A meta-analytic assessment of the DeLone and McLean IS success model: An examination of IS success at the individual level

Stacie Pettera a,⁎, Ephraim R. McLean b

⁎College of Information Science & Technology, University of Nebraska at Omaha, 1110 S. 67th Street, Omaha, NE 68182-0392, United States

1. Introduction

At the first International Conference on Information Systems, Keen [18] posed a key question that needed to be answered by the IS field to establish a coherent discipline, namely: “What is the dependent variable?” In 1992, DeLone and McLean (D&M) addressed this question by defining IS success as the dependent variable of the field. Their review of the literature resulted in a taxonomy of IS success consisting of six variables: System Quality, Information Quality, Use, User Satisfaction, Individual Impact, and Organizational Impact. The model also identified the various relationships among these success variables; but, at the same time, cautioned researchers that the model needed “further development and validation” [9].

Considerable research has been devoted to examining the effectiveness of IT. Numerous papers have been published on the topics of IS success, evaluation, effectiveness, and acceptance. Since D&M first published their model, over 1000 publications have referenced their work; and at least 150 empirical studies have examined some or all of the relationships in the model. However, the various relationships in the IS success model have found differing levels of support within the empirical literature. Some studies found high correlations among the variables, while others found either low or nonsignificant correlations. Therefore, to reconcile these conflicting results, we applied meta-analysis to examine each of the relationships.

2. The history of the D&M IS success model

2.1. The original D&M model

D&M reviewed the literature published in 1981–1987 in seven publications to develop a taxonomy of IS success. This taxonomy was based upon Mason’s modification of the Shannon and Weaver model [37] of communications which had identified three levels of information: the technical level (accuracy and efficiency of the system that produces it), the semantic level (its ability to transfer the intended message), and the effectiveness level (its impact on the receiver). Mason adapted this theory for IS and expanded the effectiveness level into three categories: receipt of information, influence on the recipient, and influence on the system [27].

D&M identified categories for system success by mapping an aspect of IS success to each of Mason’s effectiveness levels. This analysis yielded six variables of IS success: System Quality, Information Quality, Use, User Satisfaction, Individual Impact, and Organizational Impact. System Quality was equivalent to the technical level of communication, while Information Quality was equivalent to the semantic level of communication. The other four variables mapped to Mason’s subcategories of the effectiveness level. Use related to Mason’s “receipt of information.” User
Satisfaction and Individual Impact were associated with the “information’s influence on the recipient.” Organizational Impact was the “influence of the information on the system.”

D&M developed their initial taxonomy using established theories of communication adapted to IS. These theories suggested that the flow of information was linear; however, they suggested that for IS, these different measures of success were independent, but that there was interdependency among them. Fig. 1 shows the original model.

D&M suggested that researchers should use this model in a predictive manner, yet they cautioned that one must measure and/or control each of the variables in the model to ensure a complete understanding of IS success. D&M called upon others to validate their model.

2.2. D&M 10-year update

In the years that followed, several researchers altered or extended the model, while others adapted it for specific applications, such as knowledge management [e.g., 17,21] or e-commerce [e.g., 11] systems. Recognizing these potential improvements over their original model, D&M acknowledged these modifications and revised their model accordingly [10]. The updated model is shown in Fig. 2.

D&M also modified their model to address some limitations of the original model. A key addition in the updated model was the inclusion of Service Quality as an additional aspect of IS success [31]; it was added because the changing nature of IS required the need to assess service quality when evaluating IS success. D&M also recommended assigning different weights to System Quality, Information Quality, and Service Quality depending on the context and application of the model.

Another modification was the elimination of Individual Impact and Organizational Impact as separate variables, replacing them with Net Benefits. This change addressed the criticism that IS can affect levels other than individuals and organizations. Thus, the updated model accounted for benefits occurring at any level of analysis (workgroups, industries, and societies also experience IS success [29,36]); the choice of which level was to be determined by the researcher using the model.

Seddon [33] proposed a well-known respecifications of the original model; one of his concerns was that the model had elements of both process and variance models, making it, in his view, difficult to interpret and use. His change separated the process and variance components; however, D&M contended that this made the model too complicated and lacked parsimony. D&M stated that their original model, as a process model, had three components: creating and using the system, and the effects of its use. However, each of these steps was necessary, but not sufficient, for the outcome. They also supported the variance component by citing many empirical studies that fully or partially examined portions of the model.

Reflecting on this debate, another enhancement to the 2003 model was clarification of the Use construct. The authors explained this as: “Use must precede ‘user satisfaction’ in a process sense, but positive experience with ‘use’ will lead to greater ‘user satisfaction’ in a causal sense”. They felt that, given the variability of IS and their contexts, it may sometimes be appropriate to measure Intention to Use (an attitude) rather than Use (a behavior). They went on to state that if Intention to Use was a measure, then increased User Satisfaction would lead to a higher Intention to Use, which would subsequently affect Use. This resulted in the addition of Intention to Use in the updated model.

---

Fig. 1. DeLone and McLean original IS success model.

Fig. 2. DeLone and McLean updated IS success model.
These are given in Table 2. H1 to H11 are specifically stated in the updated D&M model by Seddon's model. The D&M model stood up reasonably well and outperformed the Seddon model. McGill et al. [28] also examined the full model, but found four paths in the original model to be insignificant (System Quality → Use, Information Quality → Use, Intended Use → Individual Impact, Individual Impact → Organizational Impact).

A number of researchers [e.g., 3, 12, 45], have conducted literature reviews to examine if the results of empirical studies supported the relationships posited by the original IS success model. These reviews revealed that some relationships in the model had consistently received support (i.e., significant in all) while others have received only mixed support (some found significant results while others did not). In their recent update, D&M identified studies that examined the relationships in the original success model. Consistent with other reviews, some relationships exhibited strong support with all findings significant, while others had mixed or little support.

We decided to examine the strength of the interdependent relationships among the variables that make up IS success. To do so, we developed hypotheses consistent with the updated model. These are given in Table 2. H1 to H11 are specifically stated in the updated D&M model by the use of arrows to represent relationships among the constructs. Within the updated D&M IS success model, there is a path from User Satisfaction to Net Benefits and another path from Net Benefits back to User Satisfaction. These cannot be tested as separate hypotheses in a meta-analysis; therefore, we combined them into a single hypothesis (H10) that suggests that there is a positive correlation between these two constructs. We also chose to include three additional hypotheses (H12 to H14); these were implied in the original D&M model, but are no longer part of the updated model. While D&M acknowledged that Intention to Use was a different type of measurement of the Use construct, there were many studies that choose to measure self-reported or actual use of an IS. Therefore, these three hypotheses, related to the association between the construct of Use and System Quality, Information Quality, and Service Quality, were included in our evaluation of the model.

### Table 1
IS success model constructs.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>System Quality</td>
<td>Performance of the IS in terms of reliability, convenience, ease of use, functionality, and other system metrics</td>
</tr>
<tr>
<td>Information Quality</td>
<td>Characteristics of the output offered by the IS, such as accuracy, timeliness, and completeness</td>
</tr>
<tr>
<td>Service Quality</td>
<td>Support of users by the IS department, often measured by the responsiveness, reliability, and empathy of the support organization</td>
</tr>
<tr>
<td>Intention to Use</td>
<td>Expected future consumption of an IS or its output</td>
</tr>
<tr>
<td>Use</td>
<td>Consumption of an IS or its output described in terms of actual or self-reported usage</td>
</tr>
<tr>
<td>User Satisfaction</td>
<td>Approval or likeability of an IS and its output</td>
</tr>
<tr>
<td>Net Benefits</td>
<td>The effect an IS has on an individual, group, organization, industry, society, etc., which is often measured in terms of organizational performance, perceived usefulness, and affect on work practices</td>
</tr>
</tbody>
</table>

2.3. Current understanding of the D&M IS success model

There have been few studies that attempted to validate the success model as a whole in a single study. Rai et al. [32] compared the original D&M model with Seddon's model. The D&M model stood up reasonably well and outperformed the Seddon model. McGill et al. [28] also examined the full model, but found four paths in the original model to be insignificant (System Quality → Use, Information Quality → Use, Intended Use → Individual Impact, Individual Impact → Organizational Impact).

A number of researchers [e.g., 3, 12, 45], have conducted literature reviews to examine if the results of empirical studies supported the relationships posited by the original IS success model. These reviews revealed that some relationships in the model had consistently received support (i.e., significant in all) while others have received only mixed support (some found significant results while others did not). In their recent update, D&M identified studies that examined the relationships in the original success model. Consistent with other reviews, some relationships exhibited strong support with all findings significant, while others had mixed or little support.

We decided to examine the strength of the interdependent relationships among the variables that make up IS success. To do so, we developed hypotheses consistent with the updated model. These are given in Table 2. H1 to H11 are specifically stated in the updated D&M model by the use of arrows to represent relationships among the constructs. Within the updated D&M IS success model, there is a path from User Satisfaction to Net Benefits and another path from Net Benefits back to User Satisfaction. These cannot be tested as separate hypotheses in a meta-analysis; therefore, we combined them into a single hypothesis (H10) that suggests that there is a positive correlation between these two constructs. We also chose to include three additional hypotheses (H12 to H14); these were implied in the original D&M model, but are no longer part of the updated model. While D&M acknowledged that Intention to Use was a different type of measurement of the Use construct, there were many studies that choose to measure self-reported or actual use of an IS. Therefore, these three hypotheses, related to the association between the construct of Use and System Quality, Information Quality, and Service Quality, were included in our evaluation of the model.

3. Research methodology

Literature reviews can provide insights in assessing the results of empirical studies; however, they alone cannot determine the strength of the relationships among the variables in the model. Meta-analysis is a quantitative approach that aggregates research studies and takes into account the error inherent in all quantitative studies; it is a useful technique to reconcile conflicting results among research studies and to conclude about the data that have been accumulated in a particular domain.

3.1. Identifying studies

We sought to include as many studies as possible within our examination of the model. To identify studies, we performed multiple full-text searches in online databases, such as ABI/INFORM and Business Source Premier, for papers published between 1992 (the year that the original model was published) and mid-2007 using different sets of keywords, such as “DeLone and McLean,” “IS success,” or “IS effectiveness.” We also included studies examined in similar literature reviews and/or meta-analyses [e.g., 26] if they were published in 1992 or later. This yielded nearly five hundred manuscripts.

We developed several criteria for inclusion in the meta-analysis before our examination of individual studies. We excluded a manuscript if (a) it was a literature review or purely theoretical [e.g., 5, 22]; (b) it had no quantitative empirical data to report [e.g., 7]; (c) it did not examine any of our hypotheses [e.g., 15]; or (d) it...
examined one or more of the hypotheses, but did not include enough data to perform a meta-analysis [e.g., 30].

After performing this step, we had approximately 150 studies requiring further examination. As we began looking at each study, we defined additional criteria for inclusion in our meta-analysis. First, we identified studies using the same dataset [e.g., 34,35] and retained only the study that examined the larger number of our hypotheses. Second, as we examined each study, we noted that some focused on IS success at an organizational level while others examined it at an individual level. To ensure that results were interpretable, we decided not to mix levels; since most studies concentrated on IS success at an individual level, we excluded those that examined the hypotheses using a unit of analysis other than the individual. Finally, we excluded studies if the research model was distinctly different from the D&M model, as some studies examined other models, such as TAM, TPB, or Task-Technology Fit [e.g., 4]. Therefore, studies using an alternative model were only included if the authors specifically cited one of the D&M papers as justification to include one or more variables within their study [e.g., 23]. This criterion, however, did not require that all studies had to cite one of the D&M models. If the study was not specifically analyzing another model, the study was still retained in our meta-analysis.

Applying this additional set of criteria resulted in a tally of 53 samples from 52 studies.

3.2. Coding studies

From each study, we obtained the following information: sample size, reliability of constructs (as reported using Cronbach's alpha or inter-item reliability scores), and correlations for each hypothesis. We also carefully considered the measurement of each success construct. This was critical because constructs in the original IS success model could be measured and defined differently in different studies. For example, in their original work, D&M identified User Satisfaction to be an overall satisfaction measure, a single-item measure, a multi-item measure, and/or enjoyment. Therefore, when coding a study, we used the list of possible measures provided by D&M in their original study for each of the variables; this allowed us to determine that even if the construct was not named User Satisfaction in the paper it could be used in the meta-analysis when the measure captured the concept of user satisfaction as defined by D&M.

3.3. Analyzing data

We followed the methods of Hunter and Schmidt [14] to analyze the data. Before conducting the meta-analysis, we needed a single correlation for each relationship. This was more complex in our project because several studies used multiple variables to represent a single construct (e.g., User Friendliness and Ease of Use are both dimensions of System Quality, yet by combining the different measures, we were able to achieve higher content validity in the meta-analysis [34]). When multiple variables existed for a single construct, we used the following formula to find the correlation for the relationship:

\[ r_{xy} = \frac{\sum r_{xy}}{(1)\sqrt{n + n(n-1)/r_{yy}}} \]

where \( r_{xy} \) is the correlation between the variables that accounts for multiple operationalizations of the variables, \( r_{xy} \) is the correlation between each operationalization of the variable and another construct, \( n \) is the number of operationalizations of the variable, and \( r_{yy} \) is average inter-construct correlation for the multiple variables. If the researchers did not supply enough data in their paper to use this formula (e.g., not providing a full correlation matrix), an average was used.

Next, we performed a “bare-bones” meta-analysis for each hypothesis to correct for sampling error using the following formula:

\[ \bar{r} = \frac{\sum [N_i r_i]}{\sum N_i} \]

where \( \bar{r} \) is the bare-bones meta-analysis effect size, \( N_i \) is the sample size of each study, and \( r_i \) is the correlation of each study. We performed this analysis for each of the hypotheses. A bare-bones meta-analysis considers the sample size of each study and therefore, creates a weighted average of the correlations.

We then calculated 95% confidence intervals of the bare-bones meta-analysis effect size using the standard error to determine the significance of each of the relationships. Calculating the confidence interval allows one to determine the significance of the findings. Confidence intervals that do not include zero suggest the relationship is significant [16].

While the bare-bones meta-analysis corrected for sampling error, there are additional artifacts beyond sampling error that can be corrected in a meta-analysis; we chose to correct for attenuation due to measurement error in the independent and dependent variable. Because no instrument can perfectly measure a construct, measurement error affects the correlation between two constructs. Reliability errors degrade the correlation; therefore, correcting for reliability error increases the correlations among constructs and presents a more accurate picture of the true magnitude of the relationship. To correct for measurement error, the correlation between the constructs was corrected using the formula:

\[ r_c = \frac{r_{xy}}{\sqrt{r_{xx}r_{yy}}} \]

where \( r_c \) is the meta-analysis effect size corrected for measurement error, \( r_{xy} \) is the reported correlation between the two variables, \( r_{xx} \) is the reliability estimate for the independent variable and \( r_{yy} \) is the reliability estimate for the dependent variable. If a study did not report reliability statistics, we used a weighted average of similar measures from other studies to substitute as the reliability statistics.

Another test necessary for any given meta-analysis is a homogeneity test [13,42] which determines if there are fluctuations in the correlations in a meta-analysis. These fluctuations, or heterogeneity, often suggest that an additional variable is creating the variance and affecting the effect size statistic. To determine if the studies within the meta-analysis are homogenous, credibility intervals were calculated using the corrected standard deviation. Credibility intervals are similar to confidence intervals in that there is an upper and lower bound. If the credibility interval range crosses zero or the range is large, moderators may be present [20].

A final test was performed to determine the robustness of the findings. The Fail-Safe N statistic was calculated for each of the relationships. Many journals and authors are discouraged from publishing findings that are not significant; this means that effect sizes are biased upward because only published studies with some or all significant results are able to be included in the meta-analysis. One can limit this bias, however, by performing a statistical test to determine how many studies with a correlation of zero between two variables is necessary to reduce the effect size to a trivial result, such as 0.10 correlation [24].
3.4. Meta-analysis results

Table 3 shows the results of the individual meta-analyses. Examining the 95% confidence interval of each study suggested that the majority of the relationships within the original and updated D&M models were supported. One hypothesis, H5 (Service Quality and Intention to Use), was not examined because only one study had examined it. H6 (Service Quality and User Satisfaction) and H14 (Service Quality and Use) were also found to be nonsignificant because the 95% confidence interval included zero. All other hypotheses were supported.

We also evaluated the 90% credibility intervals. If they included zero or were sufficiently large, then the studies were not homogenous, suggesting moderators (i.e., unknown variables) were present. None of the 90% credibility intervals included zero within their range; however, the size of the credibility intervals were somewhat large for many of the meta-analyses. This suggested that the population sampled in each of the studies were probably not truly homogeneous.

In our assessment of the results, we also considered the Fail-Safe N statistic. This is helpful to determining the robustness of the results and whether publication bias was a problem. A rule of thumb is that the Fail-Safe N statistic should be at least twice the number of studies in the meta-analysis (i.e., ‘K’). All of the supported hypotheses had a ratio of Fail-Safe N to the number of studies greater than 2.0, except for H7 (Use and User Satisfaction). The ratio of Fail-Safe N to K for H7 was only 1.77. Given that the magnitude of the relationship between Use and User Satisfaction was the weakest of the supported hypotheses (i.e., $\rho = 0.28$), it is not surprising that it failed the Fail-Safe N test.

4. Discussion

4.1. Empirical support for the D&M models

We found that the majority of hypotheses implied by the updated D&M models were supported. Table 4 summarizes the results of our hypothesis tests.

The three unsupported hypotheses were associated with the Service Quality construct; one was not testable due to a lack of studies examining this relationship, while the other two (Service Quality and User Satisfaction; Service Quality and Use) were examined, but were not found to be significant.

### 4.2. Strength of relationships in the D&M models

One of the benefits of meta-analysis is its ability to examine the strength or magnitude of the relationship between two constructs. Cohen and Cohen [8] established general heuristics to judge the magnitude of the effect sizes as strong (0.50), moderate (0.30), or weak (0.10). Examining the corrected correlation coefficients ($\rho$), the majority of the relationships in the original and updated IS success models were moderate or strong. To illustrate this result, we grouped and ordered the results of the individual meta-analysis using Cohen and Cohen’s heuristics.

The only weak relationship was Use and User Satisfaction. A relatively large number of studies had examined this relationship (i.e., 26 studies); and yet, surprisingly, the correlation was still weak. Our findings are consistent with other meta-analyses that have examined this relationship.

All other relationships were moderate to strong, suggesting that the relationships identified by D&M for these success variables were empirically supported.

4.3. Insights on IS success

4.3.1. Intention to Use

In this research, we found that, empirically, the inclusion of Intention to Use in the updated D&M model is appropriate and supported. The relationships between System Quality, Information Quality, User Satisfaction, and Net Benefits with Intention to Use

### Table 3

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Total N</th>
<th>$K$</th>
<th>$r$</th>
<th>Var $r$</th>
<th>95% Confidence interval</th>
<th>$\rho$</th>
<th>Var $\rho$</th>
<th>90% Credibility interval</th>
<th>Fail-Safe N</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1: System Quality and Intention to Use</td>
<td>2864</td>
<td>12</td>
<td>0.48</td>
<td>0.02</td>
<td>0.41, 0.56</td>
<td>0.54</td>
<td>0.02</td>
<td>0.33, 0.75</td>
<td>53</td>
</tr>
<tr>
<td>H2: System Quality and User Satisfaction</td>
<td>3653</td>
<td>17</td>
<td>0.54</td>
<td>0.04</td>
<td>0.44, 0.63</td>
<td>0.60</td>
<td>0.04</td>
<td>0.27, 0.93</td>
<td>85</td>
</tr>
<tr>
<td>H3: Information Quality and Intention to Use</td>
<td>1312</td>
<td>5</td>
<td>0.52</td>
<td>0.01</td>
<td>0.45, 0.58</td>
<td>0.56</td>
<td>0.01</td>
<td>0.46, 0.67</td>
<td>23</td>
</tr>
<tr>
<td>H4: Information Quality and User Satisfaction</td>
<td>2136</td>
<td>10</td>
<td>0.53</td>
<td>0.05</td>
<td>0.39, 0.66</td>
<td>0.58</td>
<td>0.06</td>
<td>0.20, 0.96</td>
<td>48</td>
</tr>
<tr>
<td>H5: Service Quality and Intention to Use</td>
<td>1650</td>
<td>9</td>
<td>0.54</td>
<td>0.04</td>
<td>0.44, 0.63</td>
<td>0.60</td>
<td>0.04</td>
<td>0.27, 0.93</td>
<td>85</td>
</tr>
<tr>
<td>H6: Service Quality and User Satisfaction</td>
<td>366</td>
<td>3</td>
<td>0.21</td>
<td>0.04</td>
<td>0.00, 0.43</td>
<td>0.24</td>
<td>0.05</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>H7: Use and User Satisfaction</td>
<td>5231</td>
<td>26</td>
<td>0.25</td>
<td>0.02</td>
<td>0.19, 0.31</td>
<td>0.28</td>
<td>0.03</td>
<td>0.03, 0.52</td>
<td>46</td>
</tr>
<tr>
<td>H8: Use and Satisfaction to Intention to Use</td>
<td>2245</td>
<td>9</td>
<td>0.65</td>
<td>0.03</td>
<td>0.54, 0.75</td>
<td>0.74</td>
<td>0.03</td>
<td>0.46, 1.00</td>
<td>57</td>
</tr>
<tr>
<td>H9: Use and Net Benefits</td>
<td>4416</td>
<td>26</td>
<td>0.36</td>
<td>0.04</td>
<td>0.28, 0.43</td>
<td>0.39</td>
<td>0.04</td>
<td>0.07, 0.72</td>
<td>76</td>
</tr>
<tr>
<td>H10: User Satisfaction and Net Benefits</td>
<td>6030</td>
<td>31</td>
<td>0.52</td>
<td>0.04</td>
<td>0.45, 0.59</td>
<td>0.58</td>
<td>0.05</td>
<td>0.23, 0.92</td>
<td>148</td>
</tr>
<tr>
<td>H11: Net Benefits and Intention to Use</td>
<td>3335</td>
<td>14</td>
<td>0.55</td>
<td>0.05</td>
<td>0.43, 0.67</td>
<td>0.63</td>
<td>0.07</td>
<td>0.20, 1.00</td>
<td>74</td>
</tr>
<tr>
<td>H12: System Quality and Use</td>
<td>2408</td>
<td>15</td>
<td>0.30</td>
<td>0.03</td>
<td>0.21, 0.39</td>
<td>0.34</td>
<td>0.05</td>
<td>0.01, 0.67</td>
<td>36</td>
</tr>
<tr>
<td>H13: Information Quality and Use</td>
<td>897</td>
<td>7</td>
<td>0.42</td>
<td>0.03</td>
<td>0.28, 0.55</td>
<td>0.49</td>
<td>0.06</td>
<td>0.09, 0.88</td>
<td>27</td>
</tr>
<tr>
<td>H14: Service Quality and Use</td>
<td>448</td>
<td>4</td>
<td>0.09</td>
<td>0.07</td>
<td>-0.18, 0.35</td>
<td>0.09</td>
<td>0.09</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Total N = total sample size for the given meta-analysis; $K$ = number of studies included in the meta-analysis; $r$ = bare-bones meta-analysis effect size (weighted average corrected only for sampling error, not reliability); Var $r$ = variance of the bare-bones effect size; 95% Confidence interval = 95% confidence interval for the bare-bones meta-analysis effect size based on the standard error; $\rho$ = effect size corrected for reliability; Var $\rho$ = variance of effect size corrected for reliability; 90% Credibility interval = test for homogeneity in the sample based on standard deviation of $\rho$; Fail-Safe N = test to evaluate robustness of the findings (i.e., number of studies required to reduce effect size to a value 0.10).
have stronger relationships than the same constructs have with Use (see Table 5). With this finding, we have some ideas and considerations for researchers interested in examining this phenomenon further.

One reason why Intention to Use (i.e., an attitude) seems to have a stronger relationship with other variables than Use (i.e., a behavior) is that we have more consistent and reliable measures for modeling Intention to Use than we have for modeling System Use. We found that Use was measured as actual use, self-reported use, depth of use, and importance of use. Each of these operationalizations of Use is addressing different aspects of the Use construct. Until we have robust, consistent, and reliable measures of Use, it may be difficult to fully understand the relationship between Use and these other variables of IS success.

Another reason why Intention to Use has a stronger relationship with success variables than Use could be that, at least in terms of the Use construct in IS, intention does not always result in behavior. While it is reasonable to assume that intentions will lead to behavior, individuals may overestimate their likelihood of using the IS.

### 4.3.2. Moderating relationships

While we found significant relationships in most of the relationships within the IS success model, and found that that most were moderate to strong, we believe that these estimates are biased downward and the relationships are actually stronger than reported. The large credibility intervals for many relationships suggest that there is a chance that moderators are biasing the correlations among the various relationships downward. Yet, the presence of moderators is still uncertain, given the relatively small number of studies examined in each meta-analysis. The large credibility intervals could be due to the relatively small size of the meta-analysis or due to unknown variables or factors affecting the results.

The first potential moderator that could be affecting each of the relationships is the voluntariness of the system. There has been discussion of whether use in the IS success model is appropriate under conditions where it is mandatory. D&M argued in their updated work that:

… whereas usage of a system may be mandatory at one level, the continued adoption and use of the system itself may be wholly voluntary, based upon management judgment, at a higher level…

Some studies found differences in technology acceptance between voluntary and mandatory systems [6]. We did examine whether the voluntariness of the system affected the relationship between Use and User Satisfaction. This relationship was easy to analyze using a moderator analysis because of: (1) the relatively large number of studies examining this relationship, and (2) the weak effect size for this relationship. Our moderator analysis did not find that voluntariness impacted the results; however, this does not provide conclusive evidence that it is irrelevant in the IS success model.

Another potential moderator is the measurement of each of the constructs. In their original paper, D&M identified multiple measures for each of the variables that make up the IS success model; however, the authors stated a need for a “significant reduction in the number of different dependent variable measures so that research results can be compared.” Yet there is still little standardization of the measures used for each of the success variables. In D&M’s work updating their success model, they concluded that “an attempt should be made to reduce significantly the number of measures used to measure IS success so that research results can be compared.” Once several studies use the same set of measures to examine the constructs within IS success model, it will be possible to determine if measurement choice is a moderator that impacts relationships of IS success.

A third potential moderator is the choice of population for a given study. Several studies included in our meta-analysis used students [e.g., 2], while others relied on IS users in organizations. Some studies drew from samples in North America, while others examined IS success in Asia [e.g., 1,19] or Europe [e.g., 40,41]. Some examined IS in general [e.g., 38], while others examined a specific IS such as a data warehouse [e.g., 43] or the World Wide Web [e.g., 25].

### 4.4. Limitations

Meta-analysis provides certain benefits as a research technique; but it still has limitations. A criticism of meta-analysis is that it can only be performed on research that performs quantitative analysis; conceptual or qualitative work must be excluded. This restricts the body of work that can be examined and included. Additionally, each of the quantitative studies in a meta-analysis must report sufficient data (such as correlation matrices for the constructs) to perform the statistical tests needed. Studies that do not include these data must be excluded and this may affect the results [e.g., 44].

Another concern is that the measures used in each study can vary, sometimes dramatically. For example, Use may be measured as extent of use, frequency of use, and duration of use. These measures may also be actual or self-reported. Should these measures be combined in the context of a meta-analysis? Other

---

**Table 5**

Magnitude of relationships.

<table>
<thead>
<tr>
<th>Relationship strength</th>
<th>Meta-analysis</th>
<th>Total N</th>
<th>K</th>
<th>ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H8: User Satisfaction and Intention to Use</td>
<td>2245</td>
<td>9</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>H11: Net Benefits and Intention to Use</td>
<td>3335</td>
<td>14</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>H2: System Quality and User Satisfaction</td>
<td>3653</td>
<td>17</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>H10: User Satisfaction and Net Benefits</td>
<td>6030</td>
<td>31</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>H4: Information Quality and User Satisfaction</td>
<td>2136</td>
<td>10</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>H3: Information Quality and Intention to Use</td>
<td>1312</td>
<td>5</td>
<td>0.56</td>
<td></td>
</tr>
<tr>
<td>H1: System Quality and Intention to Use</td>
<td>2864</td>
<td>12</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H13: Information Quality and Use</td>
<td>897</td>
<td>7</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>H9: Use and Individual Impact</td>
<td>4416</td>
<td>26</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>H12: System Quality and Use</td>
<td>2408</td>
<td>15</td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>Weak</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H7: Use and User Satisfaction</td>
<td>5231</td>
<td>26</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>Not significant</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H6: Service Quality and User Satisfaction</td>
<td>366</td>
<td>3</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>H14: Service Quality and Use</td>
<td>448</td>
<td>4</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>Not tested</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H2e: Service Quality and Intention to Use</td>
<td>n/a</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total N = total sample size for the given meta-analysis; K = number of studies included in the meta-analysis; ρ = effect size corrected for reliability.
meta-analyses on IS success have addressed the problem in the same manner as we did. There is, however, a possible benefit to using multiple measures for a given construct. The content validity of each variable of the IS success model should be quite high because we are capturing not just one aspect of a variable (such as frequency of use), but a comprehensive variable that includes the full range of the construct. The potential downside of integrating these studies with different measures is that the measurement is a moderator of the relationship between constructs. This can bias the directionality of the relationship, meaning that each of these relationships evaluated should have an even stronger magnitude than we found in our study.

Another limitation of meta-analysis is that it does not assess causal relationship between variables. We believe that a causal analysis of the IS success model would be helpful and useful. However, to perform a causal analysis, one needs perform individual meta-analysis for all possible relationships among the variables to develop a correlation matrix and not just the relationships hypothesized [39], which cannot be done with the available data.

A final limitation of our meta-analysis was that we limited the studies to those that focused on the individual level of analysis. We made this choice to prevent confounding results from different levels of analysis; however, the generalizability of the findings must therefore be questioned, primarily as they relate to the measures of the Net Benefits construct. Although we did restrict the unit of analysis, the types of IS included in this meta-analysis is quite broad. Examples of studies included decision-support systems, enterprise systems, accounting IS, knowledge management systems, and many others. However, we caution readers to interpret the findings with care.

5. Conclusion

Our work assesses the updated D&M IS success model using meta-analysis. While others have also examined some or all of the relationships in the original IS success model using meta-analysis, we evaluated the updated D&M IS success model. This allowed us to evaluate the role of intention to Use and Service Quality, which are recent additions to the model.

We have empirically evaluated the relationships within the D&M IS success model using the quantitative method of meta-analysis and found that the majority of the relationships posited in the updated D&M IS success model were supported. As businesses have become more reliant on IT in achieving success within their organizations, IS have become essential. A better understanding of what constitutes an effective IS is necessary. The D&M IS success model provides a framework that can assist in understanding this.

Acknowledgements

We would first like to thank Michael Gallivan for introducing us to the idea of conducting a meta-analysis on this topic. Also, we would like to thank Yi Ding and Legodile Kwapwa for assisting with data collection and verification. Finally, we would like to thank the many researchers that have read different drafts of this manuscript and offered their thoughtful comments throughout the writing process.

References


Additional studies included in the meta-analysis


Stacie Petter

Stacie Petter is Assistant Professor of Information Systems and Quantitative Analysis at the University of Nebraska at Omaha. Stacie received her Ph.D. and M.B.A. from the University of Nebraska at Omaha. Stacie has been published in MIS Quarterly, Journal of Information and Organizational Sciences, and Information Sciences. Stacie's research interests include software project management, knowledge management, information systems success, and research methods.