User Experience Metric and Index of Integration: Measuring Impact of HCI Activities on User Experience

Anirudha Joshi
Indian Institute of Technology Bombay
Mumbai 400076, India
+91 9820345569
anirudha@iitb.ac.in

Sanjay Tripathi
Tech Mahindra Ltd.
Pune 411004, India
+91 9922963298
stripathi@techmahindra.com

ABSTRACT
We propose two metrics to demonstrate the impact integrating human-computer interaction (HCI) activities in software engineering (SE) processes. User experience metric (UXM) is a product metric that measures the subjective and ephemeral notion of the user's experience with a product. Index of integration (IoI) is a process metric that measures how integrated the HCI activities were with the SE process. Both metrics have an organizational perspective and can be applied to a wide range of products and projects. Attempt was made to keep the metrics light-weight. While the main motivation behind proposing the two metrics was to establish a correlation between them and thereby demonstrate the effectiveness of the process, several other applications are emerging. The two metrics were evaluated with three industry projects and reviewed by four faculty members from a university and modified based on the feedback.

Categories and Subject Descriptors
D.2.8 [**Software Engineering**]: Metrics – *process metrics, product metrics.*

General Terms
Measurement, Design, Human Factors

Keywords
User experience metrics, HCI-SE integration.

1. INTRODUCTION
Large contracted software development companies with tens of thousands of employees are often involved in a wide variety of software development projects, often in an off-shore mode. Managers of user experience (UX) groups in such companies need to track progress of each project and ensure the quality of deliverables. They are often required to juggle across projects a limited resource – the time of their best UX professionals. While there are numerous usability metrics to evaluate specific projects, there are few that allow organizations to easily track progress across projects. Our first proposal is product metric (UXM) that measures the user's experience with a product. The objective is to provide a summary measure of the user experience of a product that is independent of the domain, context of use, platform or the software development process, so that the manager is able to make judgments across projects.

I-USED'08, September 24, 2008, Pisa, Italy

Another challenge faced by UX groups is integrating HCI in established SE processes. The field of HCI has a large amount of literature on user-centred design methods, techniques and processes [1], [3], [17], [23] etc. These proposals are excellent demonstrations of how user centred design can result in improved user experience design. Unfortunately, there continue to exist major gaps between HCI and SE, in academics, literature and industrial practice. The IFIP working group 2.7/13.4 on User Interface Engineering remarks that 'there are major gaps of communication between the HCI and SE fields: the architectures, processes, methods and vocabulary being used in each community are often foreign to the other community' [7]. For example, while SE literature admits that communication with the customer is an unsolved problem, even recent editions of standard text books on software engineering such as [13] and [20] do not suggest use of established user study techniques like [1] during communication. Example projects shown in [13] and [20] seem to take HCI design lightly, prematurely and without following any process. A detailed critique of SE literature from an HCI perspective is presented in [11]. There have been several proposals to integrate HCI in SE process models (for example, [5], [12], [21]) but none have become popular in the industry. One reason could be concerns about return on investments. Though there is plenty of evidence of the a return on investment of usability activities in general [2], there is no direct evidence that shows that better integration of HCI activities in SE processes will lead to better products at less cost.

Contracted software companies often promise a level of process compliance to their clients. UX managers need summary measures of process compliance of their projects to ensure that the company lives up to its promise. One way would be to measure how integrated were the HCI activities with SE processes. Our second proposal is a process metric (IoI) that would be one such measure. If validated, IoI and UXM can also be used to demonstrate the return on investment on integration of HCI with SE – if higher IoI consistently leads to higher UXM, it makes sense to invest in better integration of HCI with SE.

The main objective of this paper is to share with other metrics researchers the lessons we have learned from attempting to incorporate UXM and IoI in live industrial projects.

We begin with an introduction to different attempts done in recent years on applying metrics in HCI. Next, our metrics proposals are described. Finally, the evaluation methodology used so far to analyse the results of study is described.

2. METRICS IN HCI

Metrics are thoroughly discussed in software engineering literature. Fenton and Pfleeger [4] describe measurement as "the process by which numbers or symbols are assigned to attributes of entities in the real world in such a way as to describe them according to clearly defined rules". Pressman [20] highlights the subtle difference between measurement and metrics – measurement occurs as the result of collection of one or more data points, while a software metric tries to relate the measures in some way. IEEE Standard Glossary [6] defines a metric as "a quantitative measure of the degree to which a system, component or process possesses a given attribute".

Though the word 'metric' itself is seldom used in the practice of usability, several measures are often used. Seconds taken to withdraw money from the ATM, the number of keystrokes to enter a word in a complex script, the number of errors made to complete a banking transaction or the percent of users who abandon the shopping cart on checkout are all examples of quantitative measures of the user experience afforded by the product. However, none of these are summary measures that can be used for apple-to-apple comparison across projects varying in domains, platforms and contexts. While several research papers talk about metrics related to usability and HCI, this paper only focuses on those that give a summary measure.

Lewis [14] used a rank based system of assessing competing products based on user's objective performance measure and subjective assessment. The metric is useful for a relative comparison between like products with similar tasks but it does not result in a measure that can be used across unlike products.

McGee [18] derives a single usability scale across users by including additional reference tasks. However McGee does not suggest how to derive a single measure for usability from measures for the different tasks. Further, this work is completely dependent on the technique of usability evaluation. This is not always practical in a global contracted software company striving to move up the HCI maturity ladder. The other limitation of this method is that it relies only on perception of users and ignores perspectives of other stakeholders, particularly the goals of business stakeholders.

Lin et al [15] propose the Purdue Usability Testing Questionnaire based on eight HCI considerations to derive a single weighted average score for usability. While the approach does lead to a single usability score, the selected eight considerations (compatibility, consistency, flexibility, learnability, minimal action, minimal memory load, perceptual limitation and user guidance) seem to be a mix of usability goals and heuristics that achieve those goals. Secondly, the weightage for parameters is to be assigned by the evaluator during the evaluation without consulting stakeholders. Thirdly, the listed eight considerations and the questions listed under each of them seem to be limiting and do not leave room for project-specific goals (e.g. "do it right the first time").

Sauro et al [22] proposed a 'single, standardized and summated' usability metric for each task by averaging together four standardized values for task time, errors, completion and satisfaction. Their calculation however is based on the equal weightage. Tasks, domains, users, contexts and platforms vary a lot and it does not make sense to give equal weightage in all contexts. Moreover, the metric ignores some aspects such as learnability and ease of use, which might be important in some contexts.

Measuring the wider notion of user experience (as opposed to usability) is relatively new concept in HCI and is attracting attention of the academic as well as the industrial world. Usability parameters are typically related to the processing of information or completion of tasks. However, affective reactions and emotional consequences play important role in the overall user experience [16]. In some product contexts, we may need to consider visceral, behavioural and reflective elements [19], aesthetics [25], enjoyment [10] and creativity [24].

None of the summary metrics mentioned above measure the experience of a product with reference to all user and business goals relevant to a product. Many are too complex to compute practically on an on-going basis in the industrial practice. They lack the flexibility required to serve the needs of a wide variety of projects or to mature with the UX group. And finally, there seems to be almost no work on measuring integration of HCI activities with SE processes.

3. USER EXPERIENCE METRIC

Fenton and Pfleeger [4] emphasise the importance of goals in metrics: "a measurement program can be more successful if it is designed with the goals of the project in mind". User experience goals are very important in driving the design of interactive products. They help speed up the design process, make the design activity more tangible and help evaluate the design. User experience goals can be understood easily, even by non-UX-professionals, and they have a significant overlap with business goals. Stakeholders outline the user experience goals and UX professionals fine-tune them on the basis of their knowledge and findings from user studies. User experience goals are (and should be) available early in a project – another plus when it comes to metric calculation in a practical situation.

We propose user experience metric (UXM), a product metric that measures the quality of user experience. The motivations are:

- to measure the user experience of a product in reference to its user experience goals
- to develop a flexible metric that can be applied across a variety of projects, irrespective of domain, context, platform, process model or usability technique
- to develop a flexible metric that that will mature with the organization
- and to compute the metric with minimal additional costs and efforts.

UXM is product metric on a scale of 0-100, where 100 represents the best user experience possible and 0 represents the worst.

UXM consists of these distinctions:

Goals: High-level user experience goals guide the design of interactive systems.

Parameters: Each high-level user experience goal is broken down into a set of parameters that help the designer to achieve and measure the higher-level goal in a direct manner. For example Learnability can have parameters like Conceptual model clarity, Language understandability, Minimal training time, Consistency with earlier version etc.

Weightage: Each goal has a weightage between 0-5 where 0 represents that the goal is not important, 2 represents the typical importance and 5 represents that it is very important. Further, each parameter under a goal also has a weightage attached.

Score: Each parameter has a score between 0-100, where 0 represents the worst possible user experience on account of that parameter and 100 represents the best possible user experience.

Guidelines: The purpose of the guidelines is to help evaluators assign a score to the parameters. Guidelines let the goal-setters express themselves better and interpret goals for the context of a project – e.g. *"'Consistency with earlier version' means all frequent and critical tasks from earlier version are unchanged."* Further, guidelines tell the evaluators when to assign which score: *"The interface clearly communicates the correct conceptual model. Strongly agree = 100, Weakly agree = 75, Neutral = 50, Weakly disagree = 25 and Strongly disagree = 0".*

Goals and parameters are a way to express the desired user experience and performance in the design. Though expressing user experience goals is a common activity in HCI design, there is no standard way of doing so. There are many ways to describe the high level user experience goals. For example, ISO 9126-1 describes usability in terms of understandability, learnability, operability and attractiveness [8]. ISO 9241 on the other hand defines usability as the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use [9]. Shneiderman [23] describes goals for user interface design in terms of five human factors central to evaluation: time to learn, speed of performance, rate of errors by users, retention over time and subjective satisfaction.

We adopt goals from Shneiderman as the high-level user experience goals for a product and express them as: learnability, ease of use, speed of use, error-free use, retention and subjective satisfaction. We added 'ease of use' to the list as we thought that it is an important user experience goal distinctly different from the other factors such as speed. We also generalized the expressions of some of the goals. For example we express 'time to learn' as 'learnability' since it allows for expression of other concerns such as understandability of language or clarity with which the interface communicates the conceptual model in addition to the time users take to learn the interface. We believe that this allows the designers to express a wider set of goals.

Our proposal of an initial set of goals and parameters are listed in Table 1. However, we must highlight that this is not a prescribed, exclusive set. We give the evaluators and stakeholders freedom to derive additional, relevant parameters that express their goals. Goals and parameters could be added, removed or combined according to the context of the project, the needs of the users, the vision of the stakeholders and UX professionals and to fit the terminology that the product development team is familiar and comfortable with. The initial list is meant to give users a starting point, while the flexibility is meant to allow the metric to mature with the experience of the organization using UXM.

As Shneiderman [23] states, 'a clever design for one community of users may be inappropriate for another community' and also, 'an efficient design for one class of tasks may be inefficient for another class'. Weightages express the relative importance of goals and parameters in the context of a project. For example, a product meant to be used several times a day by a call-centre agent is likely to have higher weightage for 'speed of use'. A one-time use product like a web site for visa application for a tourist might insist on learnability and error-free use. On the other hand, a life-critical product to be used in a operation theatre is likely to rate highly error-free use and may sacrifice learnability. A game would perhaps give highest weightage to 'subjective satisfaction'.

The evaluators and stakeholders assign the weightage to set the context of the project. Goal-setters should be aware that while it may be tempting to set a high weightage to each goal, it may not be necessary, practical, or even possible to achieve such a design. The weightages should reflect the priorities of the stakeholders and users. The weightage would also help prioritize usability evaluation activity – the highest rated goals and parameters must be evaluated more thoroughly, while the lower weighted goals could be perhaps evaluated by a discount method.

The process for computing UXM for a product has these steps:

Goal Setting: Early in the project, typically just after user studies but before design, an HCI professional and stakeholders identify goals and parameters for each goal, assign weightage to each goal and parameter and decide evaluation guidelines for the parameters.

Scoring: Immediately after a usability evaluation, one or more independent HCI professionals assign a score to each parameter of each goal. The usability evaluation could be either user-based (e.g. a usability test) or review-based (e.g. heuristic evaluation).

UXM Calculation: UXM is the sum of the weighted average of the scores of all goals. $UXM = \sum (W_g \times S_g / \sum W_g)$, where W_g is the weightage of a goal and $S_g = \sum (W_p \times S_p / \sum W_p)$ where W_p is the weightage of a parameter and S_p is the score of that parameter.

Scores of some of the parameters can be directly linked to the findings of the usability evaluations (for example, % of users who did not make errors while doing benchmark tasks, or % of users who thought the product was engaging). Other parameters may not be so easily linked numerically (e.g. conceptual model confusions discovered during a think aloud test or problems identified during heuristic evaluation). In such cases, evaluators consider the guidelines and their own experience to arrive at a score for each parameter. If there are multiple evaluators, a simple average across evaluators is deemed to be the score for a given parameter. Multiple evaluators assign scores independently to begin with. If there is a significant variation in their scores, the evaluators discuss the parameter and have the opportunity to converge their scores before the average is calculated.

In case of applications with multiple user profiles, separate UXM should be calculated for each profile. Calculation of UXM could be a part of every usability evaluation of the project, but we recommend that it should certainly be a part of the final usability evaluation, beyond which no design changes are planned for.

Table 1. An example UXM calculation

Goals	Weightage	Score
Learnability	4	78.6
Speed of use	2	77.1
Ease of use	3	65.6
Error free use	3	67.5
Retention	1	75.0
Subjective satisfaction	2	78.6
UXM Value		73.3

Goal Parameters	Weightage	Score
Learnability		78.6
Conceptual model clarity	3	75
Language understandability	0	0
Minimal learning time	5	75
Consistency with earlier version	1	50
Visibility of choices and data	4	100
Consistency with other products	1	50
Speed of use		77.1
User control and freedom	3	75
No memory and cognitive load	4	100
Internal consistency	4	75
Customization	0	0
Automation and shortcuts	1	0
Ease of use		65.6
Minimal user task load	5	75
Automation of routine tasks	3	50
Error free use		67.5
Good feedback	4	75
Error tolerance	3	50
Error recovery	3	75
Retention		75.0
Retention	3	75
Subjective satisfaction		78.6
Visceral appeal	2	75
Behavioural appeal	4	75
Reflective appeal	1	100

Table 1 shows an example UXM calculated by a team for an Indic text input interface for novice users on a mobile phone. The team was given a default set of higher level goals, parameters and example parameter evaluation guidelines. The team first assigned the weightages for higher level goals (shown in the second column of the upper part of Table 1). Next, they broke down goals into parameters and assigned them weightages (shown in the second column of the lower part of Table 1). The team's experience from previous Indic text input projects and mobile phone projects helped them arrive at these weightages.

The team then evaluated the interface and assigned scores to each parameter (shown in the third column of the lower part of Table 1). A weighted average of parameter scores gave the score for each goal (shown in the light grey cells of the third column of lower part of Table 1). A weighted average of the goal scores gave the UXM value (shown in the dark grey cells of the upper part of Table 1). Parameter evaluation guidelines have not been listed in this paper due to space constraints.

4. INDEX OF INTEGRATION

We conceive Index of Integration (IoI) as an empirical process metric, nominally on a scale of 0-100, where 100 represents the best possible integration of HCI activities in the software development activities and 0 represents the worst. The metric consists of these distinctions:

Software Engineering Phases: These are the broad phases as described in a software engineering process model.

HCI Activities: These are prescribed for each phase of the software engineering process model.

Weightage: Each HCI activity will have a given weightage on the scale of 0-5 where 0 represents that the activity is not important, 3 reflects the typical importance in most projects and 5 indicates that this activity is very important in the context of that project.

Score: Each activity has a score associated with it. The score is given on a rating of 0-100, where 100 represents the best case situation where the activity was done in the best possible manner, in the most appropriate phase of software development and with the best possible deliverables. 0 represents the worst case situation where the activity was not done at all.

Activity evaluation guidelines: These spell out considerations that help the evaluation of each activity.

Software engineering phases have been extensively described in literature. For example, the phases of the waterfall process model are Communication, Planning, Modelling, Construction and Deployment [20].

On the other hand, no widely accepted industry-wide specifications of HCI activities for given SE phases have emerged so far. But there have been a few proposals. For example, [12] prescribes that the Communication phase of the waterfall model should have these HCI design activities: Contextual user studies and user modelling, Ideation, Product definition and Usability evaluation and refinement of product definition. Figure 1 summarizes the HCI activities suggested for the waterfall model phases based on these recommendations.

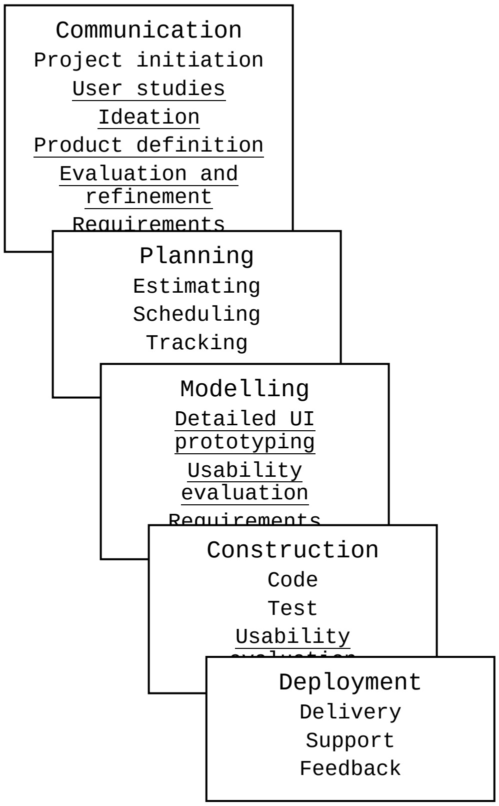

Figure 1. Integration of HCI activities with the phases of the waterfall model [12]. The HCI activities corresponding to each phase have been underlined.

Weightage of some HCI activities could vary within a range in the context of a project. For example, if the domain or users are unknown to the UX team, it may be very important to do contextual user studies in the communication phase (weightage = 4). On the other hand, if the UX team is already very familiar with the context and the domain and if they have a lot of experience designing similar products, it may be less important (weightage = 2). Table 2 summarises loosely recommended weightages for HCI activities for the waterfall model.

Guidelines may define the techniques used to carry out activities, the skills and experience levels of the people doing the activities, the deliverables and other parameters that affect the fidelity of the activity. For example, following are the guidelines for the activity of Contextual user studies and user modelling in Table 2:

1. *Organizational data gathering and user studies were done before requirements were finalised*
2. *User studies were done in the context of the users by the method of contextual inquiry*
3. *User studies were done with at least 10 users in each profile*
4. *User studies were done by people with experience in user studies in a similar domain in at least 2 projects*
5. *The findings including user problems, goals, opportunities and constraints were analyzed, documented and presented in an established user modelling methodology such as personas, work models, affinity diagram or similar*
6. *Competitive products and earlier versions of the product were evaluated for potential usability problems by using discount usability evaluation methods such as Heuristic Evaluation or better*
7. *User experience goals were explicitly agreed upon before finalizing requirements*

100 = All the above are true; 75 = At least five of the above are true, including 7, 50 = At least three of the above are true, including 7; 25 = At least two of the above are true, 0 = None of the above are true

Table 2. An example IoI calculation

SE Phases and HCI Activities	Weightage	Score
Communication		
Contextual user studies and user modelling	4	39
Ideation	2	6
Product definition	3	75
Usability evaluation and refinement of product definition	1	63
Modelling		
Detailed UI Prototyping	5	53
Usability Evaluation and Refinement of the Prototype	4	44
Construction		
Development support reviews by usability team	3	29
Usability evaluation (summative)	1	46
IoI Value		45

The process for computing IoI for a project has these steps:

Company HCI Process Prescription: The HCI group in the company prescribes what HCI activities should be done in which phase of SE process, expected deliverables from each activity, suggested weightages for each activity and suggested activity evaluation guidelines. As it often happens, a contract software development company may follow not one SE process, but several. In that case the HCI design process needs to be integrated with each SE process. The prescribed process also suggests a weightage for each HCI activity and guidelines to score each activity.

Project HCI Process Definition: After getting a project brief, the UX professional fine-tunes the weightages for the prescribed HCI activities after considering the domain, the users and project context. For example, if the HCI team has recently done contextual user studies in the same domain for a similar product and is already very knowledgeable about the context, then he may reduce the weightage of contextual user studies. On the other hand if the team is less knowledgeable, he may increase the weightage. He should consult colleagues in the development team and business stakeholders before finalizing the weightage.

Process Evaluation: After the project is over, a group of independent UX professionals review the HCI activities and evaluate them for process compliance and give a score for each activity on a scale of 0 to 100. They may reduce the score if an activity was done with lower fidelity, resulted in poor quality deliverables or was done later than prescribed. In case of multiple evaluators, an average across evaluators is deemed to be the score.

IoI Calculation: The metric is found by computing the weighted average of the scores of all activities: $IoI = \sum (W_a \times S_a / \sum W_a)$, where W_a is the weightage for a particular HCI activity, S_a is the score (from 0-100) for that activity. In case there is a lot of divergence in scores of a particular HCI activity, the activity is discussed and reviewers are given a chance to change their score before an average is taken.

Table 2 shows calculation of IoI for an example project. First, senior UX professionals defined the HCI activities, activity weightages and evaluation guidelines that the company should be following. Then a project that had recently ended was selected for retrospective review. The project manager and the UX professionals working on the project fine-tuned the weightages for the project context. The second column of Table 2 contains these weightages. A group of reviewers comprising of some project insiders and outsiders reviewed and rated the HCI activities in the project and its IoI was calculated. The third column of Table 2 contains the average scores assigned to each activity by the reviewers.

Guidelines for evaluating all HCI activities listed in Table 2 have been created. More guidelines for evaluating HCI activities as part of extreme programming process have been created as well. Both have been omitted here due to space constraints.

5. METRICS EVALUATION

The authors evaluated the metric in two ways. First, UXM and IoI metrics were computed for retrospectively projects in Tech Mahindra, a large contracted software development company. In each case, the metrics computation was done by HCI professionals from the project, independent HCI professionals and project stakeholders. At the end of metrics computation, feedback was taken from participants of each project about the metrics.

Second, the metrics were presented to a group of faculty members from a reputed university and their comments were noted. Three of these were faculty members from the computer science discipline. One was a faculty member from a design school who is also an expert in cognitive psychology.

5.1 Findings

It typically took about 3 hours to compute both IoI and UXM for each project. The time included explaining the two metrics, weightage assignment and scoring. This seemed to be optimum time, longer meetings were difficult to schedule. The projects performed similarly in IoI and UXM scores – the one project that had a high UXM value also had a high IoI value. Participants, particularly project stakeholders, were at home with the activity of metric calculation. To them, the activity seemed to bring HCI closer to SE. It seemed to create lot of buy-in for HCI activities from the project stakeholders. One project stakeholder said *"I never thought we could think so much [about user experience]."* The activity seemed to be more successful in projects where several stakeholders from the project participated as it stimulated discussion among stakeholders. While the participants appreciated the organizational perspective, the metrics seemed of less use to the projects as the projects were already over. Participants suggested that metrics should be calculated mid-way through the project while course correction was still possible.

Specifically, UXM helped the HCI designers and project stakeholders to make goals explicit. One HCI designer remarked, *"Had we done this earlier, I would have known where to focus."* The teams usually added a few goal parameters (typically 2-3 per project) and adjusted weightage to suit UXM to their project. They confirmed that this flexibility is indeed desirable. Though parameter evaluation guidelines for UXM helped, more details were desired. Participants did not make changes to the parameter evaluation guidelines except when new parameters were added. Giving examples of HCI goals (learnability, ease of use etc.) helped participants to set goal parameters and weightages. One stakeholder remarked: *"without these inputs it would have been difficult to [assign weightage and scores]."*

In case of a few UXM parameters, divergent scores emerged for some parameters in each project. Usually variations happened in parameters where the evaluation guidelines were not understood well or were interpreted differently by evaluators. In such cases, it was felt, that it was better to let participants discuss the parameter and change ratings to converge scores if they so desire. Reducing the number of steps in scoring a parameter (e.g. 0-25-50-75-100) helped reduce variation among scores. More detailed UXM parameter evaluation guidelines with examples will further help in reducing divergence.

Computing IoI was useful for project stakeholders as they could see the importance of HCI activities in the SE context. The HCI activities integrated in SE process models were acceptable as suggested. Though they were explicitly prompted, none of the project stakeholders wanted changes to the prescribed HCI activities, their weightage or evaluation guidelines. An important feedback was need for process models specifically targeted to redesign projects. Process models typically discuss new product development. Given that many industry projects are *"next version of X"* type, process models must be specifically adapted for them.

Walking through the activity evaluation guidelines helped in scoring as all stakeholders were not aware of all HCI activities. It was felt that IoI should be computed before computing UXM as this minimizes bias.

The metric descriptions presented in this paper are a result of iterative modifications that reflect the feedback and lessons learnt.

6. DISCUSSIONS

It is important to discuss the limitations and risks of the two metrics proposed. Both UXM and IoI are summary measures that leave out much information. They allow a drill-down to constituent components, but do not point to specific problems or give suggestions for improvement. But summary measures are useful in many contexts, particularly for comparison across projects. Such comparisons can help UX groups understand what works and what doesn't and improve performance year-on-year.

Perhaps most important limitation of UXM comes from the ephemeral nature of a 'user experience'. Any attempt to numerically embody such an abstract phenomenon is bound to be subjective and measures could differ according to the interpretation of the evaluators. Further, large companies are involved in software development projects for many clients, across domains, platforms, users, use contexts and task complexity, frequency and criticality.

Finally, there is a risk that because UXM measures are low cost, organizations may be tempted to sacrifice all user-facing activities (such as usability tests or field studies) in its favour. We do not recommend this at all. The purpose of UXM is not to replace these established methods but to supplement them and to help them mature.

In spite of these limitations, we believe that UXM is useful. UXM shows the extent to which user experience goals were achieved in a particular project. We found that breaking up abstract notions of user experience into specific goals and parameters helped evaluators focus on one issue at a time and reduced the subjectivity in measurement. Making the evaluation criteria explicit and averaging across several evaluators further reduced the subjectivity in judgement. The risk of variety in products (the apples-and-oranges risk) was partly mitigated by selecting goal-parameters relevant to each project and giving custom weightage to each parameter.

The main limitation of IoI is that it does not measure the absolute process quality of the project, rather how compliant was a project to the prescribed process. There are no widely-accepted integrated process models at this stage. Yet, IoI in conjunction with UXM may be used to verify the effectiveness of new process model proposals. If UXM and IoI are correlated, the new proposal seems acceptable. On the other hand if the UXM and IoI do not show a correlation, it questions the prescribed process model.

UXM and IoI have an organizational perspective and make more sense while looking across hundreds of projects rather than within each project individually. They have very low additional overheads on the process and are easy to integrate in the process.

Overall feedback indicates that UXM and IoI are useful and practical in evaluating products and processes. There was a lot of buy-in from project stakeholders calculating metrics as there was a lot of willingness to track, control the user experience of the product. The aspects that metric calculation was light-weight and independent of specific usability methods were particularly liked.

7. FUTURE WORK

In future, we plan to use metrics prospectively throughout the duration of projects and demonstrate their usefulness during the project. We will be building more elaborate tools and guidelines to improve the consistency of weightages and scores. We also propose to do a rigorous validation of the two metrics in experimental and industrial situations.

In its current form, UXM goals, parameters and weightages have to be chosen on the basis of experience of individuals. However, it is possible to design tools in future that will collate experience of several practitioners to help in choices of future goals, parameters and weightages. A similar tool for IoI can also evolve the specification of processes.

8. ACKNOWLEDGMENTS

We thank the anonymous reviewers for invaluable comments on improving our presentation of this material. In particular, we thank Ved Prakash Nirbhay and Deepak Korpal from Tech Mahindra for allowing us to interact with their team members while testing the metrics in different software projects. We also thank members of User Interaction Design Group at Tech Mahindra Ltd. for participating in User Experience metrics evaluation. We also thank Pramod Khambete for his continuous support and appreciation. We thank Prof. NL Sarda, Prof. UA Athavankar, Prof. Umesh Bellur and Prof. S Sudarshan for their continuing guidance and suggestions in developing the two metrics.

9. REFERENCES

[1] Beyer, H., Holtzblatt, K., Contextual Design: Defining Customer Centered Systems, Morgan Kaufman (1998)

[2] Bias, R., Mayhew, D. (Eds), Cost-Justifying Usability, Second Edition: An Update for the Internet Age, Morgan Kaufmann (2005)

[3] Cooper, A., Riemann, R., About Face 2.0 the Essentials of Interaction Design, Wiley (2003)

[4] Fenton, N.E., Pfleeger, S.L., Software Metrics – A Rigorous and Practical Approach, Thomsan Brooks/Cole (2002)

[5] Göransson, B., Lif, M., Gulliksen, J., Usability Design – Extending Rational Unified Process with a New Discipline. International Workshop on Interactive Systems Design, Specification, and Verification (2003)

[6] IEEE Standard Glossary of Software Engineering Terminology, IEEE, 1993

[7] IFIP working group 2.7/13.4 on User Interface Engineering, Bridging the SE & HCI Communities: http://www.se-hci.org/bridging/index.html (2004), accessed August, 2008

[8] International Organization for Standardization, ISO/IEC 9126-1:2001 Software Engineering - Product Quality (2001)

[9] International Organization for Standardization, ISO 9241-1:1997 Ergonomic requirements for office work with visual display terminals (VDTs) (1997)

[10] Jordan, P. W., Designing pleasurable products, Taylor & Francis (2000).

[11] Joshi, A: HCI in SE Process Literature, Indo-Dan HCI Research Symposium, IIT Guwahati (2006)

[12] Joshi, A., Sarda N.L.: HCI and SE: Towards a 'Truly' Unified Waterfall Process. HCI International '07 (2007)

[13] Kroll, P., Kruchten, P., The Rational Unified Process Made Easy, Pearson Education (2003)

[14] Lewis, J., A Rank-Based Method for the Usability Comparison of Competing Products. Human Factors and Ergonomics Society 35th Annual Meeting 1312--1316 (1991)

[15] Lin, H. Choong, Y. Salvendy, G. A proposed index of usability: a method for comparing the relative usability of different software systems. Behaviour & Information Technology (1997)

[16] Mahlke, S., Understanding users' experience of interaction, in Marmaras, N., Kontogiannis, T., Nathanael, D. (Eds.), Proc. EACE '05 (2005). 243-246.

[17] Mayhew, D., The Usability Engineering Lifecycle: A Practitioner's Handbook for User Interface Design; Morgan Kaufmann; 1998

[18] McGee, M., Master usability scaling: magnitude estimation and master scaling applied to usability measurement, Proc. CHI'00, ACM Press (2004) 335-342.

[19] Norman, D.A., Emotional Design: Why We Love (or Hate) Everyday Things, Basic Books (2004).

[20] Pressman, R.: Software Engineering – a Practitioner's Approach. 6th edition. McGraw Hill (2005)

[21] Pyla, P. S., Pérez-Quiñones, M. A., Arthur, J. D., Hartson, H. R.: Towards a model-based framework for integrating usability and software engineering life cycles. Interact 2003 Workshop on "Closing the Gaps: Software Engineering and Human Computer Interaction" 67--74 (2003)

[22] Sauro, J., Kindlund, E., A method to standardize usability metrics into a single score. CHI '05 401-409 (2005)

[23] Shneiderman, B.: Designing the User Interface, Strategies for Effective Human-Computer Interaction. 4th edition. Addison Wesley (2004)

[24] Swallow, D., Blyth, M., Peter, W., Grounding experience: relating theory and method to evaluate the user experience of smart-phones, Proc. 2005 Annual Conference on European Association of Cognitive Ergonomics (2005) 91-98.

[25] Tractinsky, N., Katz, A. S. & Ikar, D.. What is beautiful is usable, Interacting with Computers, 13 (2000) 127-145