The relationship between system usage and user satisfaction: a meta-analysis

Rahat H. Bokhari
Pakistan Computer Bureau, Ministry of Information Technology, Islamabad, Pakistan

Abstract

Purpose – System usage and user satisfaction are widely accepted and used as surrogate measures of IS success. Past studies attempted to explore the relationship between system usage and user satisfaction but findings are mixed, inconclusive and misleading. The main objective of this research is to better understand and explain the nature and strength of the relationship between system usage and user satisfaction by resolving the existing inconsistencies in the IS research and to validate this relationship empirically as defined in Delone and McLean's IS success model.

Design/methodology/approach – “Meta-analysis” as a research approach was adopted because of its suitability regarding the nature of the research and its capability of dealing with exploring relationships that may be obscured in other approaches to synthesize research findings. Meta-analysis findings contributed towards better explaining the relationship between system usage and user satisfaction, the main objectives of this research.

Findings – This research examines critically the past findings and resolves the existing inconsistencies. The meta-analysis findings explain that there exists a significant positive relationship between “system usage” and “user satisfaction” (i.e. $r = 0.2555$) although not very strong. This research empirically validates this relationship that has already been proposed by Delone and McLean in their IS success model. Provides a guide for future research to explore the mediating variables that might affect the relationship between system usage and user satisfaction.

Originality/value – This research better explains the relationship between system usage and user satisfaction by resolving contradictory findings in the past research and contributes to the existing body of knowledge relating to IS success.

Keywords Information systems, Customer satisfaction

Paper type Research paper

1. Introduction

The information systems (IS) discipline is primarily concerned with the successful implementation of information technology (IT) in organizations. IS are an essential component of the solutions to many of the problems faced by organizations to cope with the current challenges. In this light, it can be argued that successful IS development can be identified by certain characteristics or metrics. An IS may thus be considered successful if it meets criteria such as fulfilling user needs and organisational objectives/goals (which are in themselves both multifaceted, and partial). At the same time, a variety of factors may affect systems during their development and implementation. As a result of these factors, the evaluation of a system in terms of its “success” is an inherently complex phenomenon.

Nevertheless, the measurement of IS success has achieved prime importance among researchers. Measures of success include assessment of satisfaction as reported by key personnel, and the evaluation of the capabilities of the constructed system, described in
various terms such as system performance, effectiveness, quality, use and users’ satisfaction. For example, Miller (1989, p. 283) argues that an “effective information system is one that achieves the purposes of its users”. Garrity and Sanders (1998) define IS performance (i.e. success) in terms of its delivery on time and to budget, and the extent to which users are satisfied with it.

The tendency to define success in terms of user responses to the information systems is understandable, as users are an immediate, articulate, and often vocal provider of comments on the IS they use. As such, the most commonly used measures of IS success are “system usage” and “user satisfaction”, though it is rare that both are used simultaneously to evaluate the same system. However various studies have attempted to explore the relationship between system usage and user satisfaction but the findings about this relationship are mixed. Some studies show zero or negative correlation while others found a positive correlation between system usage and user satisfaction. As Mawhinney (1990 p. 954) notes:

... although these two variables [system usage and user satisfaction] have been widely used, researchers tend to use one or the other, but seldom use both. Little has been done to establish a relationship between these two variables.

Kim (1989) similarly argues for further research in this area, and this provides one of the motivations for the meta-analysis reported in this paper. However, I would argue that there is sufficient research in this area for a determination to be made, though the answer is not necessarily what we might expect.

This paper will first review research focused on the use of the concepts of user satisfaction, and system use in IS evaluation, before turning attention to the integration of existing research through meta-analysis. Using these meta-analytic techniques one should articulate a summary measure of the strength of the relationship between user satisfaction and system usage that consequently might be helpful to resolve the existing controversies within IS research. The author attempted to conclude with an analysis of the research findings and posit an alternative approach to the interpretation of such measures.

2. Measures of information system success
The prime purpose of introducing IS into an organisation is to improve not only individual decision-making performance but overall organisational efficiency and effectiveness. The expected gains in terms of organizational goals through IS implementation and adoption are concerned with the success of the system. While, as noted earlier, IS success is a multidimensional construct (see Delone and McLean, 1992; Saarinen, 1996), so surrogates measures have been developed and are usually in use to measure IS success. Two surrogate measures are system usage (Swanson, 1974; Ein-Dor et al., 1984; Snitkin and King, 1986) and user satisfaction (Bailey and Pearson, 1983; Ives et al., 1983; Baroudi and Orlikowski, 1988).

Delone and Mclean (1992) described a number of IS success measures after reviewing 180 studies. According to them research has focused on areas such as:

- systems quality (i.e. IS interface, availability, response time, etc.);
- information quality (i.e. characteristics of information in terms of accuracy, timeliness, reliability, format and meaningfulness, etc.);
- system usage;
• user satisfaction;
• individual impact; and
• organizational impact.

They draw on these to develop a model of IS success (see Figure 1). Their model emphasises the influences and inter-relationships between these various factors. Amoroso and Cheney (1991) also noted that system quality and information quality are two major constructs of user satisfaction. Delone and McLean’s (1992) model suggests that both system usage and user satisfaction are affected by these two factors.

Our study does not deal with how all of these factors may affect the dependents but focus on to explore and explain the relationship between system usage and user satisfaction. Delone and McLean’s (1992) model not only consolidated previous findings, but also classified the measures of success in to six plausible dimensions. However, the proposed relationships in their model were not tested empirically. Various other studies (see Delone and McLean, 2002) empirically investigated the relationship among the measures of IS success. Garrity and Sanders (1998) measures IS success at different levels – the organisational level (i.e. how a system contributes to organisational performance), the process level (i.e. efficient use of resources) and the individual level (i.e. the users’ perception of utility and satisfaction). Zmud (1979) advocated three factors related to users (i.e. system usage, user satisfaction and user performance) as measures of systems success.

Szajna (1993 p. 153) says:

While IS usage, user satisfaction, and perceptions of performance have all been employed as surrogate measures of information success, there are inconsistencies in their relationships.

Having identified the relevance and popularity of these surrogates of IS success, I shall now analyse each in detail before exploring the relationship between them.

2.1. System usage
In IS research, system usage may be defined as “either the amount of effort expended interacting with an information system or, less frequently, as the number of reports or other information products generated by the information system per unit time” (Trice and Treacy, 1988, p. 33). Delone and McLean (2002) suggest that the nature, quality, and appropriateness of the system use are also important and not just to simply measure time spent of system use. They consider system usage as the necessary condition under which IS/IT can affect organizational performance. Davis (1989, 1993)
found perceived usefulness to be significantly correlated with system usage and that perceived usefulness positively affects user satisfaction (Mahmood et al., 2000). Such research highlights the importance of system usage and user satisfaction in evaluating a system in terms of its success.

System usage is continuously considered as dependent variable in various empirical studies. A number of different system usage models have been proposed (Schewe, 1976; Robey, 1979; Lucas, 1975a). At the same time, a variety of measures have been developed and used to assess system usage in the IS field (see for example, Doll and Torkzadeh, 1998).

The use of a system depends on the users’ evaluation of that system. If the system improves the users’ task performance or decision quality, then they tend to use the system otherwise they may avoid to use a system unless its use is mandatory. Delone and McLean (2002) believe that no system use is totally mandatory. It might happen that some time the management requires employees to use the system but continued and adoption of the system itself may be voluntary based on management judgement at higher level. On the other hand, Kim and Lee (1986) caution that the degree of system usage could not be considered as an appropriate measure for IS success if use is mandatory. For this reason, some researchers prefer to use “user satisfaction” as a measure of success. Previous research (Fuerst and Cheney, 1982; Igbaria, 1990; Yoon and Guimaraes, 1995; Hendrickson and Collins, 1996) highlights two common measures of system usage – time spent in term of hours and frequency of use.

Seddon and Kiew (1994) and Seddon (1997) criticized Delone and McLeans model and also raised model related issues. They suggest “usefulness” as more meaningful as compared to “system use” as an indicator of IS success that may lead to user satisfaction. The usefulness of the system in terms of benefits that an organization may gain should be observed on its use of course. Delone and McLeans (2003) argue that systems use is an appropriate measure of IS success in most cases in previous research, so the inclusion of system use in success model is more appropriate than system usefulness suggested by Seddon. This article does not deal with model issues raised by Seddon and Kiew (1994) but focus specifically on the inter-relationship between system usage and user satisfaction (IS success dimensions mentioned in Delone and McLean’s model).

In the previous research, different factors are studied that may influence system usage (see Liang, 1986; Trice and Treacy, 1988, Mahmood and Swanberg, 2001). Among these factors are user attitude (Lucas, 1975a, b; Robey, 1979; Kaiser and Srinivasan, 1980), user expectations (Ginzberg, 1981), user involvement (Schewe, 1976; Ives and Olson, 1984; Yoon and Guimaraes, 1995), user experience (Igbaria et al., 1989; Yap et al., 1994; Loh and Ong, 1998) and user satisfaction (Baroudi et al., 1986; Joshi, 1992). One of the key challenges emerging from this review of past research is the diversity of inter-related terms and concepts that are employed.

2.2. User satisfaction

The measurement of users’ satisfaction with IS remains of prime concern to researchers. User satisfaction measures are categorized in terms of three perspectives such as:

1. user attitudes towards IS;
2. user satisfaction in terms of information quality; and
3. user satisfaction in terms of perceived IS effectiveness (see Kim, 1989).
The concept of user satisfaction can be traced back to several decades (e.g. Swanson, 1974; Nolan and Seward, 1974). According to Bailey and Pearson (1983, p. 531) “satisfaction in a given situation is the sum of one’s feelings and attitudes toward a variety of factors affecting the situation”. As with system usage, a variety of measures have been proposed for the quantification and assessment of user satisfaction (for example Bailey and Pearson, 1983; Doll and Torkzadeh, 1988; Baroudi and Orlikowski, 1988). Ives et al. (1983) identify staff and service, information product, vendor support, and knowledge and involvement as factors that underlie user satisfaction. They considered user satisfaction as “the extent to which users believe the information system available to them meets their information requirements” (Ives et al., 1983, p. 785). This definition suggests that users perceive the system irrespective of its technical quality, and relates to the fulfilment of user needs by an IS. Kim (1989) describes user satisfaction in terms of information quality, system effectiveness and user attitudes. Doll and Torkzadeh (1988, 1989) considered user satisfaction in term of system quality (i.e. system accuracy, ease of use) and information quality.

Delone and McLean (1992) included about 33 articles that address the subject of user satisfaction in their research. They conclude that user satisfaction is widely used as a measure of IS success because reliable instruments have been developed to measure satisfaction, and other measures of IS success are problematic. However, while user satisfaction has been widely used as a surrogate for systems performance and IS success, critics have questioned its general applicability (see, for example, Galletta and Lederer 1989). Some also suggests for further studies for a reliable measurement of user satisfaction (McGill et al., 2000).

3. Systems use and user satisfaction
Research regarding a relationship between systems use and user satisfaction are mixed and inconclusive. Some researchers argue that system usage leads to user satisfaction – others that user satisfaction leads to system usage. Mawhinney (1990) says that there is a lack of sufficient research effort to establish a conclusive relationship between these two constructs. Insufficient research is in some ways an easy answer, offering the opportunity for researchers to continue in well-established patterns of work until “good results” emerge. In this section I explore and statistically analyse the existing corpus of research in the area. Consequently (and for the most part, implicitly), I therefore argue that there is already sufficient research in this field to determine the existence or otherwise of such a relationship (between systems use and user satisfaction), but the answer is not necessarily what we might expect.

3.1 Alternative perspectives
The literature can be organised into three dominant perspectives – those that explore how usage impacts satisfaction; those that explore how satisfaction impacts usage; and finally, those that seek to uncover any relationship between the two. In this section I review each perspective in order to demonstrate the range of opinion, and to provide the basis for the statistical analysis in the following section.

Usage effects satisfaction. According to Delone and McLean (1992 p. 83):

… the amount of USE can affect the degree of USER SATISFACTION – positively or negatively – as well as the reverse being true.
Delone (1988) noted that duration of computer use is not associated with system success. Conrath and Mignen (1990, p. 9) question the axiomatic assumption that “positive attitudes leads to increased usage”, arguing that it is not clear-cut. Their research concludes that usage has more impact on satisfaction than vice versa and both are positively related. Conversely, Udo (1992) argued that the more a system is used, the less it is found to be effective, and that this may in turn reduce user satisfaction. Jenkins and Ricketts (1985, p. 1) state that [M]IS has no effect on decision making unless it is used, so satisfaction that results from the use of an MIS correlates with the quality of current decisions and the likelihood of future use.

Using path analysis, Lee et al. (1995) claim that system utilization leads to user satisfaction and both factors are positively related. Also using path analysis, Torkzadeh and Dwyer (1994, p. 345) found that the effect of system usage on satisfaction was slightly greater than the reverse.

Satisfaction effects usage. On the other hand, Evans (1976, cited in Thong and Yap, 1996) discusses a lower bound to satisfaction below which users may discontinue system use, which gives some direct insight into how user satisfaction influences system usage. Using path analysis, Baroudi et al. (1986) argued that user information satisfaction (UIS) could lead to system usage rather than vice versa (interpreting UIS as an attitude). Their model assumes that as a system’s use fulfils user needs, user satisfaction with the system should increase and lead to greater use. Conversely, if system usage does not meet user needs, satisfaction will decrease and restrict further use. Lucas (1975b) argued that the use of IS is dependent on user attitudes and perceptions and that both may be influenced by the quality of the system. Positive user attitudes were consistent with high level of system usage. Joshi (1992, p. 77) argues that user attitudes (as measures of user satisfaction) may influence behaviour, and that positive user satisfaction may help in obtaining acceptance and usage of IS. Schewe (1976) found no significant relationship between attitudes and system usage behaviour. He wrote:

Batch and interactive system user’s behaviour toward the information system does not appear to be influenced by their feelings of satisfaction with the use of the system (Schewe (1976, pp. 588, 589).

Bikson and Gutek (1983) also found no relationship between level of usage and level of satisfaction with new technology.

Does a causal relationship exist? As I have indicated in the sections above, some research has suggested a causal link between system usage and user satisfaction in both directions. There exists, however, a considerable body of evidence surrounding these central foci that elicit a variety of opinion.

Research findings (Lucas, 1975b; Robey, 1979) provide evidence that heavily used systems are positively correlated to user satisfaction. In stark contrast, Schewe (1976) found no significant relationship between system use and user satisfaction, and Lawrence and Low (1993) did not find this relationship to be significant. Similarly, Mawhinney (1990) found no relationship between user satisfaction and system use, and Srinivasan (1985) notes that the relationship is not always positive. Udo (1992) reviewed studies/labatory experiments that measured systems effectiveness in terms of user satisfaction and system usage. He concluded that the respondents who frequently used the system perceived it less effective in terms of increased productivity and cost efficiency.
Other interpretations of the relationship exist. Cheney and Dickson (1982, p. 175) argue that system “utilization does not necessarily mean that the users are satisfied, the [as] system may be all they have and it’s better than nothing”. Some argue that dissatisfied users may discontinue system usage and seek alternatives (Ginzberg, 1981; Sauer, 1993; Szajna and Scamell, 1993).

Some researchers have argued that system usage and user satisfaction are actually negatively correlated (Ang and Soh, 1997; Mawhinney, 1990; Mawhinney and Lederer, 1990), while (Liang, 1986; Udo and Guimaraes, 1994; Loh and Ong, 1998) showed a very weak relationship. Others (Yoon and Guimaraes, 1995; Choe, 1998; Kivijarvi and Zmud, 1993) found a positive correlation between system usage and user satisfaction.

Researchers also mention a relationship between user attitude and system usage. Such findings are similarly mixed. For example, Schewe (1976) found no significant relationship between attitude and system usage whereas Lucas (1975b) says that a favourable user attitude leads to a high use level. Al-Khaldi and Al-Jabri (1998) note that attitudes towards computers are significant determinants of behaviour and that they may influence computer use.

In conclusion, the findings about a system usage and user satisfaction interrelationship are mixed and inconclusive, with proponents on each side drawing empirical evidence in support of their argument. Such uncertainty, coupled with the weight of conventional wisdom, has led to an uneasy position in analysis of the interaction between user satisfaction and system usage.

In the following section, I describe my work to consolidate and draw insights from the large body of literature in this area. Drawing on research methods developed in the social sciences, I undertake a “meta-analysis” of these various studies, using the techniques to distil conclusions drawn from the large sample population, from which I draw implications and conclusions later in this paper.

4. Research approach
The process of research builds on the efforts of those who worked before. The overall social process involved in formalised academic publishing works to minimise the chance that researchers may repeat the past mistakes of others by drawing on a community, rather than an isolated model of work. Thus, inquiry into the existing research in a specific domain allows the construction of a comprehensive picture of the research. Research synthesis can therefore play an important part in the process of theory building and testing (Forza and Nuzzo, 1998). In recent years, meta-analysis (i.e. quantitative research synthesis) has been gaining acceptance in IS research (for example Montazemi and Wang, 1988-1989; Alavi and Joachimsthaler, 1992; Hwang and Thorn, 1999; Liu, 2004). Previously, it has been widely used in fields such as education, psychology, and public policy. Meta-analysis offers a set of tools for extracting and analysing quantitative information from pre-existing, related research. It is underpinned by several decades of theoretical development and practical experience, which provides a strong basis for its use in IS research, as I demonstrate later in this paper.

4.1 Meta-analysis
Conflicting interpretations of results on the same subject in a specific domain are not uncommon. Without some form of consolidation this can lead to unending calls for
further research. Meta-analysis may be used to review such research to aid the comprehension and interpretation of diverse research findings. It provides a procedure for combining relevant information gathered from separate studies designed to answer essentially the same research question. This technique has been successfully applied in the physical as well as the social sciences (Hunter and Schmidt, 1990), general management (Schwenk, 1990), and Information Systems (Gelderman, 1997; Hwang and Thorn, 1999; Mahmood et al., 2000). Meta-analysis of previous research not only reveals the cumulative knowledge embedded within the research corpus, but also provides clear direction about future research needs by answering the original question based on many studies rather than just one (Glass et al., 1981).

The term “meta-analysis” was first coined and applied by Glass (1976), though it is also referred to as “research integration” (Walberg and Haertel, 1980 cited in Green and Hall, 1984):

Meta-analysis refer to the analysis of analyses ... the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings (Glass, 1976, p. 3; Wolf 1986, p. 11).

Similarly, Hwang and Wu (1990) state that it allows one to analyse quantitatively large numbers of studies and to resolve differences that exist due to conflicting and contradictory research findings. It can be used to examine differences across studies, integrate previous research and test any possible relationship between variables of interest.

Meta-analysis may therefore be considered a means to review the results of previous research in a domain in quantitative terms with the intent of identifying significant relationships that exists between the variables under study. Thus, by synthesizing findings from many studies succinctly, meta-analysis can help researchers uncover consistencies in a set of seemingly inconsistent findings and to arrive at deeper insights into the problem domain (Durlak, 1995; Hunt, 1997).

Most researchers believe that the goal of cumulative research is to produce better answers than that which can be obtained in isolated studies. The growing trend of the use of meta-analysis in research has been encouraged and argued (Hunt, 1997; Lyons, 1998). It could be argued that traditional research reviews have been unable to provide definite answers in many circumstances, and meta-analysis has been successfully employed in the fields of medicine, education and psychology (Hunt, 1997). MIS researchers have also been using it to integrate previous research findings (Pettingell et al., 1988; Pervan, 1994; Gelderman, 1997). The specific techniques of meta-analysis may vary, but their underlying aim is constant – to derive a quantitative measure (effect size) of the relationship under study. Further support for the process may be found in, for example, Hunter and Schmidt (1990), who argue that better and valid inferences can be drawn by meta-analysis as compared to other traditional reviews. In summary, meta-analysis is:

... rooted in the fundamental values of the scientific enterprise: replicability, quantification, causal and correlational analysis. Valuable information is needlessly scattered in individual studies. The ability of social scientists to deliver generalizable answers to basic questions of policy is too serious a concern to allow us to treat research integration lightly. The potential benefits of meta-analysis seem enormous (Bangert-Drowns 1986).
4.2. Literature selection
The literature was searched in libraries manually and through available academic databases on computers. It encompasses major journals, conference proceedings, books and dissertations. Various papers were studied but those included are a best fit with the research question under study. The studies fall within the period of 1979 to 2000. The studies reporting empirical findings testing relationship between system usage and user satisfaction in terms of correlation “r”, or other statistics (such as $F$, $t$, $d$, $r$, or $\chi^2$) those can be converted to correlations, along with sample size are selected. Details of journal/conference proceedings and the number of studies selected are shown in Table I.

5. Data collection and analysis
The data collected from individual studies were aggregated and analysed statistically to reach informed conclusions. It is natural that different studies may report their findings using different statistics (i.e. correlations), so the results cannot be compared until those have been converted into some common metric to be useful to conduct meta-analysis. The well-known statistic measure Pearson Correlation (Pearson’s “r”) is used as an index of “Effect size” that represents the strength of relationship between system usage and user satisfaction. The reason for selecting “r” as an effect size is due to:

- availability of formulae to convert other test statistics into “r”; and
- interpretation is easy (Lipsey and Wilson, 2001).

<table>
<thead>
<tr>
<th>Journal/proceedings</th>
<th>Number of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Support Systems</td>
<td>1</td>
</tr>
<tr>
<td>The Computer Journal</td>
<td>1</td>
</tr>
<tr>
<td>Communications of the ACM</td>
<td>1</td>
</tr>
<tr>
<td>Computers in Industry</td>
<td>1</td>
</tr>
<tr>
<td>Computers in Human Behavior</td>
<td>1</td>
</tr>
<tr>
<td>Journal of Marketing Research</td>
<td>1</td>
</tr>
<tr>
<td>Journal of Information Technology</td>
<td>1</td>
</tr>
<tr>
<td>Information Processing &amp; Management</td>
<td>1</td>
</tr>
<tr>
<td>Public Opinion Quarterly</td>
<td>1</td>
</tr>
<tr>
<td>Australian Journal of Information Systems</td>
<td>1</td>
</tr>
<tr>
<td>Systems, Objectives, Solutions</td>
<td>1</td>
</tr>
<tr>
<td>Proceedings Australasian Conference on Information Systems</td>
<td>1</td>
</tr>
<tr>
<td>Proceedings Annual Meeting of Decision Sciences Institute</td>
<td>1</td>
</tr>
<tr>
<td>Behaviour &amp; Information Technology</td>
<td>1</td>
</tr>
<tr>
<td>Data Base</td>
<td>1</td>
</tr>
<tr>
<td>Information Resources Management Journal</td>
<td>2</td>
</tr>
<tr>
<td>European Journal of Information Systems</td>
<td>2</td>
</tr>
<tr>
<td>Proceedings International Conference on Information Systems</td>
<td>2</td>
</tr>
<tr>
<td>Unpublished PhD dissertation</td>
<td>2</td>
</tr>
<tr>
<td>Journal of End User Computing</td>
<td>3</td>
</tr>
<tr>
<td>Academy of Management Journal</td>
<td>3</td>
</tr>
<tr>
<td>Journal of Management Information Systems</td>
<td>4</td>
</tr>
<tr>
<td>Omega</td>
<td>4</td>
</tr>
<tr>
<td>MIS Quarterly</td>
<td>5</td>
</tr>
<tr>
<td>Information &amp; Management</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>55</td>
</tr>
</tbody>
</table>

Table I. Study sources
The value of “r” is either obtained from the publication directly or converted from other reported statistics (such as $F$, $t$, $d$, or $\chi^2$) to “r” using standard formulas (Wolf, 1986; Lyons, 1998). The findings from 55 empirical studies included 58 effect sizes that were included for meta-analysis (Table II). An effect size was calculated for each study used in the analysis, so that the correlation between system usage and users’ satisfaction could be determined. In some studies (Sanders and Courtney, 1985; Mawhinney and Lederer, 1990) more than one effect size value was available and subsequently included. Most of the studies considered the system usage construct as behaviour and user satisfaction as attitude. Details of the data and calculated effect sizes are presented in Table II.

Initial analysis of these results is presented in Figure 2. Figure 2 illustrates the spread of effect sizes across each of the studies, showing again quite a range of results. Coupled with the scatter diagram (Figure 3) showing the distribution of effect sizes against year of publication there seems little evidence of a clear-cut relationship, nor of the ongoing research leading towards a more coherent body of research.

This initial analysis showed a considerable divergence of effect size, and did not suggest that the process of research had led to a consensual position being achieved. I therefore continued the analysis, and analyse the parameters of the study in order to frame the results appropriately.

5.1 Combinations of effect sizes
Correlations provide a sound basis for handling effect sizes as they not only describe the strength of the relationship between two variables but also the direction (i.e. positive or negative association). The values of effect size collected from the studies included in this research are given in Table II. Pearson’s correlation “r” was used as the effect size and its value falls within the rage of $-1.0$ to $+1.0$. However, since the sample sizes ($N$) for the studies vary considerably, the weighted average of the correlations was used as a more appropriate estimate (see Hunter and Schmidt 1990).

We therefore took the sample size of the study as its weight for further analysis, and using standard formulae (see Hunter and Schmidt, 1990) estimated the population correlation, observed variance and the error due to sampling as follows.

Estimate of population correlation (i.e., Mean)

$$\bar{r} = \frac{\sum [N_i r_i]}{\sum N_i}$$  \hspace{1cm} (1)

where:

$\bar{r}$ = weighted average estimate of the population correlation.

$r_i$ = observed correlation for sample i.

$N_i$ = number of individuals in sample i.

$\sum N_i$ = sum of all individuals in all samples included in this study.

The mean correlation between system usage and users’ satisfaction in our study was found to be 0.2555. Lipsey and Wilson (2001, p. 147) interpreted magnitude of effect size “r” as small when “r ≤ 0.10”, medium when “r = 0.25”, and finally large when “r ≥ 0.40”. So research findings lead to the conclusion that system usage and user
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample (n)</th>
<th>Effect size “r”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schewe (1976)</td>
<td>79</td>
<td>0</td>
</tr>
<tr>
<td>Maish (1979)</td>
<td>50</td>
<td>0.356</td>
</tr>
<tr>
<td>Robey (1979)</td>
<td>66</td>
<td>0.42</td>
</tr>
<tr>
<td>Ein-Dor et al. (1981)</td>
<td>18</td>
<td>0.302</td>
</tr>
<tr>
<td>Cheney and Dickson (1982)</td>
<td>79</td>
<td>0.233</td>
</tr>
<tr>
<td>Sanders (1983)</td>
<td>264</td>
<td>0.247</td>
</tr>
<tr>
<td>Ein-Dor et al. (1984)</td>
<td>490</td>
<td>0.462</td>
</tr>
<tr>
<td>Barki and Huff (1985)</td>
<td>42</td>
<td>0.394</td>
</tr>
<tr>
<td>Sanders, and Courteney (1985)</td>
<td>132</td>
<td>0.317</td>
</tr>
<tr>
<td></td>
<td>156</td>
<td>0.238</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>0.278</td>
</tr>
<tr>
<td>Ettema (1985)</td>
<td>131</td>
<td>0.073</td>
</tr>
<tr>
<td>Srinivasan (1985)</td>
<td>29</td>
<td>−0.05</td>
</tr>
<tr>
<td>Raymond (1985)</td>
<td>464</td>
<td>0.356</td>
</tr>
<tr>
<td>Baroudi et al. (1986)</td>
<td>200</td>
<td>0.28</td>
</tr>
<tr>
<td>Liang (1986)</td>
<td>35</td>
<td>0.114</td>
</tr>
<tr>
<td>Snitkin and King (1986)</td>
<td>31</td>
<td>0.352</td>
</tr>
<tr>
<td>Ein-Dor and Segev (1986)</td>
<td>18</td>
<td>0.545</td>
</tr>
<tr>
<td>Zinkhan et al. (1987)</td>
<td>165</td>
<td>0.235</td>
</tr>
<tr>
<td>Nath (1989)</td>
<td>98</td>
<td>0.32</td>
</tr>
<tr>
<td>Igbaria (1990)</td>
<td>187</td>
<td>0.14</td>
</tr>
<tr>
<td>Igbaria and Nachman (1990)</td>
<td>104</td>
<td>0.27</td>
</tr>
<tr>
<td>Raymond (1990)</td>
<td>34</td>
<td>0.286</td>
</tr>
<tr>
<td>Mawhinney and Lederer (1990)</td>
<td>66</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>39</td>
<td>−0.02</td>
</tr>
<tr>
<td>Mawhinney (1990)</td>
<td>95</td>
<td>−0.002</td>
</tr>
<tr>
<td>Igbaria and Zviran (1991)</td>
<td>294</td>
<td>0.26</td>
</tr>
<tr>
<td>Amoroso and Cheney (1991)</td>
<td>506</td>
<td>0.18</td>
</tr>
<tr>
<td>Udo (1992)</td>
<td>201</td>
<td>0.1728</td>
</tr>
<tr>
<td>Kivijarvi and Zmud (1993)</td>
<td>159</td>
<td>0.44</td>
</tr>
<tr>
<td>Conrath and Sharma (1993)</td>
<td>167</td>
<td>0.55</td>
</tr>
<tr>
<td>Lawrence and Low (1993)</td>
<td>155</td>
<td>0</td>
</tr>
<tr>
<td>Suh et al. (1994)</td>
<td>150</td>
<td>0.29</td>
</tr>
<tr>
<td>Udo and Guimaraes (1994)</td>
<td>201</td>
<td>0.17</td>
</tr>
<tr>
<td>Igbaria et al. (1994)</td>
<td>471</td>
<td>0.24</td>
</tr>
<tr>
<td>Igbaria and Guimaraes (1994)</td>
<td>185</td>
<td>0.195</td>
</tr>
<tr>
<td>Zinatelli (1994)</td>
<td>358</td>
<td>−0.14</td>
</tr>
<tr>
<td>Lee et al. (1995)</td>
<td>236</td>
<td>0.287</td>
</tr>
<tr>
<td>Bergeron et al. (1995)</td>
<td>38</td>
<td>0.33</td>
</tr>
<tr>
<td>Yoon and Guimaraess (1995)</td>
<td>69</td>
<td>0.54</td>
</tr>
<tr>
<td>Vlahos and Ferratt (1995)</td>
<td>55</td>
<td>0.27</td>
</tr>
<tr>
<td>Leidner (1996)</td>
<td>89</td>
<td>0.15</td>
</tr>
<tr>
<td>Guimaraes et al. (1996)</td>
<td>114</td>
<td>0.5</td>
</tr>
<tr>
<td>Suarinen (1996)</td>
<td>48</td>
<td>0.28</td>
</tr>
<tr>
<td>Wilkins (1996)</td>
<td>72</td>
<td>0.28</td>
</tr>
<tr>
<td>Harrison and Rainer (1996)</td>
<td>653</td>
<td>0.11</td>
</tr>
<tr>
<td>Seddon and Kiew (1996)</td>
<td>94</td>
<td>0.572</td>
</tr>
<tr>
<td>Ang and Soh (1997)</td>
<td>133</td>
<td>0</td>
</tr>
<tr>
<td>Igbaria and Tan (1997)</td>
<td>371</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Table II. Calculated effect sizes (“r”) of the relationship between system usage and user satisfaction for 55 reported studies (continued)
Table II.

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample (n)</th>
<th>Effect size “r”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guimaraes and Igbaria (1997)</td>
<td>186</td>
<td>0.21</td>
</tr>
<tr>
<td>Choe (1998)</td>
<td>450</td>
<td>0.515</td>
</tr>
<tr>
<td>Gelderman (1998)</td>
<td>73</td>
<td>0.17</td>
</tr>
<tr>
<td>Winter et al. (1998)</td>
<td>279</td>
<td>0.26</td>
</tr>
<tr>
<td>Loh and Ong (1998)</td>
<td>84</td>
<td>0.014</td>
</tr>
<tr>
<td>Kim et al. (1998)</td>
<td>134</td>
<td>0.255</td>
</tr>
<tr>
<td>Torkzedah and Doll (1999)</td>
<td>409</td>
<td>0.29</td>
</tr>
<tr>
<td>Khalil and Elkordy (1999)</td>
<td>120</td>
<td>0.359</td>
</tr>
<tr>
<td>McGill et al. (2000)</td>
<td>79</td>
<td>0.274</td>
</tr>
</tbody>
</table>

Figure 2.
Effect size frequency histogram

Figure 3.
Effect size versus publication date
satisfaction are positively correlated. However, the strength of the relationship is rather medium. The observed variance is calculated using the weighted sum of squares formula to obtain an estimate of true variance (Formula 2).

Sample weighted variance (observed variance)

\[ s_r^2 = \frac{\sum N_i (r_i - \bar{r})^2}{\sum N_i} \]  

(2)

where:

\( s_r^2 \) = frequency weighted average square. Others terms are explained earlier.

The sampling error is function of sample size. Estimate error due to sampling is:

\[ \sigma_e^2 = \frac{((1 - \bar{r}^2)K)}{\sum N_i} \]  

(3)

where:

\( \sigma_e^2 \) = estimate of sampling error.

\( K \) = the number of results included in the analysis. Others terms as above.

The residual variance can be calculated by subtracting estimates of the variance due to sampling error from observed variance. The “residual variance” is the variance that is due to the true differences among results from different studies.

\[ \text{Residual variance} = S_r^2 - \sigma_e^2 \]  

(4)

In this study, I calculated estimates of the variance due to sampling error (Formula 3) and then subtract it from the observed variance, giving a residual variance (Formula 4). If the residual variance is sufficiently small then the difference between correlations can be considered to be sampling errors (Hunter and Schmidt, 1990). For this study, the residual variance was 0.01931. Tempering this result however, was the residual standard deviation, which is greater than 25 per cent of mean effect size (i.e. \( r = 0.2555 \)). As residual standard deviation exceeded the 25 per cent of effect size, the residual variance could not be considered as small (see Stoffelmayr et al., 1983).

This suggested that the differences between the correlations may have been due to artefacts other than sampling error, a point I return to in the later discussion section (Section 6). Different study artefacts can alter the size of the study correlation (Hunter and Schmidt, 1990). Trotman and Wood (1991) suggest testing for a moderator effect – in this case.

5.2 Fail safe N

It is unlikely that one can uncover every study concerned with the hypothesis under research. For example, it may be assumed that a number of studies with non-significant effect size are either rejected or not submitted for publication. The unpublished studies are also difficult to find as compared to published ones. Lipsey and Wilson (1993) show that on average unpublished studies have smaller mean effect size than published studies, so provide an evidence of publication bias. Smith (1980 p. 22) stated that published literature only represents 5 per cent of false positives in a population of studies supporting null hypothesis. According to Rosenthal (1979), the
studies supporting the null hypothesis of no significant results are more likely to be buried away in file drawers.

In this way, a biased conclusion may be obtained from a meta-analysis that includes the effects of a strong bias toward publishing positive results, a point identified by Kraemer and Andrews (1982, p. 405) and Wolf (1986, p. 37). Hedges and Olkin (1985) and Rosenthal (1979) suggested finding the number of studies confirming the null hypothesis that would be needed to reverse a conclusion that a significant relationship exists. Cooper (1979) called it “Fail safe N (Nfs)”. The formula proposed by Hunter and Schmidt (1990) is used (Formula 5 below) to estimate the number of “lost” studies that must exist to bring \( r_k \) to \( r_c \) (where \( r_k \) is the mean effect size \( r \) for \( k \) studies, \( r_c \) the critical mean effect size \( r \) that may be considered theoretically or practically significant):

\[
x = k(r_k/r_c - 1)
\]

(5)

where:

- \( K \) is the number of studies included in meta-analysis
- \( x \) denotes the number of studies required to bring \( r_k \) to \( r_c \).

The Fail safe N reports the number of unpublished studies having null results that may reduce the cumulated effect across studies to the point of non-significance (Lipsey and Wilson, 2001). Fail safe N calculated taking \( r_c = 0.05 \), and 0.1 for this research follows:

- Fail Safe N for critical \( r \) of 0.05 = 238;
- Fail Safe N for critical \( r \) of 0.1 = 90.

If any question is raised that observed results may be biased due to sampling then one may conclude that to reverse the conclusion (i.e. there exists a positive relationship between system usage and user satisfaction) drawn from this research would require 238 studies when \( r_c = 0.05 \), or 90 studies when \( r_c = 0.1 \) mentioning “no correlation” between system usage and user satisfaction to be included in the sample. A careful sampling and thorough search may be the best to avoid a publication bias.

The research findings are summarized as shown in Table III.

### 6. Discussion

The main objective of this research is to better understand the strength and nature of relationship between system usage and user satisfaction by resolving the existing controversies in the past research. The direct measure of IS success is difficult if not possible (Galletta and Lederer, 1989), so the system usage and user satisfaction are

<table>
<thead>
<tr>
<th>Number of studies/effect sizes</th>
<th>Sample size</th>
<th>Effect size “r” (weighted)</th>
<th>Residual variance</th>
<th>Residual standard deviation</th>
<th>Residual std. deviation &gt; 1/4 of effect size “r^2”</th>
<th>Confidence interval r_c = 0.05</th>
<th>Fail safe N r_c = 0.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>58</td>
<td>9,767</td>
<td>0.2555</td>
<td>0.01931</td>
<td>0.1389</td>
<td>Yes</td>
<td>-0.01688</td>
<td>238</td>
</tr>
</tbody>
</table>

Table III. Summary of research findings
commonly used as a surrogate measure of IS success despite criticism relating to their inconsistent theoretical definitions, lack of validation of instruments developed for their measurements leading to measurement issues (Goodhue, 1986; Kim et al., 1998). Such problems may be one of the causes for variations among the IS researchers’ findings about the phenomena (see Melone, 1990; Wilkin et al., 1999). Despite ample individual studies dealing with system use and user satisfaction, the findings about their relationship are inconclusive. A comprehensive understanding and knowledge may lead to some substantial conclusions. In order to resolve such inconsistencies/disparities among the studies “meta analysis” is used. Meta-analysis can help in integrating/cumulating the past research findings and to reach some substantive conclusion by resolving the existing conflicts/inconsistencies.

In this research, the meta-analysis of 55 studies (comprising 58 effect sizes) revealed a clearer view about the relationship between “system usage and user satisfaction”, the factors under study. This research found a positive correlation (effect size $r = 0.2555$) between system usage and user satisfaction, although not very strong but “medium” in its nature, perhaps sufficient to perpetuate the tantalizing search for a relationship between such clear-cut metrics. The meta-analysis results significantly contribute towards the knowledge and practice guidelines. Further, the results show that the residual standard deviation (i.e. 0.1389) exceeds the 25 per cent of effect size ($r = 0.2555$). So residual variance could not be considered as “small”. It leads to the conclusion that the differences among findings of past studies might be due to artefacts other than sampling error. It provides a direction for future research to find the moderators that might affect this relationship.

It has been suspected that published studies are biased towards showing statistically significant findings as compared to unpublished ones. Research studies with negative or null findings have hardly been published and are more difficult to find as compared to published studies. Someone may raise a question that observed results may be biased due to sampling if unpublished studies are missing in meta-analysis. To reduce such bias there is need for thorough search. The other way one may calculate F-safe N as discussed in section 5.2. The calculated “fail-safe N” for this study concludes that to reverse the conclusion drawn from this research (i.e. there exists a positive relationship between system usage and user satisfaction) would require 238 studies when $r_c = 0.05$, or 90 studies when $r_c = 0.1$ mentioning “no correlation between system usage and user satisfaction” to be included in the sample.

7. Conclusions
IS evaluation in terms of success metrics is widely held to be important in the assessment of whether the system meets the organizational objectives or not. The “system use” and “user satisfaction” as a metric of success has long been an elevated perspective, though, of course, systems are developed for many more reasons than happy users. Both of the measures (system use and user satisfaction) are well accepted and mostly used as surrogate measures of IS success. Researchers have made efforts to uncover a relationship that may exist between these two measures but the previous research has failed to reach an agreement about the nature and strength of its relationship. Both theoretical and methodological considerations have been offered as possible causes of these inconsistent findings.
The primary objective of this research was to explain and provide empirical evidence about the relationship between system usage and user satisfaction keeping in view controversial findings in the past research. Although the relationship between system use and user satisfaction was proposed in Delone and McLean's IS success model but it was not validated empirically. The primary findings show that there is little evidence, taken across a large body of research, that there exists a strong relationship between system usage and user satisfaction but inconsistent results might create confusion about whether there exists a relationship. The findings of meta-analysis lead to the conclusion that there exists a “medium” and significant positive relationship between system usage and user satisfaction (i.e. \( r = 0.2555 \)). The results of meta-analysis not only validate empirically the relationship between system use and user satisfaction as proposed in Delone and McLean’s IS success model but also contribute towards solving the existing controversies in the past research. The outcome of this research is very informative, though not necessarily as expected a priori. This is fascinating for many reasons, not least because it challenges “accepted wisdom” in many organisations today.

More specifically, while the identification of uni-dimensional metrics of success, satisfaction, etc., leads to well-formed research programmes, the widespread application but limited cohesive success of such covariant studies would perhaps be better addressed through a wider-ranging analysis using a portfolio of metrics taken together. Such an approach may well provide a richer, more relevant insight into this area of research, and forms part of the authors’ wider research agenda. As mentioned before, a direction for future research would be to explore moderator variables that may affect the magnitude of the relationship between system usage and user satisfaction. Further, the research would also be repeated by including more studies (i.e. after 2000) in meta-analysis and to compare the results with this study. Such efforts may further enhance the understanding of the nature of relationship and strengthen the existing research.

Different views such as “either system usage leads to user satisfaction or user satisfaction leads to system usage” exist in the past research. It may be argued that whilst system usage may lead to user satisfaction and user satisfaction may encourage users to engage in more and further use of the system, the evidence over decades of research suggests that such a clear-cut relationship does not exist. User satisfaction may act as positive feedback that adds towards greater use of the system. System use may act as evidence of system satisfaction, in certain circumstances. So a relationship may exist in both directions, but it seems far more productive, and more likely to derive strong research results, to consider such metrics within a multiple, or portfolio approach to this area of research.

References


Further reading


<table>
<thead>
<tr>
<th>Reference</th>
<th>Details</th>
</tr>
</thead>
</table>