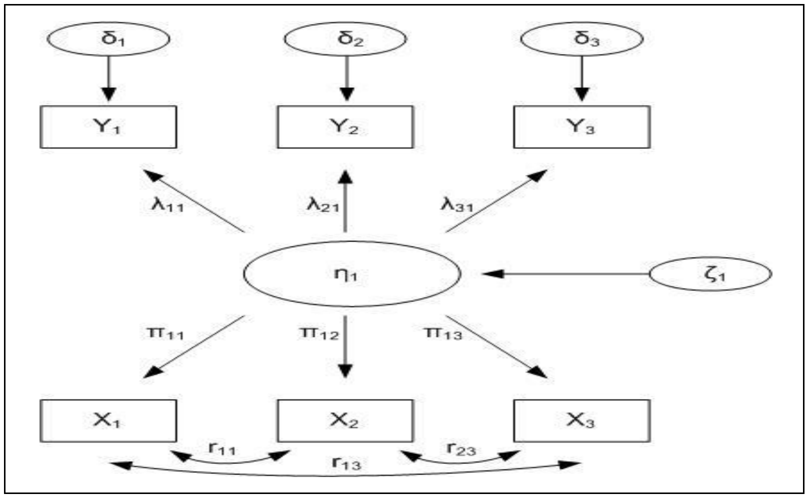


Figure 2: An example of a MIMIC model (adapted from Diamantopoulos & Winklhofer, 2001).

However, PLS does not allow the construction of a MIMIC model[7]. An alternative specification for quantifying the error terms is to use the two-construct model that integrates an additional *"phantom variable"* (Götz et al., 2010), which represents the construct's reflective operationalisation (Diamontopoulos & Winklhoffer, 2001, p: 272-274). If a strong and significant association between the construct and the phantom variable is confirmed, external validity is evidenced (Götz et al., 2010, p: 700). Figure 3 illustrates such an alternative conceptualisation of a MIMIC model.

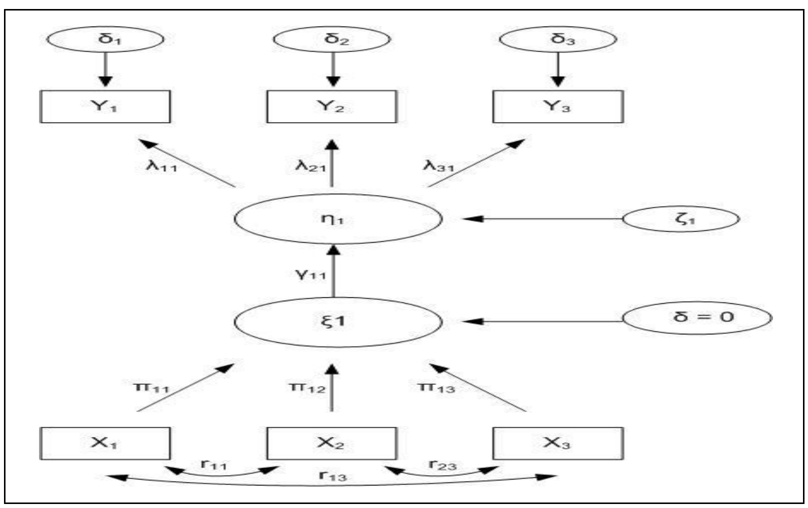


Figure 3: Alternative conseptualization of a MIMIC model using a phantom variable (adapted from Diamantopoulos & Winklhofer, 2001).

Illustrative example

Sage advice from knowledgeable methodologists, synthesized from the literature in the preceding sections, has yielded a complete and readily useable set of prescriptions for best-practice, contemporary formative construct validation. In this section, these prescriptions are instantiated through reference to a recently completed study employing the full procedures described.

[7] Currently, only SPAD-PLS supports the specification of variables by means of the MIMIC model (Götz et al., 2010, p: 700).

DOI : 02.IJIS.2015.1.2

The Referent Study

The referent study titled "Evaluating the Success of Large-scale, Integrated Information Systems through the Lens of IS-Impact and IS-Support[8]," (Rabaa'i, 2012) adopted the IS-Impact measurement model (Gable et al., 2008) as its core theory base. The referent study was conducted within the IT Evaluation Research Program (ITE-Program) at Queensland University of Technology (QUT). A goal of the ITE-Program is, "*to develop the most widely employed model for benchmarking information systems in organizations for the joint benefit of both research and practice*". The track espouses programmatic research having the principles of incrementalism, tenacity, holism and generalisability through replication and extension research strategies.

Prior work within the IS-Impact track has been consciously constrained to Financial IS for their homogeneity. The referent study adopted a context-extension strategy (Berthon, Pitt, Ewing, & Carr, 2002) with the aim "*to further validate and extend the IS-Impact measurement model in a new context - i.e. a different IS - Human Resources (HR)*". The overarching research question was: "How can the impacts of large-scale integrated HR applications be effectively and efficiently benchmarked?" The unit of analysis for the study is the IS application, 'ALESCO', an integrated large-scale HR application implemented at Queensland University of Technology (QUT), a large Australian university (with approximately 40,000 students and 5000 staff). Target respondents of the study were ALESCO key-user-groups: strategic users, management users, operational users and technical users, who directly use ALESCO or its outputs.

With the goal of addressing the above research question, IS-Impact and Satisfaction (and IS Support) were operationalised in a quantitative survey instrument[9]. SmartPLS (version 2.0) (Ringle, Wende, & Will, 2005) structural equation modelling employing 221 valid survey responses largely evidenced the validity of the commencing IS-Impact model in the HR context, with IS-Impact explaining 70% of Satisfaction (its immediate consequence in its nomological net).

Overview of the IS-Impact Measurement Model

The IS-Impact model (see Figure 4) is a multidimensional formative construct; comprised of Individual-Impact and Organizational-Impact dimensions - measuring net benefits to date; and comprised of System-Quality and Information-Quality dimensions - being the best proxy measure of probable future impacts. Gable et al. (2008) argue the need for only these four of the DeLone and McLean (1992) success constructs as dimensions in their multi-dimensional 'IS-Impact' measurement model. They define the IS-Impact of an Information System (IS) as "*a measure at a point in time of the stream of net benefits from the IS, to-date and anticipated, as perceived by all key user groups*" (Gable et al., 2008, p:381).

The IS-Impact model is conceived as formative at all of its levels - i.e. first-order formative, second-order formative; with lowest level items forming the four dimensions, which themselves form the IS-Impact construct. Figure 4 depicts the 37 items associated with the four IS-Impact dimensions (see Appendix B of (Gable et al., 2008) for a detailed list of the 37 items). It is noted that the IS-Impact model being multidimensional and formative at both levels, allows us to demonstrate validation procedures appropriate for any formative measurement model, involving any number of levels of formative dimensions.

[8] Discussion herein makes no reference to that study's attention to the 'IS Support' concept.
[9] A 7-point Likert scale was used to elicit responses to each of the 37 formative measures of the IS-Impact model and 6 criterion measures, from 1 (strongly disagree) to 7 (strongly agree).

DOI : 02.IJIS.2015.1.2

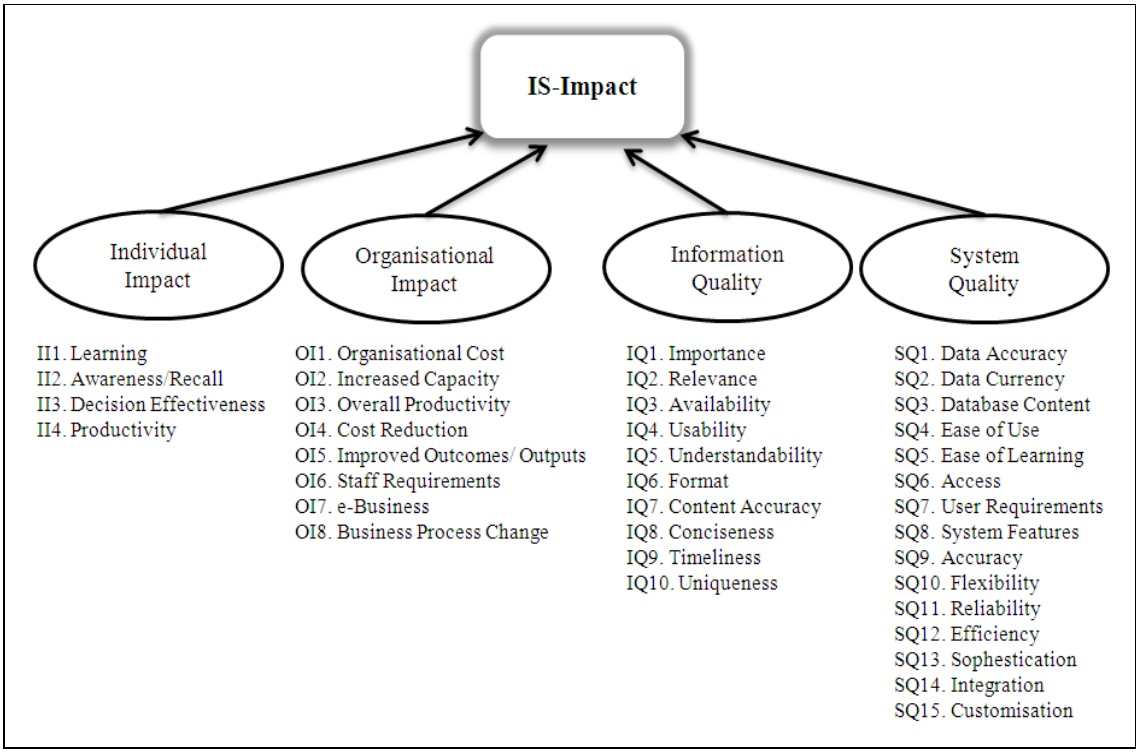


Figure 4: The IS-Impact model (adapted from Gable et al., 2008).

Overview of the Satisfaction Construct used in the Referent Study

Following Gable et al. (2008), satisfaction was conceptualized as an immediate consequance of IS-Impact, mainly with the goal of assessing IS-Impact's nomological validity (i.e. identification through structural relations)[10]. In the referent study, the satisfaction construct was measured using four indicators adopted from the overall satisfaction scale developed by Spreng et al. (1996) in the Expectation-confirmation theory **(**ECT) (Oliver, 1981) literature, which is yet considered a central theory for explaining satisfaction in marketing research (Cenfetelli, Benbasat, & Al-Natour, 2008).

The adopted Satisfaction scale was originally designed to assess users' satisfaction with camcorder use, but has since been validated in the IS context (e.g. Bhattacherjee, 2001; Bhattacherjee & Premkumar, 2004; Cenfetelli et al., 2008; Premkumar & Bhattacherjee, 2008). This adopted scale captured respondents' satisfaction levels (both in intensity and direction (Oliver, 1993, 1997)) along seven-point scales anchored between four semantic differential adjective pairs: *"frustrated/contented", "displeased/pleased", "terrible/delighted",* and *"dissatisfied/satisfied"* (Bhattacherjee, 2001).

Following, the related formative construct validation procedures are described in detail.

Assessment Results

The IS-Impact model is conceptualized as a formative first-order, formative second-order model (i.e. Type IV in Jarvis et al. (2003) specification of multidimensional constructs). Gable et al., (2008) conceptualize the four first-order dimensions (i.e. individual impact, organisational impact, information quality and system quality) as having mutually exclusive formative indicators that need not co-vary. In the organisational impact dimension, for example, an IS may result in overall productivity improvement, but it may not be cost effective. In further example, in the information quality dimension, information from an IS may be easy to understand, but it may not

[10] See Gable et al., (2008) for more information regarding the notion of having satisfaction as an immediate consequence of IS-Impact.

DOI : 02.IJIS.2015.1.2

be available in a timely fashion. Thus, individual impact, organisational impact, information quality and system quality are formative first-order dimensions.

Additionally, the impact of an IS (individual impact or organisational impact) as well as the quality of the IS (information quality or system quality), may change over time and be affected by different factors. Thus, when evaluating the impact of an IS one would be mistaken to simply equate, in example, individual impact and information quality. In further example of the mutual exclusivity of the four dimensions, a change in organisational impact of an IS does not imply a similar change in the information quality of this IS. In other words, the four dimensions that form the IS-Impact model are not interchangeable. As such, these imply that individual impact, organisational impact, information quality and system quality affect IS-Impact in a formative way. Accordingly, the IS-Impact model is a formative-second order construct (Gable et al., 2008).

Assessment of the IS-Impact model is conducted in three stages, namely: assessment of the first-order formative indicators, assessment of the second-order formative dimensions, and assessment of the IS-Impact model at the construct level.

Assessment of the First-Order Formative IS-Impact Model

This section describes validity assessment tests conducted on the first order (i.e. indicators' level) of the IS-Impact model, including: assessing the degree of multicollinearity, assessing indicator weights, assessing indicator loadings, and assessing indicator criterion validity.

Assessment of the degree of multicollinearity entails the calculation of variance inflation factors (VIF) or the tolerance values (Götz et al., 2010). Several ordinary least squares (OLS) regressions were performed, with the first-order formative indicators as the independent variables and the criterion measures of each dimension (i.e. Individual Impact, Organisational Impact, Information Quality, and System Quality) as the dependent variable, to obtain VIF and tolerance scores. Table 2 displays the results. Many researchers consider VIFs up to 10 acceptable[11] (e.g. Diamantopoulos & Siguaw, 2006; Gefen et al., 2011; Gefen et al., 2000; Götz et al., 2010; Gujarati 2003; Henseler et al., 2009; Petter et al., 2007, Rabaa'i and Gable, 2012). The largest VIF in Table 2 is 5.6, suggesting multicollinearity is not affecting the IS-Impact data in this sample.

[11] A VIF of 10 implies a Tolerance of 0.10 meaning that 90% of the variance in the item is explained by the other items. This of course also means 10% of the variance in the item isn't explained by the other items.

DOI : 02.IJIS.2015.1.2

Table 2: Tolerance, VIF, Weight, Loading and Significance level for the first-order (indicators' level) IS-Impact model[12]

Individual Impact		**Tolerance**	**VIF**	**Weight**	**Loading**	**Significance**
II1	Learning	0.313	3.195	0.239	0.815	$p < 0.05$
II2	**Awareness/Recall**	0.343	2.915	**0.044**	**0.804**	***ns***
II3	Decision Effectiveness	0.186	5.376	0.446	0.979	$p < 0.001$
II4	Individual Productivity	0.180	5.556	0.339	0.961	$p < 0.05$
Organizational Impact						
OI1	**Organisational Cost**	0.460	2.174	**0.062**	**0.725**	***ns***
OI2	Increased Capacity	0.247	4.049	0.223	0.883	$p < 0.05$
OI3	Overall Productivity	0.200	5.000	0.312	0.932	$p < 0.05$
OI4	**Cost Reduction**	0.298	3.356	**0.120**	**0.728**	***ns***
OI5	**Improved Outcomes/Outputs**	0.190	5.263	0.299	0.907	$p < 0.05$
OI6	Staff Requirements	0.352	2.841	**0.100**	**0.654**	***ns***
OI7	**e-Business**	0.240	4.167	0.359	0.912	$p < 0.01$
OI8	**Business Process Change**	0.322	3.106	**0.042**	**0.776**	***ns***
Information Quality						
IQ1	Importance	0.728	1.374	0.174	0.713	$p < 0.05$
IQ2	Relevance	0.336	2.976	0.331	0.894	$p < 0.001$
IQ3	**Availability**	0.441	2.268	**0.073**	**0.699**	***ns***
IQ4	**Usability**	0.277	3.610	**0.087**	**0.855**	***ns***
IQ5	**Understandability**	0.268	3.731	**0.119**	**0.836**	***ns***
IQ6	Format	0.202	4.950	0.295	0.905	$p < 0.001$
IQ7	**Content Accuracy**	0.937	1.067	**0.049**	**0.047**	***ns***
IQ8	**Conciseness**	0.417	2.398	**0.068**	**0.725**	***ns***
IQ9	**Timeliness**	0.427	2.342	**0.095**	**0.668**	***ns***
IQ10	**Uniqueness**	0.863	1.159	**0.044**	**0.053**	***ns***
System Quality						
SQ1	**Data Accuracy**	0.687	1.456	**-0.007**	**0.668**	***ns***
SQ2	Data Currency	0.654	1.529	0.180	0.783	$p < 0.01$
SQ3	**Database Content**	0.717	1.395	**0.005**	**0.563**	***ns***
SQ4	Ease of Use	0.195	5.128	0.229	0.839	$p < 0.05$
SQ5	Ease of Learning	0.213	4.695	0.230	0.828	$p < 0.05$
SQ6	**Access**	0.701	1.427	**0.056**	**0.553**	***ns***
SQ7	User Requirements	0.346	2.890	0.306	0.826	$p < 0.01$
SQ8	**System Features**	0.290	3.448	**0.018**	**0.787**	***ns***
SQ9	System Accuracy	0.315	3.175	0.119	0.822	$p < 0.05$
SQ10	**Flexibility**	0.523	1.912	**-0.094**	**0.654**	***ns***
SQ11	**Reliability**	0.368	2.717	**0.017**	**0.576**	***ns***
SQ12	Efficiency	0.348	2.874	0.165	0.759	$p < 0.05$
SQ13	**Sophistication**	0.579	1.727	**-0.029**	**0.573**	***ns***
SQ14	**Integration**	0.415	2.410	**0.128**	**0.692**	***ns***
SQ15	**Customisation**	0.660	1.515	**0.011**	**0.517**	***ns***
ns: non-significant						

PLS path modeling does not directly provide significance tests and confidence interval estimates of path coefficients (i.e. indicators' weight) in the research model (e.g. Rai et al., 2006). Hence, the significance of formative indicator weights can be determined by means of bootstrapping (e.g. Chin, 1998b, 2010a; Davison & Hinkley, 2003; Götz et al., 2010; Henseler et al., 2009; Tenenhaus et al., 2005; Urbach & Ahlemann, 2010). Therefore, a bootstrap analysis was performed with 200 subsamples and path coefficients were re-estimated using each of these samples. Results are presented in Table 2. While no minimum threshold values for formative indicator weights have been widely agreed (e.g. Rai et al., 2006), a high indicator weight suggests that the indicator is making a significant contribution to the formative construct

[12] Refer to (Gable et al., 2008: 405) Appendix B for complete descriptions of the 37 IS-Impact items.

DOI : 02.IJIS.2015.1.2

(Diamantopoulos, 2006). Additionally, a significance level of at least .050 suggests that an indicator is relevant for the construction of the formative construct and, thus, demonstrates a sufficient level of validity (e.g. Urbach & Ahlemann, 2010, p: 20). Moreover, it is also recommended that the path coefficients (between formative indicators and their perspective construct) should be greater than .100 (Andreev, et al., 2009; Jahner, et al., 2008) or .200 (Chin, 1998b). For the IS-Impact model, a large number of indicators' weights (21 out of the 37) (**bolded**) had non-significant path coefficients (i.e. non-significant weights).

Additionally, Table 2 presents the formative indicator loadings in the IS-Impact construct with all, except IQ7 (Content Accuracy) and IQ10 (Uniqueness) (shaded), having high loadings (i.e. zero-order bivariate correlation) on the IS-Impact construct.

To assess the formative indicator criterion validity, the formative indicators of the IS-Impact construct were correlated (using SPSS version 18) with 'global measures' (i.e. criterion measures) that summarize the essence of the dimensions and the IS-Impact construct, as suggested by (e.g. Diamantopoulos et al., 2008). The global measures used are presented in Table 3[13].

Table 3: The criterion measures used to validate the IS-Impact model

Measure		Description
II-CM	Individual Impact Criterion Measure	The impact of [the IS] on me has been positive.
OI-CM	Organisational Impact Criterion Measure	The impact of [the IS] on the faculty/division has been positive.
IQ-CM	Information Quality Criterion Measure	[The IS] system quality is satisfactory.
SQ-CM	System Quality Criterion Measure	[The IS] information quality is satisfactory.
CM1	IS-Impact Global Criterion Measure (1)	The net benefits from [the IS] to date and anticipated are substantial.
CM2	IS-Impact Global Criterion Measure (2)	The lifecycle-wide positive impacts of [the IS] are substantial.

Results in Table 4 demonstrate that all IS-Impact formative indicators except IQ7 and IQ10 (shaded), have high significant correlation with: (1) the two IS-Impact criterion measures (CM1 and CM2), and (2) their respective dimensions' criterion measure (i.e. the criterion measure that summarizes essence of each dimension) at 0.01 level.

The two formative indicators, IQ7, Content Accuracy, and IQ10, Uniqueness, have shown: (1) non-significant weights of 0.049 and 0.044 respectively (see Table 2), (2) low loadings (i.e. zero-order bivariate correlation) on the IS-Impact construct of 0.047 and 0.053 respectively (see Table 2), and (3) non-significant correlations (correlating the formative indicators with the two global IS-Impact measures and the respective dimension criterion measures (see Table 2). Hence, these two indicators were excluded from the IS-Impact model. It is noted that only with this complete, consistent, corroborating evidence were we comfortable with removing these two items. The small loadings are felt to be particularly consequential for this decision.

[13] It should be noted that the four criterion measures used to validate the first-order formative dimensions (i.e. II-CM, OI-CM, IQ-CM and SQ-CM) were adapted from (Gable et al., 2008). However, the two global criterion measures used to validate the IS-Impact construct (i.e. CM1 and CM2) are new and were not used in the Gable et al. (2008) study.

DOI : 02.IJIS.2015.1.2

Table 4: First-order indicators validity of the IS-Impact model				
	Individual Impact	**Correlation with the respective dimension CM†**	**Correlation with CM1**	**Correlation with CM2**
II1	Learning	.632**	.501**	.492**
II2	Awareness/Recall	.603**	.458**	.485**
II3	Decision Effectiveness	.728**	.567**	.555**
II4	Individual Productivity	.739**	.569**	.547**
	Organizational Impact			
OI1	Organisational Cost	.556**	.577**	.585**
OI2	Increased Capacity	.635**	.660**	.657**
OI3	Overall Productivity	.690**	.696**	.703**
OI4	Cost Reduction	.569**	.582**	.635**
OI5	Improved Outcomes/Outputs	.699**	.700**	.721**
OI6	Staff Requirements	.537**	.577**	.639**
OI7	e-Business	.691**	.655**	.716**
OI8	Business Process Change	.571**	.577**	.585**
	Information Quality			
IQ1	Importance	.393**	.373**	.387**
IQ2	Relevance	.688**	.684**	.657**
IQ3	Availability	.593**	.595**	.545**
IQ4	Usability	.649**	.619**	.592**
IQ5	Understandability	.627**	.559**	.535**
IQ6	Format	.653**	.646**	.614**
IQ7	**Content Accuracy**	**.127**	**.0480**	**.0370**
IQ8	Conciseness	.544**	.576**	.495**
IQ9	Timeliness	.612**	.583**	.506**
IQ10	**Uniqueness**	**.101**	**.0590**	**.123**
	System Quality			
SQ1	Data Accuracy	.298**	.317**	.252**
SQ2	Data Currency	.366**	.424**	.367**
SQ3	Database Content	.285**	.293**	.265**
SQ4	Ease of Use	.629**	.535**	.539**
SQ5	Ease of Learning	.622**	.473**	.478**
SQ6	Access	.266**	.210**	.176**
SQ7	User Requirements	.660**	.612**	.573**
SQ8	System Features	.641**	.619**	.596**
SQ9	System Accuracy	.685**	.669**	.654**
SQ10	Flexibility	.522**	.465**	.529**
SQ11	Reliability	.474**	.400**	.428**
SQ12	Efficiency	.546**	.491**	.480**
SQ13	Sophistication	.487**	.434**	.467**
SQ14	Integration	.555**	.572**	.566**
SQ15	Customisation	.493**	.338**	.375**

† Correlations of the indicators with the respective criterion measures (i.e. II-CM, OI-CM, IQ-CM, and SQ-CM).

** Correlation is significant at the 0.01 level (2-tailed).

DOI : 02.IJIS.2015.1.2

Assessment of the Second-Order Formative IS-Impact Model

This section discusses validity assessment tests conducted on the second order (i.e. dimensions' level) of the IS-Impact model, including: assessing the degree of multicollinearity, assessing indicator weights, assessing indicator loadings, and assessing indicator criterion validity.

Following the approach of Rai et al. (2006), the averages of the indicators used to measure each of the dimensions in the first-order (i.e. Individual Impact, Organisational Impact, Information Quality, and System Quality) were computed and used as formative indicators for the IS-Impact construct (at the second-order)[14]. To assess the degree of multicollinearity among the second-order formative indicators, OLS regressions were run, with the mean value of the indicators, of each dimension (i.e. Individual Impact, Organisational Impact, Information Quality and System Quality), as the independent variables and the criterion measures as the dependent variables to obtain VIF and tolerance scores. Table 5 depicts the results. All VIFs are less than 4.7, which suggests multicollinearity is not an issue in this data set.

Table 5: Tolerance, VIF, Weight, Loading and Significance level for the first-order (dimensions' level) IS-Impact model

	Tolerance	VIF	Weight	Loading	Significance
Individual Impact	0.410	2.439	0.341	0.7952	$p < 0.001$
Organisational Impact	0.325	3.077	**0.092**	**0.7024**	***ns****
Information Quality	0.216	4.630	0.258	0.8783	$p < 0.01$
System Quality	0.267	3.745	0.368	0.9051	$p < 0.001$
ns: non-significant					

To estimate second-order formative indicator weights (i.e. path coefficients) a bootstrap analysis was performed with 200 subsamples and path coefficients were re-estimated using each of these samples, results of which are also presented in Table 5. All second-order formative indicators, except Organisational Impact (shaded), have strong and significant path weights. Additionally, all second-order formative indicators, with no exception, show high loading (i.e. zero-order bivariate correlation) on the IS-Impact construct.

As suggested by Diamantopoulos et al., (2008) and Diamantopoulos and Winklhofer (2001), the second-order formative indicators of the IS-Impact construct were correlated (using SPSS version 18) with the 'global measures' that summarize the essence of the IS-Impact construct (i.e. CM1 and CM2) to assess second-order formative indicator validity. Table 6 illustrates the results.

Table 6: Second-order indicators validity of the IS-Impact model

	Correlation with CM1	Correlation with CM2
Individual Impact	.874**	.798**
Organisational Impact	.435**	.539**
Information Quality	.658**	.749**
System Quality	.598**	.624**

Results in Table 6 show that all second-order formative indicators have high and significant correlation, with the two global measures (CM1 and CM2), at 0.01 level, which further evidences the validity of the second-order formative indicators.

Assessment of the IS-Impact Model at the Construct Level

This section presents validity assessment tests conducted on the IS-Impact model at the construct level, including: assessing the nomological validity and the external validity.

[14] This technique was also used by several authors, for example, Law and Wong(1999) and Edwards (2001).

DOI : 02.IJIS.2015.1.2

The nomological validity of IS-Impact was assessed by linking the IS-Impact model with the Satisfaction construct in the nomological net. Figure 5 illustrates the results.

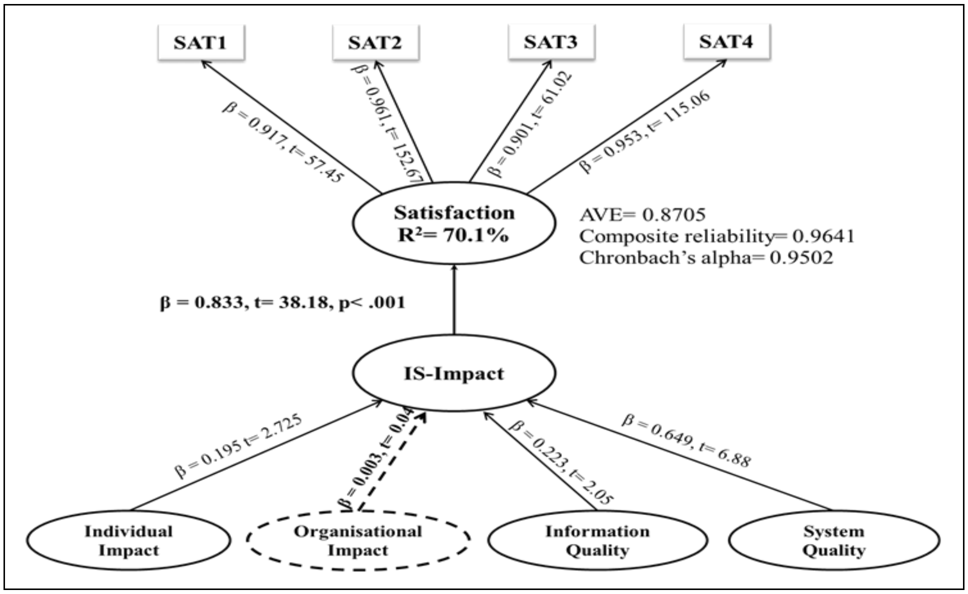


Figure 5: Nomological validity assessment of the IS-Impact model.

The results in Figure 5 show that the relationship between the IS-Impact model and Satisfaction construct is strong (β = .833, p< .001) and significant (t-value = 38.18), which supports the nomological validity of the IS-Impact model. Also, R^2 for the Satisfaction construct of 70.1% signifies that a substantial part of the variance in "Satisfaction" is explained by IS-Impact[15].

Finally, since PLS path modeling does not allow the construction of a MIMIC model, a two-construct model (Diamantopoulos & Winklhofer, 2001) was employed to assess the external validity of the IS-Impact model. Figure 6 illustrates the results.

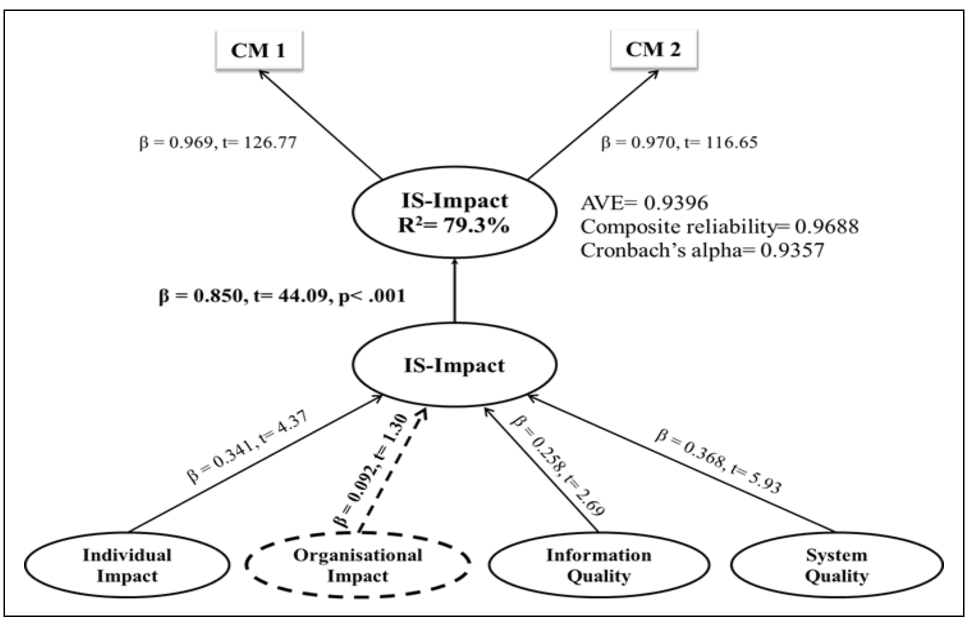


Figure 6: External validity assessment of the IS-Impact model.

[15] Discussion here and the example, include Satisfaction as an immediate consequence of the focal formative construct (the focal construct thus being in the exogenous position). Validation procedures are much the same for other possible configurations of the nomological net – e.g. where the net includes an immediate antecedent (the focal formative construct being in the endogenous position). For related discussion and examples of such correctly specified alternatives (and miss-specified alternatives), see Petter et al. (2007).

DOI : 02.IJIS.2015.1.2

The demonstrated strong (β = .850, p< .001) and significant (t-value = 44.09) connection between the formative and the reflective measurement models of IS-Impact evidences external validity. Also, R^2 for the reflective construct (phantom variable) of 79.3% indicates that a significant portion of the variance in "IS-Impact" is explained by the formative measurement model.

General Comments on Interpreting Formative Construct Validation Results

The IS-Impact model was conceptualized as a first-order formative, second-order formative multi-dimensional construct, with four formative dimensions: Individual Impact, Organizational Impact, Information Quality, and System Quality, having 4, 8, 10, and 15 indicators respectively; a total of 37 formative indicators.

Validity assessment of the IS-Impact model was conducted in three stages, namely: assessment of first-order formative indicators, assessment of the second-order formative dimensions, and assessment of the IS-Impact model at the construct level. All 37 formative indicators of the IS-Impact model were tested for possible multicollinearity. The VIF scores, which were less than the suggested threshold of 10 (see Table 2), suggest multicollinearity is not an issue in this data set.

Retaining non-significant indicators

All first-order formative indicators were assessed in term of weights, loadings, and significance (see Table 2). Twenty-one of the 37 indicators had non-significant path coefficients (i.e. non-significant weights). Since multicollinearity can be ruled out as a cause of low indicator weights, we interpret these results as a consequence of the large number of indicators (37) used originally to form IS-Impact.

The IS-Impact model consists of 37 indicators. This large number of formative indicators has *"important implications for the statistical significance and the magnitude of each indicator's weight"* (Cenfetelli & Bassellier, 2009, p: 694). For instance, a greater number of formative indicators will result in a greater likelihood that many of the indicator weights will be low in magnitude and statistically non-significant (Cenfetelli & Bassellier, 2009, p: 694). Mathieson, Peacock and Chin (2001), for example, used seven formative indicators (hardware/software, knowledge, time, financial resources, documentation, someone's help, and data) to assess their perceived resources construct. In their analysis results, four out of the seven formative indicators (e.g. financial resources and documentation), were non-significant despite their having ruled out multicollinearity (Mathieson et al., 2001, p: 108). Cenfetelli and Bassellier (2009, p: 695) argue that:

> *"Formative indicators essentially "compete" with one another to be explanatory of their targeted construct. In this competition to explain variance, only a limited number of indicators will likely be significant while the others will be nonsignificant".*

Non-significant weight of a formative indicator may lead one to conclude that an indicator has no relationship with the formative construct it measures, hence, permitting its exclusion from the model. However, MacKenzie et al. (2005, p: 712) state that:

> *"dropping a measure from a formative-indicator model may omit a unique part of the conceptual domain and change the meaning of the variable, because the construct is a composite of all the indicators".*

One main difference between reflective and formative indicators is the extent to which an indicator is required to represent the formative construct under investigation (Jarvis et al., 2003; Petter et al., 2007). As a result, formative indicators are assigned beta weights (Petter et al., 2007). Consequently, removing a non-significant formative indicator will remove the beta weight associated with it, despite how large or small it might be (Petter et al., 2007, p: 627).

DOI : 02.IJIS.2015.1.2

Hence, relying on only formative indicator weights may lead to misinterpretation of formative indicator analysis results, as one may interpret a low or non-significant formative indicator weight as the indicator is 'unimportant', despite what may be a significant zero-order correlation, thus supporting the case that the formative indicator is, indeed, 'important' for the formative construct (Cenfetelli & Bassellier, 2009, p: 697).

Cenfetelli and Bassellier (2009), in their illustration of interpreting formative measurement in IS research, use 5 formative indicators to measure Service Quality, including: Assurance, Empathy, Reliability, Responsive, and Tangibles. Their results show that both assurance and empathy have indicator weights that are not significantly different from zero (–0.08 and 0.01 respectively). However, the loadings (i.e. zero-order correlations) with Service Quality of 0.83 and 0.75, for assurance and empathy respectively, suggest that although the unique contributions of each of these indicators to Service Quality is small, in comparison to Reliability, Responsive, and Tangibles, there is a still a strong zero-order bivariate correlation between these two indicators and service quality. Cenfetelli and Bassellier (2009, p: 703) state that:

> *"Contrary to what we observe from the indicator weight results alone, these indicators are important in an absolute sense, if not a relative sense and therefore can potentially be used as a surrogate of the underlying construct if necessary".*

Formative indicator loadings of the IS-Impact model are presented in Table 2 . All, except IQ7 and IQ10, non-significant formative indicators (**bolded**) have high loadings (i.e. zero-order bivariate correlation) on the IS-Impact model. This suggests that although the unique contribution of each of these non-significant indicators to IS-Impact is small, in comparison to significant ones, there are still strong zero-order bivariate correlations between these non-significant indicators and the IS-Impact model. Hence, based on the Cenfetelli and Bassellier (2009) interpretation of formative measurement model, these results can be interpreted as follows:

> While significantly related to the IS-Impact model, [Data Accuracy], for example, does not provide additional explanatory power once other formative indicators have been taken into account, but Data Accuracy is still an important aspect of IS-Impact of its own accord.

Removing unfit indicators

All first-order formative indicators were correlated with global measures, which summarize the essence of the dimensions and the IS-Impact construct, to assess formative indicator criterion validity (see Table 3). All indicators, except IQ7 and IQ10, have high significant correlation with: (1) the two global IS-Impact measures (CM1 and CM2), and (2) the respective dimension CM at 0.01 level.

The two formative indicators IQ7, Content Accuracy, and IQ10, Uniqueness, were negatively worded items in the survey instrument; perhaps suggesting their wording may be a reason they have shown very weak correlation with (1) the Information Quality dimension, and (2) the IS-Impact construct. Varying views on and issues with negatively worded items, are reported in the literature; these have often occurred in attitudinal and perception surveys (Colosi, 2005). In social science research many researchers have reported poor performance for items that were worded negatively (DeVellis, 2003). It should be noted that three other negatively worded items SQ1 (Data Accuracy), SQ3 (Database Content) and SQ6 (Access) show high and significant correlation with (1) the System Quality dimension, and (2) the IS-Impact construct.

Nevertheless, as the two formative indicators, IQ7, Content Accuracy, and IQ10, Uniqueness, have shown: (1) non-significant weights, (2) low loadings on IS-Impact, and (3) no non-significant indicator criterion validity, they were excluded from the IS-Impact model. However, after omitting IQ7 and IQ10, we ensured, through discussion, that the remaining indicators are still measuring the entire domain of the Information Quality dimension and the content validity of the Information Quality dimension is preserved, as suggested by Petter et al. (2007), Diamantopoulos and Winklhofer (2001) and Jarvis et al. (2003).

DOI : 02.IJIS.2015.1.2

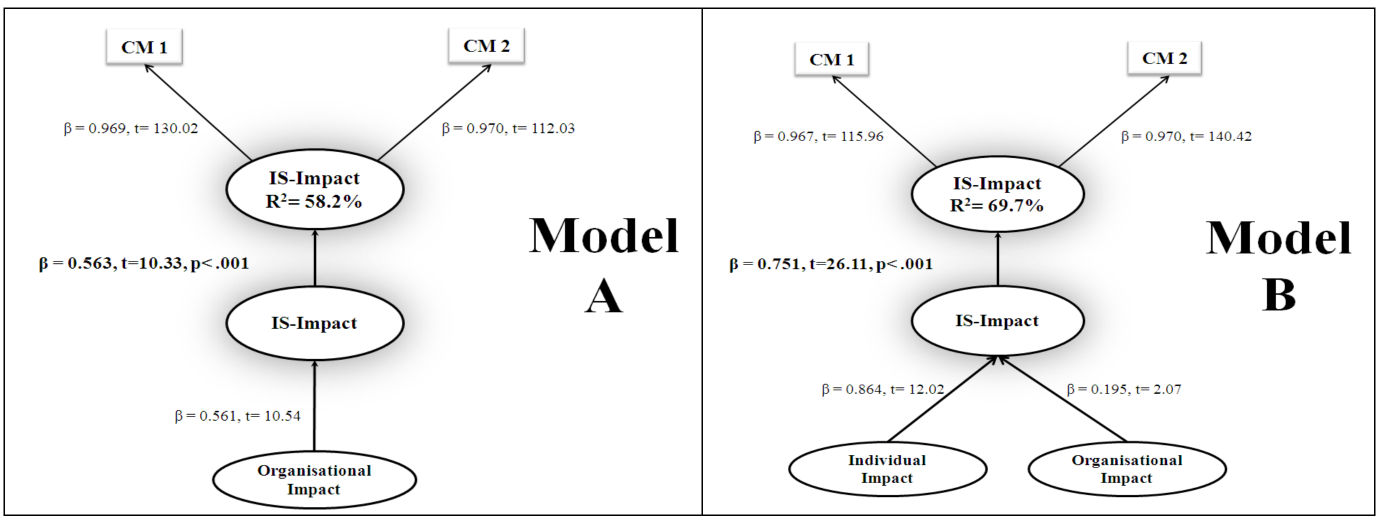
Retaining the organizational Impact dimension

To assess the second-order formative IS-Impact construct, linear composites from the indicators were created and used as formative indicators for the second-order construct. The second-order construct was assessed in terms of the degree of multicollinearity, where VIF scores are were less than 4.7 (see Table 5). Additionally, all second-order formative indicators, except Organizational Impact, have shown high weights, significant t-values, and high loadings (see Table 5).

The Organizational Impact dimension has shown low weight, non-significant path coefficient, but has shown high loading (i.e. zero-order bivariate correlation) on the IS-Impact construct (see Table 5). It is believed that the number of formative indicators used has affected the Organisation Impact weight, since multicollinearity was not an issue in this data set.

In their research article "A Theoretical Integration of User Satisfaction and Technology Acceptance", Wixom and Todd (2005), for example, introduce different antecedents of System Quality, including: Reliability, Flexibility, Integration, Accessibility, and Timeliness. In their analysis results, timeliness has a large and significant bivariate correlation (i.e. the absolute effect) with System Quality of (0.67). This absolute correlation would lead one to conclude that timeliness is highly related to system quality. Yet, when placed as an antecedent with other predictors, the relative effect of timeliness becomes non-significant, as the weight was only (0.04).

A further observation was conducted to investigate whether the presence of other dimensions is affecting the contribution (i.e. weight) of the Organizational Impact dimension to the IS-Impact model. Estimate results of PLS are depicted in Figure 7. The results demonstrate that the Organizational Impact contribution to the IS-Impact model is affected by the presence of other formative dimensions.



DOI : 02.IJIS.2015.1.2

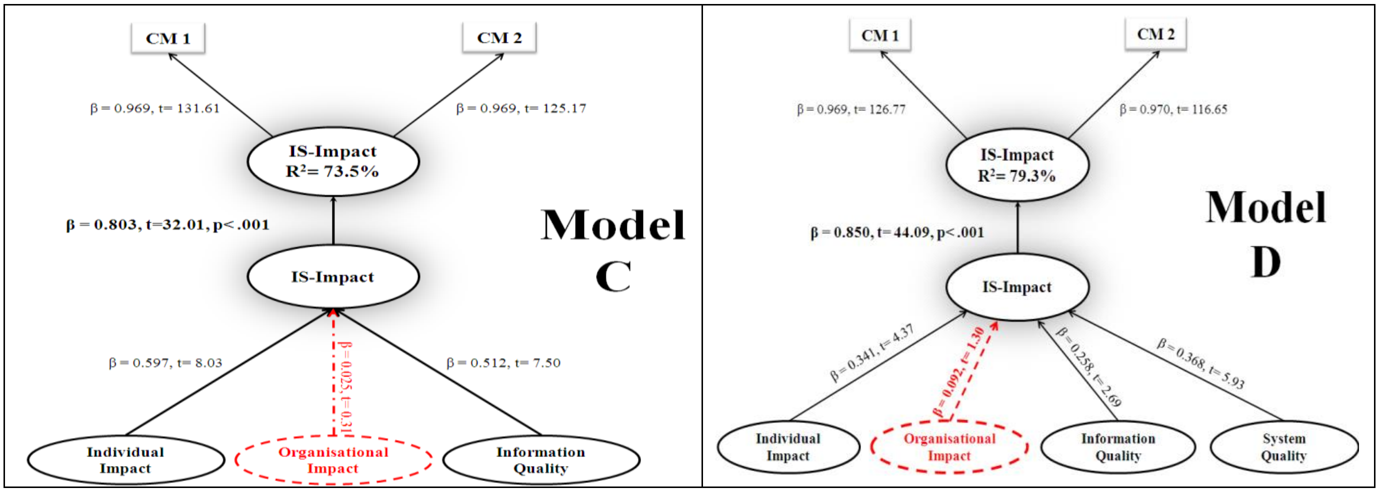


Figure 7: Organizational Impact contribution to the IS-Impact model in the presence of other formative dimensions.

In Model A, the path between Organizational Impact and IS-Impact is strong and significant with (β=0.561, t=10.54, p<.001). Also, in Model B, with the presence of Individual Impact, the path between Organizational Impact and IS-Impact is strong and significant with (β=0.195, t=2.07, p<.05). However, when placing Organizational Impact with Individual Impact and Information Quality (Model C) and Individual Impact, Information Quality and System Quality (Model D) the path between Organizational Impact and IS-Impact became non-significant.

In summary, PLS estimate results demonstrate that the Organizational Impact dimension by itself explains 58.2% of IS-Impact; however, when placed with other formative dimensions, the path weight, between Organizational Impact and IS-Impact, became statistically non-significant[16]. Hence, based on the Cenfetelli and Bassellier (2009) interpretation of formative measurement model, these results can be interpreted as follows:

> While Organizational Impact has a weight which not significantly different from zero, it has shown high loading (i.e. zero-order bivariate correlation) on the IS-Impact construct (0.7024). This absolute correlation would lead one to conclude that Organisational Impact is highly related to IS-Impact. Yet, when placed with other formative dimensions, the relative effect of Organisational Impact becomes non-significant, as the weight was only (0.092).

Moreover, all second-order formative indicators, including Organizational Impact, have high and significant correlation with the two global measures that summarize the essence of the IS-Impact construct (see Table 6) confirming the second-order indicators' validity.

The nomological validity of the IS-Impact model was assessed by connecting the IS-Impact model with the Satisfaction construct in the nomological net (see Figure 6). Results confirm the nomological validity of the IS-Impact model by (1) the strong (β = .833, p< .001) and significant (t-value = 38.18) connection between the IS-Impact and Satisfaction, and (2) R^2 value for the Satisfaction construct, of 70.1%, signifies that much of the variance in "Satisfaction" could be explained by the IS-Impact measurement model.

Finally, the external validity of the IS-Impact model was assessed through the use of two-construct model (see Figure 7), where (1) the strong (β = .850, p< .001) and significant (t-value= 44.09) connection between the formative and the reflective measurement models of IS-Impact were verified, and (2) R^2 value for the reflective phantom variable, of 79.3%, indicates that a significant part of the variance in "IS-Impact" could be explained by the IS-Impact measurement model.

[16] Actually, the fact that Organizational Impact dimension has a large loading, but non-significant weight, by definition means it overlaps with the other dimensions.

DOI : 02.IJIS.2015.1.2

Conclusion

There is increasing consciousness of formative phenomenon, a substantial proportion of the IS research community believing formative constructs serve an important purpose. Further, 'Implicit' formative construct interest may be more widespread than apparent, with researchers ostensibly employing reflective constructs, often consciously seeking item mutual exclusivity, a hallmark of formative indices. Thus the need for comprehensive and clear formative construct conceptualization and validation procedures is strong.

Structural equation modeling (SEM) has become the standard for analyzing relations between latent constructs. Component-based SEM more readily accommodates formative constructs (than does CB-SEM), a majority of proponents tending to use Partial Least Squares (PLS) in formative model testing. While validation techniques for formative constructs have been evolving, methodological guidance has been scattered, and as a community our understanding of appropriate methods of formative construct validation using PLS have evolved somewhat piecemeal, seldom offering much in the way of tangible examples.

This paper represents a synthesis of past and recent sage advice from knowledgeable methodologists, yielding a complete and readily useable set of prescriptions for best practice contemporary, quantitative formative construct validation using PLS. The prescriptions were instantiated through reference to recently published research employing the full procedures described.

The procedures and referent study (Rabaa'i, 2012) reported in this paper proceeded from the point in the construct development lifecycle at which the construct has been operationalized. The focus herein thus has been on quantitative validation employing main study data, subsequent to construct conceptualisation, specification and operationalisation; it is assumed these prior, mainly qualitative stages, have been conducted well.

This paper used the IS-Impact measurement model (see Rabaa'i, 2012; Gable et al., 2008), to empirically illustrate full procedures to be used when validating formative constructs. The referent study reports validation of a 2^{nd}-order hierarchical formative construct; formative on both levels, thereby exercising all procedures required to validate a formative construct having any number of formative levels.

Procedures are logically organized around indicators level and construct level validation. Assessing formative models at the indicators level, is in attention to the concern that each indicator contributes to the formative construct by carrying the intended meaning. Tests advocated are in attention to …

i. potential multicollinearity,

ii. formative indicator weights,

iii. formative indicator loadings, and

iv. formative indicator criterion validity.

These procedures are to be replicated at all formative levels of the focal construct, thus this paper includes a complete set of guidelines to be used when validating formative constructs whether they are multidimensional (multiple levels) or unidimensional (single level).

Construct-level validation includes tests of …

v. nomological validity, and

vi. external validity.

Assessing nomological validity involves evaluating the extent to which the formative construct behaves as expected within its network of hypotheses; those relationships between the formative construct and other of the structural model constructs which have been substantiated in prior research. Testing the external validity of a formative model entails assessing the extent to which the formative indicators in combination capture the full domain of the construct. A Multiple Indicators and Multiple Causes (MIMIC) model should be applied for the model identification procedure. As PLS does not allow the construction of a MIMIC model, an alternative is to use a two-construct model that integrates an additional *"phantom variable"*, representing the construct's reflective operationalisation.

All procedures described are instantiated through the detailed referent study. Key issues with formative construct validation are (i) the necessity of each item and (ii) the completeness of the item set. The referent study conveniently included items whose statistical indications were mixed, inviting illuminating discussion around the basis on which items should be retained or omitted. The evidence against retention of two items was consistent and strong and they were excluded. This deliberation conveniently extended to the 2nd-order dimensions of the IS-Impact model. Other examples of this decision process were cited in the literature.

It is hoped the guidance compiled herein and demonstrated through example, will be of value to both novice and more experienced researchers working with or considering formative phenomenon. While our understanding as a community will continue to evolve, it is believed the combination of procedures outlined makes clear the necessary data and tests required in order to facilitate strong formative construct validity testing.

DOI : 02.IJIS.2015.1.2

References

Agarwal, R., & Karahanna, E. (2000) "Time flies when you're having fun: cognitive absorption and beliefs about information technology usage ", MIS Quarterly, 24(4), 665-695.

Ahuja, M. K., & Thatcher, J. B. (2005) "Moving beyond intentions and toward the theory of trying: Effects of work environment and gender on post-adoption information technology use ", MIS Quarterly, 29(3), 427-459.

Albers, S. (2010). PLS and success factor studies in marketing. In V. Esposito Vinzi, H. Wang, J. Henseler & W. Chin (Eds.), Handbook of partial least squares: Concepts, methods and applications in marketing and related fields. Heidelberg: Springer.

Ali, A., Tate, M., Rabaa'i, A., & Zhang, M. (2012). A critical evaluation and comparison of two formative measures of system quality using criterion variables. Paper presented at the 23rd Australasian Conference on Information Systems, 3-5 December 2012, Geelong, Australia.

Andreev, P., Heart, T., Maoz, H., & Pliskin, N. (2009). Validating Formative Partial Least Squares (PLS) Models: Methodological Review and Empirical Illustration. Paper presented at the 13th International Conference on Information Systems (ICIS), Phoenix, Arizona.

Au, N., Ngai, E. W. T., & Cheng, T. C. E. (2008) "Extending the understanding of end user information systems satisfaction formation: An equitable needs fulfillment model approach ", MIS Quarterly, 32(1), 43-66.

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

Bassellier, G., & Benbasat, I. (2004) "Business competence of information technology professionals: Conceptual development and influence on IT-business partnerships ", MIS Quarterly, 28(2), 673-694.

Bearoch, M., Lichtenstein, Y., & Robinson, K. (2006) "Real options in information technology risk management: An empirical validation of risk-option relationships ", MIS Quarterly, 30(4), 827-864.

Berthon, P., Pitt, L., Ewing, M., & Carr, C. (2002) "Potential research space in MIS: A framework for envisioning and evaluating research replication, extension, and generation ", Information Systems Research, 13(4), 416-427.

Bhattacherjee, A. (2001) "Understanding Information Systems Continuance: An Expectation-Confirmation Model ", MIS Quarterly, 25(3), 351-370.

Bhattacherjee, A., & Premkumar, G. (2004) "Understanding Changes in Belief and Attitude toward Information Technology Usage: A Theoretical Model and Longitudinal Test ", MIS Quarterly, 28(2), 229-254.

Bollen, K. A. (1984) "Multiple indicators: Internal consistency or no necessary relationship? ", Quality & Quantity, 18(4), 377.

Bollen, K. A. (1989). Structural equations with latent variables. New York, NY: Wiley.

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

Bullock, H. E., Harlow, L. L., & Mulaik, S. A. (1994) "Causation Issues in Structural Equation Modeling Research ", Structural Equation Modeling, 1(3), 253-267.

Burton-Jones, A., & Straub, D. W. (2006) "Reconceptualizing System Usage: An Approach and Empirical Test ", Information Systems Research, 17(3), 228-246.

Campbell, D. T., & Fiske, D. W. (1959) "Convergent and discriminant validation by the multitrait-multimethod matrix ", Psychological Bulletin, 56(2), 81-105.

Cassel, C. M., Hackl, P., & Westlund, A. H. (2000) "On measurement of intangible assets: A study of robustness of partial least squares ", Total Quality Management, 11(7), 897-907.

Cenfetelli, R. T., & Bassellier, G. (2009) "Interpretation of Formative Measurement in Informations Systems Research ", MIS Quarterly, 33(4), 689-707.

DOI : 02.IJIS.2015.1.2

Cenfetelli, R. T., Benbasat, I., & Al-Natour, S. (2008) "Addressing the What and How of Online Services: Positioning Supporting-Services Functionality and Service Quality for Business-to-Consumer Success ", Information Systems Research, 19(2), 161-181.

Chin, W. (1995) "Partial Least Squares is to LISREL as Principle Components Analysis is to Common Factor Analysis ", Technology Studies, 2, 315-319.

Chin, W. (1998a) "Issues and opinion on structural equation modeling ", MIS Quarterly, 22(1), 7-16.

Chin, W. (1998b). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), Modern Methods for Business Research (pp. 295-358). Mahwah, NJ: Lawrence Erlbaum Associates.

Chin, W. (2010a). Bootstrap Cross-validation Indices for PLS Path Model Assessment. In V. E. Vinzi, H. Wang, J. Henseler & W. Chin (Eds.), Handbook of partial least squares: Concepts, methods and applications in marketing and related fields. Heidelberg: Springer.

Chin, W. (2010b). How to Write Up and Report PLS Analyses. In V. Esposito Vinzi, H. Wang, J. Henseler & W. Chin (Eds.), Handbook of partial least squares: Concepts, methods and applications in marketing and related fields. Heidelberg: Springer.

Chin, W., & Newsted, P. R. (1999). Structural Equation Modeling Analysis with Small Samples Using Partial Least Squares. Thousand Oaks, CA: Sage Publications.

Choudhury, V., & Karahanna, E. (2008) "The relative advantage of electronic channels: A multidimensional view ", MIS Quarterly, 32(1), 179-200.

Chwelos, P., Benbasat, I., & Dexter, A. S. (2001) "Empirical test of an EDI adoption model ", Information Systems Research, 12(3), 304-321.

Colosi, R. (2005). Negatively worded questions cause respondent confusion. Proceedings of the Survey Research Methods Section. Retrieved January, 7, 2010, from http://www.amstat.org/sections/srms/Proceedings/

Cronbach, L. J. (1951) "Coefficient alpha and the internal structure of tests ", Psychometrika, 16(3), 297-334.

Davison, A. C., & Hinkley, D. V. (2003). Bootstrap Methods and their Application (2 ed.). New York: Cambridge University Press.

DeLone, H., & McLean, E. (1992) " Information systems success: the quest for the dependent variable ", Information Systems Research, 3(1), 60-95.

DeVellis, R. F. (2003). Scale Development: theory and applications (2 ed.). Thousand Oaks: Sage Publications.

Diamantopoulos, A. (1999) "Export performance measurement: reflective versus formative indicators ", International Marketing Review, 16(6), 444-457.

Diamantopoulos, A. (2006) "The error term in formative measurement models: Interpretation and modeling implications ", Journal of Modelling in Management, 1(1), 7-17.

Diamantopoulos, A., Riefler, P., & Roth, K. P. (2008) "Advancing formative measurement models ", Journal of Business Research, 61(12), 1203-1218.

Diamantopoulos, A., & Siguaw, J. A. (2006) "Formative Versus Reflective Indicators in Organizational Measure Development: A Comparison and Empirical Illustration ", British Journal of Management 17(4), 263-282.

Diamantopoulos, A., & Winklhofer, H. M. (2001) "Index Construction with Formative Indicators: An Alternative to Scale Development ", Journal of Marketing Research, 38(2), 269-277.

Dibbern, J., Goles, T., Hirschheim, R. A., & Jayatilaka, B. (2004) "Information systems outsourcing: A survey and analysis of the literature ", The Data Base for Advances in Information Systems, 35(4), 6-102.

Duarte, P. A. O., & Raposo, M. L. B. (2010). A PLS Model to Study Brand Preference: An Application to the Mobile Phone Market. In V. E. Vinzi, H. Wang, J. Henseler & W. Chin (Eds.), Handbook of

DOI : 02.IJIS.2015.1.2

partial least squares: Concepts, methods and applications in marketing and related fields. Heidelberg: Springer.

Edwards, J. R. (2001) "Multidimensional Constructs in Organizational Behavior Research: An Integrative Analytical Framework ", Organizational Research Methods, 4(2), 144-192.

Edwards, J. R., & Bagozzi, R. (2000) "On the nature and direction of relationships between constructs and measures ", Psychological Methods, 5(2), 155-174.

Esposito Vinzi, V., Chin, W. W., Henseler, J., & Wang, H. (2010). Editorial: Perspectives on Partial Least Squares. In V. Esposito Vinzi, H. Wang, J. Henseler & W. W. Chin (Eds.), Handbook of partial least squares: Concepts, methods and applications in marketing and related fields. Heidelberg: Springer.

Fornell, C., & Robinson, W. T. (1983) "Industrial organization and consumer satisfaction/dissatisfaction ", Journal of Consumer Research, 9(4), 403-412.

Gable, G., Sedera, D., & Chan, T. (2008) "Re-conceptualizing Information System Success: the IS-Impact Measurement Model ", Journal of the Association for Information Systems, 9(7), 377-408.

Garson, G. D. (1998). Statnotes: Topics in Multivariate Analysis. Retrieved July 21, 2011, from http://faculty.chass.ncsu.edu/garson/PA765/structur.htm

Gefen, D., Rigdon, E., & Straub, D. (2011) "EDITOR'S COMMENTS: An Update and Extension to SEM Guidelines for Administrative and Social Science Research ", MIS Quarterly, 35(2), iii-xiv.

Gefen, D., Straub, D. W., & Boudreau, M.-C. (2000) "Structural Equation Modeling and Regression: Guidelines for Research and Practice ", Communications of the AIS, 4(7), 1-70.

Goodhue, D., Lewis, W., & Thompson, R. (2006). PLS, Small Sample Size, and Statistical Power in MIS Research. Paper presented at the 39th Hawaii International Conference on System Sciences, Kauai, Hawaii.

Goodhue, D., Lewis, W., & Thompson, R. (2012) "Comparing PLS to Regression and LISREL: A Response to Marcoulides, Chin, and Saunders ", MIS Quarterly, 36(3), 703-716.

Götz, O., Liehr-Gobbers, K., & Krafft, M. (2010). Evaluation of structural equation models using the partial least squares (PLS) approach. In V. Esposito Vinzi, H. Wang, J. Henseler & W. W. Chin (Eds.), Handbook of partial least squares: Concepts, methods and applications in marketing and related fields. Heidelberg: Springer.

Grewal, R., Cote, J. A., & Baumgartner, H. (2004) "Multicollinearity and measurement error in structural equation models: Implications for theory testing ", Marketing Science, 23(4), 519-529.

Hair, J., Ringle, C., & Sarstedt, M. (2011) "PLS-SEM: Indeed a Silver Bullet ", Journal of Marketing Theory and Practice, 19(2), 139-151.

Hauser, R. M., & Goldberger, A. S. (1971). The treatment of unobservable variables in path analysis. In H. L. Costner (Ed.), Sociological Methodology (pp. 81-117). London: Jossey-Bass.

Henseler, J., Ringle, C., & Sinkovics, R. (2009) "The use of partial least squares path modeling in international marketing ", Advances in International Marketing, 20, 277-320.

Higgins, C. A., Duxbury, L. E., & Irving, R. H. (1992) "Work-family conflict in thedual-career family ", Organizational Behavior and Human Decision Processes, 51(1), 51-75.

Hulland, J. (1999) "Use of Partial Least Squares (PLS) in Strategic Management Research: A Review of Four Recent Studies ", Strategic Management Journal, 20(2), 195-204.

J., H., Black, W. C., Babin, B. J., Anderson, R., & Tathum, R. (2006). Multivariate Data Analysis (6 ed.). Upper Saddle: Prentice Hall.

Jahner, S., Leimeister, J. M., Knebel, U., & Krcmar, H. (2008). A Cross-Cultural Comparison of Perceived Strategic Importance of RFID for CIOs in Germany and Italy. Paper presented at the 41st Annual Hawaii International Conference on System Sciences.

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003) "A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research ", Journal of Consumer Research, 30(2), 199-218.

DOI : 02.IJIS.2015.1.2

Jöreskog, K., & Goldberger, A. S. (1975) "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable ", Journal of the American Statistical Association, 70, 631-639.

Kanawattanachai, P., & Yoo, Y. (2007) "The impact of knowledge coordination on virtual team performance over time ", MIS Quarterly, 31(4), 783-808.

Kim, G., Shin, B., & Grover, V. (2010) "Investigating two contradictory views of formative measurement in information systems research ", MIS Quarterly, 34(2), 345-365.

Komiak, S. Y. X., & Benbasat, I. (2006) "The effects of personalization and familiarity on trust and adoption of recommendation agents ", MIS Quarterly, 30(4), 941-960.

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 Information Systems, 25(2), 143-160.

Liang, H., Saraf, N., Hu, Q., & Xue, Y. (2007) "Assimilation of enterprise systems: The effect of institutional pressures and the mediating role of top management ", MIS Quarterly, 31(1), 59-87.

Limayem, M., Hirt, S. G., & Cheung, C. M. K. (2007) "How habit limits the predictive power of intention: The case of information systems continuance ", MIS Quarterly, 31(4), 705-737.

Loch, K. D., Straub, D. W., & 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.

MacCallum, R. C., & Browne, M. W. (1993) "The use of causal indicators in covariance structure models: Some practical issues ", Psychological Bulletin, 114, 533-541.

MacKenzie, S. B., Podsakoff, P. M., & Jarvis, C. B. (2005) "The Problem of Measurement Model Misspecification in Behavioral and Organizational Research and Some Recommended Solutions ", Journal of Applied Psychology, 90(4), 710-730.

Malhotra, A., Gosain, S., & El Sawy, O. A. (2007) "Leveraging standard electronic business interfaces to enable adaptive supply chain partnerships ", Information Systems Research, 18(3), 260-279.

Marcoulides, G. A., & Saunders, C. (2006) "Editor's Comments: PLS: A Silver Bullet? ", MIS Quarterly, 30(2), iii-ix.

Mathieson, K., Peacock, E., & Chin, W. (2001) "Extending the Technology Acceptance Model: The Influence of Perceived User Resources ", Data Base for Advances in Information Systems, 32(3), 86-112.

Moores, T. T., & Chang, J. C.-J. (2006) "Ethical decision making in software piracy: Initial development and test of a four component model ", MIS Quarterly, 30(1), 167-180.

Nunnally, J., & Bernstein, I. (1994). Psychometric Theory (3rd ed.). New York: McGraw Hill.

Oliver, R. (1981) "Measurement and evaluation of satisfaction processes in retail settings ", Journal of Retailing, 57(3), 25-48.

Oliver, R. (1993) "Cognitive, Affective, and Attribute Bases of the Satisfaction Response ", Journal of Consumer Research, 20(3), 418-430.

Oliver, R. (1997). Satisfaction : A Behavioral Perspective on the Consumer. New York: McGraw Hill.

Pavlou, P., Liang, H., & Xue, Y. (2007) "Understanding and mitigating uncertainty in online exchange relationships: A principal-agent perspective ", MIS Quarterly, 31(1), 105-136.

Petter, S., Straub, D., & Rai, A. (2007) "Specifying Formative Constructs in Information systems research ", MIS Quarterly, 31(4), 623-656.

Premkumar, G., & Bhattacherjee, A. (2008) "Explaining Information Systems Usage: A Test of Competing Models ", Omega, 36(1), 64-75.

Rabaa'i, A. (2012). Evaluating the success of large–scale, integrated information systems through the lens of IS–Impact and IS–Support. Unpublished PhD thesis, Queensland University of Technology (QUT), Brisbane, Queensland - Australia.

DOI : 02.IJIS.2015.1.2

Rabaa'i, A., & Gable, G. (2012). IS service quality as a multi-dimensional formative construct. Paper presented at the 16th Pacific Asia Conference on Information Systems, 11-15 July, 2012. Ho Chi Minh city, Vietnam.

Rai, A., Patnayakuni, R., & Seth, N. (2006) "Firm performance impacts of digitally enabled supply chain integration capabilities ", MIS Quarterly, 30(2), 225-246.

Reinartz, W., Krafft, M., & Hoyer, W. D. (2004) "The customer relationship management process: Its measurement and impact on performance ", Journal of Marketing Research, 41(3), 293-305.

Ringle, C., Sarstedt, M., & Straub, D. (2012) "Editor's Comments: A Critical Look at the Use of PLS-SEM in MIS Quarterly ", MIS Quarterly, 36(1), iii-xiv.

Ringle, C., Wende, S., & Will, A. (2005). SmartPLS Version 2.0 M2. from http://www.smartpls.de

Rossiter, J. R. (2002) "The C-OAR-SE procedure for scale development in marketing ", International Journal of Research in Marketing, 19(4), 305-335.

Sambamurthy, V., & Chin, W. (1994) "The effect of group attitudes towards alternative gdss design on the decision-making performance of computer-supported groups ", Decision Sciences, 25(2), 215-242.

Spreng, R., MacKenzie, S., & Olshavsky, R. (1996) "A Reexamination of the Determinants of Consumer Satisfaction ", Journal of Marketing Research, 60(3), 15-32.

Straub, D., Boudreau, M.-C., & Gefen, D. (2004) "Validation guidelines for IS positivist research ", Communications of the AIS, 13(24), 380-427.

Tenenhaus, M. (2008) "Component-based Structural Equation Modelling ", Total Quality Management and Business Excellence, 19(7), 871 - 886.

Tenenhaus, M., Vinzi, V. E., Chatelin, Y.-M., & Lauro, C. (2005) "PLS path modeling ", Computational Statistics & Data Analysis, 48, 159-205.

Urbach, N., & Ahlemann, F. (2010) "Structural Equation Modeling in Information Systems Research Using Partial Least Squares ", Journal of Information Technology Theory and Application, 11(2), 5-40.

Wilcox, J. B., Howell, R. D., & Breivik, E. (2008) "Questions about formative measurement ", Journal of Business Research, 61(12), 1219-1228.

Wixom, B. H., & Todd, P. A. (2005) "A Theoretical Integration of User Satisfaction and Technology Acceptance ", Information Systems Research, 16(1), 85-102.

DOI : 02.IJIS.2015.1.2