Enterprise Systems Success: A Measurement Model

Guy Gable  
*Queensland University of Technology*

Darshana Sedera  
*Queensland University of Technology*

Taizan Chan  
*Queensland University of Technology*

Follow this and additional works at: [http://aisel.aisnet.org/icis2003](http://aisel.aisnet.org/icis2003)

Recommended Citation


This material is brought to you by the International Conference on Information Systems (ICIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICIS 2003 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
ENTERPRISE SYSTEMS SUCCESS:
A MEASUREMENT MODEL

Guy G. Gable, Darshana Sedera, and Taizan Chan
Center for Information Technology Innovation
Queensland University of Technology
Brisbane, Australia
g.gable@qut.edu.au  d.sedera@qut.edu.au  t.chan@qut.edu.au

Abstract

This paper presents a validated measurement model and instrument for assessing enterprise systems success from multiple perspectives. The final validated study model employs 27 measures of the four dimensions: information quality, system quality, individual impact, and organizational impact. The model is empirically tested with survey data gathered from 27 public sector organizations that implemented SAP R/3 in the late 1990s. The study consists of an exploratory inventory survey (model building) to identify the salient success dimensions and measures, followed by a confirmatory weights survey, for testing model validity (model testing). Test results demonstrate the discriminant validity of the four dimensions, as well as their convergence on a single higher-order phenomenon: enterprise systems success (ESS). Criterion validity testing further demonstrates the additivity of the four dimensions of success, and the completeness of the resultant overarching, second-order measure of ESS.

Introduction

Information systems investments are under increasing scrutiny and pressure to justify their value and contribution to the productivity, quality, and competitiveness of organizations. The importance of assessing the value of IS, underpins key issues reported by organizational executives around the world (Ball and Harris 1982; Brancheau and Wetherbe 1987; Dickson et al. 1984). Evidence of IS success has been mixed with some studies showing positive impacts of IS in organizations (e.g., Barua et al. 1991; Barua and Lee 1997; Brynjolfsson and Hitt 1996; Lehr and Lichtenberg 1999; Mukherjee 2001), while others have shown nil or detrimental impacts (e.g., Attewell and Rule 1984; Brynjolfsson and Yang 1996; Quinn and Cameron 1988; Wilson 1993). These conflicting results may be attributable to (1) incomplete or inappropriate measures of success, (2) lack of theoretical grounding of causal and process models of IS success, (3) myopic focus on financial performance indicators, or weaknesses in (4) survey instruments employed (e.g., constructs lacking in validity) or (5) data collection approach (e.g., asking the wrong people).

Despite the substantial investments made by organizations around the world in enterprise systems (ES),1 systematic attempts to measure their success have been few (e.g., Baer 1999; Davis 1989; Deloitte Consulting 2000; Knowles et al. 2000; Sedera et al. 2001; Shang and Seddon 2000). The impacts resulting from ES are arguably difficult to measure (Baer 1999; Davis 1989; Deloitte Consulting 2000; Knowles et al. 2000; Sedera et al. 2001). An enterprise system entails many users ranging from top executives to data entry operators; many applications that span the organization; and a diversity of capabilities and functionality. These contemporary IS characteristics (along with other issues discussed in the literature review section) suggest that existing models of IS success (that were developed for a more traditional IS context) may not be entirely appropriate for measuring ES success. This paper presents and validates a measurement model for assessing enterprise system success.

---

1In this paper, the terms ERP, enterprise resource planning, and the more contemporary enterprise systems (ES) are used interchangeably. For an in-depth discussion on ERP, see Klaus et al. (2000).
process nature of six IS success constructs: (1) system quality, (2) information quality, (3) use, (4) user satisfaction, (5) individual impact, and (6) organizational impact. While it is unclear whether the process paths proposed by Delone and McLean were originally intended to suggest causality, many researchers have sought to test these as causal paths and have found them to be broadly valid (e.g., Rai et al. 2002; Seddon and Kiew 1994). According to Seddon and Kiew and to Myers et al., the main contributions Delone and McLean make to our understanding of IS evaluation are: (1) the constructs of the model provide a broadly valid (e.g., Rai et al. 2002; Seddon and Kiew 1994). According to Seddon and Kiew and to Myers et al., the main contributions Delone and McLean make to our understanding of IS evaluation are: (1) the constructs of the model provide a
classifies potential ERP benefits into 21 lower level measures organized around five main categories: operational benefits, managerial benefits, strategic benefits, IT infrastructure benefits, and organizational benefits. Their framework has yet to be operationalized.2

**Gaps in Existing IS Success Studies**

The development of IS success models, such as the Delone and McLean model, has been an important contribution toward our improved understanding of IS management. However, several issues in IS success models remain.

**Mutual exclusivity and additivity of success measures:** While some feel the various success categories studied (e.g., information quality vs. system quality) offer surrogate, or perhaps alternative, measures of success (Bailey and Pearson 1983; Doll and Torkzadeh 1988; Ives et al. 1983; Saarinen 1996), other researchers have suggested they represent distinct dimensions of a complex, higher-order phenomenon (Chandler 1982; Ein-Dor and Segev 1978; King and Rodriguez 1978). An analogous example of the latter view, to which we subscribe, is Gable’s (1996) study of 150 computer system selection projects involving external consultants, wherein he tested a multidimensional model of consultant engagement success. Gable argues that his dimensions can be usefully combined to yield an overarching measure of success. An important criteria that our proposed model therefore aims to satisfy is to insure that each measure in our model not only addresses an important aspect of IS success, but also does so in such a manner that it does not overlap with another measure.

**Model completeness:** Table 1, reflects 45 success measurement studies since 1992 (refer to Gable [1996] for a similar comparison of studies on the Delone and McLean model up to 1992). It shows that only two studies have considered all six of the Delone and McLean dimensions of success (none prior to 1992). The employment of only one or a subset of the dimensions of success as a surrogate for overall success may be one of the reasons for mixed results reported in the literature regarding the antecedents of success (e.g., Barki and Hartwick 1989; Gatian 1994; Ginzberg 1981; Hawk and Aldag 1990; Ives and Olson 1984, Myers et al. 1997).

Gable (1996, p. 1177) notes that “the completeness of the model becomes critical as adding good and bad, high and low, positive and negative, or hot and cold effects may otherwise mask, neutralize, or distort results.”

**Choice of IS success dimensions:** Delone and McLean (1992) suggest that in order to develop a comprehensive measurement model/instrument for a particular context, the constructs and measures should be systematically selected considering contingency variables, such as the organizational structure, size, or technology, and the individual characteristics of the system. Yet, most studies in this arena do not elaborate on the rationale for their choice of success dimensions and success measures employed. Further, as was stated in the previous section, in order to fully account for potentially countervailing measures and dimensions of success (e.g., high quality but poor cost-effectiveness), model completeness becomes critical. Through a review of alternative models from the literature, Melone (1990) highlights the subjectivity inherent in the selection of a single effectiveness measure. This suggests that where the aim is to gain a full, overarching view of success, it is critical that the complete set of success dimensions be employed, not a selected subset.

**Theoretical basis for causal/process paths:** Although the Delone and McLean model has been a valuable contribution to our improved understanding of IS success, their taxonomy is presented without sufficient explanation of its underlying theory and epistemology, with many questioning the suggested causal/process nature of the model (e.g., Ballantine et al. 1996; Myers et al. 1997). Seddon and Kiew (1994) were the first to empirically test the causal structure (part of the structure). Their investigation supported some of the model paths but not others. Other researchers have since tested other causal relationships between the six

---

2Other frameworks that were considered but found to be less suitable include MIT’90s IT impacts framework (Allen and Scott Morton 1992; Scott Morton 1990) and balanced scorecard (Kaplan and Norton 1992).
variables of the Delone and McLean model yielding mixed results (Bonner 1995; Hunton and Flower 1997). While Rai et al. (2002) attempted to provide a theoretical underpinning for the causal model, they managed to do so only for the paths leading to the USE construct. This lack of theoretical grounding, combined with a weak explanation for causality and mixed results from empirical studies, raises concerns about the causality of the Delone and McLean model and the utility of the suggested relationships.

**Excessive emphasis on quantitative (financial) measures:** Although IS investments are in many ways comparable to traditional investments such as production equipment, they entail a strong organizational element as well. It is a common tendency to measure ES (and IS generally) only in terms of financial criteria. However, it is widely acknowledged that ES result in considerable intangible impacts in addition to more tangible impacts. Thus, use of traditional financial measures alone may not account for evidence of IS payoffs (Ballantine et al. 1996; Brynjolfsson 1993; Kaplan and Norton 1992). Also, economic evaluations and quantitative measures tend to be difficult to obtain and easy to manipulate.

**The nature of the contemporary IS environment:** The transition from indirect batch-oriented use of IS to more direct, on-line and integrated IS has changed the way organizations produce and manage information. The modern IS trend is toward changed organizational structures and behaviors that facilitate interorganizational activities. New measures and evaluation models are required to measure success with contemporary IS (Ishman 1996; Sederer et al. 2003c). Yet, most IS success studies continue to rely on instruments and measures that were validated with what are now outdated information systems (Jurison 1996; Saarinen 1996).

**Multiple stakeholder perspectives:** The respondents’ perspective on measurement is another important design consideration in IS evaluation. The importance of analyzing IS success at multiple levels within organizations has been discussed among academics for several decades (e.g., Cameron and Whetten 1983; Leider and Elam 1994; Quinn and Rohrbaugh 1983; Sederer et al. 2003a; Tallon et al. 2000; Thong and Yap 1996; Yoon and Guimaraes 1995). Yet, IS studies have, in the main, attempted to quantify the impacts (benefits and drawbacks) of IS by analyzing data collected mostly at very senior levels of the firm only.

**Model Building**

**The Inventory Survey**

Having identified possibly relevant models and frameworks from our review of the literature, the main purpose of the initial exploratory inventory survey was to identify a salient set of ES success dimensions to include in the *a priori* ES success measurement model. In September 2001, the exploratory survey was conducted to inventory impacts of the SAP R/3 system, as perceived by staff at all levels of 27 government agencies in the state of Queensland, Australia. The inventory survey was non-anonymous with three main instrument sections querying (1) respondent demographics (e.g., name, position, number years with agency, years with Queensland Government) and a brief description of their involvement with the SAP R/3 system (note that based on these and other demographics, respondents were carefully classified into three cohorts: management, users, and technical); (2) specific impacts of the SAP system; and (3) any past, in-progress, or pending initiatives for increasing positive impacts from the SAP system, as well as suggestions for further possible improvements.3

The specific question posed in section (2) was “What do you consider have been the impacts4 of SAP in your agency since its implementation?” With many, if not most, of these large system implementations, emotions run high, with much proselytizing for the pro- or con-camps. Although some time had passed since the initial implementation and the situation had become more settled, we were loath to ask any questions that probe levels of success directly. We positioned the study as one of impacts rather than levels of success, feeling this more likely to elicit objective responses rather than emotional party-line responses. A total of 137 responses were received, citing a total of 485 impacts (i.e., averaging 3.6 per respondent).

---

3For the purposes of this paper, evidence from (3) served solely to aid interpretation of evidence from (2).

4It should be highlighted that the word *impacts* in the exploratory survey round was used in the broadest sense, to encompass impacts on individuals, the organization, information, the system, etc.
Reasons for dropping the Shang and Seddon (2000) framework include overlaps between the constructs and measures; its strong emphasis on top managerial perspective (not a holistic view); and its somewhat narrow emphasis on organizational performance.

Seddon (1997) argues that the true underlying construct IS researchers have been trying to gauge is usefulness, not usage or use. The expert workshops and review of related literature, suggested that the usefulness of a system derives from such factors as the quality of the system and the information it produces. When attempting to adapt Seddon and Kiew’s measures of usefulness to the study purpose, it was noted that most of these measures had already been addressed by measures previously adapted from the Meyers et al. framework for the system quality and

Mapping

Synthesizing this morass of qualitative evidence into a useful, meaningful, and coherent classification of success dimensions and measures was a critical and complex stage of the study. The objectives of this exercise were to yield a framework that is (1) simple and generalizable beyond the current study, while also being (2) intuitive to the study respondents. These two aims at times were diametrically opposed.

The literature suggests two main approaches for data coding and synthesis: (1) a bottom-up, data driven, open coding approach or (2) a top-down, structured coding, framework approach (see details in Chan et al. 2000). The top-down approach employs deduction, and starts with a logical framework or model to categorize the responses. The bottom-up approach employs induction, starting with the data in hand, that is arranged into a logical classification. Given the relative advantages and disadvantages of these approaches (for a detailed discussion, see Chan et al. 2002), it was decided that the top-down approach first be attempted, and that a bottom-up approach only be adopted given poor fit of the data with candidate frameworks (Sedera et al. 2002).

Selecting a Framework

An attempt was made to map the first-round survey impacts into both the Delone and McLean (1992) IS success model supplemented with the Myers et al. (1997) IS assessment selection model, as well as into the Shang and Seddon (2000) ERP benefits framework. In order to minimize individual errors of judgment, three academics and two senior business analysts from surveyed organizations participated in the mapping exercise. Participants mapped a sample of citations from the exploratory round into the two selected frameworks and compared results. Any discrepancies were discussed until consensus on the mappings was reached.

The synthesis process identified the constructs and underlying measures of the Delone and McLean model and the associated measures from Myers et al. as the most suitable taxonomy of ES success. Having omitted use (discussed in the following section), the remaining five constructs of the Delone and McLean model were all well instantiated during the round 1 mapping exercise, and accommodated all impacts cited by respondents. On this basis, the Delone and McLean constructs, and measures, served as the basis of the starting ES success model. The constructs/measures of the Delone and McLean model provided a holistic view across the organization—from a top management perspective, to that of data entry officers—and provided a detailed categorization of success dimensions. Furthermore, the mutual exclusivity of the measures allowed a more logical synthesis of the inventory-round survey results than did the Shang and Seddon ERP benefits framework.

Adapting and Operationalizing the Selected Framework

Operationalization of the beginning model proceeded as follows: (1) revisit relevance of the Delone and McLean use construct; (2) revisit relevance of user satisfaction; (3) define a more expansive organizational impacts construct; (4) introduce further ES-related measures, and (5) remove measures that are inappropriate for this study.

Revisit use construct: A main criticism of the Delone and McLean model has centered on the use construct, which many feel to be an inappropriate measure of IS success (e.g., Barki and Huff 1985; Gelderman 1998; Seddon 1997; Young 1989; Yuthas and Young 1998). Delone and McLean (1992, p. 68) themselves suggest that “usage, either perceived or actual, is only pertinent when such use is not mandatory.” When use of a system is mandatory, the extent of use of a system conveys little information about the success of the system (Robey 1979; Welke and Konsynski 1980). As the ES under investigation is mandatory, the construct use was omitted from the a priori model. In further support of this action, it is noted that only 12 of the 485 citations (2 percent) of inventory survey round mapped into the use construct and its measures.

[5] Reasons for dropping the Shang and Seddon (2000) framework include overlaps between the constructs and measures; its strong emphasis on top managerial perspective (not a holistic view); and its somewhat narrow emphasis on organizational performance.

[6] Seddon (1997) argues that the true underlying construct IS researchers have been trying to gauge is usefulness, not usage or use. The expert workshops and review of related literature, suggested that the usefulness of a system derives from such factors as the quality of the system and the information it produces. When attempting to adapt Seddon and Kiew’s measures of usefulness to the study purpose, it was noted that most of these measures had already been addressed by measures previously adapted from the Meyers et al. framework for the system quality and
Revisit user satisfaction: User satisfaction is possibly the most extensively used single measure for IS evaluation (e.g., Delone and McLean 1992; Doll and Torkzadeh 1988; Etezadi-Amili and Farhoomand 1996; Gatian 1994; Igbaria and Nachman 1990; Igbaria and Tan 1997; Lucas 1975). There exist several widely cited studies and standard instruments that measure satisfaction (e.g., Bailey and Pearson 1983; Baroudi and Orlikowski 1988; Doll and Torkzadeh 1988). Early satisfaction constructs (e.g., user information satisfaction—Bailey and Pearson 1983) have been found to mix measures of multiple dimensions of success (e.g., quality and impact) rather than satisfaction. Not surprisingly, many survey items used in prior satisfaction-only studies were analyzed during operationalization of this study model and found to readily map into other constructs (e.g., system quality, information quality7). Having developed a comprehensive model and instrument for measuring success, our expectation ultimately was that pure satisfaction items alone do not reflect a separate dimension of success, but rather measures of overall success. This view is consistent with the findings of Teo and Wong (1998), who, having studied the impact of IT investment and performance impact measures, concluded that satisfaction was not a distinct dimension. Although we clearly had concerns with the validity of the satisfaction construct as a dimension of success and with its inclusion in the a priori model, in order to test its discriminant validity, it remained.

A more holistic organizational impacts construct: The mapping exercise revealed that Delone and McLean’s measures of organizational impacts, which are primarily financial, did not adequately account for potential organizational impacts of an enterprise system (Ballantine et al. 1996). While economic success is crucial for survival, an organization may be considered successful in many other ways (Ballantine et al. 1996; Kaplan and Norton 1992). It was, therefore, decided that a more holistic organizational impacts construct should include measures from the Shang and Seddon (2000) ERP benefits framework, as well as from the balanced scorecard (Kaplan and Norton 1992). The former account for ES-related measures, such as business process improvement, increased capacity, and e-government readiness,8 while the latter provide a balanced view of IS success and include measures such as staff requirements reduction and cost reduction (Sedera et al. 2001).

Other Enterprise System-related measures: The mapping exercise also facilitated identification and inclusion in the model of other, new measures related to ES. For example, citations from the exploratory survey revealed that an important measure of system quality for ES is the extent to which functionality and data have been integrated. In addition, the amenability of the ES to customization was also identified as an important aspect of the quality of the enterprise system. As a result, four measures, customization (system quality), increased capacity, e-government, and business process change (organizational impact), have been added.10

Removal of inappropriate measures: Several measures were considered unsuitable and were omitted from the a priori model because (1) they appeared redundant, overlapping with other similar measure(s), or (2) they were not perceptual and thus incongruent with the instrument design which sought a holistic view across the organization. Measures added or removed and the rationale for these changes are summarized in Figure 2.11
The A Priori Model

Having started with the Delone and McLean constructs and measures (supplemented by Myers et al.), and having adapted their framework through review of the literature, the round 1 survey, and a series of expert workshops, we propose the following a priori model of ES success. Figure 2 shows how the a priori model deviates from Delone and McLean and Myers et al. with omitted measures italicized and new measures bolded. Further note, as indicated previously, the satisfaction construct was included in the a priori model with some trepidation.

Unlike the original Delone and McLean model, the a priori model (Figure 2) is simply a measurement model for assessing the multidimensional phenomenon of ES success using five separate dimensions of success (constructs): system quality, information quality, satisfaction, individual impact, and organizational impact. The model does not purport any causality among the dimensions. Rather, the dimensions are posited to be correlated and additive measures of the same multidimensional phenomenon—ES success.

Figure 2. The A Priori Model a

<table>
<thead>
<tr>
<th>System Quality</th>
<th>Information Quality</th>
<th>Satisfaction</th>
<th>Individual Impact</th>
<th>Organizational Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>SQ1 Data accuracy</td>
<td>IQ1 Importance</td>
<td>S1 Information</td>
<td>II1 Learning</td>
<td>OI1 Organizational costs</td>
</tr>
<tr>
<td>SQ2 Data currency</td>
<td>IQ2 Availability</td>
<td>S2 Systems</td>
<td>II2 Awareness / Recall</td>
<td>OI2 Staff requirements</td>
</tr>
<tr>
<td>SQ3 Database contents</td>
<td>IQ3 Usability</td>
<td>S3 Overall</td>
<td>II3 Decision effectiveness</td>
<td>OI3 Cost reduction</td>
</tr>
<tr>
<td>SQ4 Ease of use</td>
<td>IQ4 Understandability</td>
<td>S4 Enjoyment</td>
<td>II4 Individual productivity</td>
<td>OI4 Overall productivity</td>
</tr>
<tr>
<td>SQ5 Ease of learning</td>
<td>IQ5 Relevance</td>
<td>decision making^^ problem identification++ specifications^^ task performance+++</td>
<td>OI5 Improved outcomes/outputs</td>
<td></td>
</tr>
<tr>
<td>SQ6 Access</td>
<td>IQ6 Format</td>
<td>accurate interpretation+</td>
<td>OI6 Increased capacity**</td>
<td></td>
</tr>
<tr>
<td>SQ7 User requirements</td>
<td>IQ7 Content Accuracy</td>
<td>OI7 e-government**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQ8 System features</td>
<td>IQ8 Conciseness</td>
<td>OI8 Business Process Change**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQ9 System accuracy</td>
<td>IQ9 Timeliness</td>
<td>application portfolio^ no. of critical applications^ increased work volume^^</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQ10 Flexibility</td>
<td>IQ10 Uniqueness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQ11 Reliability</td>
<td>usefulness#</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQ12 Efficiency</td>
<td>completeness#</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQ13 Sophistication</td>
<td>informative#</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQ14 Integration</td>
<td>currency#</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQ15 Customization**</td>
<td>reliability### readability### clarity### appearance###</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

** ERP-related measures identified in the exploratory round (or more generally - contemporary IS related)
# Found to overlap and thus collapsed into single measure of Relevance (IQ5) as all are necessary to determine the relevance of Information.
## This measure overlapped with the measure Content Accuracy (IQ7) as only information that is accurate is reliable and vice versa.
### Found to overlap, thus collapsed into single measure of Format (IQ6) as the format of information determines its readability, clarity & appearance.
* Deemed a consequence of information quality as one's ability to make accurate interpretation is affected directly by the quality of the information used.
** Not a general measure applicable only to senior level management and not throughout the organizational levels.
*** Was felt that Task Performance is largely defined by the measures of Decision Effectiveness (II3) and Individual Productivity (II4).
^-perceptual measures that are incongruent with the rest of the measures in the model and the aim to derive a holistic view across the organization.
^ Compound measures that are reflected directly by other measures (Specifications by SQ7, SQ8, Decision Making by S1, S2, and Increased Work Volume by SQ4, SQ5, SQ7).

The 41 items were carefully operationalized to yield an instrument appropriate for all three sample cohorts (management, users, technical).
Model Testing

The Weights Survey

The purpose of the weights survey was to validate the a priori model. A survey instrument was designed to operationalize the 41 measures of the five constructs (Figure 2). Where possible, corresponding question items were drawn from previously validated instruments, and adapted to the current study context and the respondent cohorts. Items were scored on a seven-point Likert scale with the end values (1) strongly disagree and (7) strongly agree, and the middle value (4) neutral. The draft survey instrument was pilot tested with a selected sample of staff of the Queensland Government Treasury Department. Feedback from the pilot round respondents resulted in minor modifications to survey items. Dissemination of the survey instrument was through (1) a Web survey facility and (2) an MS Word instrument attached to e-mail. The same 27 public sector organizations from the exploratory round were again surveyed. In order to gauge ES success from multiple stakeholder perspectives, responses were sought from all employment levels. In all, 310 valid responses were received (nine were not included in the analysis due to missing data or perceived frivolity).

Construct Validity

Following the weights survey, the study model and related instrument items were tested for construct and criterion validity and reliability. The 41 items were included in an exploratory factor analysis. In order to attain a more interpretable and parsimonious factor solution, of the 15 system quality items and 10 information quality items (see Figure 2), six and four items were dropped respectively.

The inclusion of the satisfaction items in the main factor analysis had them loading along with the system quality items. This was not surprising considering that system quality was the first factor extracted, explaining a full 50 percent of the model variance (both before and after ultimately excluding the Satisfaction items).\(^\text{12}\) Given our earlier reservations and speculations regarding the satisfaction construct, it thus appeared logical that the satisfaction items load on that factor which explained most of the factor model variance.

On the basis of our proposition that satisfaction is not a dimension of success, we excluded the satisfaction items from the exploratory factor analysis, which resulted in a clean and logical four factor solution (Table 2), all items loading as anticipated\(^\text{13}\) explaining 67 percent of model variance, with all factors having Cronbach alphas greater than 0.9.\(^\text{14}\)

Table 3 is a summary of numbers of measures considered, dropped and retained across main study phases. Literature Review: Delone and McLean (1992) suggest 112 possible measures of the six constructs in their model. Myers et al (1997) suggest eight further measures, yielding 120 measures. Before Mapping: Critical review of the 120 starting measures, identified 24 measures considered to be redundant, nine non-perceptual measures, and 34 measures ill-suited to the enterprise systems study context. All use measures were removed at this stage for reasons suggested earlier. Mapping: 16 measures were dropped during the mapping exercise and four new measures were added. Model Testing: Factor analysis suggested the further exclusion of 10 measures that did not clearly load on a single factor.

\(^{\text{12}}\)Perhaps suggesting that system quality is the least misleading, single-factor surrogate measure of overall success, although we are not advocating its use for this purpose.

\(^{\text{13}}\)Highly similar results were produced in separate factor analyses for each of the sample cohorts.

\(^{\text{14}}\)Further analysis of the first-order exploratory factor solution was conducted using LISREL 8.53. \(t\) values above 61.96 and large \(R^2\) values for all observed variables suggest acceptable reliability. The 3.5 ratio of \(\chi^2\) to degrees of freedom is within the recommended range. Researchers have recommended using ratio of \(\chi^2\) to the degree of freedom, starting with ratios of as little as 2 to as high as 5 (Marsh and Hocevar 1985).

It must be acknowledged that the satisfaction items in this study may have been somewhat idiosyncratic.
Table 2. The Final Exploratory Factor Solution—Varimax Rotation

<table>
<thead>
<tr>
<th>Items</th>
<th>SQ</th>
<th>OI</th>
<th>IQ</th>
<th>II</th>
</tr>
</thead>
<tbody>
<tr>
<td>SQ4</td>
<td>0.74</td>
<td>0.16</td>
<td>0.28</td>
<td>0.28</td>
</tr>
<tr>
<td>SQ5</td>
<td>0.72</td>
<td>0.14</td>
<td>0.23</td>
<td>0.21</td>
</tr>
<tr>
<td>SQ7</td>
<td>0.52</td>
<td>0.35</td>
<td>0.33</td>
<td>0.31</td>
</tr>
<tr>
<td>SQ8</td>
<td>0.57</td>
<td>0.20</td>
<td>0.37</td>
<td>0.20</td>
</tr>
<tr>
<td>SQ9</td>
<td>0.57</td>
<td>0.16</td>
<td>0.37</td>
<td>0.05</td>
</tr>
<tr>
<td>SQ10</td>
<td>0.68</td>
<td>0.19</td>
<td>0.25</td>
<td>0.14</td>
</tr>
<tr>
<td>SQ13</td>
<td>0.62</td>
<td>0.17</td>
<td>0.32</td>
<td>0.13</td>
</tr>
<tr>
<td>SQ14</td>
<td>0.55</td>
<td>0.20</td>
<td>0.24</td>
<td>0.11</td>
</tr>
<tr>
<td>SQ15</td>
<td>0.66</td>
<td>0.26</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>OI1</td>
<td>0.16</td>
<td>0.73</td>
<td>0.18</td>
<td>0.23</td>
</tr>
<tr>
<td>OI2</td>
<td>0.16</td>
<td>0.82</td>
<td>0.13</td>
<td>0.11</td>
</tr>
<tr>
<td>OI3</td>
<td>0.14</td>
<td>0.83</td>
<td>0.17</td>
<td>0.15</td>
</tr>
<tr>
<td>OI4</td>
<td>0.30</td>
<td>0.68</td>
<td>0.20</td>
<td>0.43</td>
</tr>
<tr>
<td>OI5</td>
<td>0.43</td>
<td>0.62</td>
<td>0.15</td>
<td>0.42</td>
</tr>
<tr>
<td>OI6</td>
<td>0.25</td>
<td>0.54</td>
<td>0.26</td>
<td>0.37</td>
</tr>
<tr>
<td>OI7</td>
<td>0.21</td>
<td>0.62</td>
<td>0.33</td>
<td>0.15</td>
</tr>
<tr>
<td>OI8</td>
<td>0.37</td>
<td>0.59</td>
<td>0.25</td>
<td>0.34</td>
</tr>
<tr>
<td>IQ1</td>
<td>0.29</td>
<td>0.36</td>
<td>0.63</td>
<td>0.21</td>
</tr>
<tr>
<td>IQ2</td>
<td>0.31</td>
<td>0.24</td>
<td>0.63</td>
<td>0.13</td>
</tr>
<tr>
<td>IQ3</td>
<td>0.35</td>
<td>0.24</td>
<td>0.80</td>
<td>0.17</td>
</tr>
<tr>
<td>IQ5</td>
<td>0.35</td>
<td>0.18</td>
<td>0.78</td>
<td>0.25</td>
</tr>
<tr>
<td>IQ6</td>
<td>0.34</td>
<td>0.14</td>
<td>0.80</td>
<td>0.18</td>
</tr>
<tr>
<td>IQ8</td>
<td>0.33</td>
<td>0.23</td>
<td>0.58</td>
<td>0.20</td>
</tr>
<tr>
<td>II1</td>
<td>0.11</td>
<td>0.17</td>
<td>0.18</td>
<td>0.83</td>
</tr>
<tr>
<td>II2</td>
<td>0.24</td>
<td>0.25</td>
<td>0.19</td>
<td>0.82</td>
</tr>
<tr>
<td>II3</td>
<td>0.23</td>
<td>0.30</td>
<td>0.19</td>
<td>0.82</td>
</tr>
<tr>
<td>II4</td>
<td>0.28</td>
<td>0.34</td>
<td>0.21</td>
<td>0.79</td>
</tr>
<tr>
<td>Alpha</td>
<td>0.90</td>
<td>0.92</td>
<td>0.91</td>
<td>0.93</td>
</tr>
</tbody>
</table>

KMO Sampling Adequacy 0.94
Bartlett’s Test of Sphericity

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Start</th>
<th>Drop</th>
<th>Remain</th>
<th>Start</th>
<th>Drop</th>
<th>Remain</th>
<th>Add</th>
<th>Remain</th>
<th>Drop</th>
<th>Remain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systems Quality</td>
<td>18</td>
<td>4</td>
<td>14</td>
<td>14</td>
<td>1</td>
<td>15</td>
<td>15</td>
<td>4</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Information Quality</td>
<td>25</td>
<td>7</td>
<td>18</td>
<td>8</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>4</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Individual Impact</td>
<td>19</td>
<td>12</td>
<td>7</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Organizational Impact</td>
<td>22</td>
<td>14</td>
<td>8</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>7</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Use</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>120</td>
<td>67</td>
<td>53</td>
<td>16</td>
<td>37</td>
<td>4</td>
<td>41</td>
<td>14</td>
<td>27</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Summary of Measures Retained Across Study Phases

<table>
<thead>
<tr>
<th>Research Phase</th>
<th>Literature Review</th>
<th>Model Building</th>
<th>Model Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension</td>
<td>Start</td>
<td>Drop</td>
<td>Remain</td>
</tr>
<tr>
<td>Systems Quality</td>
<td>18</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>Information Quality</td>
<td>25</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>Individual Impact</td>
<td>19</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>Organizational Impact</td>
<td>22</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>7</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Use</td>
<td>29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>120</td>
<td>67</td>
<td>53</td>
</tr>
</tbody>
</table>
The four factors were next included in a second-order factor analysis. This analysis had two main aims: (1) to confirm that all dimensions load on the same, higher-order latent construct (convergent validity), and (2) to test the validity of a single overall measure of success derived from the component dimensions (additivity). As predicted, all dimensions loaded on a single higher-order factor.\(^\text{15}\)

**Criterion Validity**

Besides items referenced thus far (Table 2), the survey instrument elicited criterion measures of overall success in response to each of two statements: (A) “Overall, the impact of SAP on the agency has been positive” and (B) “Overall, the impact of SAP on me has been positive.” With the objective of further assessing the content, construct, and criterion validity of the factor solution, two composite measures of overall success were next computed as follows: (C) **criterion average** is the simple average of the two criterion items; and (D) **dimensions average** is the simple average of the four success dimensions. Table 4 shows results of correlating the three criterion measures (A), (B), and (C), with the four dimensions and their average (D). From Table 4 we can make several broad observations.

The extent to which each dimension or the dimensions average correlates with the criterion scores is evidence of their criterion validity.\(^\text{16}\) All correlations in the table are significant at the .001 level suggesting strong correspondence between the criterion measures and success dimensions. The three largest correlations in Table 4 are (A), (B), and (C) with (D), yielding r’s of 0.80, 0.83, and 0.86 respectively. These large correlations further increase our confidence in the validity of each of the four dimensions. The largest correlation (r = 0.86) is between (C) criterion average and (D) dimensions average, suggesting that (C) and (D) may be our strongest measures of overall success.\(^\text{17}\)

That the (D) dimensions average yields the largest correlation with all the criteria further supports the view that the dimensions are additive, and thus when combined yield a stronger overall measure of success than possible from any single dimension; in other words, this further evidences the validity of the dimensions average.

\(^{15}\)With loadings of -0.504, -0.206, 0.476, and 0.691 for SQ, OI, IQ, and II respectively.

\(^{16}\)This method of validation assumes the criterion measures are valid (Kerlinger 1988).

\(^{17}\)With the objective of yet further assessing the content, construct, and criterion validity of the factor solution, two further composite measures of overall success were computed: (E) satisfaction is the simple average of the four satisfaction items and (F) combined criterion is the simple average of the four success dimensions. Table 4 shows results of correlating the three criterion measures (A), (B), and (C), with the four dimensions and their average (D). From Table 4 we can make several broad observations.

Evidence of the validity of combining the four satisfaction items with the two criterion items, to yield (F) combined criterion, includes (1) satisfaction correlates strongly with (C) criterion average (r = 0.81), (2) (F) has the largest correlation with (D) dimensions average (r = 0.90), and (3) (C) has strong face validity. With reference to the last listed point, following is the exact wording of these six items from the survey instrument. The face validity of the construct is strong in that together these six items address each of the four success dimensions, as well as SAP overall. It is further noted that these are the only five of the 43 (41 + 2 criterion) items intended to measure success that proceed with the wording “Overall, …”: “Overall, the impact of SAP on the agency has been positive,” “Overall, the impact of SAP on me has been positive,” “Overall, the SAP system quality is satisfactory,” “Overall, the SAP information quality is satisfactory,” “Overall, SAP is satisfactory,” “SAP is enjoyable to use.”

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Criterion Average</th>
<th>Satisfaction</th>
<th>Combined Criterion (C+3)/2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C</td>
<td>E</td>
<td>F</td>
</tr>
<tr>
<td>1 Information Quality</td>
<td>0.67</td>
<td>0.75</td>
<td>0.74</td>
</tr>
<tr>
<td>2 System Quality</td>
<td>0.75</td>
<td>0.86</td>
<td>0.85</td>
</tr>
<tr>
<td>3 Individual Impact</td>
<td>0.76</td>
<td>0.65</td>
<td>0.75</td>
</tr>
<tr>
<td>4 Organization Impact</td>
<td>0.77</td>
<td>0.69</td>
<td>0.77</td>
</tr>
<tr>
<td>D Dimensions Average</td>
<td>0.86</td>
<td>0.86</td>
<td>0.90</td>
</tr>
</tbody>
</table>

*Equal weight is given to the two criterion items and the four satisfaction items.*
Table 4. Correlations between Criteria and Dimensions

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Impact on Agency A</th>
<th>Impact on Individual B</th>
<th>Criterion Average C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Information Quality</td>
<td>0.63</td>
<td>0.64</td>
<td>0.67</td>
</tr>
<tr>
<td>2 System Quality</td>
<td>0.72</td>
<td>0.72</td>
<td>0.75</td>
</tr>
<tr>
<td>3 Individual Impact</td>
<td>0.64</td>
<td>0.79</td>
<td>0.76</td>
</tr>
<tr>
<td>4 Organization Impact</td>
<td>0.76</td>
<td>0.71</td>
<td>0.77</td>
</tr>
<tr>
<td>D Dimensions Average</td>
<td><strong>0.80</strong></td>
<td><strong>0.83</strong></td>
<td><strong>0.86</strong></td>
</tr>
</tbody>
</table>

The Revised Model

Figure 3 depicts the revised model. It has the four quadrants—(1) individual impact, (2) organizational impact, (3) information quality, and (4) system quality—representing four distinct but related dimensions of the multidimensional phenomenon: enterprise systems success. When evaluating an enterprise system, measures of these dimensions represent a snapshot of the organization’s experience of the enterprise system at a point in time. The impact dimensions are an assessment of benefits that have followed (or not) from the system. The quality dimensions reflect future potential. Together, these four dimensions reflect an ostensibly complete view of the enterprise system—an overarching measure of enterprise systems success.

The revised model for ES success deviates from the traditional Delone and McLean model in the following ways: (1) it depicts a measurement model and does not purport a causal/process model of success, (2) it omits the use construct, (3) satisfaction is treated as an overall measure of success, rather than as a dimension of success, (4) new measures were added to reflect the contemporary IS context and organizational characteristics, and (5) it includes additional measures to probe a more holistic organizational impacts construct.

Discussion

This paper has presented a validated model and instrument for measuring enterprise system success from multiple perspectives. Given past IS success studies have lacked theoretical grounding, the selection of model constructs in this study was grounded in an exploratory (model building) survey aimed at confirming the relevance and completeness of the most widely cited IS success
model. This analysis and later validation of the model constructs (model testing) suggest the existence of four distinct and individually important dimensions of success that the authors believe are applicable to any IS evaluation. The constructs are positively associated and when combined yield a single valid measure of overall success.

**Implications for Research**

The study addresses several concerns with past IS studies:

- deflects concerns over lack of theoretical justification by conceiving dimensions of a measurement model rather than constructs in a causal or process model
- further diffuses this concern through clear statement of the rationale for choice of dimensions, grounded in the exploratory survey and related mapping exercise
- clearly states the rationale for selection of the success measures
- presents empirical evidence of the irrelevance of use in the study context
- evidences the redundancy of a usefulness construct, given a complete set of measures of the four study model dimensions
- presents a strong rationale for conceiving satisfaction as an overarching measure of success rather than as a dimension
- validates the final model from multiple stakeholder perspectives: management, user, and technical

In addition, the study reflects

- close attention to the mutual exclusivity of the dimensions
- the most extensive and complete set of IS measures tested in a single IS success study
- first operationalization and test of the Myers et al. (1997) IS success framework
- first test of the completeness and relevance of the six Delone and McLean (1992) constructs
- first partial empirical test of the Shang and Seddon (2000) benefits framework
- evidence of the additivity of the four model dimensions, and the validity of an overall measure of success based in their combination

The correlations in Table 4 (and footnote 17), combined with results of second-order factor analysis, although far from conclusive, offer consistent and, we feel, compelling evidence of the additivity of the four model dimensions. We further note that we have in essence reduced all dimensions to the common denominator impacts, with individual-impact and organizational-impact referring to impacts to date, and system-quality and information-quality referring to expected future impacts, thereby lending further strength to the argument of additivity.18

Important findings for the ES context include

- validates the model constructs and measures in a contemporary ES context

---

18Note also that, of all measures of ESS reported herein, the dimensions-average (D) has the largest correlations with antecedents of ESS (e.g., client knowledge, consultant knowledge) to be reported elsewhere (Sedera et al. 2003b).
first comprehensive empirical, quantitative assessment of ES success reported in the academic press

- illustrates the relative importance of nonfinancial measures of ES success

Implications for Practice

The final model and the survey instrument can be used to benchmark organizations/departments with ES. Although, the study was conducted in a public sector environment, we feel that except for the organizational impact measure of e-government, which has a direct counterpart in e-commerce or e-business, the model is generalizable and appropriate for measuring ES success in the private sector. Thus, we believe that the survey instrument developed in this study can be successfully used as-is or with minor variation in a profit-oriented organization to evaluate ES success. Our sense is that the model and instrument are readily adaptable to other ES by simply replacing the term SAP with some reference to an alternate ES (i.e., another generic ES, such as ORACLE, or an internal acronym, such as QGFMS, Queensland Government Financial Management System).

The study model and instrument offer a practical means for organizations to evaluate the success of complex, contemporary information systems like ES. A two-stage approach might be employed, the first round again aiming to instill increased confidence in model completeness, and to aid in adapting the model/instrument as necessary. Alternatively, given only minor changes to the model/instrument, round 1 might be dispensed with based on strong model validity observed in this study.

This paper has focused on model validity. Having validated the model, the study focus now shifts to descriptive and comparative analysis of the model data. As suggested earlier, in order gain a fuller appreciation, it is important to analyze IS success at multiple levels within the organization, from multiple perspectives. The study model and instrument have explicitly been designed to work for all levels of the organization. Having gathered these multiple perspectives, it becomes possible from the model item and dimension mean scores to usefully describe and compare these various perspectives.

Mean scores on the dimensions and measures establish a benchmark against which further, future executions of the survey can be compared. Differences across the respondent cohorts (or across other demographics, such as agencies) suggest either differing expectations or differing experiences, both of which suggest possible value from management action. Differing expectations may suggest the need for better conditioning of expectations of particular groups, or increased training. Differing experiences of the ES may suggest value in closer attention to the source of these differences, and may point to problem organizational entities, modules, versions, processes, cultures, or even individuals.

References


