2009

Formative and Reflective Measurement and Validation Mismatch in Survey Research: An Archival Analysis of Information Systems Success Constructs 1985-2007

Guy Grant Gable
Queensland University of Technology, g.gable@qut.edu.au

Darshana Sedera
Queensland University of Technology, d.sedera@qut.edu.au

FORMATIVE AND REFLECTIVE MEASUREMENT AND VALIDATION MISMATCH IN SURVEY RESEARCH:

AN ARCHIVAL ANALYSIS OF INFORMATION SYSTEMS SUCCESS CONSTRUCTS 1985-2007

Research-in-Progress

Guy Grant Gable
Queensland University of Technology
Brisbane – Australia
g.gable@qut.edu.au

Darshana Sedera
Queensland University of Technology
Brisbane – Australia
d.sedera@qut.edu.au

` **Abstract**

The generic IS-success constructs first identified by DeLone and McLean (1992) continue to be widely employed in research. Yet, recent work by Petter et al (2007) has cast doubt on the validity of many mainstream constructs employed in IS research over the past 3 decades; critiquing the almost universal conceptualization and validation of these constructs as reflective when in many studies the measures appear to have been implicitly operationalized as formative. Cited examples of proper specification of the Delone and McLean constructs are few, particularly in light of their extensive employment in IS research. This paper introduces a four-stage formative construct development framework: Conceive > Operationalize > Respond > Validate (CORV). Employing the CORV framework in an archival analysis of research published in top outlets 1985-2007, the paper explores the extent of possible problems with past IS research due to potential misspecification of the four application-related success dimensions: Individual-Impact, Organizational-Impact, System-Quality and Information-Quality. Results suggest major concerns where there is a mismatch of the Respond and Validate stages. A general dearth of attention to the Operationalize and Respond stages in methodological writings is also observed.

Introduction

While many studies have addressed specialized systems and various related, more specific measures of success, the six generic IS-success constructs first identified by DeLone and McLean (1992) continue to be widely employed in research (typically as the dependent variable). DeLone and McLean (2003) suggest a 7th construct yielding: (1) Individual-Impact, (2) Organizational-Impact, (3) System-Quality, (4) Information-Quality, (5) Satisfaction, (6) Use, and (7) Service-Quality. Nonetheless, recent work by Petter et al (2007), following from work by Jarvis et al. (2003), has cast doubt on the validity of many mainstream constructs employed in IS research over the past 3 decades. They critique the almost universal conceptualization and validation of these constructs as reflective, when in many studies the measures appear to have been implicitly conceived as formative. They are politic in not citing specific infractions, but rather they list a range of studies and example constructs that have been 'properly' specified as reflective or formative. It is noteworthy that no examples of the proper specification of either the Individual-Impact or Organizational-Impact constructs are cited (recognizing that their list is not intended to be comprehensive). Cited examples of the proper specification of other of the DeLone and McLean constructs are few, particularly in light of their extensive employment in IS research (e.g. only one example each of System-Quality and Information-Quality, both from the same study (Wixom et al. 2005))[1]. With the aim of exploring the extent of possible problems due to misspecification of mainstream constructs in IS research, this paper introduces a four-stage formative construct development framework: Conceive > Operationalize > Respond > Validate (CORV). Given the centrality of IS-success in important IS research nomological nets, this study employs the CORV framework in an archival analysis of past studies involving the empirical measurement of IS-success, as reported in top-tier IS outlets over the period 1985-2007.

Conceptual Framework: Conceive>Operationalize>Respond>Validate (CORV)

A key aim of the intended archival analysis is to assess the mutual exclusivity of items and completeness of constructs that are implicitly conceived as formative, but which are not subsequently fully treated as formative. Completeness or content validity is a concern at all levels of a formative construct; at the item level, and at the sub-construct level if the construct has multiple dimensions. Unlike reflective constructs, where close correspondence between items associated with a given construct is desired, with formative constructs mutual exclusivity of the items is sought in the interests of parsimony and to avoid possible multicollinearity. We suggest a 4-stage formative construct measurement framework and propose that, in order for a formative construct to be valid; it must be (1) Conceived, (2) Operationalized, (3) Responded to, and (4) Validated, all as formative. Decisions are made by the researcher and respondent, either implicitly or explicitly, regarding the formative or reflective nature of the construct, at each of these four stages. How the items are validated (Stage 4), by definition too ultimately determines how the items are subsequently combined in a construct (e.g. sum, average, loadings, weights) and analyzed in relation to other constructs or for descriptive statistics. While stages (1), (2) and (4) are decisions by the researcher, (3) is a decision of the respondent. Recent literature (e.g. (Jarvis et al. 2003; Petter et al. 2007)) provides guidance on conceiving a construct (stage 1) as a formative variable. Much of that same literature too addresses formative construct validation tests (stage 4). There are however no guidelines that address and interrelate all stages. There too is a dearth of detailed discussion on the possible consequences of inappropriate responses (for whatever reason) to ostensibly formative items (stage 3[2]). Following, each of the four CORV framework stages is briefly described.

Stage 1: How items are conceived (Conceive)

Researchers (Diamantopoulos et al. 2001; Jarvis et al. 2003) point out several characteristics of a formative model, which make it sharply distinct from the reflective model. *First*, the indicators characterize a set of distinct causes which are not interchangeable, as each indicator captures a specific aspect of the construct's domain (see also Jarvis et al., (2003); and Rossiter, (2002)); indeed, omitting an indicator potentially alters the nature of the construct (Bollen et al. 1991). *Second*, there are no specific expectations about patterns or magnitude of correlations between the indicators; formative indicators may correlate positively or negatively or lack any correlation with other measures (for a detailed discussion see (Bollen 1984)). *Third*, formative indicators have no individual measurement error terms; that is, they are assumed to be error-free in a conventional sense (Edwards et al. 2000). The error term (ζ) is specified at the construct level (MacCallum et al. 1993) and does not constitute measurement error (Diamantopoulos, 2006). *Fourth*, a formative measurement model, in isolation, is under-identified and, therefore, cannot be estimated (Bollen 1989). In contrast, reflective measurement models with three or more indicators are identified and can be estimated (e.g., see Long, (1983)). Conceptualization of a

[1] The authors do not agree to this viewpoint.
[2] Though the possible occurrence of CMV and related discussion in the literature might be considered attention to stage (3).

formative construct entails deciding that it is a composite index, then what parts make it complete. It is crucial at this stage to carefully define the construct, ideally based on a theoretical premise, as without a precise definition it is impossible to define the parts and assess completeness. Items are not fully or carefully 'worded' at this stage, but rather, the component items are identified and defined, as is their relationship with the index.

Stage 2: How items are presented (Operationalize)

Once a construct has been conceptualized, it is then operationalized in a survey instrument. Operationalization in this study refers to how items are worded, how they are presented, and what combination of constructs is employed (those that make up the model to be tested)[3]. Guided by the conceptualization from stage 1, the survey items may be designed anew, adopted as-is from prior work, or adapted from prior work. Note that assessing whether an item in any given study has been 'worded' formatively, requires that its wording be considered relative to the construct with which it is associated. In example the construct socio-economic status is typically conceived as a combination of education, income and occupation (Hauser et al. 1971). An item intended to measure Education may be reflective as regards the concept of Education (at the first-order), but formative if intended to measure socio-economic status (at the second-order). Items that are blocked within the survey instrument, preceded by a definition of the 'reflective' construct to which the researchers intend the items pertain, may be considered to have been presented more 'reflectively'. Operationalization also refers to how the instrument is administered and who is sampled.

Stage 3: How items are answered (Respond)

It is suggested that one should consider the formative or reflective nature of the 'response', as well as the formative or reflective nature of the measures. Wilcox et al, (2008) suggest that the same list of items might, depending on the wording of the general instructions, be conceptualized as either formative or reflective. In example, they depict items used by Gaski and Nevin (1985) to measure 'perceived coercive power' (actions a supplier might take to coerce) as both reflective and formative. They state … "*If the general instructions involve future actions, the responses might reflect a general capability by the supplier. Since the instructions refer to hypothetical actions the respondents are likely to reply based on some general notion of supplier capability instead of specific actions. Conversely, if the general instructions are pointing to past behavior a formative measurement model might be more applicable.*"(Wilcox et al. 2008). We herein introduce the notion of 'reflectiveness'; the extent to which responses are reflective of the construct, rather than being specific evaluations of each component item. It is argued that the distance of the respondent from the phenomenon of interest may influence the reflectiveness of their response. In example, Gable et al. (2008) note, that their Strategic respondents place relatively greater emphasis on Organizational-Impact in their overall evaluation of IS-Impact; whereas Technical respondents place relatively greater emphasis on System-Quality. To the extent that the items are worded and presented formatively (e.g. are mutually exclusive), and the respondent is 'expert' on the phenomenon (they have an informed and relevant perspective), they are more likely to respond formatively, clearly differentiating and responding separately to each item. It is nonetheless possible that even where these circumstances exist, the respondent will answer more reflectively, for example due to cognitive fatigue or lack of motivation. Regardless of whether the items are worded/presented formatively, should the respondent not be 'expert' they may have no option but to respond reflectively (if at all). In this situation, they are inadequately informed to address the more granular items, and consequently reflect on the higher-order, perhaps more abstract notion of the construct in formulating their response.

Stage 4: How items are validated (Validate)

Regardless of how items are conceived, operationalized and answered, they may be validated as either reflective or formative. Reflective construct validation focuses on anticipated covariance among the reflective items. The most commonly employed Exploratory Factor Analysis (EFA), Correlation and Reliability tests in such studies are covariance based approaches and are inappropriate for establishing validity of a formative construct. Formative construct validation places much emphasis on face validity, subsequently employing techniques of identification through measurement relations and identification through structural relations (Howell et al. 2007; Jarvis et al. 2003; Petter et al. 2007; Williams et al. 2003).

[3] While most studies entail multiple constructs configured as a model, for simplification this paper addresses the less usual example of a single construct.

CORV Scenarios

This study is primarily concerned with past studies that have employed items that are explicitly or implicitly conceived as formative, but which are not subsequently fully treated as formative. Table 1 depicts all possible such scenarios, the bottom row indicating whether the scenario is concerning.

			Most Likely Scenarios				Unlikely Scenarios		
	Stage		(a)	(b)	(c)	(d)	(e)	(f)	(g)
Measure	Conceive	(1)	F	F	F	F	F	F	F
	Operationalize	(2)	R	F	R	F	R	R	F
	Respond	(3)	R	R	F	F	R	F	R
Test	Validate	(4)	R	R	R	R	F	F	F
	Issues?		?	?	Yes	Yes	?	No	?

Table 1: CORV scenarios

In scenarios (a) through (d) the researcher, though having conceived the construct as formative (in the study sample almost always unknowingly), has validated it as reflective.

(a) FRRR – Items are conceived formatively but presented reflectively, answered reflectively[4] and analyzed reflectively. Since the items are operationalized as reflective, and responded to as reflective, model completeness may be compromised, with respondents addressing only a part of the formative construct. Employing covariance based validation methods, such as Exploratory Factor Analysis, in formative construct validation may result in the dropping of important items.

(b) FFRR – In this scenario, though the items have been conceived formatively and presented formatively, they have been answered and validated reflectively. This scenario may arise where the respondent is essentially unable to answer the items at the lower, more granular formative level, possibly due to their lacking the understanding, exposure, experience, or motivation. The respondents in this scenario, however are inclined to respond (e.g. because the survey has been mandated, the instrument does not allow missing items, there is no 'not applicable' option), and therefore respond at a higher, more abstract level, which may reflect their perception of the composite phenomenon (e.g. through meetings, hallway chatter, news broadcasts, etc.).

(c) FRFR – This scenario is in some sense the opposite of scenario (b) in that even though the items may have been presented as reflective (e.g. blocked with reflective definition), the respondent is very close to the data being sought and is uninfluenced by this, answering each question distinctly and formatively. The model validation entails reflective testing procedures employing covariance based approaches, which, given the formative responses, may result in the inappropriate exclusion of possibly valid items thereby compromising content validity (similar consequences to scenario (a)).

(d) FFFR – This is the classic case suggested in more recent literature on misconceiving formative as reflective. Here, the items have been conceived, presented and answered formatively, then quite inappropriately validated as reflective. This paper suggests however that scenario (a) in fact is the more commonplace, where items, though conceived formatively, have been presented, answered and validated reflectively.

Scenarios (e) through (g) are included for completeness; these scenarios being unlikely, mainly due to the dearth of use of formative constructs. These are scenarios in which the researcher has intentionally both conceived and analyzed the data as formative.

(e) FRRF – as with scenarios (a) and (c) (and f), the formatively conceived items are presented reflectively, (e.g. blocked with reflective definition), thus increasing the likelihood that they may be answered reflectively, consequences of which in analysis are beyond the scope of this paper (and being assessed in continuing research).

(f) FRFF – though items are presented reflectively, they are otherwise treated as formative. One must question whether, given formative wording, answers and analysis, the presentation can be considered reflective. This scenario assumes that respondents are appropriately close to the data and able and willing to answer formatively.

(g) FFRF – In this scenario, the researcher has done everything right, except perhaps choice of respondent. The items are worded, presented and analyzed as formative, yet the respondent has answered reflectively, either due to inability (distance from the phenomenon – e.g. due to lack of relevant expertise or experience) or

[4] It is assumed, where an item is answered reflectively, it is most likely reflective of the overarching construct, but could be reflective of some component or sub-construct or dimension.

laziness (it is implicit that 'answering' formatively is cognitively more demanding than answering reflectively, because the respondent is having to essentially answer separate and distinct questions).

The bottom row of Table 1 flags those scenarios having a concerning combination of stages. For scenario (f) we are reasonably confident, ceteris paribus, that there are no related concerns (deriving from the combination of stages as discussed herein[5]). In scenarios (c) and (d) there is concern due to reflective construct validation of formatively answered items; this may well result in the exclusion of valid formative items due to their lack of covariance with other of the items, and thus consequent concerns with completeness (note however, that should the items naturally co-vary e.g. due to 'common cause', there may be no problem) (Wilcox et al. 2008). The extent of concern with scenarios (a) and (b) is less clear, in which scenarios; formatively worded items have been answered and validated reflectively. Neither is it clear what the implications are of (e) and (g) where items answered reflectively are validated formatively. It is generally noted that concerns arise primarily where there is a mismatch between how the items are answered and how they are analyzed (Stages 3 and 4).

Referent Theory: the IS-Impact Measurement Model

To evaluate completeness and mutual exclusivity of past study constructs and items, a master set of sub-constructs and items are required for comparison. This study adopts the IS-Impact measurement model (Gable et al. 2008) for this purpose. IS-Impact purportedly represents a complete set of mutually exclusive items pertaining to the four constructs of interest in this study. Gable et al. (2008) argue the need for only four of the DeLone and McLean (2003) 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. 2008:381). They conceptualize IS-Impact as a bicameral, formative index; the impact half, comprised of Individual-Impact and Organizational-Impact, measuring net benefits to date; the quality half, comprised of System-Quality and Information-Quality, being our best proxy measure of probable future impacts. The IS-Impact model depicted in is conceived as formative at all of its levels; with lowest level items forming the four dimension (see Gable et al. (2008:405) for a detailed list of the 37 questions), the pairs of dimensions forming the halves, and the halves forming the IS-Impact index. It is this ostensibly complete pool of 37 ostensibly mutually exclusive items that is employed following for mapping purposes.

The Study Approach

In order to contain the study effort, the archival analysis was constrained to the period 1985-2007 and the following top-tier outlets: MIS Quarterly (MISQ), Information Systems Research (ISR), Management Science (MS), Journal of MIS (JMIS), Journal of the AIS (JAIS), Decision Sciences (DS), Information & Management (I&M) and European Journal of Information Systems (EJIS), as well as the International Conference on Information Systems (ICIS). Papers were sought that reported the empirical measurement of at least one of the four IS-impact model dimensions. Each of the 43 papers was content analyzed for: (1) respondent cohort/s sampled, (2) 'blocking' or not of items in the survey instrument, and (3) model validation test/s.

Discussion

Stage 1 – Conceive:

Table 3 maps measures employed in each of the 43 studies against the 37-item pool. Table 3 specifically addresses the completeness of measures in formative construct validation. A good formative index is one that exhausts the entire domain of the construct completely, meaning that the items should collectively represent all the relevant aspects of the construct of interest (Bagozzi et al. 1982a; Bagozzi et al. 1982b; Fornell et al. 1982). Given the composite nature of formative constructs, 'completeness' is centrally important in conceptualization. Without considering the intent of each study, each of which may have been quite specific and constructs and measures employed entirely appropriate, it is noted that the substantial white space in Table 3 suggests potential content validity issues (83% empty … 274 populated cells / 1591 total cells = 17%; where 43 studies x 37 items = 1591). Gable et al. (2008) synthesize their final pool of 37 ostensibly mutually exclusive items from a starting

[5] This discussion has focused on the 'combination' of stages. Though we indicate 'No' issues with scenario (f) in Table 1, it is acknowledged that where a construct is 'implicitly' conceptualized as formative, even though component items may thus be mutually exclusive, implicit rather than explicit conceptualization may well result in problems with completeness.

set of 119 measures, suggesting substantial redundancy in the starting 119. Given this redundancy, it is possible that in studies reported in Table 3, researchers have selected from the 119, multiple, synonymous 'reflective' measures of some 'part' of one or more of the four constructs being evaluated (none of the 119 items serves well as a single overarching criterion measure of any of the four constructs). This was not the case. Thus, while such mapping activities undoubtedly entail some subjectivity, it is revealing to note that in none of the studies did multiple items map to a single item in the set of 37 (all cells in Table 3 are either blank or a '1'). This is strong evidence that broadly, the 43 studies sought to achieve mutual exclusivity of items, thus implying a formative conceptualization.

Stage 2 - Operationalize

Important aspects of Operationalization include item wording, item position, item introduction/definition, and survey administration. The archival analysis addressed only item position, and more specifically, whether items were scattered or logically blocked by construct. Petter et al (2007:636) suggest "*One viable explanation for high reliabilities for this particular operationalization of the construct, for instance, is the blocking of items and the likelihood of huge common methods bias in the typical administration of this instrument*" (Burton-Jones et al. 2006). The specific wording of the items too can be consequential and may influence cross-item responses. This aspect is not addressed herein. How the instrument is administered too is consequential in operationalization; neither is this addressed in this study. From Table 2 we observe that 33 (77%) of the 43 studies blocked items for presentation in the survey instrument. Though admittedly simplistic, for the sake of discussion, we further assume that all studies that have blocked related items in the survey instrument, have 'presented' the items reflectively (the extent of covariance observed across the studies too suggests this).

Stage 3 - Respond

It is acknowledged that what was in the mind of the respondent when they answered each question is difficult to assess even by the original researchers (let alone in archival analysis), though gathering various demographics and conducting related analyzes may provide relevant evidence, as can pilot and follow-on research with respondents through focus groups, interviews or other surveys. Nonetheless, we were able to identity for each study in Table 2, the explicit or implicit respondent cohorts, as one or more of: Strategic, Management, Operational, Technical or External (these cohorts having been identified by Gable et al (2008)). It is thus possible to consider the appropriateness of the specific respondent cohorts sampled in each study, in combination with the specific constructs measured. Though it is difficult to make conclusive observations from Table 2, a diversity of construct and cohort combinations is apparent. Generally, it is suggested that Strategic respondents often will not be well placed to evaluate System-Quality; whereas Technical respondents are often not well placed to evaluate Organizational-Impact. Well informed respondents are likely to respond more formatively, and less informed respondents are more likely to answer more reflectively. Though we cannot know from the published studies how respondents actually 'answered' the items, we expect that most often, when the items are presented 'reflectively', they are answered reflectively. Given that all 43 studies in Table 2 employed covariance-based statistical tests (and given the quality of the publication outlets sampled), we further assume that strong covariance was observed in all studies[6] as anticipated by the respective researchers (e.g. generally large Cronbach Alphas >0.7 were reported). Logically, for such covariance to exist there would seem to be three possibilities: (1) respondents have responded reflectively at a higher level of abstraction, (2) the items have a strong common cause, or (3) common method variance (CMV). With regards to (3) it is observed that the 43 studies generally evidence discriminant validity[7] between their model constructs (e.g. through EFA/CFA factor loadings) thus arguing against the existence of CMV between constructs. As regards (1) and (2), though it is not possible from study data to differentiate instances of 'reflectiveness' from 'common cause', it is assumed that some portion of the studies experienced reflective responses.

Stage 4 - Validate

Commonly employed reflective construct validation techniques include: Exploratory Factor Analysis (EFA), Correlation and Reliability tests. These methods are covariance based and are inappropriate for establishing validity of a formative model. Though 16 of the 43 studies employed Structural Equation Modeling (SEM) in data analysis, none of these specifically addressed formative construct validation or employed SEM appropriately for formative tests.

---

[6] This was not explicitly coded in the archival analysis, but is being addressed through follow-on work.
[7] Though not coded explicitly for each study, this is being addressed in follow-on work.

Table 2: The 43 Studies[8]

				Conceive					Operation-alize	Respond					Validate	
				Implied	Dimension				Items	Employment Cohorts						N-Net
#	Study	Year	Outlet	Formative	II	OI	SQ	IQ	Blocked	S	M	O	T	X	Tests Conducted	Tested
1	Sanders and Courtney	1985	MISQ	Y	Y			Y	Y	-	Y	Y	-	-	Alpha,Regre,F Test,Cor	No
2	Raymond	1985	MISQ	Y			Y	Y	No	Y	Y	Y	-	-	EFA	No
3	Barki and Huff	1985	I&M	Y	Y				Y	Y	Y	Y	-	-	Cor,Alpha	Y
4	Mahmood and Medewitz	1985	I&M	Y			Y	Y	Y	-	-	-	-	Y	VA	No
5	Srinivasan	1985	MISQ	Y			Y	Y	Y	Y	Y	Y	-	-	Cor	No
6	Miller and Doyle	1987	MISQ	Y			Y	Y	Y	Y	Y	Y	-	-	EFA,DC,VA	No
7	Raymond	1987	I&M	Y			Y	Y	No	-	Y	-	-	-	EFA	No
8	Tait and Vessey	1988	MISQ	Y				Y	Y	-	Y	Y	-	Y	Alpha	Y
9	Raymond	1990	JMIS	Y			Y	Y	Y	Y	Y	-	-	-	Cor	No
10	Guimaraes et al.	1992	DSS	Y	Y			Y	Y	-	Y	-	-	-	EFA,VA,R2	No
11	Doll, Xia and Torkzadeh	1994	MISQ	Y			Y	Y	Y	-	-	Y	-	-	CFA,GFI	No
12	Sethi and King	1994	DSS	Y			Y	Y	Y	Y	Y	Y	-	-	VA,GFI,EFA	Y
13	Gatian	1994	I&M	Y	Y		Y		Y	-	-	Y	-	-	RMSR,GFI	Y
14	Rainer and Watson	1995	JMIS	Y	Y		Y	Y	Y	-	-	-	-	Y	EFA,Cor,Alpha,	No
15	Brown, Gatian and Hicks	1995	JMIS	Y		Y			Y	-	Y	Y	Y	Y	Cor,T-tests	No
16	Stylianou et al.	1996	I&M	Y		Y			Y	-	Y	-	-	-	Cor	No
17	Law and Gorla	1996	I&M	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	-	SEM	No
18	Saarinen	1996	I&M	Y			Y	Y	Y	-	Y	-	-	-	Cor	No
19	Bajwa, Rai and Brennen	1997	DSS	Y		Y			Y	Y	Y	Y	Y	-	SEM	Y
20	Guimaraes	1997	DSS	Y	Y		Y	Y	Y	-	Y	-	-	-	Alpha,DC,Cor	Y
21	Li	1997	I&M	Y			Y	Y	Y	Y	Y	Y	-	-	chi square ave	No
22	Lu and Wang	1997	I&M	Y			Y	Y	No	-	Y	-	-	-	Cor	No
23	Essex, Magal, Masteller	1998	JMIS	Y			Y	Y	Y	-	Y	Y	-	-	EFA,Cor,	No
24	Goodhue	1998	DSS	Y			Y	Y	No	-	-	Y	-	-	DC,CEFA,Alpha	Y
25	McHaney and Cronan	1998	DSS	Y			Y	Y	Y	-	Y	Y	Y	Y	EFA,Cor,GFI,Alpha	No
26	Mirani and Lederer	1998	DSS	Y		Y	Y	Y	No	-	Y	Y	-	-	GFI,CFA,Cor	No
27	Gelderman	1998	I&M	Y	Y		Y	Y	No	-	Y	Y	-	-	Alpha,Cor	No
28	Yuthas and Young	1998	I&M	Y			Y	Y	No	-	-	-	-	Y	Cor,T-tests	No
29	Santhanam et al.	2000	DSS	Y		Y		Y	No	-	Y	Y	-	-	EFA,Cor	No
30	Liu & Arnett	2000	I&M	Y			Y	Y	Y	-	-	-	Y	-	EFA,Alpha	No
31	Wixom & Watson	2001	MISQ	Y			Y		Y	Y	Y	Y	-	-	AVE	Y
32	Chae & Kim	2001	ICIS	Y				Y	No	-	-	Y	-	Y	SEM (EFA,GFI,AGFI,NFI)	Y
33	McKinney et al.	2002	ISR	Y			Y	Y	Y	-	-	-	-	Y	EFA,Alpha,DC	No
34	Rai et al.	2002	ISR	Y				Y	Y	-	-	Y	-	-	EFA,Alpha,GFI	Y
35	Gable et al.	2003	ICIS	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	EFA,Alpha	No
36	Sedera & Gable	2004	ICIS	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	CFA,Alpha	Y
37	McGill & Klobas	2005	DSS	Y			Y		No	-	-	Y	Y	-	Alpha,GFI	Y
38	Wixom & Todd	2005a	ISR	Y			Y	Y	Y	-	Y	Y	-	-	EFA,AVE,Cor,PLS	Y
39	Bradley et al.	2006	JMIS	Y				Y	Y	Y	-	Y	Y	-	Alpha,AVE	Y
40	Nicolaou & McKnight	2006	ISR	Y				Y	Y	Y	Y	Y	Y	-	Alpha,AVE,EFA,PLS	Y
41	Kositanurit et al.	2006	EJIS	Y			Y		Y	-	-	Y	Y	-	EFA,RA,	No
42	Sabherwal et al.	2006	MS	Y			Y		Y	-	Y	Y	Y	-	SEM	Y
43	Hartono et al.	2007	DSS	Y			Y	Y	Y	-	Y	Y	-	-	Pearson's R	No
				43					33	13	30	31	12	8		16
				100%					77%	30%	70%	72%	28%	19%		37%

Tests Conducted Legend for abbreviated terms
EFA = Exploratory Factor Analysis; CFA = Confirmatory Factor Analysis; Alpha = Cronbach Alpha; Regre = Regression; Cor = Correlation; DC = Discriminant and Convergent Validity; VA = Variance tests; R2 = r-squared

Finally, based on 'tests conducted', it would seem that all 43 studies validated their construct(s) as reflective. There exist formative construct validation techniques aimed at assisting with pruning of items (to achieve parsimony and avoid multicollinearity) e.g. Tolerance, Correlation with criterion measures, MIMIC … but there are those who argue against pruning. Certain techniques can give insights into completeness (e.g. identification through structural relations). These techniques may too give some insight into which hierarchical structure yields best SEM model fit. Having said this, rigorous conceptual and theoretical arguments are the best approach, and qualitative techniques must be used to evaluate utility and intuitiveness. None of these techniques are employed in any of the 43 studies.

Conclusion

Application of the CORV framework to past empirical studies of the four IS-Impact dimensions has revealed: major concerns where there is a mismatch of Respond and Validate (Stages 3 and 4) and a general dearth of attention to Stages 2 and 3 in methodological writings. It has enabled clearer differentiation between various scenarios, suggesting that the classic case reported in recent literature on misconceiving formative as reflective,

[8] All of the 'tests conducted' were covariance-based and none were formative construct validity tests

where the items have been conceived, presented and answered formatively, and then inappropriately validated as reflective (FFFR) is less commonplace.

Table 3: The 43 Study Items Mapped to the 37-Item Pool

	System Quality																Information Quality											Organizational									Individual					
No*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	T	1	2	3	4	5	6	7	8	9	10	T	1	2	3	4	5	6	7	8	T	1	2	3	4	T	GT
1																0					1						1								0				1	1	2	3
2						1										1					1		1		1		3								0						0	4
3																0											0								0				1	1	2	2
4				1												1					1		1				2								0						0	3
5				1		1					1	1				4				1	1	1	1	1	1		6								0						0	10
6						1	1			1		1				4					1		1		1		3								0						0	7
7						1						1				2					1		1		1		3								0						0	5
8																0					1		1				2								0						0	2
9						1						1				2					1		1		1		3								0						0	5
10																0					1						1								0				1	1	2	3
11				1												1						1	1		1		3								0						0	4
12							1									1			1	1	1						3								0						0	4
13												1				1											0								0				1		1	2
14				1		1						1				3		1			1		1	1	1		5								0				1	1	2	10
15																0											0	1		1	1		1		4						0	4
16																0											0					1			1						0	1
17				1	1	1								1		4					1		1				2			1					1				1	1	2	9
18				1								1				2					1	1	1		1		4								0						0	6
19																0											0					1			1						0	1
20				1												1					1		1		1		3								0	1				1	2	6
21						1				1		1		1		4					1	1	1		1		4								0						0	8
22						1				1		1				3					1		1		1		3								0						0	6
23				1							1					2					1	1	1	1	1		5								0						0	7
24				1		1				1	1					4		1			1	1	1		1		5								0						0	9
25				1					1							2					1	1	1		1		4								0						0	6
26						1				1		1				3						1	1	1			3				1					1						0	7
27			1													1						1	1		1		3								0					1	1	5
28				1		1					1	1				4				1	1	1	1	1	1		6								0						0	10
29																0			1	1							2			1	1	1			3						0	5
30	1			1		1								1		4					1		1		1		3								0						0	7
31										1				1	1	3											0								0						0	3
32																0						1	1		1		3								0						0	3
33				1		1		1					1			4				1	1		1				3								0						0	7
34																0					1	1	1	1			4								0						0	4
35	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15	1	1	1	1	1	1	1	1	1	1	10	1	1	1	1	1	1	1	1	8	1	1	1	1	4	37
36				1	1		1	1	1	1			1	1	1	9		1	1	1	1	1			1		6	1	1	1	1	1	1	1	1	8	1	1	1	1	4	27
37				1	1					1						3											0								0						0	3
38	1	1		1		1				1	1			1		7		1				1	1		1		4								0						0	11
39																0		1			1		1		1		4								0						0	4
40																0					1		1		1		3								0						0	3
41	1	1	1	1			1			1						6											0								0						0	6
42				1		1					1					3											0								0						0	3
43								1								1	1										1								0						0	2
	4	3	3	19	4	17	5	4	3	11	7	12	3	7	3	105	2	6	4	7	27	15	28	8	22	1	120	3	2	5	5	5	3	2	2	27	3	2	8	9	22	274

* see Appendix B Gable et al. (2008:405) for a full list of the 37 items

It is rather suggested that the more commonplace scenario is where items, though conceived formatively, have been presented and answered and validated reflectively (FRRR), the consequences of which though unclear and beyond the scope of analysis presented herein, would appear to be relatively less concerning. IS-Impact was adopted as the referent theory, it being argued that its pool of 37 items represents an ostensibly complete set of mutually exclusive items against which to map items employed in past research. It is observed that study findings further evidence the validity of IS-Impact; namely the completeness and mutual exclusivity of its component items; thereby too further endorsing its adoption as the referent theory in this study. More specifically, from Table 3 we observe evidence of IS-Impact: a) Completeness – the 37 IS-Impact items accommodate all 274 items employed in the sample 43 studies; and b) Mutual Exclusivity – only 1 item from each of the 43 studies mapped into any one of the 37 items in the pool. The study has many limitations. Restricted time, resources and access have not allowed close attention to many potentially revealing specifics in relation to the 43 studies analyzed. Further efforts in this direction are warranted and ongoing. Retrospective or contemporary case studies applying the CORV framework, may suggest useful prescriptions for researchers, particularly with regard to Stage 2 - Operationalization of formative constructs. If Stages 1 and 2 are done well, Stage 3 – Respond, is likely to elicit the appropriate response. It is further noted that the 43 studies analyzed account for only 28% (43/155) of relevant studies identified from the literature (due to inadequate detail in the other 112), thus the representativeness of the 43 is unclear. We conclude with several questions for the reader and for further research. Might the same respondent answer some items formatively and others reflectively on the same construct? If so, what are the implications? Might the same items be answered reflectively by some respondents and formatively by others? If so, what are the implications? Are there different severities of CMV? Might 'reflectiveness' be construed as CMV? How serious or not is this? What are the implications? What research and analyzes might be conducted in attention to these questions?

Acknowledgements: The authors acknowledge the extensive and conscientious efforts of our colleagues, Mr. Felix Ter Chian Tan and Ms. Nor Hidayati Zakariya for their assistance in data collection.

References

Bagozzi, R., and Fornell, C. "Theoretical Concepts Measure-ments, and Meaning," in: *A Second Generation of Multi-variate Analysis,* C. Fornell (ed.), New York 1982a, pp. 24-38.

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

Bajwa, D.S., Rai, A., and Brennan, I. "Key antecedents of Executive Information System success: a path analytic approach," *Decision Support Systems* (22:1) 1997, pp 31-43.

Barki, H., and Huff, S.L. "Change, attitude to change, and decision support system success," *Information & Management* (9:5) 1985, pp 261-268.

Bollen, K.A. "Multiple Indicators: Internal Consistency or No Necessary Relationship," *Quality and Quantity* (8) 1984, pp 377-385.

Bollen, K.A. *Structural Equations With Latent Variables* Wiley, New York, 1989, pp. xiv, 514.

Bollen, K.A., and Lennox, R. "Conventional Wisdom on Measurement: A Structural Equation Perspective," *Psychological Bulletin* (110) 1991, pp 305-314.

Bradley, R.V., Pridmore, J.L., and Byrd, T.A. "Information Systems Success in the Context of Different Corporate Cultural Types: An Empirical Investigation," *Journal of Management Informaton Systems* (23:2), Fall 2006, pp 267-294.

Brown, R.M., Gatian, A.W., and Hicks, J.O., Jr. "Strategic information systems and financial performance," *Journal of Management Information Systems* (11:4) 1995, p 215.

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

Chase, M., and Kim, J. "Information Quality for Mobile Internet Services: A Theoretical Model with Empirical Validation," Twenty-Second International Conference on Information Systems, 2001, pp. 43-53.

DeLone, W.H., and McLean, E.R. "Information Systems Success: The Quest For The Dependent Variable," *Information Systems Research* (3:1) 1992, pp 60-95.

DeLone, W.H., and McLean, E.R. "The Delone And Mclean Model Of Information Systems Success: A Ten Years Update," *Journal of Management Information Systems* (19:4) 2003, pp 9-30.

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

Doll, W.J., Xia, W., and Torkzadeh, G. "A confirmatory factor analysis of the end-user computing sat," *MIS Quarterly* (18:4) 1994, p 453.

Edwards, J.R., and Bagozzi, R.P. "On the nature and direction of relationships between constructs and their measures," *Psychol Methods* (5) 2000, pp 155–174.

Essex, P., A., Magal, S., R. , and Masteller, D., E. "Determinants of information center success," *Journal of Management Information Systems* (15:2) 1998, p 95.

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

Gable, G., Sedera, D., and Chan, T. "Enterprise Systems Success: A Measurement Model," Proceedings of the 24th International Conference on Information Systems, Association for Information Systems, Seattle, Washington, 2003, pp. 576-591.

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

Gaski, J.F., and Nevin, J.R. "The differential effects of exercised and unexercised power sources in a marketing channel," *Journal of Marketing Research* (22) 1985, pp 130-142.

Gatian, A.W. "Is user satisfaction a valid measure of system effectiveness?," *Information & Management* (26:3) 1994, pp 119-131.

Gelderman, M. "The relation between user satisfaction, usage of information systems and performance," *Information & Management* (34:1) 1998, pp 11-18.

Goodhue, D., L. "Development and measurement validity of a task-technology fit instrument for user evaluations of information systems," *Decision Sciences* (29:1) 1998, p 105.

Guimaraes, T., and Igbaria, M. "Client/server system success: Exploring the human side," *Decision Sciences* (28:4) 1997, p 851.

Guimaraes, T., Igbaria, M., and Lu, M.-T. "The Determinants of DSS Success: An Integrated Model," *Decision Sciences* (23:2) 1992, p 409.

Hartono, E., Santhanam, R., and Holsapple, C.W. "Fators That Contribute To Management Support System Success: An Analysis of Field Studies," *Decision Support Systems* (43) 2007, pp 256-268.

Hauser, R.M., and Goldberger, A.S. "The treatment of unobservable variables in path analysis," in: *Sociological Methodology,* H.L. Costner (ed.), Jossey-Bass, San Francisco, 1971, pp. 81-117.

Howell, R.D., Breivek, E., and Wilcox, J.B. "Reconsidering formative measurement," *Psychol Methods* (12) 2007, pp 205-218.

Jarvis, C.B., MacKenzie, S.B., and Podsakoff, P.A. "A critical review of construct indicators and measurement model misspecification in marketing and consumer research," *journal of consumer research* (30) 2003, pp 199-216.

Kositanurit, B., Ngwenyama, O., and Osei-Bryson, K.-M. "An Exploration of Factors That Impact Individual Performance in an ERP Environment: An Analysis Using Multiple Analytical Techniques," *European Journal of Information Systems* (15) 2006, pp 556-568.

Law, D., and Gorla, N. "Exploring factors underlying effective office information systems," *Information & Management* (31:1) 1996, pp 25-35.

Li, E.Y. "Perceived importance of information system success factors: A meta analysis of group differences," *Information & Management* (32:1) 1997, pp 15-28.

Liu, C., and Arnett, K.P. "Exploring the Factors Associated With Web Site Success in the Context of Electronic Commerce," *Information & Management* (38) 2000, pp 23-33.

Long, J.S. *Confirmatory Factor Analysis: A Preface To LISREL* Sage, Beverly Hills, CA, 1983.

Lu, H.-P., and Wang, J.-Y. "The relationships between management styles, user participation, and system success over MIS growth stages," *Information & Management* (32:4) 1997, pp 203-213.

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

Mahmood, M.A., and Medewitz, J.N. "Impact Of Design Methods On Decision Support System Success: An Empirical Assessment," *Information & Management* (9:3) 1985, pp 137-151.

McGill, T.J., and Klobas, J.E. "The Role of Spreadsheet Knowledge in User-developed Application Success," *Decision Support Systems* (39) 2005, pp 355-369.

McHaney, R., and Cronan, T.P. "Computer simulation success: On the use of the end-user computing satisfaction instrument: A comment," *Decision Sciences* (29:2) 1998, p 525.

McKinney, V., Yoon, K., and Zahedi, F.M. "The Measurement of Web-Customer Satisfaction: An Expectation and Disconfirmation Approach," *Information Systems Research* (13:3), September 2002, pp 296-315.

Miller, J., and Doyle, B.A. "Measuring the Effectiveness of Computer-Based Information System in the Financial Services Sector," *MIS Quarterly* (11:1) 1987, p 107.

Mirani, R., and Lederer, A., L. "An instrument for assessing the organizational benefits of IS projects," *Decision Sciences* (29:4) 1998, p 803.

Nicolaou, A.I. "Perceived Information Quality in Data Exchanges: Effects on Risk, Trust, and Intention to Use," *Information Systems Research* (17:4), December 2006, pp 332-351.

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

Rai, A., Lang, S.S., and Welker, R.B. "Assessing the Validity of IS Success Models:
An Empirical Test and Theoretical Analysis," *Information Systems Research,* (13:1), March 2002, 2002, pp 50-69.

Rainer, J.K.R., and Watson, H.J. "The Keys to Executive Information System Success," *Journal of Management Information Systems* (12:2) 1995, pp 83-99.

Raymond, L. "Organizational Characteristics And MIS Success In The Context Of Small Business," *MIS Quarterly* (9:1) 1985, pp 37-52.

Raymond, L. "Validating and applying user satisfaction as a measure of MIS success in small organizations," *Information & Management* (12:4) 1987, pp 173-179.

Raymond, L. "Organizational Context and Information Systems Success: A Contingency Approach," *Journal of Management Information Systems* (6:4), Spring 1990, pp 5-20.

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

Sabherwal, R., Jeyaraj, A., and Chowa, C. "Information System Success: Individual and Organizational Determinants," *Management Science* (52:12) 2006, pp 1849-1864.

Sanders, G.L., and Courtney, J.F. "A Field Study of Organizational Factors Influencing DSS Success," *MIS Quarterly* (9:1) 1985, p 77.

Santhanam, R., Guimaraes, T., and George, J.F. "An Empirical Investigation of ODSS Impact on Individuals And Organizations," *Decision Support Systems* (30) 2000, pp 51-72.

Sedera, D., and Gable, G. "A Factor and Structural Equation Analysis of the Enterprise Systems Success Measuement Model," Twenty-Fifth International Conference on Information Systems, 2004, pp. 449-463.

Sethi, V., and King, W.R. "Development Of Measures To Assess The Extent To Which An Information Technology Application Provides Competitive Advantage," *Management Science* (40:12), December 1994, pp 1601 - 1627.

Sirinivasan, A. "Alternative Measures Of System Effectiveness: Associations And Implications," *MIS Quarterly* (9:3) 1985, pp 243-253.

Stylianou, A.C., Jeffries, C.J., and Robbins, S.S. "Corporate mergers and the problems of IS integration," *Information & Management* (31:4) 1996, pp 203-213.

Tait, P., and Vessey, I. "The Effect of User Involvement on System Success: A Conting," *MIS Quarterly* (12:1) 1988, p 91.

Wilcox, J.B., Howell, R.D., and Breivik, E. "Questions about formative measurement," *Journal of Business Research*) 2008.

Williams, L., Edwards, J.R., and Vandenberg, R.J. "Recent advances in causal modeling methods for organizational and management research," *Journal of Management Information Systems* (29:6) 2003, pp 903-936.

Wixom, B.H., and Todd, P.A. "A theoretical integration of user satisfaction and technology acceptance.," *Information Systems Research* (16:1) 2005, pp 85-102.

Wixom, B.H., and Watson, H.J. "An Empirical Investigation of the Factors Affecting Data Warehousing Success," *MIS Quarterly* (25:1), March 2001, pp 17-41.

Yuthas, K., and Young, S.T. "Material matters: Assessing the effectiveness of materials management IS," *Information & Management* (33:3) 1998, pp 115-170.