Requirements Eng (1996) 1: 4–26

Requirements Engineering

User Participation in the Requirements Engineering Process: An Empirical Study

Khaled El Emam[a], Soizic Quintin[b] and Nazim H. Madhavji[a,c]
[a]Centre de Recherche Informatique de Montreal, Canada; [b]Genie Informatique, Université de Technologie de Compiegne, France; and [c]School of Computer Science, McGill University, Montreal, Canada

In the development of information systems, user participation in the requirements engineering (RE) process is hypothesised to be necessary for RE success. In this paper we develop a theoretical model which predicts that the interaction between user participation in the RE process and uncertainty has an impact on RE success. This theory is empirically tested using survey data. We develop instruments to measure user participation and uncertainty. An existing instrument for measuring RE success was used. This instrument covers two dimensions of RE success: (a) the quality of RE service, and (b) the quality of RE products. The results indicate that as uncertainty increases, greater user participation alleviates the negative influence of uncertainty on the quality of RE service, and that as uncertainty decreases, the beneficial effects on the quality of RE service of increasing user participation diminish. Furthermore, we did not find that the interaction between user participation and uncertainty had an impact on the quality of RE products. Based on these results, we make recommendations for managing user participation in the RE process, and provide directions for future research.

Keywords: User participation; Contingency theory; Empirical study; Uncertainty; Requirements engineering success

This work has been supported, in part, by the IT Macroscope Project and NSERC Canada.

Correspondence and offprint requests to: Khaled El Emam, Centre de Recherche Informatique de Montreal (CRIM), 1801 McGill College Avenue, Suite 800, Montreal, Quebec, Canada H3A 2N4.

1. Introduction

The development and implementation of information systems (IS) is a form of organisational change [1–5]. It can even be argued that change is a desirable consequence of IS implementation; for example, one study found a positive relationship between perceived success of IS projects and the amount of change IS brings about [6]. A requirement for overcoming resistance to change is high user participation [7]. Furthermore, user participation and influence are expected to increase the likelihood of user acceptance of the solution and of improved system quality [6,8–11]. Mumford [12] argues that due to the increasing complexity of organisational life, it is difficult for analysts alone to design a system that will meet user requirements, and therefore user participation in system development is critical. In addition, the lack of user participation can lead to 'many faults and economic disadvantages' [13] and was considered to be at least partially responsible for considerable increases in the costs of developing systems at ITT and IBM [14]. Moreover, the process of developing the AS/400 system was deemed successful, at least in part, due to intensive early user participation [15,16].

A considerable amount of empirical research on the relationship between user participation and IS success has been conducted. A large part of this work was reviewed in Ives and Olson [9] and Pettingell et al. [17]. The literature review by Ives and Olson [9] concludes

[1] Meta-analysis is a quantitative approach for the synthesis and integration of the empirical results from multiple research studies [101]. This approach has considerable advantages when compared to subjective literature reviews, especially in terms of the reliability and objectivity of the reviews' results [101,50].

that no convincing results have been shown supporting the above relationship. The meta-analytic[1] review by Pettingell et al. [17] demonstrates that there *is* a (weak to moderate) relationship.[2] Given the axiomatic status of the user participation – IS success relationship, this apparent lack of agreement amongst the empirical studies in the literature and the general dearth of strong empirical evidence highlight the need for further research on the user participation construct.[3]

In this paper we report on a quantitative study of the relationship between user participation in the requirements engineering (RE) process and RE success. A number of qualitative studies have previously been reported [18–20], however, to our knowledge this relationship has not been studied quantitatively before. Given that RE success is considered to be a prerequisite for the success of software systems [21], the factors that have an impact on RE success, such as user participation, must be investigated quantitatively. Such quantitative empirical investigations will provide evidence as to whether the relationship holds, under what conditions, and its strength and nature.

As suggested by Ives and Olson [9], we assume that the relationship between user participation and success is conditional. The moderating construct is uncertainty. This amounts to a contingency theory of user participation in the RE process. Uncertainty has been previously hypothesised as a moderating construct in contingency theories of RE success [22,23].

Briefly, our study utilised survey data from 39 RE processes in different organisations. The results indicate that the magnitude of the relationship between user participation and the quality of RE service (defined as having two dimensions: user satisfaction and commitment, and fit of the recommended solution with the organisation) increases as uncertainty increases. Furthermore, we found no relationship between user participation and the quality of RE products (defined as having two dimensions: the quality of the cost/benefits analysis, and the quality of the architecture). Based on these results, we propose an empirically grounded selection model of user participation at different levels of uncertainty that can be utilised by RE practitioners.

[2] For example, the average correlation between user participation and attitudes towards the system ranged from 0.23 to 0.34, and the average correlation between user participation and reported system use was 0.12.

[3] A construct is defined as a meaningful conceptual object that is neither observable nor directly measurable [102]. For example, *requirements engineering success* and *user participation* are constructs. For ease of presentation we will use the terms *construct* and *variable* interchangeably in this paper.

In the next section of this paper we present the theoretical basis for our work. Section 3 describes the research method that was followed. In Section 4, we present the results, interpret them and discuss their implications for practice. Section 5 concludes the paper with directions for future research.

2. Background

The background covers two issues. First, we define the main terms used in this paper. Second, we present the theoretical model that is being empirically tested in our study.

2.1. Definitions

2.1.1. RE Process

The definition of the RE process that will be used here is based on the kinds of activities that are performed during that process. Davis [24] makes a distinction between two types of such activities: problem analysis and product description. During problem analysis the objective is to gain the most knowledge about the problem at hand. During product description, the objective is to construct the specification for the software system, which includes behavioural and non-behavioural requirements specifications.

To these two types of activities, we add a third type: *recommending a solution*. During this third type of activity, the objective is to examine the organisational and technical feasibility of the proposed system, as well as its profitability. The specific activities that fall under recommending a solution include: analysing the alternatives (e.g. build a system or buy a package), analysing the impact that the software system will have on the user organisation, performing a cost/benefits analysis, and forming system development and system implementation plans. These activities are standard recommended practices in the IS development literature [25,26]. Thus, the RE process is defined as the one in which the activities of problem analysis, product definition, and recommending a solution are performed.

2.1.2. RE Success

RE success is defined as the extent to which the outcomes of the RE process, as defined above, serve the needs of, and provide a basis for ensuring the success of, all subsequent activities, individually and in aggregate, related to the software system throughout the software system's lifetime. These subsequent activities include: design, coding, testing, putting into operation, and postdeployment evolution. Previous research [21] has identified that RE success is a multi-dimensional concept; the two most important dimensions are the quality of RE service and the quality of RE products.

2.1.3. User Participation

In their article, Barki and Hartwick [27] argue for making a distinction between the terms *user participation* and *user involvement* in IS development. User participation is defined as referring to the behaviours and activities that the users perform during IS development. User involvement is defined as referring to the subjective psychological state of the users, and consists of 'the importance and personal relevance that users attach to a particular system or IS in general'. Their argument is based on an integration of related research in four disciplines: management information systems, psychology, marketing, and organisational behaviour. In this paper, the concern is with user participation.

2.1.4. Uncertainty

Naumann et al. [28] define uncertainty as 'the state of knowledge of "real" user information needs' and Daft et al. [29] define it as the 'absence of information'. Uncertainty, however, is a multi-faceted concept. For instance, Davis [22] identifies three dimensions of RE process uncertainty: (1) existence and stability of a usable set of requirements, (2) users' ability to specify requirements, and (3) ability of the analysts to elicit requirements. Another useful conceptualisation of uncertainty is given by Gorry and Scott Morton [30]. They make a distinction between structured and unstructured decisions that must be supported and/or automated by IS. The former is where the decision rules of users are routine, well defined and can be formalised. The latter is where the user problems are not well defined and there are no routine and formalised procedures for dealing with them. We are concerned with an element of the RE process that has an impact on RE process uncertainty, namely the *utilising system* [22]. This includes the notion of structuredness mentioned above. Characteristics of the utilising system affecting RE process uncertainty include the stability of business processes and stability of management. A more recent extension to the Davis [22] framework adds the notion of equivocality [23]. Equivocality is defined as the existence of conflicting requirements and is concerned with ambiguity [29]. Therefore, our definition of RE uncertainty is symbolised by the extent of business process and management stability (the availability of firm requirements), and conflicting requirements (equivocality or ambiguity).

2.2. Theoretical Model

A high-level view of the theoretical model being tested in this paper is shown in Fig. 1(a). This shows that the fit or *congruence* between uncertainty and user participation has an impact on RE success: the greater the congruence, the greater RE success. The two most important dimensions of RE success have been found to be the quality of RE service and the quality of RE products [21]. A justification and a detailing of this model are given below.

IS researchers have recognised the important role of the project and organisational context in determining IS success, and have made theoretical arguments stating that IS development efforts should *fit* the project and organisational environment in order to ensure IS success [31–33]. Based on such a perspective, some authors have developed decision rules, for example, for selecting life cycle process models [34] and for deciding whether or not to prototype [35]. When expressed as a

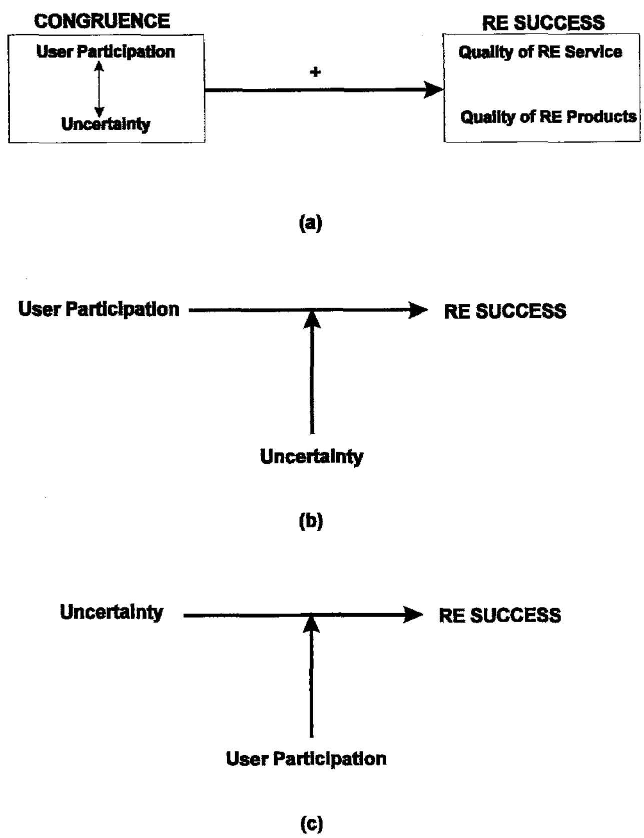

Fig. 1. The contingency model of user participation in the RE process: (a) overview of the model; (b) moderating effect of uncertainty; and (c) buffering effect of user participation.

theoretical model, such sets of decision rules are called a *contingency model.* A contingency model acknowledges that an interaction amongst independent variables has an impact on the dependent variable.[4]

In the context of user participation, it has been noted that there is not a simple bivariate relationship between user participation and IS success, for example [9,36,37]. Others have subscribed to this point of view, and have developed contingency models of user participation [6,9,38–41]. The model that we propose in this paper is also a contingency model; however, our outcome variable is RE success and not IS success.

Previous research has investigated the effect of the interaction between user participation and uncertainty on IS success. It is expected that when there is an absence of information on system requirements (i.e., high uncertainty), increased user participation will lead to the acquisition of this information. Furthermore, it is expected that when there is high equivocality, an exchange of views between analysts and users will lead to a resolution of conflicts and disagreements.

For instance, one empirical study found no relationship between user participation during the early processes of IS development and IS success when uncertainty was low, but a strong relationship when uncertainty was high [38]. This suggests that user participation has an increasingly positive influence on IS success as uncertainty increases. Another study concluded that low uncertainty was related to IS success irrespective of the level of user participation at the front end of the development process [37]. This suggests that user participation has little influence when uncertainty is low. A recent empirical study by McKeen et al. [41] that investigated a contingency model of user participation found that when uncertainty and ambiguity were low then the relationship between user participation and user satisfaction was weaker than when uncertainty and ambiguity were high.

Existing theoretical models suggest that the amount of appropriate user participation in the RE process depends upon the level of uncertainty. For instance, Alter and Ginzberg [42] present a model that maps key uncertainties at each stage of the development process with strategies for coping with such uncertainties. All of these uncertainties occur at the early stages of the life cycle. One of the suggested coping strategies for the majority of the identified uncertainties is obtaining user participation and user commitment. This suggests that user participation at the early stages is more important when uncertainty is high than when it is low. Schonberger [43] has developed a contingency model hypothesising an interaction between user participation and uncertainty. Specifically, at higher levels of uncertainty, greater user participation is recommended. Furthermore, Davis [22] asserts that there should not be a single approach to requirements determination that is applied to all projects. He subsequently presents a contingency model that relates four requirements elicitation strategies with different levels of uncertainty. At increasingly higher levels of uncertainty, he recommends synthesis from characteristics of the utilising system and an iterative discovery approach, both of which require increasingly extensive analyst–user interaction. Similar selection models are presented by Naumann et al. [28] and Naumann and Davis [44]. These also suggest that user participation becomes more important as uncertainty increases.

Based on this previous work, the specific predicted relationships of our contingency model can be specified:

P1: Uncertainty *moderates* the effect of user participation on RE success. Such a contingency relationship is shown in Fig. 1(b). This means that increases in user participation become more effective (in terms of RE success) as uncertainty increases.

P2: User participation *buffers* the effect of uncertainty on RE success. Such a contingency relationship is shown in Fig. 1(c). This means that as uncertainty increases, greater user participation reduces the negative consequences of high uncertainty on RE success.

The remainder of this paper describes an empirical study to test the above predictions.

3. Research Method

There have been a number of different methods that have been used by researchers in the past to study user participation in IS development. Newman and Robey [45] identify two types of research approaches: factor research models and process research models. With the former, one defines predictor and outcome variables and their relationships, then empirically tests those relationships. With the latter, one longitudinally studies a sequence of events in order to understand and explain how particular outcomes are arrived at; thus the focus is on the dynamics of the process. Newman and Robey [45] further note that these two research models could be considered as complementary approaches to researching a particular phenomenon.

Given the nature of our theoretical model and our predictions, it would seem most appropriate for us to

---

[4] A good discussion of contingency models, from the perspective of organisational theory, is given in Fry and Smith [103].

follow the factor research model. While this approach would not empirically explain the predictions from our theory, if our study establishes the predicted relationships, then future research can study the sequence of events that would explain the predictions [45].

The details of our research method are presented in Appendix A of this paper. In this section we provide only a brief overview of the method that was followed.

3.1. Source of Data

The context of our study was the RE phase of a software system development method (henceforth Method X). Method X has been developed and is marketed by an IS consultancy firm based in Canada with clients worldwide (henceforth Company Y). In this text we will use the term *RE process* to denote the RE phase of Method X.

We collected data on the RE process from 39 software development projects in different organisations. Therefore, the unit of analysis is the RE process in a software development project. All of these projects used Method X. All data were collected using questionnaires that were filled in by employees of Company Y consulting for client IS organisations. The characteristics of all respondents, the RE processes, and the IS organisations that they consulted for are summarised in Fig. 2. The questionnaires measured the three variables in our model: (1) RE success, (2) user participation in the RE process, and (3) uncertainty.

For measuring RE success, we used a previously developed instrument [21]. The instrument has two dimensions: (1) quality of RE service, which has two sub-dimensions covering user satisfaction and commitment, and the fit of the recommended solution with the organisation, and (2) quality of RE products, which has two sub-dimensions covering the quality of the architecture, and the quality of the cost/benefits analysis. These dimensions are shown in Fig. 3. The interpretation of the dimensions and sub-dimensions is based on the data collected from practitioners [21]. The characteristics of this instrument are presented in Appendix A.

For the measurement of user participation and uncertainty, we developed two new instruments. These are described in detail in Appendix A.

3.2. Data Analysis Method

In order to test our contingency model, a form of interaction analysis is necessary. The approach we employed uses product terms in multiple regression analysis [46]. This is most appropriate[5] for testing our theory since it allows us to investigate differences in slope. Our theory predicts that the change in RE success for a change in user participation, holding all other variables constant, is a function of uncertainty; also, that a change in RE success for a change in uncertainty, holding all other variables constant, is a function of user participation. This *mutual interaction* can be expressed by the following estimating equation [47]:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_1X_2 + \varepsilon$$

where Y = RE success, X_1 = user participation and X_2 = uncertainty.

The analysis method described in Jaccard et al. [46] allows us to answer three questions: (1) 'is there an interaction effect?'; (2) 'if so, what is the strength of this effect?'; and (3) 'if so, what is the nature of this effect?' A brief description of how each of these questions can be answered is given below.

To determine whether there is an interaction effect we utilise a hierarchical F test for the multiple regression equation without the product term and one with the product term. If the F ratio is statistically significant, then we can interpret that as meaning that there is an interaction effect.

To determine the strength of the effect, we compare the squared multiple correlation coefficient (R^2) for the equation without a product term with the equation with the product term. This informs us of the amount of variance in RE success that is accounted for by the interaction effect.

To determine the nature of the effect, we calculate the slope of user participation on RE success at specific values of uncertainty, and the slope of uncertainty on RE success at specific values of user participation. We calculate two slopes because we are testing a mutual dependence and not simply a moderator effect. Thus, the slope of user participation at specific values of uncertainty is given by:

$$Y = (b_0 + b_2X_2) + b'_1X_1$$

[5] One commonly used approach would be to employ one of the two independent variables as a contingency variable for grouping (e.g., around the mean), and compare correlation coefficients for the two groups. This approach, however, can provide misleading results [104]. Furthermore, our theory hypothesises differences in slopes, not differences in the proportion of RE success that can be explained, making this approach inappropriate. Alternative approaches include using ANOVA or contingency table analysis. Like the first approach, these require a dichotomisation (or polychotomisation) of at least one of the independent variables. As well as leading to loss of information [46], the large effort we have put into developing measurement instruments and ensuring their reliability and validity is a strong argument against reducing the variables to two-point (or three-point) scales.

Characteristic	Value	%/Mean
Location of organisation	Canada	59%
	USA	30.8%
	Australia	10.2%
Main business of organisation	Government	33.33%
	Communications/Aerospace	20.5%
	Retail/Distribution/Transport	20.5%
	Insurance/Financial/Banking	7.7%
	Other	17.95%
Annual sales (or budget) of the organisation	<=CA$99 million	12.8%
	CA$100 m–CA$149 m	5.1%
	CA$150 m–CA$199 m	2.6%
	CA$200 m–CA$249 m	7.7%
	CA$250 m–CA$999 m	17.95%
	>=CA$1 billion	53.85%
Number of people working in the IS department	<=10	5.1%
	11–29	7.7%
	30–49	23.1%
	50–99	15.4%
	100–499	30.8%
	>=500	17.95%
Position of respondents in the RE Process[a]	Management	44.7%
	Technical	39.5%
	Coaching/Auditing	31.6%
Years of experience of the project manager during the RE process		13.74 years
Years of experience of the lead architect		11.27 years
Number of analysts taking part in the RE process		7 analysts
Number of users participating in the RE process		10.7 users
Number of user departments affected by the system represented in the RE process		3.8 depts
Number of user sites represented in the RE process[b]		44 sites
Proportion of overall system development effort devoted to the RE process[c]		14%

[a] In some cases, a respondent had more than one position. For example, a lead architect would double as a part-time project manager. Therefore, the total here does not add up to 100%.

[b] When two outlier values are removed, the mean number of user sites reduces to 4.5. This is a truer reflection of the demographics of the RE processes studied.

[c] In 23.1% of the responses received, the respondents could not answer this question. One possible reason is that the system has not been developed yet and therefore they do not know the total system development effort. It should also be noted that during pilot testing of the questionnaire it became evident that it is 'general knowledge' that the RE process should consume around 15% of total effort; therefore some project managers estimate the total system development effort from the actual RE process effort. To the extent that this approach was used, it contaminates the meaningfulness of the number presented in this table.

Fig. 2. Characteristics of organisations, respondents, and RE processes.

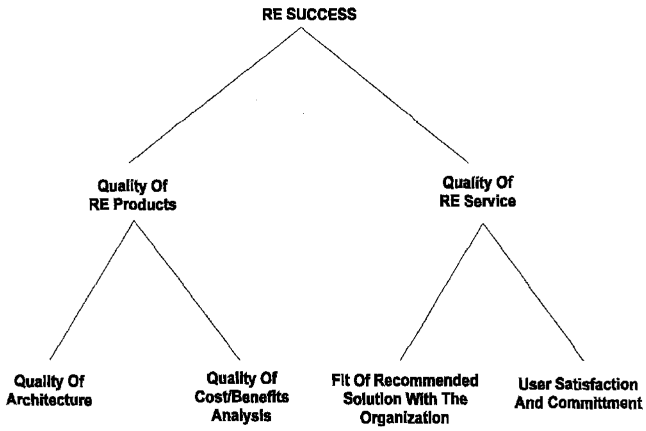

Fig. 3. The dimensions of RE success.

where $b'_1 = b_1 + b_3X_2$ (at specific values of X_2).

According to our prediction P1 (in Section 2.2), we would expect the slope b'_1 to increase positively as the values of X_2 increase. Similarly, the slope of uncertainty at specific values of user participation is given by:

$$Y = (b_0 + b_1X_1) + b'_2X_2$$

where $b'_2 = b_2 + b_3X_1$ (at specific values of X_1).

According to our prediction P2 (in Section 2.2), we would expect the slope b'_2 to increase positively as the values of X_1 increase. The significance of the size of the slope can be calculated using a t-statistic which utilises the standard error at the specific value of the independent variable [46].

4. Results and Discussion

In presenting the results, we proceed by answering the three questions that were posited in the previous section. Given that there are two dependent variables (quality of RE service and quality of RE products), we present the two resultant models separately.

For the quality of RE service model, the results of the multiple regression for the model with no multiplicative term are shown in Fig. 4, and the results for the model with the multiplicative term are shown in Fig. 5.

To determine whether there is an interaction effect, we look at the significance of the b_3 coefficient in Fig. 5. This gives the same conclusion as the hierarchical F test. This is significant at $\alpha = 0.05$, and therefore suggests that there is an interaction effect.

To determine the strength of the effect, we look at the difference in the R^2 values. The difference is 6.15%. This is the amount of variation in the quality of RE service that is accounted for by the interaction effect.

To determine the nature of the effect, the slope of one independent variable on the dependent variable is evaluated at different levels of the other independent variable. To determine the levels, we use a high value of the mean plus one standard deviation, and a low value of the mean minus one standard deviation. The slopes are shown in Figs 6 and 7. In Fig. 6, the slope of user participation and the quality of RE service at low uncertainty is –0.179, and at high uncertainty, it is 0.485. The former is not significantly different from zero. The data in Fig. 7 show that the slope of uncertainty on the quality of RE service at low user participation is –5.087, and at high user participation it is –2.818. Both are significantly different from zero. These results are shown conceptually in Fig. 8.

In terms of our predictions P1 and P2, the quality of RE service model matches them. There is a moderating effect of uncertainty on the user participation and quality of RE service relationship. The slope increases positively with increases in uncertainty as predicted from our theory (P1). There is also a buffering effect of user participation on the uncertainty and quality of RE service relationship. The slope decreases with increases in user participation as predicted from our theory (P2).

We have thus found that the interaction between user participation and uncertainty has a significant impact on the quality of RE service. As uncertainty increases, the importance of user participation also increases. Therefore, greater user participation seems to be a good strategy for alleviating the negative consequences of uncertainty on the quality of RE service. Increased user participation seems to be conducive towards greater user consensus and also it helps them reason about what their business processes should be like and what they want the IS to do. Furthermore, as uncertainty decreases, the importance of user participation decreases. When uncertainty is low, changes in user participation have no impact on the quality of RE service. What is evident here, using the framework in [48], is a *blocking effect*: low uncertainty blocks the effect of user participation. Therefore, added participation has little benefit. This does *not* mean that no participation is necessary, only that increases in participation do not bring about added benefits. Furthermore, it may be that when there is low uncertainty, users resent increases in participation when they feel that it does not contribute substantially. This resentment may bring about reductions in quality of service as user participation increases.

Also, it should be noted that, while the slope of user participation on the quality of RE service is not significantly different from zero at low uncertainty, it is negative. If we plot the slope at *very* low uncertainty the

Variable	*b* Coefficient	Std. Error
Intercept	78.93617*	2.11863
User Participation	0.38736*	0.13560
Uncertainty	–3.91389*	0.46348
F (2, 36) = 53.40449	R^2 = 0.74791	p <0.00000

Fig. 4. The model with the quality of RE service as a dependent variable and no product term (*$p < 0.05$).

Variable	*b* Coefficient	Std. Error
Intercept	80.71760*	1.94191
User Participation	0.15267	0.13848
Uncertainty	–3.95313*	0.40889
Interaction	0.06773*	0.02016
F (3, 35) = 49.54375	R^2 = 0.80940	p <0.00000

Fig. 5. The model with the quality of RE service as a dependent variable and the product term (*$p < 0.05$).

Uncertainty	b'_1	Std. Error
LOW	–0.179	0.2195
HIGH	0.485*	0.1288

Fig. 6. The slope of user participation on the quality of RE service for different levels of uncertainty (*significance at $\alpha = 0.05$ using Bonferroni corrected α levels).

User Participation	b'_2	Std. Error
LOW	–5.087*	0.5376
HIGH	–2.818*	0.5228

Fig. 7. The slope of uncertainty on the quality of RE service for different levels of uncertainty (*significance at $\alpha = 0.05$ using Bonferroni corrected α levels).

magnitude of the negative slope increases, and at some point it will become significantly different from zero. The interpretation of such a negative slope is that, at very low uncertainty, increases in user participation reduce the quality of RE service. This interpretation is consistent with one of the scenarios given in Naumann et al. [28]. The scenario was for a project where uncertainty was considerably low, and there was an inappropriately high amount of user participation. This led the user to complain of 'unnecessary system work, excessive documentation, and undue delay' [28], clearly a case of low user satisfaction (and hence low quality of RE service).

The 'main effects' in the model (the b_1 and b_2 coefficients) can be interpreted as reflecting conditional relationships [46]. The regression coefficient of the uncertainty variable can be interpreted as the influence of uncertainty on the quality of RE service when user

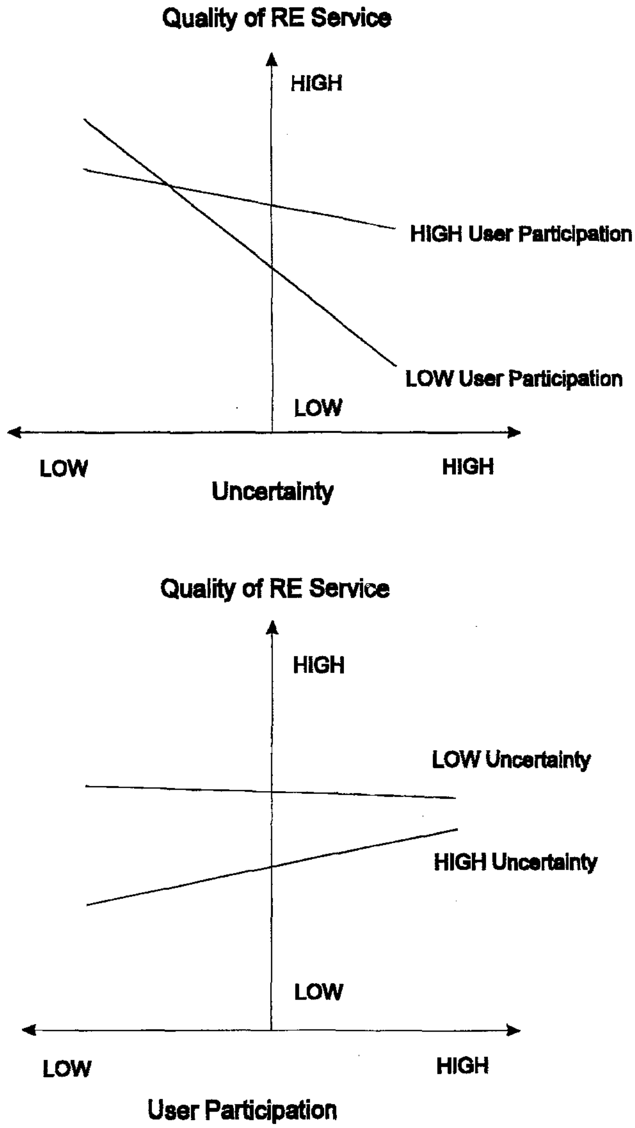

Fig. 8. Conceptualising the interaction between user participation and uncertainty and its effect on the quality of RE service.

participation is zero,[6] and a similar interpretation of the user participation coefficient. Without centring, zero values for these two coefficients would not be within the observed range in the data, and therefore represent extrapolations [49]. With centring, however, the zero values are the mean, and therefore these coefficients represent the slope at the 'average' value. Given that we have centred our variables, the latter is the interpretation of the b_1 and b_2 coefficients.

For the quality of RE products, the multiple regression models had relatively low R^2 values, and none of the regression coefficients were significant. This leads to the conclusion, based on our data, that there are neither interaction nor main effects when the quality of RE products is the dependent variable.

From a methodological perspective, this may be explained in two ways: first, the comparatively low variation in the quality of RE products measure (when compared to quality of RE service); this restriction in range would lead to a reduction in the relationship between user participation and the quality of RE products; second, the fact that IS staff rated the questionnaire instead of users may have underestimated the relationship between user participation and the quality of RE products. From a substantive perspective, this can be explained by the fact that the major RE documents are produced mainly by the analysts, not the users. Therefore, the quality of these documents would be more dependent on the capability of the analysts than that of the users and the knowledge that they can bring.

While our results could not explain the quality of RE products, a previous study has identified the quality of RE service as being a more important dimension of RE success than the quality of RE products [21]. Therefore, we are able to make recommendations for improving RE practices based on our results.

Within the limitations of a single cross-sectional study, our results support the following conclusions:

[6] To overcome some potential difficulties in the data analysis, the uncertainty and user participation variables were 'centred'. This means that the average was subtracted from each value. This results in the new average of each variable being zero. The rationale for this transformation is discussed in Appendix A.

		Uncertainty	
		LOW	HIGH
User Participation	Increasing	NO EFFECT (N)	YES
	Decreasing	NO EFFECT (Y)	NO

Fig. 9. Selection model for user participation as uncertainty changes.

- As uncertainty increases, greater user participation alleviates the negative consequences of uncertainty on the quality of RE service. This recommendation is shown in the selection model in Fig. 9.
- As uncertainty decreases, the beneficial effects of increasing user participation on the quality of RE service diminishes. However, user participation does consume resources; therefore our recommendation is to reduce user participation when uncertainty is low so as to be more cost-effective (see Fig. 9).
- Changes in the levels of user participation and uncertainty and their interaction do not have an effect on the quality of RE products. This suggests that alternative theories have to be developed (and tested) to explain variation in the quality of RE products.

5. Future Research

A number of avenues for future research are suggested by our results. We only focus on three, however, because we believe that they have the most epistemological value: (1) replication of our study, (2) using different empirical research methods to study the phenomenon of user participation, and (3) examining the different methods that can be used to enhance user participation. The former two are concerned with further research on the contingency model we have presented. Further research is necessary as [50]: 'Scientists have known for centuries that a single study will not resolve a major issue. Indeed, a small sample study will not even resolve a minor issue. Thus, the foundation of science is the cumulation of knowledge from the results of many studies'. The latter avenue for research is concerned with detailing our model so as to provide more precise recommendations for practice.

Replication[7] of empirical studies is an important part of any scientific discipline [51]. Replication provides for external validity to the extent that the replication results are similar to the original [52], as well as providing confirmatory power. Furthermore, replication is particularly important when one considers the number of possible researcher biases that could have influenced research results [53]. Another phenomenon that begs for replication is the 'file drawer problem' [54]. This problem occurs when there is a reluctance by journal editors to publish, and hence a reluctance by researchers to submit, research results that do not show statistically significant relationships. Therefore, published works are considered to be a biased sample of the empirical studies that are actually conducted. By combining the results from a large number of replications that show significant relationships, one can assess the number of studies showing no significant relationships that would have to be published before our overall conclusion of there being a significant relationship is put into doubt [54]. This assessment would allow the RE community to determine how much confidence they can place in their theories.

As is common in many empirical research studies, a necessary trade-off exists in the selection of a particular research strategy (e.g., longitudinal studies vs. cross-sectional studies, or field experiments vs. surveys vs. laboratory experiments). McGrath [55] makes the point clearly: 'all research strategies are "bad" (in the sense of having serious methodological limitations); none of them are "good" (in the sense of being even relatively unflawed). So, methodological discussions should not waste time arguing about which is the right strategy, or the best one; they are all poor in an absolute sense.' One possible approach for alleviating such concerns is to follow a multi-method empirical research strategy. The logic of the multi-method strategy is [56] 'to attack a research problem with an arsenal of methods that have non-overlapping weaknesses in addition to their complementary strengths.' This strategy effectively addresses mono-method bias in the results of a study. Rarely does a single research study satisfy a complete multi-method strategy, but the collective results of an emerging discipline can use multiple methods. In our study, the research method was a cross-sectional survey. However, to gain greater confidence in our results, the phenomenon of user participation has to be studied using other methods. Field studies of RE processes have been conducted [18,20,21], but these had a wider scope than just user participation in the RE process. In the domain of IS, Jenkins [57] and Galliers and Land [58] provide overviews of possible methods that can be employed by researchers. Furthermore, future research that would longitudinally investigate the dynamics of user participation in the RE process using the process research model [45] should be encouraged.

Olson and Ives [59] have identified 44 activities[8] that can increase user participation. One study found a relationship between one of the activities, user sign-off of project phase documentation, and user satisfaction [60]. However, in general, little empirical work has

[7] It has been noted that in the software community replications are 'not valued as important research contributions' [105]. Furthermore, the same authors call for the need to publish replicated results that do not agree with the previous results – a statement that indicates, at least, a perceived difficulty in publishing unsuccessful replications.

[8] Only 13 of these have been mapped to what we interpret to be activities in the RE process, which Olson and Ives [59] refer to as *System Definition*.

been done to evaluate these activities comparatively under different conditions. In order to provide more precise recommendations to practitioners about which activities to use and when, further research is needed in this area.

Appendix A: Details of the Research Method

The research method description covers four activities: first, the sampling procedure and the response rate; second, the measurement method and the instruments that were developed; third, an analysis of non-response bias in the results; and, finally, the data analysis method by which we test our theoretical model.

A.1. Sampling Procedure and Response Rate

For this study, the target population was the 200 client organisations worldwide[9] that have licensed Method X. Company Y employees were chosen for scoring the instruments. These employees can be considered as external assessors of the client organisations' practices.

The sample frame for this study consists of RE processes following Method X in clients of Company Y (and hence there existed a consultant with good knowledge of these RE processes). An initial list of senior consultants that work for Company Y and that have used the RE process in Method X was formulated.[10] It was considered by senior management and researchers of Company Y that the consultants whose names are on the list were involved with a representative cross-section of their clientele (they covered all business regions and sectors in Canada and outside Canada and all organisation sizes that Company Y did business with). Systematically, consultants were selected from the list and contacted (by telephone, face-to-face, and/or by electronic mail) and requested to participate in the study. Some consultants refused to participate. The non-refusals constituted our sample. All consultants in the sample were highly familiar with the organisations through their consulting assignments with their IS department and had used Method X extensively in the field. All data that were collected from each consultant were from an RE process that he or she was recently involved in.

In total, 86 questionnaires were sent out to the senior consultants. Of the 86 questionnaires, 7% were sent to consultants who had since left Company Y. Of the remaining 80 questionnaires, we received 41 responses (including late respondents).[11] This gives a response rate of 51.25%. Of all the responses, two were unusable due to extensive missing data relevant for this study, leaving a total of 39 usable responses. This sample is considered to represent approximately 19.5% of the target population.

A total of 18 responses were received before the response deadline, representing 46% of all responses received. All non-respondents after the response deadline were contacted and reminded to fill out the questionnaire. When non-respondents and late respondents were contacted and asked why they had not yet responded, their primary stated reason was that they were too busy. Thus, we consider that to be the main reason for non-response.

A.2. Measurement

In this subsection we describe the method that was followed in developing the instruments used in our study, and the details of developing each of the instruments. The instruments that were developed measured: (a) RE success, (b) user participation in the RE process, and (c) uncertainty.

A.2.1. Instrument Development

Instrumentation Method

The instrumentation method[12] followed in the study presented here draws from both the normative and the descriptive literature. The normative literature prescribes the procedures to be used in instrument development, for example [61–65]. The descriptive literature specifies the procedures used by particular authors for developing instruments, for example [66–70].

---

[9] Since our study spanned more than one calendar year, this figure is only approximate and represents information we obtained from Company Y's annual report at the time the sampling was initiated.

[10] In an attempt to construct a stratified sample, the only population characteristics that were available to us were gross revenue (of Company Y) by region and by industrial sector. However, such information was deemed to be highly misleading since a sizeable percentage of gross revenue is obtained from a small number of client organisations, and hence would not appropriately characterise the population.

[11] Questionnaires that were received but that were left completely blank are not included in calculating the response rate.

[12] All informants in all activities of the instrumentation method had a common understanding of the RE process since RE process objectives and activities are defined in Method X.

Four primary criteria[13] for developing and evaluating the instruments were utilised [71,72]: content validity, reliability, construct validity, and effectiveness. Based on our instrumentation method and collected data, we can provide direct and indirect evidence useful for evaluating the extent to which each criterion has been satisfied.

Content validity is defined as the representativeness or sampling adequacy of the content of a measurement instrument [61]. Ensuring content validity depends largely on a literature review and expert judgement.

Reliability is defined as the extent to which an experiment, test, or any measuring procedure yields the same results on repeated trials, and is concerned with the problem of random measurement error [62]. While there are alternative and complementary ways of evaluating reliability, the reliability of instruments in this study was evaluated using the Cronbach alpha coefficient [73]. IS researchers tend to report the Cronbach alpha coefficient most frequently [74], and some researchers consider it to be the most important reliability evaluation approach [75]. The computation of the Cronbach alpha coefficient partitions total variance on a scale into signal and noise (i.e., random measurement error). The proportion of total variation that is signal equals alpha. A Cronbach alpha coefficient of 0 means perfect unreliability, and a coefficient of 1 means perfect reliability. Thus, as the noise or error variance approaches 0, reliability, as computed using Cronbach alpha, approaches 1. A Cronbach alpha coefficient of 0.8 is considered to be sufficient for research purposes [62].

Construct validity is an operational concept that asks whether the items chosen are describing the true construct(s) [62]. In this study, the constructs are RE success, user participation, and uncertainty. Construct validity includes two other concepts: convergent validity and discriminant validity. Convergent validity determines whether the items chosen are measuring one underlying construct. Discriminant validity determines whether an item differentiates between constructs. Construct validity of the instrument was evaluated using factor analysis and item–total correlations [61,62].

Factor analysis is considered to be 'a powerful and indispensable method of construct validation' [61]. Factor analysis is a multivariate technique for 'clustering' variables. If the emerging 'clusters' match the expected dimensions, then this is evidence of construct validity. The logic behind using factor analysis is that variation among a number of questions that form a cluster can be attributed to variation amongst projects on one common underlying factor (e.g., user participation or RE success). In the presentation of the results, factor values that are greater than 0.5 are considered significant. Factor values that are not included in the presentation of our results are all less than the cut-off value.

For item–total correlations, each item score was subtracted from the total to avoid a spurious part–whole correlation [76], and the correlation of each item with the new total score was computed. The logic behind item–total correlations is that we assume that the total score is valid. The extent to which an individual item measures the same thing as the total score is an indicator of the validity of that item [61].

Effectiveness refers to the extent to which an item is measuring a construct relative to the other scales that are measuring the same construct. Attaining a reasonable level of effectiveness is important so as not to have a lengthy instrument. Multiple criteria were utilised to determine effectiveness (i.e., the results of the validity and reliability analyses). The purpose was to eliminate concepts from the instrument without negatively affecting its reliability and validity.

Another criterion that has been used by IS researchers for evaluating an instrument is predictive validity [66,68]. Predictive validity can be demonstrated by correlating the score on an instrument with another measure of the same construct. In this study, these other measures were summary scales of the construct. When used, the manner in which this summary scale was developed is described for the construct.

Furthermore, for all the instruments described below, three small pilot studies were conducted. The purpose of these was to identify ambiguities, inconsistencies, bad wording and to generally get feedback on the usability and clarity of the instruments. During the first study, the instruments were given to two senior practitioners working for Company Y and they were asked to review them and provide comments on any ambiguities, inconsistencies, and their understandability, usability, and clarity. During the second study, two other senior practitioners working for Company Y were requested to rate an RE process in an interview setting with one of the authors of this paper present. Each interviewee was requested to talk out loud while rating, indicating what he interprets each question to mean and the rationale for the ratings. During the third study, two researchers from the Software Engineering Laboratory at McGill University reviewed the instruments. Based on the comments from the three pilot studies, the instruments were revised.

[13] These criteria were recognised in software engineering more than a decade ago by Curtis [106] and Shaw [107]. However, their application in mainstream software engineering research is almost non-existent.

The relationship between the users and the RE staff was:

harmonious |__|__|__|__|__|__|__| dissonant

score values -> **7 6 5 4 3 2 1**

Fig. 10. Example of a semantic differential scale and its score values as used in our study.

The comments from the pilot studies identified some question wording problems where some questions were ambiguous and/or were being interpreted differently by different respondents or differently from the interpretation that we had intended. Also, the comments helped us reorganise the questionnaire to make it easier to follow and resulted in a reduction in its size.

Scaling Method

The scaling model that we have used in defining our three instruments is the Likert or summative scaling model [77]. Using this model, a number of questions measuring each construct are developed. The scores on each question are summed to obtain the overall score for that variable.

All the questions[14] that we have used in our study utilised a semantic differential scale [78]. A semantic differential scale consists of a concept followed usually by a seven-point scale anchored by bipolar adjective pairs. An example of a semantic differential scale is shown in Fig. 10. The responses on this scale were scored as shown in Fig. 10. Thus, for this example, if the response is at the extreme left category, it was assigned the value of 7. Similarly, if the response is at the extreme right category, it was assigned the value of 1. For our three variables, the scoring scheme for each question was set up such that higher scores mean: (a) greater RE success, (b) greater user participation, and (c) greater uncertainty.

A.2.2. Requirements Engineering Success Instrument

The characteristics of this instrument based on the data from the current study are shown in Fig. 11. The reliability for each dimension is considerably high (0.96 and 0.89). Evidence for construct validity is shown in terms of the factor structure, which matches the expected dimensions, and in terms of the relatively high item–total correlations. However, the item measuring 'the adequacy of the diagnosis of the existing system' had a factor loading less than 0.5 and its correlation with the total scale was small and not statistically significant. A possible reason for this is that the IS staff do not spend sufficient time performing this activity because it may be perceived as politically dangerous (i.e., because the users may not appreciate effort being spent modelling and diagnosing the current system knowing that it will be replaced) [19,79]. Therefore, this item may not covary with other indicators of RE product quality.

For predictive validity, two summary questions for each dimension were used. These summary questions were on a seven-point semantic differential scale [78]. Each summary question pertained to one sub-dimension. The values for both questions were summed for each dimension. The coefficients for predictive validity are high (0.94 and 0.71) and significant ($p < 0.000$).

A.2.3. User Participation Instrument

Given the relatively long history of IS research on the user participation construct, there exist a number of user participation instruments in the literature. Many of these, however, are unsuitable for direct use in our study. For example, some of the published instruments evaluate overall user participation in an IS department [60,70], but we are concerned with individual projects; instruments measuring overall user participation in IS development for individual projects have been developed [6,40]; however, we are interested in measuring user participation only in the RE process. Other researchers have measured user participation in the RE process using single items [38]; these, however, tend to have low reliability [62,80].

Multiple-item instruments that measure user participation in the RE process for individual projects have been developed [37,41,69,81,82]. These are based on sets of generic RE activities. Conversely, in the context of our study, there are specific RE activities defined in Method X. Therefore, instead of using generic activities, our instrument can use the Method X activities. These activities are described briefly in Appendix B.

For each activity, we wanted to know three things:[15] (1) the extent to which there was consultation of users,

[14] There is one exception, namely in the user participation instrument. We developed a dichotomous scale, but this was eliminated from the analysis because of the small variation in responses.

[15] Hirschheim [108,109] identifies two elements of participative design that ought to be considered: (1) the content of participation, and (2) user involvement (this is Hirschheim's terminology, which is similar to our definition of *user participation* given in Section 2.1). Our measurement of user participation focuses on the second element and not the first. The content of participation as defined by Hirschheim is concerned with the subject matter of participation, which includes both technical and social considerations. In Method X the *scope* of user participation is well defined and explicitly includes social considerations (e.g., an analysis of how the proposed system will fit into and have an impact on the user organisation is part of one of the formal Method X deliverables). We therefore anticipated that if we attempted to measure the content of participation there would be little variation across projects following Method X.

(a)

Quality of Requirements Engineering Service

Mean	Std. Dev.	Cronbach α	Predictive Validity
79.783	25.29729	0.96525	0.9416*

Scales Used	Factor 1	Factor 2	Item–Total Correlations
• The awareness of the users of the business changes required in order to implement the recommended solution	0.772840	REMAINDER OF FACTOR 2 IS SHOWN IN FIG. 11(b)	0.653988*
• The users defending the recommended solution in front of executive management	0.852156		0.784147*
• The users' consensus on the recommended solution	0.881427		0.852906*
• The relationship between the users and the requirements engineering staff	0.851432		0.834333*
• The fit between the recommended solution and the strategic orientation of the organisation	0.589543		0.672874*
• The ability of the user organisation to make the necessary changes to implement the recommended solution	0.691609		0.721003*
• The willingness of the user organisation to make the necessary changes to implement the recommended solution	0.846621		0.825082*
• The degree of top management support for changes necessary to implement the recommended solution	0.835191		0.784211*

(b)

Quality of Requirements Engineering Products

Mean	Std. Dev.	Cronbach α	Predictive Validity
80.0106	16.31702	0.890762	0.7095*

Scales Used	Factor 1	Factor 2	Item–Total Correlations
• The coverage of the costs in the cost/benefits analysis	REMAINDER OF FACTOR 1 IS SHOWN IN FIG. 11(a)	0.777055	0.537158**
• The coverage of the benefits in the cost/benefits analysis		0.757643	0.604415*
• The amount of benefits that are expected to be brought to the organisation by implementing the recommended solution compared to alternative solutions		0.402690	0.369213***
• The accuracy of the cost estimates for the recommended solution compared to the accuracy the organisation is accustomed to		0.501920	0.555295**
• The clarity of the links between the (process and data) models and the system objectives		0.594607	0.583304*
• The clarity of the links between the (process and data) models and the key issues		0.537871	0.540921**
• The adequacy of the diagnosis of the existing system		0.407199	0.172328
• The clarity of the links between the weaknesses and strengths of the existing system and the weaknesses and strengths of the recommended solution		0.664985	0.576977*

KEY
All tests are one-tailed
*$p<0.0000$
**$p<0.000$
***$p<0.00$

Fig. 11. Characteristics of the measure of the quality of (a) RE service and (b) RE products.

(2) the extent to which users had responsibilities, and (3) whether the users validated and signed off the deliverables of each activity. For each of the former two, a seven-point semantic differential scale [78] was developed for each activity; for the latter a dichotomous scale was developed for each activity.

After the data were collected, it was found that there was little variation in the items assessing whether the users validated and signed off the deliverables of each activity. In general, the users did validate and sign off deliverables for all activities. Therefore, these items were removed from the final instrument.

The characteristics of the final instrument are shown in Fig. 12. As can be seen, the composite reliability is high (0.93133). There is also good evidence of construct validity, with the factor structure matching the expected two dimensions, and overall high item–total correlations. For predictive validity, a single summary item asking about overall user participation in the RE process was used. The significant ($p < 0.000$) 0.5809 correlation provides further evidence of validity.

In general, there will only be two types of personnel knowledgeable about the extent of user participation in an IS project: users and IS staff. Previous research has found that there is a weak or no relationship between user participation as rated by users compared to the ratings by IS staff. For instance, Olson and Ives [60] found a zero correlation (non-parametric) between overall user participation ratings made by users and IS managers. This, however, they attribute to the fact that the IS managers were rating the typical user while users were rating their own level of participation. Hawk and Aldag [82] found only a 0.31 ($p < 0.01$) correlation between project-specific user participation ratings made by users and those made by systems analysts. In previous research, however, where disagreements between the two perspectives were found, the IS rating was used [10].

When either of those types of personnel rate user participation there is also the danger of self-serving bias [82]. This means that individuals are likely to attribute success to themselves, and to minimise self-blame in the case of failure. If the users self-rate their participation in the IS development process, then this would tend to inflate the relationship between user participation and success. One study that investigated this phenomenon in the context of user participation measurement found that a user-rated user participation measure is more strongly related to success (using a user satisfaction instrument) than when user participation is rated by analysts [82]. Such evidence supports the self-serving bias hypothesis.

The alternative to user-rated instruments is to have IS staff rate user participation. However, the IS staff's perceptions of user participation may still be influenced by self-serving bias. Low user participation implies increased IS staff responsibility, and vice versa [82]. Such a bias would lead to an understatement of the relationship between user participation and success.

In our study, the IS staff rated user participation. Therefore, the results of our study would tend to be conservative since the extent of user participation is expected to be understated for successful RE processes and overstated for unsuccessful RE processes. It should be noted that the alternative, of users rating their participation, could still introduce some bias into our findings, resulting in inflated relationships between user participation and RE success.

A.2.4. Uncertainty Instrument

A number of researchers have conducted empirical studies where they defined and measured uncertainty based on the notion of structuredness. Munro and Davis [83] describe an experiment to compare the effectiveness of two methods for determining information requirements under different situations. One of the situational characteristics that they varied was the structuredness[16] of management decisions. They subjectively determined whether a decision was relatively structured or relatively unstructured. Based on their experience, they assert that there is no objective way to evaluate the structuredness of a decision and call for more research in this area. Another approach was used in Edstrom's study [6]. He makes a distinction between users in the technological core of the organisation (e.g., production and inventory control) and users working at the boundary of the organisation. The former 'would tend to have programmed conceptual frameworks, since their task environment is relatively well structured', and the latter 'will tend to have heuristic frameworks, since they will have to cope with more uncertainty and must be able to adapt to changes in the environment.' Kim and Lee [38] use structuredness[17] as a moderating variable in a model of user participation. Structuredness was divided into structure of the procedure and periodicity of the job interval. Each was measured using a single item. Franz and Robey [37] developed five items to measure the structure of decisions. These evaluate the extent to which decisions are routine and frequent and how routine the decision-making procedures were.

---

[16] In that paper, the authors use the term *programmed*. This is only a difference in terminology since, in the Gorry and Scott Morton [30] framework, use of the terms *structured* and *unstructured* are based on Simon [110], which is the same source cited by Munro and Davis [83]

[17] The authors also use the term *task complexity* in their paper.

User Participation	Composite Mean	Composite Std. Dev.	Composite Reliability	Predictive Validity
	64.13949	16.75642	0.93133	0.5809**

Dimension Name	Scales Used	Factor 1	Factor 2	Item-Total Correlations	Cronbach α	Mean	Std. Dev.
User Responsibility	• Examination of the current system	0.667929		0.655845*	0.88841	20.5667	9.5284
	• Define system context and objectives	0.656569		0.630411*			
	• Build the conceptual process model	0.879654		0.824277*			
	• Build the conceptual data model	0.890194		0.783208*			
	• Establish basic system concepts	0.805042		0.753793*			
	• Draft functional process model	0.920133		0.851660*			
	• Determine technical feasibility	0.739786		0.550720**			
	• Determine cost effectiveness	0.425729		0.402872***			
User Consultation	• Examination of the current system		0.717492	0.626204*	0.91883	43.5728	10.6369
	• Define system context and objectives		0.883336	0.859157*			
	• Build the conceptual process model		0.883352	0.840381*			
	• Build the conceptual data model		0.827543	0.786034*			
	• Establish basic system concepts		0.880585	0.834332*			
	• Draft functional process model		0.901477	0.860740*			
	• Determine technical feasibility		0.504056	0.469873***			
	• Determine cost effectiveness		0.729175	0.636198*			

KEY

All tests are one-tailed

*$p < 0.0000$
**$p < 0.000$
***$p < 0.00$

Fig. 12. Characteristics of the measure of user participation.

Based on the more general notion of characteristics of the utilising system of Davis [22] and the more recent extension of this work to include the notion of equivocality [23], we developed a four-item measure of requirements stability and reliability. One item was dropped because it reduced overall reliability and showed little validity (this item was concerned with the overall stability of the business process). The remaining three items are shown in Fig. 13. These three items match the definition of uncertainty given in Mathiassen and Stage [84]: 'the availability and reliability of the information that is relevant in a given situation'. Thus, the extent of user commitment to their decisions and the strength of leadership items are concerned with having clear and firm requirements available from the users. The conflicting requirements item is concerned with the reliability of the requirements that are gathered.[18]

We hypothesised another dimension to uncertainty, namely technology uncertainty. While this does not enter into our theoretical model, it was used to provide evidence of validity. All items covering the two dimensions of uncertainty were presented to seven experienced practitioners in an interview setting. The items were actually the concepts that made up the semantic differential scale that we used (these are the bullet points in Fig. 13). All items were written on slips of paper approximately 1 × 8.5 inches in size. The interviewees were requested to put the items into two or more categories such that the items in each category were most similar to each other and most dissimilar from the other categories. The items were randomly ordered before each interview. The interviewees were subsequently requested to provide an interpretation of each category. Based on the interviewee's categorisation, a distance matrix was constructed. The distance matrix was derived from an incidence matrix as follows: $d_{ij} = 1 - s_{ij}$, where d_{ij} is the distance between items i and j, and s_{ij} is the similarity between items i and j. Similarity is defined as the proportion of interviewees who placed items i and j in the same category. Subsequently, hierarchical agglomerative clustering algorithms [85] were used to identify the various dimensions of uncertainty as perceived by the interviewees.[19] The results of this clustering were the two hypothesised dimensions, and the interviewee interpretations were consistent with our expectations.

[18] Measures of project size and the number of users are not included in our uncertainty instrument because, as elaborated in Burns and Dennis [111], these are concerned with complexity as opposed to uncertainty.

[19] Miller [112] shows that such a matrix is a metric distance matrix, and hence satisfies the criteria desirable for the application of numeric cluster analysis algorithms [85].

Further evidence of validity and reliability is shown in Fig. 13 based on the survey data. As can be seen, two distinct factors emerged as expected. The Cronbach alpha reliability was 0.8361, and item–total correlations are reasonably high and all significant at $p < 0.001$.

A.3. Non-Response Bias

Late respondents are considered to provide a good measure of the characteristics of non-respondents [86]. To test for non-response bias, early respondents were compared with late respondents with respect to their demographic characteristics. The demographic characteristic frequencies were tabulated in $r \times c$ tables, where r was always 2 (early respondents/late respondents) and c was 2 or greater. Given the relatively small sample, a decision rule had to be employed in choosing the most appropriate test [87]. For 2 × 2 tables, if all expected frequencies are greater than 5, then a chi-square test was used. Otherwise, the Fisher exact test was used.[20] For $2 \times c$ tables, where $c > 2$, if less than 20% of the expected frequencies were less than 5 and if no cell had an expected frequency of less than 1, then the chi-square test was used. Otherwise, cells were combined by merging columns,[21] and the appropriate decision rule was applied.

All tests of non-response bias were two-tailed and were conducted at an alpha level of 0.05. For the following characterisations no bias was found (i.e., the null hypothesis could not be rejected): by sales (over 1 billion/less than 1 billion), number of employees in the IS department (more than 100/less than 100), business sector (government/communications – aerospace/retail – distribution – transportation/insurance – finance – banking/other), number of analysts taking part in the RE process (low/high determined around the mean), percentage of total effort expended on the RE process (low/high determined around the mean), the number of user departments affected by the system (low/high

[20] The approximation of the X^2 statistic to a chi-square distribution assumes that expected frequencies are not too small. However, Cochran [113] has pointed out that the rule utilising the number five as a minimal expected frequency may be too restrictive. Furthermore, it has been suggested that the conventional chi-square statistic may be used for $2 \times c$ tables where expected frequencies may be as low as 1 [114]. In any case, in our data set, for 2 × 2 tables, there was no difference in the resulting interpretation of bias between the conventional chi-square test and the Fisher exact test.

[21] Combining categories is generally not considered to be good practice. This leads to loss of information, affects the sampling characteristics in manners whose consequences are unknown, and the different possible ways in which the pooling of categories is done can lead to different results. In this study, where pooling was performed, pooled and unpooled results were compared. Where there was no difference, the unpooled results are presented. On only one occasion was there a difference.

Uncertainty

Dimension Name	Scales Used	Factor 1	Factor 2	Item–Total Correlations	Cronbach α	Mean	Std. Dev.
Requirements Stability & Reliability	• The extent to which the users are committed to the decisions they make	0.9439		0.8208*	0.8361	10.5306	4.9024
	• The strength of leadership in the user organisation	0.9369		0.8227*			
	• The extent to which there are conflicts in the requirements	0.7177		0.4899**			
Technology Uncertainty	• The extent to which the application technology is state-of-the-art		0.9075	0.7901*	0.8564	11.6842	4.4546
	• The stability of the application technology[9]		0.8733	0.7231*			
	• The extent to which the application technology is fragmented[b]		0.8560	0.6901*			

9 This refers to the extent to which the technology that will be used to develop the application is changing.

b Fragmentation refers to the extent to which there are multiple different technologies, on multiple platforms, and provided by multiple vendors.

KEY

All tests are one-tailed

$^{*}p < 0.0000$
$^{**}p < 0.000$

Fig. 13. Characteristics of the measure of uncertainty.

	User Participation	Uncertainty	Interaction	R^2 (raw)	R^2 (centred)
User Participation	1.00			0.8641	0.3347
Uncertainty	–0.3282*	1.00		0.9202	0.1084
Interaction	0.3834*	0.7067*	1.00	0.9237	0.2706

Fig. 14. Bivariate correlations amongst the independent variables without centring, and R^2 values when regressing each independent variable on the other two for raw and centred values (*$p < 0.05$).

determined around the mean), number of user sites (low/high determined around the mean),[22] position of respondents (manager/technical/other), project manager's years of experience (low/high determined around the mean), and location (Canada/outside Canada). However, for the characterisation by the number of users involved in the RE process (low/high determined around the mean) and by the lead architect's years of experience (low/high determined around the mean) the null hypothesis could be rejected. A closer examination showed that lead architects tend to have more years of experience in the early respondents group than the late respondents group. Also, the number of users involved tend to be smaller in the late respondents group. Hence, on demographic characteristics there seems to be a slight bias.

Finn et al. [88] have argued that demographic differences between respondents and non-respondents do not automatically signal bias. To investigate this possibility, respondents and non-respondents were compared in terms of the four main variables being investigated in this study (user participation/uncertainty/quality of RE service/and quality of RE products). All four variables were dichotomised around the mean. The null hypothesis could not be rejected at an alpha level of 0.05 for a two-tailed test for all variables. Hence, there seems to be no significant bias in responses between respondents and non-respondents, even though a slight demographic difference was identified.

A.4. Details of the Data Analysis Method

In regression models where multiplicative terms are used, there is usually a high degree *multicollinearity*.[23] This is because the multiplicative term exhibits strong correlations with its component parts.[24]

[22] We repeated this test after removing outliers. The null hypothesis could not be rejected at $\alpha = 0.05$.

In our data set, the bivariate correlations amongst the independent variables are shown in Fig. 14. It is clear that the correlation between the interaction term and uncertainty is rather high. A better approach to detect multicollinearity is to regress on each independent variable all the other independent variables. This allows one to detect the case where an independent variable depends on more than one other independent variable. The results of this regression are also shown in Fig. 14. This clearly indicates high multicollinearity (because of the relatively high values of R^2).

One approach to deal with multicollinearity is to transform the independent variables by centring them. This means subtracting the mean for that variable from each raw observation [89]. The centred values are then used to estimate the regression coefficients. Using this transformation, the new (centred) R^2 values are shown in Fig. 14. This has resulted in a substantial reduction in multicollinearity as indicated by the relatively low R^2

[23] One of the requirements of the linear regression model that we have employed in our analysis is that no independent variable is perfectly linearly related to one or more of the other independent variables. This situation is referred to as *perfect* multicollinearity. When there is perfect multicollinearity then the regression surface is not even defined. Perfect multicollinearity is not common. However, the larger the multicollinearity, the greater the standard errors of the coefficient estimators. This means that t-statistics for significance tests tend to be small, and confidence intervals tend to be wide [115]. One implication of this is that the conclusions drawn about the relative impacts of the independent variables based on coefficient estimates from the sample are less stable.

[24] Such high multicollinearity has led some authors to discourage the use of multiplicative terms in multiple regression [115]. The multiple regression equation with a multiplicative term, however, specifies a conditional relationship. The standard errors of the regression coefficients are not general as in an additive model (i.e., with no multiplicative term), but are conditional upon certain values of X_1 and X_2. Therefore the standard errors of $b_1$and b_2 are the conditional coefficients when X_2 and X_1 are zero respectively [49]. In the presence of interaction, the conditional standard errors of the coefficients in the interactive model will be lower than those in the additive model for some values of X_1 and X_2 [116]. Furthermore, a Monte Carlo simulation demonstrated the reliability and stability of these coefficients [49]. However, multicollinearity has another possible effect: that of computational errors [89]. Therefore, steps still have to be taken to reduce it.

values. Furthermore, the largest of the variance inflation factors [90] for the three X's is 1.5 (the average value is 1.33), which is considerably smaller than the maximum acceptable value of 10.

Another one of the assumptions that the multiple regression model is based upon, as stated by some authors [91], is that all the variables should be measured at least on an interval scale. This assumption is based on the mapping originally developed by Stevens [92] between scale types and 'permissible' statistical procedures. In our context, this raises two questions. First, what are the levels of our measurement scales? Second, to what extent can the violation of this assumption have an impact on our results?

The scaling model we have used in the measurement of our constructs is the summative (or 'Likert', as it is also referred to) model [77]. Some authors state that summative scaling produces interval-level measurement scales [77], while others argue that this leads to ordinal-level scales [93]. In general, however, our scales are expected to occupy the grey region between ordinal and interval level measurement.

Given the proscriptive nature of Stevens' mapping, the permissible statistics for scales that do not reach an interval level are distribution-free (or non-parametric) methods (as opposed to parametric methods, of which multiple regression is one) [87]. Such a broad proscription is viewed by Nunnally as being 'narrow' and would exclude much useful research [62]. Furthermore, studies that investigated the effect of data transformations on the conclusions drawn from parametric methods (e.g., F ratios and t tests) found little evidence supporting the proscriptive viewpoint [94–97]. Suffice it to say that the issue of the validity of the above proscription is, at best, debatable. As noted by many authors, including Stevens himself, the basic point is that of pragmatism: useful research can still be conducted even if, strictly speaking, the proscriptions are violated [91,92,98,99]. A detailed discussion of this point that supports our argument is given in Briand et al. [100].

Appendix B: Activities of the RE process in Method X

The following are the activities of the RE process as stipulated in Method X.

Examine the Existing System

This gives the analysts an understanding of the user environment where the new system will be implemented. During this activity, users provide the relevant information about their existing manual and automated processes. Users should ensure that all statements based on this examination are based on fact and not only on impressions.

Define System Context and Objectives

The objectives of the system are the results to be achieved. The context defines the boundaries of the system. The key issues identify the factors critical to the success of the system. Users participate in this activity to ensure that the objectives are realistic and supported by the findings of the activity above (examining the existing system). Also, users must be involved in deciding on the boundaries of the system to ensure that they agree to it and that it matches their needs.

Build the Conceptual Process and Data Models

These two activities ought to be based on interactions and consultations with the users to ensure that all the users see the same thing the same way and also that the models reflect their business processes. Users must understand these models, and may even develop the models themselves.

Establish Basic System Concepts

Basic system concepts guide the technical choices made in the system architecture. Each one must be fully justified. The users must determine what is important and what is 'nice to have'.

Draft Functional Process Model

During this activity the functions of the system are defined and divided into automated and manual functions. Users should provide input in developing these models as well as ensure that they understand and agree to their contents.

Determine Technical Feasibility

During this activity, a decision is made on whether to adopt a package, to custom build the IS, or do both (where the package is part of the solution). The future direction and strategies of the user organisation must be considered when making this choice.

Determine Cost Effectiveness

The cost of the system (costs of development, acquisition, implementation, operation, and maintenance) must be compared to its anticipated benefits. The users must be convinced and committed to the cost and benefits estimates that are made during this activity.

Acknowledgements

The authors wish to thank Lionel Briand for his constructive comments on an earlier version of this paper. Also, we wish to thank the anonymous referees for their valuable comments and suggestions on an earlier version of this paper.

References

1. Ginzberg M. Steps towards more effective implementation of MS and MIS. Interfaces 1978; 8(3): 57–63
2. Ginzberg M. A study of the implementation process. In: TIMS studies in the management sciences, vol 13. North-Holland, Amsterdam, 1979, pp 85–102
3. Zmud R, Cox J. The implementation process: a change approach. MIS Q 1979; June: 35–43
4. Keen P. Information systems and organizational change. Commun ACM 1981; 24(1): 24–33
5. Bostrom R, Heinen J. MIS problems and failures: a socio-technical perspective. MIS Q 1977; September: 17–32
6. Edstrom A. User influence and the success of MIS projects: a contingency approach. Hum Relations 1977; 30(7): 589–607
7. Zand D, Sorenson R. Theory of change and the effective use of management science. Admin Sci Q 1975; 20: 532–545
8. Berry D. Involving users in expert system development. Expert Syst 1994; 11(1): 23–28
9. Ives B, Olson M. User involvement and MIS success: a review of research. Manage Sci 1984; 30(5): 586–603
10. Vanlommel E, de Brabander B. The organization of electronic data processing (EDP) activities and computer use. J Business 1975; 48(2): 391–410
11. Torkzadeh G, Doll W. The test–retest reliability of user involvement instruments. Inform Manage 1994; 26: 21–31
12. Mumford E. Participation: from Aristotle to today. In: Bemelmans T (ed.). Beyond productivity: information systems development for organizational effectiveness. Elsevier, Amsterdam, 1984
13. Sack K. User participation in software development: what is it, why, and how? In: Briefs U, Tagg E (eds) Education for system designer/user cooperation. Elsevier, Amsterdam, 1985
14. Mumford E. Defining system requirements to meet business needs: a case study example. Comput J 1985; 28(2): 97–104
15. Pine B. Design, test and validation of Application System/400 through early user involvement. IBM Syst J 1989; 28(3): 376–385
16. Sulack A, Lindner R, Dietz D. A new development rhythm for AS/400 Software. IBM Syst J 1989; 28(3): 386–406
17. Pettingell K, Marshall T, Remington W. A review of the influence of user involvement on system success. In: Proceedings of the 9th international conference on information systems, 1988, pp 227–236
18. Lubars M, Potts C, Richter C. A review of the state of the practice in requirements modeling. In: Proceedings of the IEEE international symposium on requirements engineering, 1993, pp 2–14
19. El Emam K, Madhavji NH. A field study of requirements engineering practices in information systems development. In: Proceedings of the second IEEE international symposium on requirements engineering, 1995, pp 68–80
20. Curtis B, Krasner H, Iscoe N. A field study of the software design process for large systems. Commun ACM 1988;31(11): 1268–1286
21. El Emam K, Madhavji NH. Measuring the success of requirements engineering processes. In: Proceedings of the second IEEE international symposium on requirements engineering, 1995, pp 204–211
22. Davis G. Strategies for information requirements determination. IBM Syst J 1982; 21(1): 4–31
23. Fazlollahi B, Tanniru M. Selecting a requirements determination methodology: contingency approach revisited. In Proceedings of the 9th international conference on information systems, 1988, pp 55–64
24. Davis A. Software requirements: objects, functions, and states. Prentice-Hall, Englewood Cliffs, NJ, 1993
25. Davis G, Olson M. Management information systems: conceptual foundations, structure, and development, McGraw Hill, New York, 1985
26. Palvia P, Palvia S. The feasibility study in information systems: an analysis of criteria and contents. Inform Manage 1988; 14: 211–224
27. Barki H, Hartwick J. Rethinking the concept of user involvement. MIS Q 1989; March: 52–63
28. Naumann J, Davis G, McKeen J. Determining information requirements: a contingency method for selection of a requirements assurance strategy. J Syst Software 1980; 1: 273–281
29. Daft R, Lengel R, Trevino L. Message equivocality, media selection, and manager performance: implications for information systems. MIS Q 1987; September: 355–366
30. Gorry G, Scott Morton M. A framework for management information systems. Sloan Manage Rev 1971; Fall: 55–70
31. Ein-Dor P, Segev E. Organizational context and the success of management information systems. Manage Sci 1978; 24(10): 1064–1077
32. Ginzberg M. An organizational contingencies view of accounting and information systems implementation. Accounting Organ Soc 1980; 5(4): 369–382
33. Wetherbe J, Whitehead C. A contingency view of managing the data processing organization. In MIS Q 1977; March: 19–25
34. Alexander L, Davis A. Criteria for selecting software process models. In: Proceedings of the 15th international IEEE COMPSAC, 1991, pp 521–528
35. Janson M, Smith L. Prototyping for systems development: a critical appraisal. MIS Q 1985; December: 305–316
36. Swanson E. Management information systems: appreciation and involvement. Manage Sci 1974; 21(2): 178–188
37. Franz C, Robey D. Organizational context, user involvement, and the usefulness of information systems. Decision Sci 1986; 17: 329–356
38. Kim E, Lee J. An exploratory contingency model of user participation and MIS use. Inform Manage 1986; 11: 87–97

39. Tait P, Vessey I. The effect of user involvement on system success: a contingency approach. MIS Q 1988; March: 91–108
40. Amoako-Gyampah K, White K. User Involvement and user satisfaction: an exploratory contingency model. Inform Manage 1993; 25: 1–10
41. McKeen J, Guimaraes T, Wetherbe J. The relationship between user participation and user satisfaction: an investigation of four contingency factors. MIS Q 1994; December: 427–451
42. Alter S, Ginzberg M. Managing uncertainty in MIS implementation. Sloan Manage Rev 1978; Fall: 23–31
43. Schonberger R. MIS design: a contingency approach. MIS Q 1980; March: 13–20
44. Naumann J, Davis G. A contingency theory to select an information requirements determination methodology. In: Proceedings of the second software life cycle management workshop, 1978, pp 63–65
45. Newman M, Robey D. A social process of user–analyst relationships. MIS Q 1992; June: 249–266
46. Jaccard J, Turrisi R, Wan C. Interaction effects in multiple regression. Sage, Beverly Hills, CA, 1990
47. Fisher G. Problems in the use and interpretation of product variables. In: Long J (ed). Common problems/proper solutions: avoiding error in quantitative research. Sage, Beverly Hills, CA, 1988
48. Joyce W, Slocum J Jr, von Glinow M. Person–situation interaction: competing models of fit. J Occup Behav 1982; 3: 265–280
49. Friedrich R. In defence of multiplicative terms in multiple regression equations. Am J Polit Sci 1982; 26(4): 797–833
50. Hunter J, Schmidt F, Jackson G. Meta-analysis: cumulating research findings across studies. Sage, Beverly Hills, CA, 1982
51. Lamal P. On the importance of replication. In: Neuliep J (ed). Replication research in the social sciences. Sage, Beverly Hills, CA, 1991
52. Lindsay R, Ehrenberg A. The design of replicated studies. Am Statist 1993; 47(3): 217–228
53. Rosenthal R. Experimenter effects in behavioral research. Irvington, 1976
54. Rosenthal R. Replication in behavioral research. In: Neuliep J (ed). Replication research in the social sciences. Sage, Beverly Hills, CA, 1991
55. McGrath J. Dilemmatics: the study of research choices and dilemmas. In: McGrath J, Martin J, Kulka R(eds). Judgement calls in research. Sage, Beverly Hills, CA, 1982
56. Brewer J, Hunter A. Multimethod research: a synthesis of styles. Sage, Beverly Hills, CA, 1989
57. Jenkins A. Research methodologies and MIS research. In: Mumford E et al (eds). Research methods in information systems. Elsevier, Amsterdam, 1985
58. Galliers R, Land F. Choosing appropriate information systems research methodologies. Commun ACM 1987; 30(11): 900–902
59. Olson M, Ives B. Measuring user involvement in information systems development. In: Proceedings of the international conference on information systems, 1980, pp 130–143
60. Olson M, Ives B. User involvement in system design: an empirical test of alternative approaches. Inform Manage 1981; 4: 183–195
61. Kerlinger F. Foundations of behavioral research. Holt, Rinehart & Winston, 1986
62. Nunnally J. Psychometric theory. McGraw Hill, New York, 1978
63. Straub D. Validating instruments in MIS research. MIS Q 1989; June: 147–169
64. Ghiselli E, Campbell J, Zedeck S. Measurement theory for the behavioral sciences. Freeman, 1981
65. Allen M, Yen W. Introduction to measurement theory. Brooks/Cole, 1979
66. Bailey J, Pearson S. Development of a tool for measuring and analyzing computer user satisfaction. Manage Sci 1983; 29(5): 530–545
67. Davis F. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989; 13(3): 319–340
68. Ives B, Olson M, Baroudi J. The measurement of user information satisfaction. Commun ACM 1983; 26(10): 785–793
69. Doll W, Torkzadeh G. The measurement of end-user software involvement. OMEGA: Int J Manage Sci 1990; 18(4): 399–406
70. Baroudi J, Olson M, Ives B. An empirical study of the impact of user involvement on system usage and information satisfaction. Commun ACM 1986; 29(3): 232–238
71. El Emam K, Moukheiber N, Madhavji NH. An empirical evaluation of the G/Q/M method. In: Proceedings of CASCON '93. Centre for Advanced Studies, IBM Canada Ltd Laboratories, October 1993, pp 265–289
72. El Emam K, Madhavji NH. A method for instrumenting software evolution processes and an example application. Paper presented at the International workshop on software evolution processes and measurements, Victoria, Canada, 1994
73. Cronbach L. Coefficient alpha and the internal consistency of tests. Psychometrika 1951; September: 297–334
74. Subramanian A, Nilakanta S. Measurement: a blueprint for theory-building in MIS. Inform Manage 1994; 26: 13–20
75. Sethi V, King W. Construct measurement in information systems research: an illustration in strategic systems. Decis Sci 1991; 22: 455–472
76. Cohen J, Cohen P. Applied multiple regression/correlation analysis for the behavioral sciences. Erlbaum, Hillsdale, NJ, 1983
77. McIver J, Carmines E. Unidimensional scaling. Sage, Beverly Hills, CA, 1981
78. Osgood C, Suci G, Tannenbaum P. The measurement of meaning. University of Illinois Press, 1967
79. Yourdon E. Modern structured analysis. Prentice-Hall, Englewood Cliffs, NJ, 1989
80. Zeller R, Carmines E. Measurement in the social sciences: the link between theory and data. Cambridge University Press, Cambridge, 1980
81. Barki H, Hartwick J. Measuring user participation and involvement. Technical Report, Ecole des Hautes Etudes Commerciales and the Faculty of Management, McGill University, 1992
82. Hawk S, Aldag R. Measurement biases in user involvement research. OMEGA: Int J Manage Sci 1990; 18(6): 605–613

83. Munro M, Davis G. Determining management information needs: a comparison of methods. MIS Q 1977; June: 55–67
84. Mathiassen L, Stage J. The principle of limited reduction in software design. Inform Technol People1992; 6(2–3): 171–185
85. Aldenderfer M, Blashfield R. Cluster analysis. Sage, Beverly Hills, CA, 1984
86. Armstrong J, Overton T. Estimating nonresponse bias in mail surveys. J Marketing Res 1977; XIV: 396–402
87. Siegel S, Castellan J. Nonparametric statistics for the behavioral sciences. McGraw Hill, New York, 1988
88. Finn D, Wang C-K, Lamb C. An examination of the effects of sample composition bias in a mail survey. J Marketing Res Soc 1983; 25(4): 331–338
89. Cronbach L. Statistical tests for moderator variables: flaws in analysis recently proposed. Psychol Bull 1987; 102: 414–417
90. Neter J, Wasserman W, Kutner M. Applied linear statistical models: regression, analysis of variance, and experimental designs. Irwin, 1990
91. Bohrnstedt G, Carter T. Robustness in regression analysis. In: Costner H (ed). Sociological methodology. Jossey-Bass, San Francisco, 1971
92. Stevens S. Mathematics, measurement, and psychophysics. In: Stevens S (ed). Handbook of experimental psychology. Wiley, New York, 1951
93. Galletta D, Lederer A. Some cautions on the measurement of user information satisfaction. Decis Sci 1989; 20: 419–438
94. Labovitz S. The assignment of numbers to rank order categories. Am Sociol Rev 1970; 35: 515–524
95. Labovitz S. Some observations on measurement and statistics. Social Forces 1967; 46(2): 151–160
96. Baker B, Hardyck C, Petrinovich L. Weak measurements vs. strong statistics: an empirical critique of S. S. Stevens' proscriptions on statistics. Educ Psychol Meas 1966; 26: 291–309
97. Arnold H, Evans M. Testing multiplicative models does not require ratio scales. Organ Behav Hum Perform 1979; 24: 41–59
98. Gardner P. Scales and statistics. Rev Educ Res 1975; 45(1): 43–57
99. Velleman P, Wilkinson L. Nominal, ordinal, interval, and ratio typologies are misleading. Am Statist 1993; 47(1): 65–72
100. Briand L, El Emam K, Morasca S. On the application of measurement theory in software engineering. Empirical Software Eng 1996; 1(1) (to appear) (also as International Software Engineering Research Network technical report ISERN-95-04)
101. Wolf F. Meta-analysis: quantitative methods for research synthesis. Sage, Beverly Hills, CA, 1986
102. DeVellis R. Scale development: theory and applications. Sage, Beverly Hills, CA, 1991
103. Fry L, Smith D. Congruence, contingency, and theory building. Acad Manage Rev 1987; 12(1): 117–132
104. Wright G Jr. Linear models for evaluating conditional relationships. Am J Polit Sci 1976; XX(2): 349–373
105. Votta L, Porter A. Experimental software engineering: a report on the state of the art. In: Proceedings of the 17th international conference on software engineering, 1995
106. Curtis B. Measurement and experimentation in software engineering. In: Proc IEEE 1980; 68(9): 1144–1197
107. Shaw M. When is 'good' enough? Evaluating and selecting software metrics. In: Perlis A, Sayward F, Shaw M (eds). Software metrics: an analysis and evaluation. MIT Press, Cambridge, MA, 1981
108. Hirschheim R. Assessing participative systems design: some conclusions from an exploratory study. Inform Manage 1983; 6: 317–327
109. Hirschheim R. User experience with and assessment of participative systems design. MIS Q 1985; December: 295–304
110. Simon H. The new science of management decision. Harper & Row, New York, 1960
111. Burns R, Dennis A. Selecting the appropriate application development methodology. Data Base 1985; Fall: 19–23
112. Miller G. A psychological method to investigate verbal concepts. J Math Psychol 1969; 6: 169–191
113. Cochran W. Some methods for strengthening the common chi-square tests. Biometrics 1954; 10: 417–451
114. Lewontin R, Felsenstein J. The robustness of homogeneity tests in 2 × N tables. Biometrics 1965; 21: 19–33
115. Pedhazur E. Multiple regression in behavioral research. Harcourt Brace Jovanovich, 1982
116. Kmenta J. Elements of econometrics. Macmillan, New York, 1986