UNIT 5
The UX Design Process

UX Evaluation and Improve UX Goals, Metrics and Targets, UX Evaluation Techniques. - Formative vs summative, types of formative and informal summative evaluation methods, types of evaluation data, some data collection technics, variations in formative evaluation results, informal summative dada analysis, formative data analysis, feedback to process, evaluation report

UX GOALS

- UX goals are high-level objectives for an interaction design, stated in terms of anticipated user experience. UX goals can be driven by business goals and reflect real use of a product and identify what is important to an organization, its customers, and its users.
- They are expressed as desired effects to be experienced in usage by users of features in the design and they translate into a set of UX measures. A UX measure is a usage attribute to be assessed in evaluating a UX goal.
- We can define the primary high-level UX goals for the ticket buyer to include:
- Fast and easy walk-up-and-use user experience, with absolutely no user training
- Fast learning so new user performance (after limited experience) is on par with that of an experienced user
- High customer satisfaction leading to high rate of repeat customers
- Some other possibilities:
- High learnability for more advanced tasks
- Draw, engagement, attraction
- Low error rate for completing transactions correctly, especially in the interaction for payment

UX METRICS

- A UX metric describes the kind of value to be obtained for a UX measure. It states what is being measured. There can be more than one metric for a given measure.
- Most commonly, UX metrics are objective, performance-oriented, and taken while the participant is doing a benchmark task. Other UX

educlash Result / Revaluation Tracker
Track the latest Mumbai University Results / Revaluation as they happen, all in one App
Visit educlash.com for more

metrics can be subjective, based on a rating or score computed from questionnaire results.

- Typically, subjective UX metrics will represent the kind of numeric outcome you want from a questionnaire, usually based on simple arithmetic statistical measures such as the numeric average.
- Remember that you are going only for an engineering indicator of user experience, not for statistical significance.
- We hope you will explore many other possibilities for UX metrics, extending beyond what we have mentioned here, including:
- percentage of task completed in a given time
- ratio of successes to failures
- time spent moving cursor (would have to be measured using software instrumentation, but would give information about the efficiency of such physical actions, necessary for some specialized applications)
- for visibility and other issues, fixations on the screen, cognitive load as indicated by correlation to pupil diameter, and so on using eye-tracking

UX TARGET

- The target level for a UX metric is the value indicating attainment of user experience success. It is a quantification of the UX goal for each specific UX measure and UX metric. UX metrics for which you have not yet achieved the target levels in evaluation serve as focal points for improvement by designers.
- Because "passing" the user experience test means meeting all your target levels simultaneously, you have to ensure that the target levels for all UX measures in the entire table must be, in fact, simultaneously attainable.
- That is, do not build in trade-offs of the kind where meeting one target level goal might make it much more difficult to meet another related target level.
- So how do you come up with reasonable values for your target levels? As a general rule of thumb, a target level is usually set to be an improvement over the corresponding baseline level.
- Why build a new system if it is not going to be better? Of course, improved user performance is not the only motivation for building a new system.

educlash Result / Revaluation Tracker
Track the latest Mumbai University Results / Revaluation as they happen, all in one App
Visit educlash.com for more

- For initial performance measures, you should set target levels that allow enough time, for example, for unfamiliar users to read menus and labels, think a bit, and look around each screen to get their bearings.

UX Evaluation Techniques
Formative vs summative evaluation

In simplest terms, formative evaluation helps you form the design and summative evaluation helps you sum up the design. A cute, but apropos, way to look at the difference: "When the cook tastes the soup, that's formative; when the guests taste the soup, that's summative"

- Formative evaluation is primarily diagnostic; it is about collecting qualitative data to identify and fix UX problems and their causes in the design.
- Summative evaluation is about collecting quantitative data for assessing a level of quality due to a design, especially for assessing improvement in the user experience due to formative evaluation.
- Formal summative evaluation is typified by an empirical competitive benchmark study based on formal, rigorous experimental design aimed at comparing design hypothesis factors. Formal summative evaluation is a kind of controlled hypothesis testing with an m by n factorial design with y independent variables.
- Summative evaluation is used, as a partner of formative evaluation, for quantitatively summing up or assessing UX levels using metrics for user performance (such as the time on task), for example, as indicators of progress in UX improvement, usually in comparison with pre-established UX target levels.

TYPES OF FORMATIVE AND INFORMAL SUMMATIVE EVALUATION METHODS

In practice, there are two orthogonal dimensions for classifying types of formative UX evaluation methods:

- empirical method vs. analytic method
- rigorous method vs. rapid method

Rigorous method vs. rapid method

Formative UX evaluation methods can be either rigorous or rapid. We define rigorous UX evaluation methods to be those methods that maximize

effectiveness and minimize the risk of errors regardless of speed or cost, meaning to refrain from shortcuts or abridgements.

- Rigorous and rapid methods exist mainly as quality vs. cost trade-offs.
- Choose a rigorous empirical method such as lab-based testing when you need effectiveness and thoroughness, but expect it to be more expensive and time- consuming.
- Choose the lab-based method to assess quantitative UX measures and metrics, such as time-on-task and error rates, as indications of how well the user does in a performance- oriented context.
- Choose lab-based testing if you need a controlled environment to limit distractions.
- Choose empirical testing in the field if you need more realistic usage conditions for ecological validity than you can establish in a lab.
- Choose a rapid evaluation method for speed and cost savings, but expect it to be (possibly acceptably) less effective.
- Choose a rapid UX evaluation method for early stages of progress, when things are changing a lot, anyway, and investing in detailed evaluation is not warranted.
- Choose a rapid method, such as a design walkthrough, an informal demonstration of design concepts, as a platform for getting initial reactions and early feedback from the rest of the design team, customers, and potential users.

Analytic Method vs. Empirical Method

- Empirical methods employ data observed in the performance of real user participants, usually data collected in lab-based testing.
- Analytical methods are based on looking at inherent attributes of the design rather than seeing the design in use.
- Many of the rapid UX evaluation methods, such as design walkthroughs and UX inspection methods, are analytic methods. Some methods in practice are a mix of analytical and empirical.
- Empirical methods are sometimes called "payoff methods" because they are based on how a design or design change pays off in terms of real observable usage.

educlash Result / Revaluation Tracker
Track the latest Mumbai University Results / Revaluation as they happen, all in one App
Visit educlash.com for more

- Analytical methods are sometimes called "intrinsic methods" because they are based on analyzing intrinsic characteristics of the design rather than seeing the design in use.
- Emotional impact factors can also be evaluated analytically.
- Emotional impact factors can also be evaluated empirically.

TYPES OF EVALUATION DATA

Fundamentally, UX data can be objective or subjective and it can be quantitative or qualitative.

Objective Data vs. Subjective Data

Objective UX data are data observed directly by either the evaluator or the participant. Subjective UX data represent opinions, judgments, and other subjective feedback usually from the user, concerning the user experience and satisfaction with the interaction design.

Quantitative Data vs. Qualitative Data

- Quantitative data are numeric data, such as data obtained by user performance metrics or opinion ratings. Quantitative data are the basis of an informal summative evaluation component and help the team assess UX achievements and monitor convergence toward UX targets, usually in comparison with the specified levels set in the UX targets
- Qualitative data are non-numeric and descriptive data, usually describing a UX problem or issue observed or experienced during usage.
- Qualitative data are usually collected via critical incident (also coming up later) and/or the think- aloud technique (see later) and are the key to identifying UX problems and their causes.

SOME DATA COLLECTION TECHNIQUES

Critical incidents

- Despite numerous variations in procedures for gathering and analysing critical incidents, researchers and practitioners agree about the definition of a critical incident.

- A critical incident is an event observed within task performance that is a significant indicator of some factor defining the objective of the study.
- The best kind of critical incident data are detailed, observed during usage, and associated closely with specific task performance.
- Critical incidents are observed directly by the facilitator or other observers and are sometimes expressed by the user participant.
- Critical incident data about a UX problem should contain as much detail as possible, including contextual information, such as:
  - the user's general activity or task
  - objects or artifacts involved
  - what happened instead
  - as much as possible about the mental and emotional state of the user
  - additional comments or suggested solutions to the problem

The Think-Aloud Technique
- The think-aloud technique is a qualitative data collection technique in which user participants, as the name implies, express verbally their thoughts about their interaction experience, including their motives, rationale, and perceptions of UX problems.
- The think-aloud technique is simple to use, for both analyst and participant. It is useful for when a participant walks through a prototype or helps you with a UX inspection.
- The think-aloud technique is also effective in assessing emotional impact because emotional impact is felt internally and the internal thoughts and feelings of the user are exactly what the think-aloud technique accesses for you.

Questionnaires
- A questionnaire is the primary instrument for collecting subjective data from participants in all types of evaluations. It can be used to supplement objective (directly observable) data from lab-based or other data collection methods or as an evaluation method on its own.
- A questionnaire can contain probing questions about the total user experience. Although post- session questionnaires have been used primarily to assess user satisfaction, they can also contain effective questions oriented specifically toward evaluating broader emotional impact and usefulness of the design.

educlash Result / Revaluation Tracker
Track the latest Mumbai University Results / Revaluation as they happen, all in one App
Visit educlash.com for more

- **The Questionnaire for User Interface Satisfaction (QUIS)**
- The QUIS is organized around such general categories as screen, terminology and system information, learning, and system capabilities. Within each of these general categories are sets of questions about detailed features.

Data Collection Techniques Especially for Evaluating Emotional Impact

- "Emotion is a multifaceted phenomenon which people deliver through feeling states, verbal and non-verbal languages, facial expressions, behaviors, and so on." Therefore, these are the things to "measure" or at least observe or ask about.
- To access these emotional reactions, we must tap into the user's subjective feelings; one effective way to do that is to have the user or participant do the reporting. Thus, verbal participant self-reporting techniques are a primary way that we collect emotional impact indicators.

Data Collection Techniques to Evaluate Phenomenological Aspects of Interaction

- The new perspective that the phenomenological view brings to user experience requires a new kind of evaluation.
- Thus, phenomenological usage is not about tasks but about human activities. Systems and products with phenomenological impact are understood through usage over time as users assimilate them into their lifestyles.
- These longer term user experience studies are, in some ways, similar to contextual inquiry and even approach traditional ethnography in that they require "living with the users."
- Regardless of which technique is used for phenomenological data collection, the objective is to look for occurrences within long-term usage that are indicators of:
  - ways people tend to use the product
  - high points of joy in use, revealing what it is in the design that yields joy of use and opportunities to make it even better
  - problems and difficulties people have in usage that interfere with a high-quality user experience
  - usage people want but is not supported by the product

educlash Result / Revaluation Tracker
Track the latest Mumbai University Results / Revaluation as they happen, all in one App
Visit educlash.com for more

- ○ how the basic mode of usage changes, evolves, or emerges over time
- ○ how usage is adapted; new and unusual kinds of usage people come up with on their own

VARIATIONS IN FORMATIVE EVALUATION RESULTS

- The most important reason for this effect is due to the individual differences among people. Different people see usage and problems differently.
- Different people have different detection rates. They naturally see different UX problems in the same design. Also in most of the methods, issues found are not questioned for validity. This results in numerous false positives, and there is no approach for scrutinizing and weeding them out.
- Further, because of the vagueness of the methods, intra-evaluator variability can contribute as much as inter-evaluator variability. The same person can get different results in two successive evaluations of the same system.
- No one has the resources to look everywhere and test every possible feature on every possible screen or Web page in every possible task.
- You are just not going to find all the UX problems in all those places. One evaluator might find a problem in a place that other evaluators did not even look. Why are we surprised that each evaluator does not come up with the same comprehensive problem list? It would take a miracle.

FEEDBACK TO PROCESS

- Now that you have been through an iteration of the UX process lifecycle, it is time to reflect not just on the design itself, but also on how well your process worked. If you have any suspicions after doing the testing that the quantitative criteria were not quite right, you might ask if your UX targets worked well.
- For example, if all target levels were met or exceeded on the very first round of evaluation, it will almost certainly be the case that your UX targets were too lenient.
- Even in later iterations, if all UX targets are met but observations during evaluation sessions indicate that participants were frustrated and performed tasks poorly, your intuition will probably tell you that

educlash Result / Revaluation Tracker
Track the latest Mumbai University Results / Revaluation as they happen, all in one App
Visit educlash.com for more

the design is nevertheless not acceptable in terms of its quality of user experience

- Finally, assess how well the overall process worked for the team. You will never be in a better position to sit down, discuss it, and document possible improvements for the next time.

INFORMAL SUMMATIVE (QUANTITATIVE) DATA ANALYSIS

- They use simple "descriptive" statistics (such as mean and standard deviation) to make an engineering determination as to whether the interaction design has met the UX target levels.
- If the design has not yet met those targets, qualitative analysis will indicate how to modify the design to improve the UX ratings and help converge toward those goals in subsequent cycles of formative evaluation. Iteration can seem to some like a process going around in circles, which can be scary to managers.
- The first step in analyzing quantitative data is to compute simple descriptive statistics (e.g., averages) for timing, error counts, questionnaire ratings, and so on, as stated in the UX targets.
- Be careful about computing only mean values, though, because the mean is not resistant to outliers and, therefore, can be a misleading indicator.
- Because we are not doing formal quantitative analysis, the small number of participants typical in formative evaluation can lead to a mean value that meets a reasonable UX target and you can still have serious UX problems.
- Next, by directly comparing the observed results with the specified UX goals, you can tell immediately which UX targets have been met, and which have not, during this cycle of formative evaluation.

FORMATIVE (QUALITATIVE) DATA ANALYSIS

Formative analysis of qualitative data is the bread and butter of UX evaluation. The goal of formative data analysis is to identify UX problems and causes (design flaws) so that they can be fixed, thereby improving product user experience. The process of determining how to convert collected data into scheduled design.

- **Get an Early Jump on Problem Analysis**
- *Keep a participant around to help with early analysis*

educlash Result / Revaluation Tracker
Track the latest Mumbai University Results / Revaluation as they happen, all in one App
Visit educlash.com for more

- In a typical way of doing things, data collection is "completed," the participant is dismissed, and the team does high fives and cracks open the bubbly before turning its attention to data analysis. We suggest bringing in the problem analyst as early as possible, especially if the analyst is not on the data collection team.
- *Early UX problem data records*
- The earlier you can get your raw critical incident notes packaged as data records, the more expedient the transition to subsequent data analysis.
- *Clean up your raw data before your memory fades*
- It is essential for data collectors to clean up the raw data as soon after data collection as time and evaluator skills permit to capture as complete a record of each critical incident as possible while perishable detailed data are still fresh.
- In this transition from data collection to data analysis, experienced data collectors will anticipate the need for certain kinds of content later in problem analysis.
- *Clarify and amplify your emotional impact data*
- UX problems involving emotional impact are, by nature, usually broader in scope and less about details than usability problems.
- **Sources of Raw Qualitative Data**
- We are talking primarily about data from lab-based UX testing here, but critical incident data can come from other sources such as expert UX inspections.
- It is our job to sort through these, often unstructured, data and extract the essential critical incident and UX problem information.
- Regardless of the source of the raw data, much of the data analysis we do in this chapter is essentially the same. Some sources are less detailed and some are more synoptic.

- **Isolate Individual Critical Incident Descriptions**
- On occasion, participants can experience more than one distinct UX problem at the same time and a single critical incident comment can refer to all of these problems.
- The first step in the sequence for refining raw data into UX problem reports is to scan the raw critical incident notes, looking for such

educlash Result / Revaluation Tracker
Track the latest Mumbai University Results / Revaluation as they happen, all in one App
Visit educlash.com for more

notes about more than one UX problem, and separate them into multiple critical incident notes, each about a single UX problem.

- **UX Problem Instances**
- *Content:* To begin with, in whatever scheme you use for maintaining UX data, each UX problem instance should be linked back to its constituent critical incident data parts.
- The next thing to do is to give the problem a name so people can refer to this problem in discussions. Next we want to include enough information to make the UX problem instance as useful as possible for data analysis. You need enough information to accomplish the main goals:
  ○ understand each problem
  ○ glean insight into its causes and possible solutions
  ○ be conscious of relationships among similar problems
- *Project context:* Project context is a rather voluminous and largely uninteresting (at least during the session) body of data that gives a setting for UX data within administrative and project-oriented parameters.
- Project context data can include:
  ○ organization (e.g., company, department)
  ○ project (e.g., product or system, project management, dates, budget, personnel)
  ○ version (e.g., design/product release, version number, iteration number)
  ○ evaluation session (e.g., date, participants, evaluators, associated UX target table)
  ○ task run (e.g., which task, associated UX targets)