Acceptance Issues in Metrics Program Implementation

Medha Umarji, Henry Emurian
Dept. of Information Systems, Univ. of Maryland Baltimore County
Baltimore, Maryland, USA.
medha1, emurian @umbc.edu

Abstract

Process improvement initiatives such as metrics programs have a high failure rate during their assimilation in a software organization. Lack of attention to social issues and lack of communication are some of the factors affecting their acceptance. Social Psychology includes study of such acceptance issues in the adoption of technology in the workplace. Although metrics implementations are often more intrusive and their benefits are indirect compared to workplace technology in general, both are introduced for similar reasons and induce organizational change. Therefore, we apply technology adoption concepts to metrics program implementation to develop a predictive model that takes as input organizational culture, practitioner mindset, and the nature of the metrics program. Our model relates these inputs to the likelihood of a successful metrics program implementation and to areas that may need improvement. This paper includes the results of a pilot study. The predictor variables Ease of Use, Usefulness, Control, and Attitude were significantly correlated with the dependent variable, the respondent’s Intention to perform metrics activities.

1. Introduction

Software metrics and measurement programs are becoming increasingly important in organizations today. Software organizations constantly grapple with problems related to software engineering practices and there is a need for continuous improvement of software processes. However, it is problematic to improve a process without measuring it first. Therefore, measures of software engineering processes and products are collected and analyzed and feedback for process improvement is communicated to the practitioners and managers [18].

Metrics and metrics programs have been studied extensively in the literature. Although their merits are acknowledged universally, it is found that most metrics programs (50-80%) do not continue beyond the second year [23]. Metrics programs are perceived as negative by most software organizations [6] and two in three metrics implementations fail [20].

A majority of the metrics programs fall slack after the initial assimilation phase when the success of the program hinges on the communication of improvements through the feedback loop. The problem lies in disseminating the rationale of measurement in a manner specific to the organizational context, personnel, and work practices. We believe that it is important to be able to predict where the adoption process can go wrong and focus on these problem areas so that huge investments of time and resources will not be futile.

We draw a parallel from the domain of Social Psychology, particularly the study of acceptance of technology in the workplace. The Technology Acceptance Model (TAM) [5], Theory of Planned Behavior (TPB) [1], Perceived Characteristics of Innovating (PCI) [14], and other models have addressed various problems faced when a new system is introduced into the workplace. The TAM, especially, has been applied in many usage scenarios and has performed consistently well [9], [21].

We believe that constructs from these models, such as Usefulness, Ease of Use, and Attitude, are applicable to the context of software metrics programs. Accordingly, we treat metrics programs as a new technology, select the relevant constructs and tailor them to model the adoption of metrics programs, with inputs from existing case studies and anecdotal evidence. Some extensions to these constructs are also proposed, by taking into account the unique nature of software organizations, the pervasive changes caused by software process improvement, and the unique characteristics of software practitioners as a user group. We have designed and conducted a pilot study to test and validate the model. The results of this study
are presented here, along with some interpretations and proposed modifications to the model.

The rest of the paper is organized as follows: Section 2 reviews the related literature on metrics programs and technology acceptance theories in Social Psychology. Section 3 gives a detailed explanation of the proposed solution to the problem (i.e., the Metrics Acceptance Model). Section 4 describes the variables, test items, and the research methodology of our study. Section 5 discusses the results from statistical analyses, and Section 6 presents our conclusions and suggestions for future work.

2. Related literature

Based on the existing literature, it can be said that software measurement and metrics programs should be implemented by every organization that wants to improve its processes. However, actual industry experience in this area tells a different story [4], [15], [16]. Fifty to 80% of all metrics initiatives have a life span of around two years [23]. This section presents the literature related to metrics programs and outlines common causes of failure.

2.1 Metrics programs

In a seminal study, Jeffrey and Berry [11] developed an evaluation framework for prediction of success in metrics programs consisting of four categories: organizational context, input resources, the metrics implementation process, and products such as benefits and feedback. They analyzed metrics programs across three organizations in order to validate this framework, and concluded that all four aspects are important and their interactions should be studied in more detail through empirical research.

Hall and Fenton [8] compared metrics programs across two organizations, and reported that managers are generally more enthusiastic about metrics than practitioners. Practitioners said that metrics are useful, but when asked to participate, they found many reasons why their work must be exempt. This phenomenon of positive perceptions and negative actions warrants more research. Hall and Fenton also found that practitioners lacked motivation to collect metrics data because practitioners disbelieve the accuracy of the resulting data.

Pfleeger [19] presented lessons learned while building a metrics program for Contel Corporation. This paper asserted that the biggest problem in establishing a metrics program was convincing practitioners that it is worth their time and effort. Also, people felt threatened by metrics because metrics can be used as a commentary on their performance; and so, if practitioners do not want to collect metrics, it is risky to make them do so. Automated and non-intrusive data collection was highly recommended. It was found to be important to keep the metrics “close” to the practitioners by involving them in analyzing and acting upon metrics data. At Contel Corporation, metrics advocates were groomed to spread success stories to motivate people, a technique that proved very successful.

Gopal et al. [7] conducted an industry wide survey of metrics experts to determine the success factors for metrics programs. They reported that uses of metrics in decision-making and improved organizational performance are key indicators of success. Also, success factors discussed in the literature—such as ease of data collection, metrics quality, management support, and communication and feedback—are supported by the survey results.

Some factors responsible for failure are discussed next. One common cause for failure is a communication gap between practitioners and management. At the start of any initiative, management tends to make higher claims than they can actually fulfill, perhaps intending to get people enthusiastic and involved in the process [10]. Problems in communication may also cause practitioners to perceive the organization’s stance as more intimidating than it actually is [19]. Another problem is the lack of alignment between the goals of the practitioners and those of the organization [7]. Data fabrication is also an issue [8]. Under the pressure to show process improvement, practitioners (in some cases encouraged by their managers) report data that does not reflect reality, but instead portrays what the management wants to see. This information is actually detrimental to the organization and may give a false picture of well-being. Lack of funding is another concern. If a certain project is funded, this helps to increase the credibility of the project. If metrics programs have been started as a result of downsizing and cost cutting, there is a possibility that insufficient funds will be available [16].

In order to resolve some of the issues mentioned above, metrics programs should be treated with an organizational change perspective. Hence, we also explored another body of literature related to the diffusion of new technology in an organization.

2.2 Social psychology literature

The Technology Acceptance Model (TAM) is a parsimonious and powerful model, based in Social Psychology, for predicting the usage of an Information System (IS). TAM claims that Ease of Use and
Usefulness can predict usage of a system. Attitude and behavioral Intention act as the mediating variables in this model. Ease of Use is “the degree to which the … user expects the target systems to be free of effort” [5, p.985]. Usefulness is “the user’s subjective probability that using a specific application system will increase his or her job performance within an organizational context” [5, p.985]. Attitude is defined as the user’s evaluation of desirability of using the system [22]. In the TAM, Usefulness is influenced by Ease of Use and both have an effect on Attitude. Also, Attitude and Usefulness together influence Intention to use the system. Intention has a direct impact on the actual system usage.

TAM has been tested in numerous situations with different types of information systems. Over time, TAM has been extended to account for the influence of perceived user resources, affective and cognitive dimensions of attitude, social influence processes and cognitive instrumental processes [13], [14], [24]. All these factors have also been found to significantly influence user acceptance.

The Theory of Planned Behavior (TPB) is a more general theory that has been applied to many domains, including IS, to predict actual behavior based on behavioral intention. The TPB suggests that Intention is directly determined by three factors: Attitude toward the behavior, the Subjective Norm, and Perceived Behavioral Control.

Previous research has found that when constructs of these or similar models are combined, they show better predictability compared to that of each model alone [9], [14], [21]. If the model can be made context-specific, its predictive power will improve. Recently, empirical studies have been conducted to test these models in software organizations to determine how well they can predict software practitioner acceptance of methodologies such as the object-oriented paradigm [9], [17].

3. Metrics acceptance model

We have created a predictive model that takes as input the organizational culture, practitioner mindset, and nature of the metrics program and outputs the probability of success of a metrics program along with areas that need improvement.

The model that we have developed (refer to Figure 1) captures all the main drivers of success, recommended by the literature, and assesses the organization against these factors from a practitioner viewpoint. The term practitioner mainly refers to those who are active in the software development process and are providing data about their activities and the software they produce to the metrics program. We envision that our model would be applied in the initial planning or decision-making phases of the metrics program. A questionnaire, based on our metrics acceptance model, would be administered to the software practitioners expected to participate in the metrics program and would capture their thoughts and impressions about the organization, the changes in work practices that will be caused by the metrics program, and their willingness to adapt to the change. The results would indicate the probability of acceptance of the metrics program by the practitioners, as well as what issues are likely to be obstacles to acceptance.

We term this model as the “Metrics Acceptance Model” (MAM). The MAM includes some constructs from the models in Social Psychology, as well as concepts suggested by the literature on software metrics. In addition, based on our understanding and experience of the metrics domain, we have also added some components to the MAM. We have divided the constructs under four main categories: Usefulness, Ease of Use, Attitude, and Control. A brief description of each is provided along with a list of the parameters that help to determine each category.

3.1 Usefulness

In our case, Usefulness can be viewed from two perspectives: organizational Usefulness and personal Usefulness. Organizational Usefulness is the practitioner’s perception of how their participation in the metrics program will be beneficial to the organization (and indirectly to the practitioner). Personal Usefulness is the practitioner’s perception of how performing metrics activities will enhance their
job satisfaction and career prospects. We adopt the claim that Usefulness will be a significant determinant of Intention to perform metrics activities. We describe below some of the indicators of Usefulness that we hypothesize will eventually affect the Intention to adopt and sustain the metrics program.

Effective communication of goals and practice: This helps employees in better understanding the scope and importance of the program. Their doubts and apprehensions have to be addressed in order to convince them to participate, which can only be done through constant communication. Reasons for introducing the metrics program should be strong and the goals of doing so should be well-defined. To foster support, practitioners have to be convinced that such a program is necessary and is in their best interests [10], [11].

Descriptive norm: This is one component of Subjective Norm, which is defined in the TPB, and suggests that social pressure can influence an individual’s intention to use a system [22]. Also, Subjective Norm, which has to do with the attitudes of a group of co-workers, plays a significant role in metrics programs. If one practitioner feels that the metrics program is not effective or worthwhile, they may influence people around them and this attitude begins to spread through the group. This would potentially cause a slack in the general energy and interest.

Injunctive norm: This is the other component of Subjective Norm, and it is the notion that if people within the social network of a person want him/her to perform a certain behavior that they consider to be useful, eventually his/her motivation for performing that behavior will be positively influenced. Also, people tend to believe their managers, their superiors, and experienced colleagues more than theory or books [22].

Importance of work: If the practitioners believe that the project they are working on is important in terms of its impact on the organization and team benefits, their idea of Usefulness will be enhanced. They should be made to believe that data about their project will help the organization learn something useful. This construct is not defined in prior literature and we introduce this concept here because we feel, based on experience, that it could explain, to some extent, a practitioner’s impression of Usefulness.

Image and visibility: These constructs refer to the social status of a person. In an organizational setting these issues are very important to a person’s career. TAM2, an extension of the TAM [24], suggests that “an individual may perceive that using a system will lead to improvements in his or her job performance indirectly due to image enhancement”.

Reward structure: This is an important determinant of personal Usefulness because one major problem with a metrics program is that the efforts are immediate and the rewards are long-term. Thus, if immediate efforts are rewarded immediately, it would help to motivate people to participate in the metrics program. It is important to note, however, that these perquisites should not just be offered at the start of the program; it should be a continuous effort.

Fear of adverse consequences: This is a factor that can have a negative influence on personal Usefulness. People do not like to report low productivity about themselves or their colleagues as this might affect relationships with them. Also, they are afraid of reporting any failings on the part of their managers as it might affect their own career opportunities. There is also the issue that people do not like to be measured and evaluated by others, especially if this is done by new people who are brought in for this task.

Based on the above discussion, the expectation is that communication, descriptive and injunctive norms, importance of work, image and visibility, reward structures, and fear of adverse consequences will account for a major portion of the Usefulness construct.

3.2. Ease-of-Use

Ease of Use refers to the ease with which the practitioners can adapt to the change in work practices caused by the metrics program and use the tool or set of tools required in the implementation of the metrics program.

Compatibility: Compatibility of the metrics tool with the existing systems and tools can make it easier for the practitioners to adapt to its use. Also, compatibility of the changes induced by the metrics program with the existing work practices will have an effect on the Ease of Use.

Easy to use tool: An intuitive and user-friendly tool is essential for users to interact effectively and efficiently, and may also have an impact on the self-efficacy of the users. The tool should be non-intrusive in nature [19]. Another issue here is customizability, because it is important to know whether the tool can be easily customized to the needs of each individual project or not.

Learning: The amount of learning required to work with the tool is another issue that is closely tied to the Ease of Use of the tool. If a lot of training is needed then practitioners are likely to think that the metrics tools are difficult to use.

Ease of data collection: Most prior studies have insisted that the data collection should be automated [8], [19]. The frequency with which data is collected
can also impact the Ease of Use because it may be tedious to collect data very often. A variety of data collection modes—such as automatic, manual data entry, interviews, and surveys—may make the data collection process time-consuming and tedious. This also closely ties in with the effort expended on collecting the data.

Hence, we hypothesize that compatibility, intuitive and user-friendly tool, amount of learning required and ease of data collection all contribute to the Ease of Use construct in our model.

3.3 Control

This signifies the control that the practitioner has over the measurement program. It can be accounted for by taking the following factors into consideration:

Volitional control: It is the degree of voluntariness of using a system. Ideally, a person should participate in the program by his/her own volition. Mandatory activities are part of organizational policy and outside the control of the employees [24].

Perceived behavioral control (PBC): This refers to both personal and situational impediments to performing a behavior [2]. PBC includes both external and internal factors. External factors are the user resources such as relevant documentation, availability of a help facility, reliable and experienced personnel, adequate time, financial stability, and enough documentation. Internal factors refer to the personal ease or difficulty of performing the behavior, based on one’s prior experiences, skills gained through learning, and intelligence. PBC also has an impact on the overall feeling of control and will eventually affect the Intention to use/actual usage of the system. However we do not measure internal PBC factors explicitly in this model, because we measure self-efficacy, which proxies for measures of internal PBC factors [2], [3].

Control over the measurement process: The control that the user has over the metrics tool in terms of the access to metrics data and overall tailorability of the tool will also have an impact on the overall feeling of Control. The suggestions that each person can give while designing and implementing the program and the degree to which that feedback is incorporated into the measurement process will also impact the feeling of Control over the program, and this may eventually tap into the Intention to participate in the program.

Self-efficacy: Self-efficacy is defined formally as “the belief in one’s capabilities to organize and execute the sources of action required to manage prospective situations” [3]. Self-efficacy is introduced into this study to include a measure of the individual practitioner’s assessment of his/her capability of performing metrics activities. Self-efficacy is a sub-component of the Perceived Control concept of the Theory of Planned Behavior and is an optimal predictor of intention [3].

It is hypothesized that all these factors affect the Intention to participate in the metrics programs through the Control construct.

3.4 Attitude

Attitude “is the overall evaluation of performing the behavior, by the individual” [22, p. 130]. Attitude has two sub-components: affective (e.g., happy-sad) and cognitive (e.g., beneficial-harmful). In this model, we explore the influence of affective Attitude on Intention, because the cognitive component is already addressed in other concepts, such as Usefulness. It has been shown in previous literature that Attitude is a significant determinant of Intention to perform a behavior [5].

3.5 Intention

“Intentions are conceptualized to capture the motivational factors that influence a behavior” [22, p.129]. Behavioral Intention is the closest determinant of actual behavior, and since we have a predictive model, we chose to have Intention as our dependent variable, because actual performance would be difficult to measure before actually implementing the metrics program.

By addressing these issues, we try to capture all possible factors that can hamper the software measurement process. The model is then converted into a survey instrument, with items for measuring each of the constructs described above.

4. Methodology

Our research plan calls for a survey methodology as a vehicle for testing the metrics acceptance model. However, we have a number of items and new constructs, suggesting that a pilot study must be conducted to validate the model and the associated questionnaire.

The pilot study sample consisted of two groups of people. Group 1 consisted of 9 software practitioners working in one organization that had an already well-defined and operating metrics program in place. Group 2 had 15 practitioners from different organizations. Some of the organizations within Group 2 had a metrics program in place and some did not.

In order to provide a relevant context for the subjects who were not exposed to a metrics program in
and are expected to show high inter-question correlation. We tested for the reliability of the measures and report Cronbach’s alpha. We report the alpha value $\alpha$ for each sub-construct in the set of tables 1a through le in the next section.

The goal of the pilot study was to test and refine the questionnaire, judge the effectiveness of the methodology, and decide which analysis techniques are most appropriate for this work. In the pilot study, we work on testing that each construct is sound and well-defined by its indicators. In order to do this, confirmatory factor analysis is employed with principal components analysis as the method for extracting the factors and varimax rotation for facilitating interpretation of the results.

We would have liked to validate the entire metrics acceptance model based on the small number of data points that we have. In behavioral research, the norm is to have the ratio of data points to variables as 10:1 or a minimum of 100 data points [7]. Since we have only 24 subjects, it may be difficult to confirm the model based on the data because the model may lack stability, with variables switching factors as new data is added. Therefore, our deductions and inferences are more local, with respect to our proposed constructs and latent variables. However, we do perform correlation and regression analysis to test the relationships between the composite constructs and the dependent variable. The results are very encouraging and are presented in the next section.

5. Results and discussion

As is evident from the preceding discussion and in Figure 1, there are four main constructs in the Metrics Acceptance Model: Usefulness, Ease of Use, Control, and Attitude. Each has many sub-constructs that contribute to it and each sub-construct is measured by a group of indicators or “observed variables”. Many of these indicators are borrowed from the literature, but others are introduced by us to infuse metrics program knowledge into the model. Therefore, we run confirmatory factor analysis on the indicators of each construct to test whether the factor loadings conform to our hypotheses about the sub-constructs and their implications on the main construct. Each sub-construct has two to five items (i.e., survey questions) that are used to measure it. We use principal components to confirm the loadings of items on these proposed theoretical constructs. Reliability testing of the items within each factor is performed, and the alpha values are reported in the following tables (a). The factor loadings are shown in Table 1(a through e).
Table 1a. Usefulness items and factor loadings

<table>
<thead>
<tr>
<th>Communication (comm) a = 0.790</th>
<th>(a = 0.790)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) The management has conveyed all their policies related to metrics openly and clearly</td>
<td>.878</td>
</tr>
<tr>
<td>b) The management takes care to address any doubts about the metrics program</td>
<td>.772</td>
</tr>
<tr>
<td>c) There is a lot of discussion in the office, regarding the metrics program</td>
<td>.677</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Descriptive Norm (org_policy) a = 0.752</th>
<th>(a = 0.752)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) The management has valid, well-defined reasons for introducing the metrics program</td>
<td>.773</td>
</tr>
<tr>
<td>b) The suggestions made by people about metrics activities are received well by the management</td>
<td>.729</td>
</tr>
<tr>
<td>c) The team responsible for the metrics program is neutral and unbiased</td>
<td>.693</td>
</tr>
<tr>
<td>d) I know the people of the team responsible for the metrics program well</td>
<td>.622</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Injunctive Norm (inj_norm) a = 0.702</th>
<th>(a = 0.702)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) The opinion of the people in my team/workplace is important to me</td>
<td>.675</td>
</tr>
<tr>
<td>b) The metrics program involves more interaction with my colleagues and co-workers</td>
<td>.681</td>
</tr>
<tr>
<td>c) People that are important to me believe that the metrics program is useful and want me to participate in the metrics program</td>
<td>.868</td>
</tr>
</tbody>
</table>

Table 1b. Usefulness items and factor loadings

| Importance of Work (imp_work) a = 0.903 |
|----------------------------------------|-------------|
| a) My project is important to the organization. | .956 |
| b) Metrics data about my project will help the organization to learn something useful | .877 |

| Image and Job Security (a = 0.703) |
|------------------------------------|-------------|
| a) People in my organization who perform the metrics activities have more prestige than those who do not. | -.74 |
| b) I felt less secure about my job after hearing about this metrics program | .863 |

| Organizational Usefulness (org_use) a = 0.733 |
|-----------------------------------------------|-------------|
| a) The benefits to the organization of my participation in the metrics program would substantially outweigh the costs to the organization | .750 |
| b) My participation in such a metrics program will make my organization more successful | .881 |

| Fear of Adverse Consequences (work_prac) a = 0.775 |
|---------------------------------------------------|-------------|
| a) My role in the metrics program might affect my relationship with my co-workers | .694 |
| b) I may not be willing to report data that may be harmful to someone, including myself. | .866 |
| c) I have to make some major changes in my way of working to adapt to the metrics tool | .794 |

| Fear of Adverse Consequences (metrics_fail) |
|---------------------------------------------|-------------|
| Failure to perform metrics activities well will affect my career | .837 |

5.1 Usefulness

This construct has 7 sub-constructs corresponding to each of the issues relevant to the Usefulness, as discussed above. The 22 items for this construct loaded on 9 factors.

The three items for communication loaded on one principal component, \(\text{comm}\), as we had hypothesized. The factor \(\text{comm}\) is about how effective the management policies are and how open management is to suggestions and doubts. Communication is also facilitated by discussion about the metrics program amongst the employees.

The three items for importance of work loaded on two principal components. Importance of a project and metrics data related to it loaded highly on one factor \((\text{imp_project})\). This supports our claim that if the practitioners perceive their project work to be important to the organization, they will be more than willing to share information about it. However, one item about the uniqueness of their work did not load on this factor, indicating that, if a respondent’s project is unlike any other project in the organization, the respondent will be reluctant to share data about it. This uniqueness item loaded into a factor by itself \((\text{unique})\), and was negatively correlated with Usefulness, suggesting that the projects are very similar across organizations and practitioners are aware of this.

However, this question may be biased by the fact that respondents may not know enough about other projects in their organization. Since the construct is sufficiently described by the two other items, the next version of our survey will not include the uniqueness item and hence it is not displayed in the set of items above.

Two items for reward structure did not load highly on any factors and need to be explored further, with more data. Three items under fear of adverse consequences loaded on one factor \((\text{work_prac})\). All three items are related to how the changes in work practices will affect relationships and the way of working and this is correlated with Usefulness, thus supporting our idea. Also, the fear that failure to perform metrics activities will influence career growth loads highly on one factor that we call fail_metrics. This supports our premise that career consequences are correlated with Usefulness.

Prestige associated with metrics activities was negatively correlated with the job security, which can be interpreted as follows: people who are secure and happy about their jobs do not think that there will be any increase in status whether or not they perform metrics activities. Prestige is a part of the image and visibility construct related to personal Usefulness and
job security is a part of fear of adverse consequences but both contribute to the Usefulness.

Next, the items for descriptive norm, reasons for introducing the metrics program and the metrics team loaded on one factor that we call org_policy. as all of them have to do with organizational policies. For example, the metrics team is appointed by the management and reflects their interest. Also, if the organization can state valid and well-defined reasons for introducing the metrics program, it will help to convince the practitioners of the need and utility of this program.

The items for injunctive norm loaded highly on one factor (inj_norm), which means that they are cohesive and highly correlated with Usefulness. The factor supports the claim that social influence plays a role in increasing Usefulness. The two items related to organizational Usefulness loaded highly on one factor (org_use). This also supports our claim.

Based on the above discussion, we can see that most of the items related to Usefulness loaded high on very relevant factors and this supports our model. There are some other items that did not have factor loadings greater than 0.6 and we did not consider them for our analysis. This is because we have very few data points, and the weak loadings may shift between factors with more data. Therefore, it is premature to make generalizable statements based on them.

5.2 Ease-of-Use

The 11 items for this construct loaded highly on four factors. These factors correspond to the theoretical sub-constructs we had designed, but there are some minor modifications. Compatibility of the metrics program and the metrics tool with existing applications and work practices comprised one construct. We found from the factor analysis that both the items related to compatibility load highly on one construct that we term comp.

There were four items for determining the ease-of-use, flexibility and customizability of the metrics tool, and three of them loaded highly on the factor ease_tool. The item for testing the clarity of the interaction with the metrics tool did not load highly, and we suspect this may be an artifact of the low number of data points we have available.

The items corresponding to the amount of learning required and ease of data collection both loaded on one factor (learning_dc). We think this may be because learning to operate the metrics tool is synonymous with learning to collect data, since most metrics tools are used for collecting data. Automation of the data collection further facilitates both these processes.

<table>
<thead>
<tr>
<th>Table 1c Ease of Use items and factor loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compatibility (comp) ( a = 0.664 )</td>
</tr>
<tr>
<td>a) The metrics program activities are compatible with the existing work practices in the organization</td>
</tr>
<tr>
<td>b) The metrics tool is compatible and well integrated with the existing tools and applications</td>
</tr>
<tr>
<td>Metrics Tool (ease_tool) ( a = 0.837 )</td>
</tr>
<tr>
<td>a) It is difficult to get the tool tailored to my type of work</td>
</tr>
<tr>
<td>b) I find the metrics tool flexible to interact with</td>
</tr>
<tr>
<td>c) I find the metrics tool easy to use</td>
</tr>
<tr>
<td>Ease of Learning &amp; Data collection (learning_de) ( a = 0.831 )</td>
</tr>
<tr>
<td>a) Learning to operate the metrics tool is easy for me</td>
</tr>
<tr>
<td>b) It is easy for me to become skilled at using the metrics tool</td>
</tr>
<tr>
<td>c) The data collection process is automated</td>
</tr>
<tr>
<td>d) I agree that the data is collected with the appropriate frequency</td>
</tr>
<tr>
<td>Effort expended on Data collection (Effort_dc)</td>
</tr>
<tr>
<td>The proportion of effort on DC is appropriate</td>
</tr>
</tbody>
</table>

5.3 Control

The 12 items under this construct load highly on four factors. Again, this confirms our modeling of the control construct.

Volitional control is the degree of voluntariness of usage. There is only one item for measuring this and it loads highly on one factor along with the five items for self-efficacy. These two concepts seem to tie in neatly, because both reflect different aspects of internal control that a person feels over metrics activities.

For perceived behavioral control, we found that access to well qualified people, access to data and some contact to ask general questions all load highly on one factor org_resources. The items for measuring the schedule and funding both loaded highly on one factor. We interpret this result to reflect the reality that schedule and cost may not always be within the control of the organization. They are more often determined by the customer demand and market fluctuation. Therefore we recognize this as a separate factor and label it critical_resources.

The items relevant to control over the measurement process load highly on one factor and we call it control_process. We claim that the above five factors determine the Control construct.

5.4 Attitude

During factor analysis, the items for attitude loaded on two factors. The wording of the questionnaire should have been closer to the practitioner state of mind, like “Relaxed/Stressed” instead of “Good/Bad”. Another possible explanation can be the lack of enough data points. We plan to explore this construct more thoroughly.
5.5 Intention

Intention is the dependent variable and is measured by the above items. The relations of the independent variables with Intention are discussed in the previous sections.

5.6 Correlation and Regression Analysis

Table 2 Pearson’s Correlation Coefficient

<table>
<thead>
<tr>
<th></th>
<th>Usefulness</th>
<th>EaseOfUse</th>
<th>Control</th>
<th>Attitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intention</td>
<td>0.555**</td>
<td>0.557**</td>
<td>0.618**</td>
<td>0.409*</td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

We compute correlations between the four composite constructs and the dependent variable, Intention. All the constructs (Usefulness: $r = 0.555$, Ease Of Use: $r = 0.557$ Control: $r = 0.618$ and Attitude: $r = 0.409$) are significantly correlated with Intention.

With the above discussion, we have justified the contents of each construct and its relationship with the dependent variable and shown that this is consistent with our hypotheses. Each of the constructs (Usefulness, Ease of Use, Attitude and Control) that we have used here has been studied extensively in the literature and has been shown to have a significant impact on the Intention to use a system. Our results lend support for this approach within the metrics implementation program domain.

6. Summary

Many empirical studies in metrics implementation urge organizations to motivate and seek active involvement of practitioners, automate data collection, and integrate the measurement process seamlessly into the work culture. We would assume that all organizations work towards these goals, but evidently there is still a lot to be done. This study is a positive step in this direction that will help an organization to gauge their progress in these areas and also to involve the relevant stakeholders in the metrics initiative by soliciting their opinion. For the purposes of analysis, our survey instrument is entirely quantitative. This limits respondents’ choices in their answers to questions, and does not allow conditional, or contingent, answers (e.g. the response to a question might depend on other factors outside the scope of the survey). Thus, in practice, the use of the survey should be augmented with interviews or other more qualitative methods to gain information of this kind.
The results from this study support the proposed model, with minor modifications. It was found after the reliability analysis that some questions were ambiguous and need to be reworded. We maintained most questions because we do not have enough cases to decide that an item is not required. On the whole, the model remained stable.

We have learned that each of the variables (Usefulness, Ease of Use, Control, and Attitude) is strongly, positively correlated with Intention. We know the factors that determine each variable and, to some extent, the strength of these relationships. Therefore, we can point out, in a particular organization, which correlations are not present or not strong enough, with some certainty. Consequently, we have a predictive and a problem-finding mechanism for metrics programs from a behavioral standpoint. Our next step is to conduct a full scale, longitudinal study that will help in making the model more robust and generalizable. Also, we plan to make this questionnaire available for organizations, packaged with suggestions for remediying the problems that can be predicted by our model.

References