….

The Importance of User Involvement in Successful Systems: A Meta-analytical Reappraisal

by Detmar W. Straub and Jonathan K. Trower

Curtis L. Carlson School of Management
Management Sciences Department
University of Minnesota
271 19th Avenue South
Minneapolis, MN 55455
(612) 625-1012
BITNET: DSTRAUB@UMNACVX
JTROWER@UMNSOM

March, 1988

[This paper is a slightly upgraded version of an MISRC (Carlson School of Management) working paper. It was converted from an old version of Word Perfect and the figures and tables did not transfer well. Sorry about that, but many people have asked for the working paper version and this is as much as we are able to provide without having the time to completely rework the formatting of the paper.]]

Copyright c Detmar W. Straub and Jonathan K. Trower

All rights reserved.

The Importance of User Involvement in Successful Systems: A Meta-analytical Reappraisal

ABSTRACT

Understanding the effect of user participation on the success of systems is critical both for MIS researchers and for practitioners. User participation in the systems development effort results in significant time and cost commitments, and mistakes in allocating these resources can seriously hamper organizational productivity.

In their evaluation of research on user involvement and system success, Ives and Olson (1984), using a tally form of meta-analysis, concluded that the benefits of user involvement had not been demonstrated. The present meta-analytical reappraisal of this research, which utilizes both tally and statistical techniques, has found, to the contrary, that **user involvement does impact the successful implementation of information systems**.

Overall, this study reinforces the value of meta-analytical approaches to interpretations of an entire research tradition, such as that conducted by Ives and Olson (1984). Among other factors which were considered in the present study, however, were statistical power, mean correlations, weighting for sample sizes, and measurement error across tests.

Search terms: User involvement in system development; user participation in System Development Life Cycle; MIS Success; system success; meta-analysis; tally techniques

~~Table of Contents~~

~~List of Tables and Figures~~

1.0. Introduction

One of the most critical issues in MIS research is the relationship of user involvement to the successful implementation of systems. If it can be established that increased involvement of users in the system development process makes a significant difference in how effectively systems are integrated into organizations, I/S managers can more easily justify the huge outlay of user time and energy required for user participation on project teams. Proof of the effectiveness of user participation, moreover, would argue that the long run benefits of increased user involvement, 1) improved communication, 2) lessened resistance to new systems, 3) decreased implementation time, and 4) increased productivity (Hirscheim, 1985), greatly outweigh short term personnel costs.

1.1. Meta-analysis of the User Involvement Thesis

Numerous MIS researchers have analyzed the role of users in the systems development process. They have identified and hypothesized behavioral and technical factors associated with successful systems (e.g., Lucas, 1975; Tait and Vessey, 1988) and have tested relationships between variables. As articulated in this literature, the **user involvement thesis** is straightforward. Acceptance of an information system by an organization and its ultimate success is believed to be ~~directly related~~ to the nature and extent of user commitment to the project, especially for systems that require heavy usage by end users.

What is known at this point in time about the strength of this underlying relationship in the user involvement thesis? Meta-analysis, or the analysis of findings across studies, has played an important role in evaluating this thesis. In their meta-analysis of the user involvement literature, Ives and Olson (1984) reviewed twenty-two studies. Their tally counted eight studies with positive results, seven with "mixed" results, and seven with negative or nonsignificant results (p. 600). They also used this tally technique to review the relationship between user involvement and various factors of system success (e.g., User Information Satisfaction [UIS]), finding that 10 studies of different system success measures were positive, 14 negative or nonsignificant, and 7 mixed (p. 597). They conclude that evidence in favor of the user involvement thesis "is not strong" (p. 599).

Such an evaluation, however, must itself be closely examined for methodological robustness. The impact of an overall non-supportive finding on the management of MIS is potentially great, and a questioning or refutation of the user involvement thesis, therefore, must be scrutinized very carefully. In practice, the belief that user participation should influence system outcomes continues in spite of this apparent ambiguity in the literature (Hirscheim, 1985; Ives and Olson, 1984). An intuitive sense that there is substance to the claim persists, empirical evidence notwithstanding.

Using a modified tally technique that: 1) employs the more appropriate test-level unit of analysis, 2) discounts tests with low statistical power, and 3) weights tests by sample size, we have found that **research in favor of the user involvement thesis actually dominates the literature**. This effect is also present, albeit to a lesser extent, when statistical meta-analysis--namely correlational means, sample size correction, and measurement correction--is performed. The present reappraisal of this literature, which includes empirical work performed since Ives and Olson's review, has found that **user involvement does, in fact, impact MIS success.**

2.0. Need for Reappraisal

When accumulating results across studies, standard voting or tally methods limit the analysis in several crucial respects. First, they do not consider the impact of low statistical power (Baroudi and Orlikowski, 1986) or sample size on the tallies. In this paper, we argue that insignificant findings with low power are inconclusive, and should not be tallied. The only results that should be tallied are significant positive findings, significant negative findings, and nonsignificant findings with high power. Therefore, a modified form of tally analysis that takes statistical power into account is more appropriate. Second, the differential effect of studies that use larger samples should also be considered in the tally analysis. This technique is desirable because simple voting schemes count small sample studies just as heavily as large sample studies, a methodology which does not fairly assess the accumulative tradition (Glass, 1978).

It is both possible and desirable to utilize sophisticated statistical meta-analysis which corrects cumulative statistics for sample size and measurement error (Hunter et al., 1982). As noted above, simple voting approaches do not consider effect of sample size nor do they assess the adverse effect of measurement error, which can greatly suppress observed correlations, and, hence, significant results. Hunter et al. (1982) suggest a technique that corrects for this kind of attenuation, and its application to the user involvement literature is now timely.

3.0. Methodology

3.1. Sample for the Modified Tally Analysis

As in Ives and Olson (1984), the sample for this study was confined to a subset of the population of all research discussing user involvement.[1] Selection for the sample was based on the following criteria:

1) the study appeared in Ives and Olson's listing (1984) or empirically examines the relationship between user involvement and system success and appeared since Ives and Olson's study (1984), and
2) the researchers did not express serious reservations about the correctness of their own findings.

The first criteria is simply a recognition of the value and importance of the Ives and Olson meta-analysis. It also permits inclusion of studies completed since the Ives and Olson review. Thus, there is a large overlap of samples in the two meta-analyses (83%--20 of 24--appear in both samples) which may allow comparison of these markedly different interpretive approaches. The second criteria was designed as a quality control to ensure that researchers' qualitative appraisals of the meaningfulness of their own findings could be recognized in the meta-analysis. Studies affected here include Lucas (1976) and Sartore (1976).

A full listing of studies included in the current meta-analysis is found later in the paper, in Table 3, "Tally of Supportive and Non-Supportive Findings" and Table 4, "Inconclusive Findings." There were 24 studies and 56 tests that satisfied the criteria for inclusion in the modified tally analysis.

3.1.1. Subset of this Sample for Statistical Meta-analysis

Statistical meta-analysis is performed across studies on statistics that can be accumulated, such as correlations, covariances, and slopes (Rosenthal, 1978; Hunter et al. 1982; Glass et al., 1981). By combining correlations reported in the research tradition into a single "mean correlation," the meta-analysis can evaluate the significance of one value rather than many, and assess the strength of the critical variable relationship. Subsetting of the sample for the present statistical meta-analysis was based on the following criteria:

1) Spearman or Pearson correlations or Kendall Taus were reported in the study, or
2) correlations could be derived from data reported in the study, or
3) sample sizes were reported from which unreported correlations could be interpolated, and
4) correlations were not duplicates of tests on the same data set.

The first criteria reflects the fact that the preponderance of statistical tests in the tradition are correlational. Correlations could be derived from data in the Alter (1978) study,[2] and interpolated in the cases of correlations that were carried out but unreported in Powers and Dickson (1973) and Gallagher (1974). These studies were included in the sample on the basis of criteria two and three. To satisfy criteria four, one test from Igersheim (1976) and one from Kaiser and Srinavasan (1980) were omitted to ensure that there were no redundant tests on the same data set.

Tests that qualified for the statistical meta-analysis are indicated by an "*" in the complete listing of the meta-analysis sample (Tables 3 and 4). There were 30 tests in this subsample.[3]

3.2. Level of Analysis

Level of the analysis may be thought of as the level of generality or granularity from which the researcher has chosen to view the phenomenon (Blalock, 1969). As Cook and Campbell (1979) point out, researchers view the phenomenon from a high, conceptual or "molar" level to a low, concrete or "micromediation[al]" level (pp. 27, 62-63).[4] In testing the user involvement hypothesis, researchers typically do not vary the level of analysis for the independent variable, user involvement, as much as for the dependent variable, system success. Figure 1a shows a broad conceptualization of research constructs ranging from the molar to the micromediational levels of analysis.

Researchers have studied the system success construct at all levels of analysis. At the micromediational level, researchers use measures as surrogates for system success or assumed components of system success. Figure 1b illustrates a micromediational level of analysis such as that used by Swanson (1974) in studying user appreciation of the MIS/360 system. Other studies that approach the system success construct at the micromediational level include Nolan and Seward (1974), Debons et al. (1978), and Neumann and Segev (1980).

At the intermediate level, researchers factor measures into components of a higher level construct. Figure 1c illustrates factors derived for UIS by researchers such as Ives, Olson, and Baroudi (1983). Bailey and Pearson (1983) also analyze the system success construct at the intermediate level.

Finally, researchers at the molar level investigate interrelationships between factors of system success. Figure 1d represents interrelationships tested in works like Baroudi, Ives, and Olson (1986). Ives and Olson's meta-analysis (1984) is another clear example of research at the molar level of analysis.

1a 1b

1c 1d

Figure 1. Levels of Analysis in System Success Measures

The present study also utilizes the molar level of analysis. Since the goal of meta-analysis is to evaluate the results of the entire research tradition across variations in method, in construct operationalization, and in level of analysis, analysis at the molar level is appropriate (Glass et al., 1981). This is not to say that meta-analysis at an intermediate level, like UIS, would not also be valuable. Unfortunately, though, the statistical power is too low at this time to warranty insignificant results at any level but the molar level.[5] UIS meta-analysis, therefore, cannot be performed until new UIS studies have been undertaken.

Molar level analysis is frequently used in meta-analysis of relationships between systemic variables and performance. McEvoy and Cascio (1987) and Wagner and Gooding (1987) are recent examples of meta-analysis executed at the molar level.

3.3. Unit of Analysis

Ives and Olson's tally technique (1984) makes individual assessments of each factor variable set (e.g., UIS), making the factor-level result one of their basic units of analysis. But factor-level analyses, used also in meta-analyses like Locke and Schweiger (1979), may not be able to categorize findings clearly. Besides categories of: a) significant positive and b) insignificant or negative significant, for instance, Ives and Olson (1984), used a c) "mixed" category to represent contradictory test results among related factors. This category only succeeds in obfuscating results because it neither unambiguously confirms nor disconfirms the user involvement thesis.

This meta-analysis, therefore, employs only the test or judgment as the unit of analysis, as suggested by Hunter et al. (1982) and Glass et al. (1981) and as employed in meta-analyses such as those by Wagner and Gooding (1987) and Smith and White (1988). This approach has several advantages over factor-level analysis. First, analysis at the test level eliminates artifactual categories like "mixed." Researchers test for support or nonsupport, at a given alpha protection level, or, in case and other non-quantitative empirical work, judge the results to demonstrate support or not (Cf. Ives and Olson [1984], p. 597). For this study, therefore, supportive findings will simply be tests which are positive significant at a given alpha. Nonsupportive findings, on the other hand, are:

1) findings in the right direction (i.e., r > 0), that are not significant at a given alpha, but do have high statistical power, 2) negative findings that are significant, and 3) negative findings that are insignificant with high statistical power.[6] The different coding schemes used by Ives and Olson and by the present study are shown in Table 1.

Meta-analysis Categories	Ives & Olson	Present Study
Positive significant	X	X
Insignificant (High power)	X	X
Negative significant	X	X
Mixed	X	
Inconclusive (Low power)		X

Table 1. Coding Schemes for Meta-analytical Studies

3.4. Modified Tally Technique

3.4.1. Tally Excluding Inconclusive Results

The argument in this paper is that a careful appraisal of insignificant results can contribute to a more accurate tallying of the results in the literature. When tallies are utilized in assessing the overall meaning of a research tradition, it is important that the statistical power of insignificant study tests be considered. Statistical power is the probability that the null hypothesis has been correctly rejected, as shown in Table 2.

Proper rejection is closely associated with sample size so that tests with larger sample sizes are less likely to reject the null hypothesis improperly (Cohen, 1969; Baroudi and Orlikowski, 1986; Kraemer and Thiemann, 1987). It is also statistically related to alpha, the standard error, or reliability of the sample results, and the effect size, or degree to which the phenomenon has practical significance (Cohen, 1969). Nonsignificant results from tests with low power, i.e., a probability of less than 80% that the null hypothesis has been correctly rejected (Cohen, 1969), are ~~inconclusive~~ and do not indicate that the effect is truly not present. By applying the most conservative standards to each study, we can assume that a hypothesis test would not miss a large effect size (Cohen, 1977); initial analysis is based, therefore, on a large effect size.

		Decision	
		Reject H_0	**Fail to Reject**
True State of Nature	**H_0 true**	**TYPE I ERROR (á)**	**CORRECT**
	H_0 false	**CORRECT (1 - ß = POWER)**	**TYPE II ERROR (ß)**

Table 2. Relationship of Power to Decision on Hypothesis Tests

In excluding tests with power of less than .80 from the tally, the comparison between supportive and nonsupportive findings becomes more meaningful. In the present meta-analysis, tests will be separated into those that suffered from low power, and are therefore inconclusive, and those that were either positive significant, nonsignificant with high power, or significant but negative.[7]

3.4.2. Tally Weighted by Sample Size

One conceptual problem remains with regard to the tally technique, even after inconclusive tests have been excluded. To gauge the overall effect of the tradition, studies with larger sample sizes should be counted more heavily than those with smaller sample sizes. In the present study, the calculation is made by weighting each test by its sample size. This method, in effect, gauges the magnitude of the effect for or against the user involvement thesis.

Whereas this procedure makes statistical sense, it may not make sense in the case where a small-sample, carefully done study is compared to a large- sample, methodologically flawed study. Meta-analysis in no way dismisses the researcher from being responsible for a meticulous design and execution (Hunter et al., 1982), but qualitative differences in the research are extremely difficult to assess. Therefore, ~~ceteris paribus~~, methodological error may be assumed to randomize over studies in heterogenous settings and differing measures of the same constructs (Glass et al., 1981; Cook and Campbell, 1979).

3.5. Statistical Meta-analysis

The accumulation of results across studies can also be achieved by means of a statistical meta-analysis. Once the correlational mean of all tests has been corrected for sampling and measurement error, the corrected mean correlation that results will be the best estimate of the true correlation between user involvement and systems success. The simple mean is first corrected for variations in sample size by test, and then corrected for measurement error.

3.5.1. Simple Mean of Correlations

A simple mean of study correlations gives an initial assessment of the overall strength of the relationship between the hypothesized independent variables and dependent variables. This measure of central tendency improves on the tally technique by allowing the relationship to be expressed as a single value. As Glass (1978) points out, however, it is overly simplistic because it does not correct for varying sample size among studies. As we shall see, there are additional corrections that need to be made to produce a more accurate estimate of the true mean.

3.5.2. Weighting Correlations by Sample Size

Correlations can be accumulated across studies by weighting each correlation according to its sample size, as expressed in Equation 1 (based on Hunter et al., 1982), below:

$$\bar{r}_s = \frac{\Sigma\,[n_i * r_i]}{\Sigma\, n_i} \qquad \text{(Equation 1)}$$

where: $\bar{r}_s$ **= mean correlation corrected for sample size**
n_i **= size of the sample in test i**
r_i **= correlation in test i**

The mean correlation results from a weighting of each correlation by sample size; it is, in effect, a mean correlation with sampling variation removed.

3.5.3. Correcting Correlations for Attenuation

It is well known that measurement error can attenuate, or lower, correlations in a systematic fashion (Hunter et al., 1982). The formula for correction of attenuation is shown below as Equation 2 (based on Hunter, et al., 1982):

$$\bar{r}_c = \frac{\bar{r}_s}{\frac{\Sigma \sqrt{r_{xx}}}{n_a} \frac{\Sigma \sqrt{r_{yy}}}{n_b}} \qquad \textbf{(Equation 2)}$$

where: $\bar{r}_c$ = mean correlation corrected for attenuation
$\bar{r}_s$ = mean correlation corrected for sample size
r_{xx} = reliability of independent variables
r_{yy} = reliability of dependent variables
n_a = number of tests reporting r_{xx}
n_b = number of tests reporting r_{yy}

Because not all of the studies gauge validity or reliability of measures, it is necessary to estimate distributions of measurement error in the dependent variables from studies that do report them. Table 7 later in the paper gives the reliabilities reported for tests in our sample. The subsequent measurement error adjustment to the mean correlation removes the effects of measurement error, and produces the final best estimate of the mean.

3.6. Testing Procedure

Testing the mean correlation of studies follows Hunter et al.'s (1982) procedure. Since it is our contention here that theory and prior work actually favor the user involvement thesis, the mean correlation is expected to be greater than zero. The level of confidence used to construct confidence intervals to test this proposition has been set at 95%; confidence intervals at 90% will also be reported.

4.0. Data Analysis

4.1. Modified Tally Analysis

4.1.1. Tally of Supportive versus Nonsupportive Tests, Excluding Inconclusive, Low Power Tests

Ives and Olson (1984) list 10 studies reporting significant results versus 14 that report nonsignificant or significant negative results. The ratio between these groupings is .71 : 1. Overall, only 42% of the tests were supportive of the user involvement thesis. But, as noted above, this meta-analysis does not take into account the fact that many of these nonsignificant findings are simply inconclusive because of low power.

In the current analysis, tests that are subject to low power (.80 or less with a large effect size) have been removed from the tally. These tests, and their related power, are reported in Table 4, "Inconclusive Tests."

It is important to note that individual tests have been categorized as significant if they met a protection level criterion of .10. Since many of the authors chose this level themselves, it seemed best to be consistent and apply this same standard to all studies. It should be noted, however, that this cutoff gives a greater benefit of doubt to the test.

From Table 5, it is clear that the ratio of supportive versus non-supportive has changed markedly from the 1984 Ives-Olson tally. The effect of low power, nonsignificant findings has been removed and the more recent studies added, and the corrected ratio of supportive to nonsupportive findings is now 1.8 : 1, or 64% supportive results. This analysis alone suggests that the tradition has more underlying agreement than previously thought.

Author	N	Test	Statistic	Sig. p-value	Neg./Non-Sig. p-value
* 1. Alter (1978)	55	Spearman r	.210	<.050	
* 2. _____	55	Spearman r	.253	<.050	
3. Baronas & Louis (1988)	92	F-test	5.51	<.01	
* 4. Baroudi et al.(1986)	200	Correlation	.180	.025	
* 5. ________________	200	Correlation	.280	.010	
6. Boland (1978)	10	T-test	2.400	.027	
7. Cronan & Means (1984)	60	Chi-square	4.580	.099	
* 8. Edstrom (1977)	13	Spearman r	.510	.050	
* 9. Franz (1979)	148	Correlation	.240	.010	
* 10. _____	148	Correlation	.320	.005	

* 11. _____	148	Correlation	.310	.005
12. Fuerst (1979)	-	Conditional prob. & F	---	>.100
* 13. Gallagher (1974)	74	Correlation	<.195	<.100
14. Guthrie (1972)	1991	Kruskal-Wallis	4.120	>.100
* 15. Igersheim (1976)	49	Pearson r	.440	.001
* 16. _________	54	Pearson r	.360	.005
* 17. _________	30	Pearson r	.390	.025
* 18. _________	41	Pearson r	.340	.025
* 19. _________	51	Pearson r	.550	.001
* 20. Joshi (1986)	226	Correlation	.516	<.001
* 21. Kaiser & Srin.(1980)	29	Pearson r	.506	.005
* 22. ____________________	16	Pearson r	.927	.005
23. King & Rodriguez (1978, 1981)	45	Mann-Whitney U Test	---	.092
24. ____________________	45	Mann-Whitney U Test	---	.467
* 25. Lucas (1975)	616	Pearson r	.150	.001
* 26. _____	683	Pearson r	.220	.001
* 27. Maish (1979)	56	Kendall's Tau	.279	.010
* 28. Olson & Ives (1981)	83	Spearman r	-.030	>.100
* 29. ______________	83	Spearman r	-.070	>.100
* 30. ______________	23	Spearman r	.010	>.100
* 31. Powers & Dickson(1973)	20	Kendall's Tau	<.377	<.100
* 32. ________________	20	Kendall's Tau	<.377	<.100

Table 3. Tally of Supportive and Non-Supportive Findings

Legend: "---" indicates that the test statistic was not reported
"*" indicates studies which were included in the statistical meta-analysis

~~Author~~	~~N~~	~~Test~~	~~Statistic~~	Sig. ~~p-value~~	Neg./ Non-Sig. ~~p-value~~
33. Robey & Farrow (1979)	63	Between-Phase correlations	---	.060	
34. ________	63	Path coefficient	-.270		<.100
35. ________	63	Path coefficient	-.210		<.100
36. ________	63	Path coefficient	-.380		<.100
37. ________	63	Path coefficient	.440	<.100	
38. ________	63	Path coefficient	.440	<.100	
39. ________	63	Path coefficient	.460	<.100	
40. Schewe (1976)	41	Beta	---		>.100
41. ______	38	Beta	---		>.100
42. ______	41	Beta	---		>.100
43. ______	38	Beta	---		>.100
* 44. Spence (1978)	23	Spearman r	-.377		.038
* 45. ______	65	Spearman r	.205	.051	
46. ______	65	Chi-square	2.239		.135
47. Swanson (1974)	37	Chi-square	7.890	.005	
48. _______	37	Chi-square	9.790	.002	
49. Tait & Vessey (1988)	84	Beta	.195		>.050
50. Thurston (1959)	--	Case studies	---	+	
51. Vanlommel & Debrabander (1975)	20	Contingency Coefficients	.110		>.050
52. _________	20	Contingency Coefficients	.100		>.050
53. _________	20	Contingency Coefficients	.150		>.050
54. _________	20	Contingency Coefficients	.240	<.100	
55. _________	20	Contingency Coefficients	.470	<.050	
56. _________	20	Contingency Coefficients	.04		>.050

Total supportive results.............................. 36
Total nonsupportive results.................................... 20

Ratio = 1.8 : 1 or 64% supportive findings

Table 3. Tally of Supportive and Non-Supportive Findings (continued)

Legend: "+" indicates that the author judged the relationship to be significant
"---" indicates that the test statistic was not reported
"*" indicates studies which were included in the statistical meta-analysis

4.1.2. Weighted Tally Analysis

A weighted tally analysis was performed on the data in Table 3, excluding the large sample Guthrie study. Guthrie's study was excluded because his conclusions varied so markedly from his data analysis and it was difficult to know how to classify the results. Weighting tests of user involvement results in 3570 cases that are supportive versus 833 cases that are non-supportive. This is a ratio of 4.3 : 1, or 81% supportive findings.[8] Both modified and weighted tally techniques favor the user involvement thesis. The weighted analysis strongly favors the thesis, indicating that, according to the weighted research results, four out of five projects are successful when users are involved in the development process. The results of this test and other tally tests are shown in Table 5, "Summary of Tally Analyses."

Author	Test	Statistic	N	Power
1. Boland (1978)	T-test	1.270	10	.29
2. Spence (1978)	Chi-square	.114	13	< .20
3. ______	Chi-square	.000	12	< .20
* 4. ______	Spearman r	-.156	15	.30
5. ______	Chi-square	.256	23	.48
6. ______	Chi-square	.045	15	< .20
* 7. ______	Spearman r	-.157	13	.27
* 8. ______	Spearman r	.143	12	.25

Table 4. Inconclusive Tests

4.2. Findings of the Statistical Meta-analysis

4.2.1. Simple Mean Correlational Analysis

The mean correlation of the 30 tests in our sample of tests is .250 with a standard deviation of .151. The 95% confidence interval is -.047 < rho < .547, indicating a wide range of possible values, including rho = 0.[9] Key statistics for the simple mean of the correlations are shown in Table 6, "Statistics for Mean Correlational Techniques."

		Tally Ratios*	Percentage of Supportive Findings
	Ives and Olson	**.71 : 1**	**42%**
Present Study	**High power tally**	**1.8 : 1**	**64%**
	Weighted tally	**4.3 : 1**	**81%**

Table 5. Summary of Tally Analyses

***These are tally ratios of supportive to nonsupportive findings**

4.2.2. Weighted Sample Size Analysis

The mean correlation of these studies, weighted for sample size, was .239 as summarized in Table 6. The effect of sampling error on the overall variance in the correlations is considerable. It accounts for 36% of the observed variance in the correlations. If the corrected value were the true correlation, we would not be justified in concluding that there is a significant relationship between these constructs, since the 95% confidence interval is -.057 < rho < .535. This confidence interval suggests that there is still a considerable range of possible values the true value could take, including zero.

~~Method~~ ~~Statistic~~	Simple Mean Method	Mean Corrected for Sample Size Method	Mean Corrected for Attenuation Method
Mean	**.250**	**.239**	**.282**
Standard Deviation	**.151**	**.121**	**.142**
95% Confidence Interval	**-.047<rho<.547**	**-.057<rho<.535**	**.004<rho<.560**
90% Confidence Interval	**.002<rho<.498**	**.040<rho<.438**	**.048<rho<.516**
Additional Variance Explained	**--**	**36%**	**2%**

Table 6. Statistics for Mean Correlational Techniques

4.2.3. Correction for Attenuation Analysis

Measurement error has an impact on the accumulative correlation across tests in the user involvement literature. Utilizing the reliabilities reported in the literature, as shown in Table 7, "Reported Reliabilities in Studies", the correlation corrected for attenuation is .282, with reliability error accounting for 2% of the variance in correlations across studies (cf. Table 6). With a correlation of .282, the 95% confidence interval is .004 < rho < .564, showing that the true score could vary within quite a large range, all of which, however, are greater than zero.

Author	r_{xx}	r_{yy}
1. Baroudi, Olson, & Ives (1983)	.900	.970
		.740
2. Igersheim (1976)		.878
3. Joshi (1986)	.802	.908
4. Olson and Ives (1983)	.740	.870
	.850	
5. Tait and Vessey (1988)	.970	.972

Table 7. Reported Reliabilities in Studies

4.2.4. Consideration of Reporting Error

In all likelihood, the true correlation across studies is even higher than calculated here. Attenuation for measurement error was slight, amounting to only 2% of the total variance. This is because published reliabilities in this literature were high (average for r_{xx} = .85, for r_{yy} = .89) and their variability from study to study low. It is very possible that low reliabilities do not appear in the literature because of the minimum acceptance levels formally or informally set by journals. Consider the effect of only 3 additional, hypothetical studies averaging .75 (.70,.75, and .80) for r_{xx} and r_{yy}. The correlation for user involvement and system success becomes approximately .300 rather than .282. And, given that many research instruments in this literature utilize only one measure per construct, it is not unreasonable to suppose that at least one half of the reliabilities would fall near the .75 level.[10]

4.2.5. Chi-square Test of Correlational Significance

The simple mean correlation can be submitted to a Chi-square test for deviations from the expected distribution (Hunter et al., 1982). In this case, the correlation of .250 across 30 tests (d.f. = 29) and a total sample size of 3249 yields a Chi-square statistic of 81.19, which is highly significant
(p-value < .001).

This significant result, however, may not necessarily indicate that departures from the mean have practical significance. Like all Chi-square tests, relatively trivial deviations across studies or large deviations for certain outliers can generate large Chi-squares (Hunter et al., 1982). Nevertheless, the test does reinforce an emerging pattern of support for the user involvement thesis.

5.0. Discussion

Although the final corrected correlation of user involvement and system success across studies (.282) does not indicate a strong relationship, it is significantly different from zero in a positive direction and does show that some relationship in all likelihood exists. The modified tally results, on the other hand, strongly support the underlying significance of this relationship. The fact that the relationship between user involvement and system success was supported across studies justifies the belief that users should be actively engaged in systems projects during critical phases of the project life cycle. The fact that in the statistical meta-analysis the relationship is not ~~strong~~ as in a correlation of .75 to 1.0 or ~~moderate~~ as in a correlation of .5 to .74 suggests that there are other factors that can be elicited to explain the successful implementation of systems, as argued by Ives and Olson (1984). These exogenous factors call for further research to determine the magnitude of alternative effects and to develop a more satisfactory explanatory model for system success.

6.0. Directions for Further Research

Further studies on the effect of user participation in systems development teams and in the implementation effort needs to examine other factors that impact success (Swanson, 1988). There may be intervening and moderator variables that affect this process, as Ives and Olson (1984) have argued. For example, Baroudi, Olson, and Ives (1986) have found that UIS may moderate the relationship between user involvement and system usage, but that there is probably not a reciprocal causal relationship between UIS and system usage. Tait and Vessey (1988) have found that system resource constraints, such as limitations in the technology available, influence how new systems are received, and additional studies need to confirm or disconfirm this result.

New dimensions and perspectives on the variable relationships may also be explored. It is possible and perhaps even likely that the relationship between user involvement and successful systems implementation actually varies by organization, project, and environment (Naumann, Jenkins, and Wetherbe, 1984). If this is accurate, support for the thesis offered in this paper may simply represent an average response across settings. Variability between studies may not reflect variability in methodologies and measurements so much as underlying variance in unmeasured or exogenous variables. Markus (1983), Keen and Gerson (1977), and others have argued, for example, that systems development in a highly charged political environment will be more successful when **user involvement is held to a minimum or is tightly controlled**. Some systems development efforts, on the other hand, may be heavily dependent on user commitment and participation. For example, Ives and Olson (1984) argue that decision support systems, by their very nature, require high levels of user manager participation.

Other intervening and moderator variables uncovered in empirical work include: 1) system resource constraints, such as limitations in the technology available (Tait and Vessey, 1988), 2) project size, 3) project structuredness, 4) stability of projects, 5) proficiency of users, and 6) proficiency of analysts (Naumann, Jenkins, and Wetherbe, 1984).

There is no question but that rigorous methodological standards will raise the quality of research in this area, as Ives and Olson (1984) have asserted. Because there are such a large number of hypothetical causative factors that, unmeasured, could confound results, research must be carefully designed to reduce variation and isolate effects. To localize strong, weak, non-existent, or negative effects of user involvement on systems success, researchers need to strictly control setting. As Figure 2 shows, projects may be selected which are similar on most dimensions except level of user involvement. Generalization will be limited by such designs, but the research stream will be advanced by researchers discovering the differential effects of setting.

From the standpoint of cumulating findings in the research stream, MIS researchers should provide zero order correlations between key variables to aid in meta-analysis. When the data is ordinal, a Spearman correlation or a Kendall's Tau is appropriate; otherwise, a Pearson correlation is required. These statistics may be supplemental to the chosen statistical technique for the study.

Encouraged by enlightened journal policies, researchers should report reliabilities for instruments even when they do not indicate highly reliable measurement. Journal editors also need to sponsor the reporting of insignificant findings, giving preference, however, to studies with power above the .80 level.

Figure 2. Testing the User Involvement Thesis in Controlled Settings

7.0. Conclusion

Understanding the effect of user participation on the success of systems is critical both for MIS researchers and for practitioners. User participation in the systems development effort results in large time and cost commitments, and mistakes in allocating these resources can seriously hamper organizational productivity.

In their evaluation of studies in the user involvement research, Ives and Olson (1984), using a tally form of meta-analysis, concluded that the benefits of user involvement had not been convincingly demonstrated. Our meta-analytical reappraisal of this literature has found that **user involvement is a factor that must be considered in explaining the success of information systems**.

Overall, this study reinforces the value of meta-analytical approaches to interpretations of an entire research tradition, such as that conducted by Ives and Olson (1984). This form of analysis is clearer, however, when statistical power, mean correlations, sample sizes, and measurement error across tests are taken into account.

References

Alter, Steven (1978). "Development Patterns for Decision Support Systems," MIS Quarterly, Vol. 2, No. 3.

Baronas, Ann-Marie K. and Meryl Reis Louis (1988). "Restoring a Sense of Control During Implementation: How User Involvement Leads to System Acceptance," MIS Quarterly, Vol. 12, No. 1 (March), 111-124.

Baroudi, Jack J., Margrethe H. Olson, and Blake Ives (1986). "An Empirical Study of the Impact of User Involvement on System Usage and Information Satisfaction," Communications of the ACM, Vol. 29, No. 3 (March), 232-238.

Baroudi, Jack J., and Wanda J. Orlikowski (1986). "Misinformation in MIS Research: The Problem of Statistical Power," Center for Research on Information Systems (New York University) Working Paper CRIS #125, GBA #86-62.

Bailey, James E. and Sammy W. Pearson (1983). "Development of a Tool for Measuring and Analyzing Computer User Satisfaction," Management Science, Vol. 29, No. 5 (May), 530-545.

Blalock, Hubert M. Jr. (1969). Theory Construction: From Verbal to Mathematical Formulation. Prentice-Hall: Englewood Cliffs, NJ.

Boland, R.J. (1978). "The Process and Product of System Design," Management Science, Vol. 24, No. 9, 887-898.

Cohen, Jacob. (1969). Statistical Power Analysis for the Behavioral Sciences. Academic Press: New York.

Cohen, Jacob. (1977). Statistical Power Analysis for the Behavioral Sciences. Revised Edition. Academic Press: New York.

Cook, Thomas D., and Donald T. Campbell (1979). Quasi-Experimentation: Design & Analysis Issues for Field Settings. Houghton Mifflin Company: Boston.

Cronan, Thomas P. and Thomas L. Means (1984). "System Development: An Empirical Study of User Communication," Database (Spring), 25-27.

Culnan, Mary J. (1983). "Chauffeured Versus End User Access to Commercial Databases: the Effects of Task and Individual Differences," MIS Quarterly, Vol. 7, No. 1 (March), 55-67.

Debons, A., W. Ramage, and J. Orien (1978). "Effectiveness Model of Productivity" in Research on Productivity Measurement Systems for Administrative Services, eds. L. Hanes and C.H. Kriebel, Vol. 2 (July), NSF Grant APR-20546.

Edstrom, A. (1977). "User Influence and the Success of MIS Projects," Human Relations, Vol. 30, 589-606.

Franz, C.R. (1979). "Contingency Factors Affecting the User Involvement Role in the

Design of Successful Information Systems," Ph.D. Dissertation, University of Nebraska.

Fuerst, W.L. (1979). "Characteristics Affecting DSS Usage," National AIDS Conference Proceedings, 172ff.

Gallagher, C.A. (1974). "Perceptions of the Value of a Management Information System," Academy of Management Journal, Vol. 17, No. 1, 46-55.

Glass, Gene V. (1978). "Integrating Findings: The Meta-analysis of Research" in Review of Research in Education, Volume V, ed. L. Shulman. Ithasca, IL: Peacock.

Glass, Eugene, Barry McGaw, and Mary Lee Smith (1981). Meta-analysis in Social Research. Beverly Hills, CA: Sage.

Guthrie, A. (1972). A Survey of Canadian Middle Managers' Attitudes Toward Management Information Systems. Carleton University Press: Ottawa, Ontario.

Guzzo, Richard A., Susan E. Jackson, and Raymond A. Katzell (1987). "Meta-Analysis Analysis," ***Research in Organizational Behavior***, Vol. 9, 407-442.

Hirscheim, R.A. (1985). "User Experience with and Assessment of Participative System Design," MIS Quarterly, Vol. 9, No. 4 (December), 295-305.

Hunter, John E., and Frank L. Schmidt and Gregg B. Jackson. (1982). MetaAnalysis: Cumulating Research Findings Across Studies. Beverly Hills, CA: Sage.

Igersheim, R.H. (1976). "Management Response to an Information System," AFIPS Conference Proceedings, 877-882.

Ives, Blake and M. Olson (1984). "User Involvement and MIS Success: A Review of Research", Management Science, Vol. 30, No. 5, 586-603.

Ives, Blake, Margrethe H. Olson, and Jack J. Baroudi (1983). "The Measurement of User Information Satisfaction," Communications of the CACM, Vol. 26, No. 10 (October), 785-793.

Joshi, Kailash (1986)."An Investigation of Equity and Role Variables as Determinants of User Information Satisfaction," Ph.D. dissertation, Indiana University.

Kaiser, K.M. and A. Srinivasan (1980). "The Relationship of User Attitudes Toward Design Criteria and Information Systems Success," National AIDS Conference Proceedings, 201-203.

Keen, Peter G.W. and Elihu M. Gerson (1977). "The Politics of Software System Design," Datamation (November), 80-84.

King, W.R. and J.I. Rodriquez (1978). "Evaluating MIS," MIS Quarterly, Vol. 2, No. 3, 43-52.

King, W.R. and J.I. Rodriquez (1981). "Participative Design of Strategic Decision Support Systems: An Empirical Assessment," Management Science, Vol. 27, No. 6, 717-726.

Kraemer, Helena Chmura and Sue Thiemann (1987). How Many Subjects? Statistical Power Analysis in Research. Newbury Park, CA: Sage.

Locke, E.A. and D.M. Schweiger (1979). "Participation in Decision-Making: One More Look," Research in Organizational Behavior, Vol. 1, 265-339.

Lucas, H.C., Jr. (1975). Why Information Systems Fail. Columbia University Press: New York.

Lucas, H.C., Jr. (1976). The Implementation of Computer-Based Models. National Association of Accountants: New York.

Maish, A.M. (1979). "A User's Behavior Toward His MIS," MIS Quarterly, Vol. 3, No. 1, 39-52.

Markus, M. Lynne (1983). "Power, Politics, and MIS Implementation," Communications of the ACM, Vol. 26, No. 6 (June), 430-444.

McEvoy, Glenn M., and Wayne F. Cascio (1987). "Do Good or Performers Leave? A Meta-Analysis of the Relationship Between Performance and Turnover," Academy of Management Journal, Vol. 30, No. 4, 744-762.

Naumann, Justus D., A. Milton Jenkins, and James C. Wetherbe (1984). "An Empirical Investigation of Systems Development Practices and Results," Information and Management, Vol. 7, No. 2 (April).

Neumann, S. and E. Segev (1980). "Evaluate your Information System," Journal of Systems Management, Vol. 31 (March).

Nolan, R. and H. Seward (1974). "Measuring User Satisfaction to Evaluate Information Systems" in Managing the Data Resource Function, ed. R.L. Nolan. West Publishing: Los Angeles, CA.

Olson, M.H., and B. Ives (1981). "User Involvement in System Design: An Empirical Test of Alternative Approaches," Information and Management, Vol. 4, No. 4, 183-196.

Powers, R.F. and G.W. Dickson (1973). "MIS Project Management: Myths, Opinions, and Reality," California Management Review, Vol. 15, No. 3, 147-156.

Rivard, Suzanne and Sid L. Huff (1988). "Factors of Success for End-User Computing," Communications of the ACM, Vol. 31, No. 5 (May), 552-561.

Robey, D. and D. Farrow (1979). "Information Systems Development: Some Dynamics of User Involvement," National AIDS Conference Proceedings (November), 149-151.

Rosenthal, Robert (1978). "Combining Results of Independent Studies," Psychological Bulletin, Vol. 85, No. 1, 185-193.

Rosenthal, Robert (198?). Book on metaanalysis

Sartore, A. (1976). "Implementing a Management Information System: The Relationship of Participation, Knowledge, Performance, and Satisfaction in an Academic Environment," Ph.D. dissertation, University of California at Irvine.

Schewe, C.D. (1976). "The MIS User: An Exploratory Behavioral Analysis," Academy of Management Journal, Vol. 19, No.4 (December), 577-590.

Smith, Mark and Michael C. White (1987). "Strategy, CEO Specialization, and Succession," Administrative Science Quarterly, Vol. 32 (June), 263-280.

Spence, J.O. (1978). "A Case Study Analysis of Organizational Effectiveness Between User-Managers and Information Service Department Personnel," Ph.D. dissertation, Texas Tech University.

Swanson, E.B. (1974). "Management Information Systems: Appreciation and Involvement," Management Science, Vol. 21, 178-188.

Swanson, E. Burton (1988). Information System Implementation. Irwin: Homewood, IL.

Tait, Peter and Iris Vessey (1988). "The Effect of User Involvement on System Success: A Contingency Approach," MIS Quarterly, Vol. 12, No. 1 (March), 91-110.

Thurston, P.H. (1959). Systems and Procedures Responsibility. Harvard University Press: Cambridge.

Vanlommel, E. and B. DeBradander (1975). "The Organization of Electronic Data Processing," Journal of Business, Vol. 48, No. 2, 391-410.

Wagner, John A., III, and Richard Z. Gooding (1987). "Effects of Societal Trends on Participation Research," Administrative Science Quarterly, Vol. 32, 241-262.

Wolf, Frederic M. (1986). ***Meta-Analysis***. Beverly Hills, CA: Sage.

Endnotes

1.The effect of omitted samples is small when all of the variation across studies is due to sampling error (Hunter, et al., 1982, p. 29). As will be shown later, almost all of the variation in our sample was due to sampling error.

2.The sample size of 55 reflects the total of the reported cells. Since this total does not match the column, row, or grand totals--all of which indicate a sample of 56--it is clear that one of the cell values is missing in Alter's published results.

3.In the meta-analysis, we proceeded by assuming that the subsample is not systematically biased, in accord with the usual convention (Hunter, et. al, 1982). Statistical meta-analysis requires a common test statistic, in this case correlations. Of the 30 correlational tests in the user involvement literature, 23 were supportive and 7 non-supportive, meaning that non-correlational tests were composed of 13 supportive and 21 non-supportive findings. Supportive and non-supportive findings, therefore, were not equally represented in the statistical meta-analysis. A Chi-square analysis indicates that the two subsamples are independent (Chi-square=9.57, p-value=.002). From the standpoint of unweighted tally, supportive findings, in short, have been sampled disproportionately. There is a possibility, however, that ~~magnitude of effect~~ was not equally represented, and magnitude of effect could weight findings in either direction. There may also be systematic bias from whether researchers choose to utilize correlations or not. In sum, statistical meta-analysis may be drawing tests from the population in a systematic, but unobservable and immeasurable fashion. These uncertainties cannot be resolved, though, because we do not have a single test statistic from the population of all studies that have been conducted.

4. Blalock (1969) speaks of these as "levels of generality." The intermediate level here is equivalent to Blalock's "middle range" (p. 141).

5.The power for examining UIS as dependent variable is .73 (n=20). For the other success factors considered together, the power is only .46 (n=10).

6.Insignificant, negative, and high power results are another possibility, but researchers have not yet posited a hypothesis of negative results, so this has yet to occur in the literature.

7.No tests in the literature proved to be insignificant in the negative direction but with high power.

8.Although Guthrie concludes that user involvement is important for successful implementations, he says he cannot demonstrate the relationship empirically. When the Guthrie data is included in the non-supportive column, the weighted ratio drops to 1.3 : 1 (3575 versus 2824), or 56% supportive findings. Even in this circumstance, the evidence still favors the user involvement thesis. When the Guthrie study is included as a supportive test, the weighed ratio rises to 6.7 : 1 (5561 versus 833), or 87% supportive! It should be noted that Ives and Olson categorize this study as supportive. In the modified tally analysis, we classify it as non-supportive to give the benefit of doubt to the non-supportive position.

9.The data was tested for normality with the Shapiro-Wilk test. With a test statistic of .964 (p-value = .468), we concluded that the population of values was probably normally distributed. Therefore, it was fitting to use the appropriate z-values in the construction of the confidence intervals.

10.Examples of Cronbach alphas averaging below .75 that, unfortunately, are not ordinarily reported in the journals are found in Rivard and Huff's study of User Information Satisfaction in the UDA (User Developed

Application) environment. They report Cronbach alphas of .873, .628, .739, .620, .728, .630, and .803, ~~averaging .717~~. Culnan's (1983) study of end users and commercial database use also reports lower Cronbach's alphas, viz. values of .54, .79, .65, .75, .75, and .50; ~~the average of these is .66~~.