

Rigorous Evaluation

Usability Testing

SWEN-444

To Review - What is Usability?

- A **measure** of the **quality** of the **user's experience** when **interacting** with a product or system
- **How usable** is the interface?
 - Learnable
 - Memorable
 - Efficient
 - Understandable
 - Satisfying

Usability Measures

- Ease of learning (learnability)—how fast can a user learn to accomplish basic tasks?
- Ease of remembering (memorability)—can a user remember enough to be effective the next time?
- Efficiency of use—how fast can an experienced user accomplish tasks?
- Error frequency and severity (understandability/comprehensibility) - how often do users make errors, how serious are they, and how do users recover from them?
- Subjective satisfaction—how much does the user *like* using the system? Emotional impact

What is Usability Testing?

- Formal and rigorous testing using a **structured process**
- **Validate** adherence to interaction **requirements**
- “Actual” **users** who perform realistic and representative tasks
- Utilize a **functional prototype**
- **Quantitative and qualitative** usability measures

Constraints on Usability Testing

- **Time to ...**
 - Design, prepare, and administer the test
 - Analyze the results
- **Financial**
 - Equipment and software
 - Laboratory time
 - Recording media
 - [Participant compensation]
- **Space**—to perform the usability test
 - A dedicated laboratory or room is recommended.

Awareness of Regulations

- **Human Subjects Protocols**
 - You must be fully aware of the regulations imposed by the various institutions and regulatory bodies that pertain to your experimental design
 - **Health and well being of subjects**
 - The U.S. Department of Health and Human Services Web site
 - <http://www.hhs.gov/ohrp/>
- **Informed consent form** – all participant users should read and sign

Usability testing- Ethics

- Pressures on a user – being observed, perhaps videoed
 - Performance anxiety
 - May feel like it is an intelligence test, feeling stupid in front of observers
 - Compare self with other subjects, compete
- Treat the user with respect
 - Don't waste the users time – eliminate unnecessary tasks
 - Make the user comfortable – one task at a time, first task easy, breaks, relaxed atmosphere
 - Protect user's privacy
 - User can stop at any time

Advantages and Limitations of Usability Testing

- **Advantages**
 - Discover usability issues before deployment
 - Particularly important for a market driven product
 - Begin to build user loyalty
 - Gain knowledge for future releases
- **Disadvantages**
 - Artificial context
 - No guarantee of product acceptance
 - Result skew if true user demographic missed
 - May not be the most efficient and cost effective method for usability evaluation

Results from Usability Tests

- Quantitative data:
 - Performance data - times, error rates, etc.
 - Subjective ratings, from post test surveys
- Qualitative data:
 - Participant comments from notes, surveys, etc.
 - Test team comments, notes, logs
 - Background participant data from user profiles, recruiting survey, pretest questionnaire
 - Any video or audio recordings, etc.
- List of problems (known and/or suspected)

What About Beta Testing?

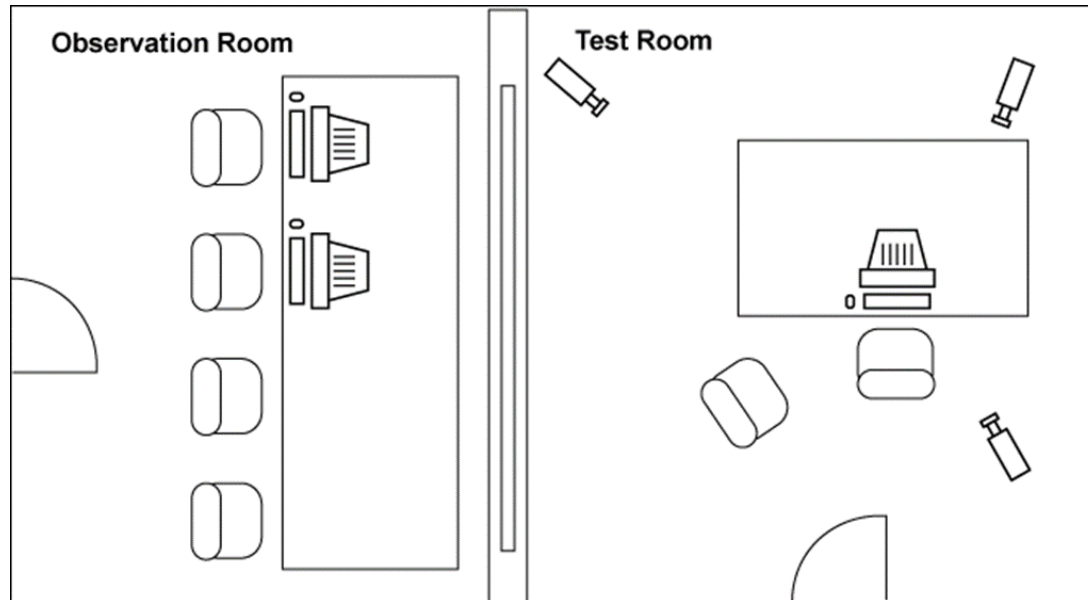
- Beta testing – give real users pre-release products to do real tasks in real environments
- Why not do beta testing instead of expensive usability testing?
 - Late in the process when rework is most expensive
 - Beta testers don't have to use the product
 - Feedback is unsystematic, spotty problem reporting
 - No direct observation of user interaction
 - Users choose the tasks – sufficient coverage?
 - Undesirable side effects for customer satisfaction and product reputation

Test Plan – Design the Test (5W+H)

- **Business case** – why, the purpose; value justifies cost, concerns, goals
- UX design **goals** and **concerns**
- Relevant user **tasks** by role
 - Critical, new, problematic, frequent (80/20 rule), typical
- **Task scenarios** – how will tasks be used in the user environment?

Test Plan –Design the Test

- Schedule
- Resources – people and equipment
- Location



Test Plan - Design the Test

- **Who:** Select Participants, Testers, and Observers
- Whenever possible, participants should be real users
- You don't need a large sample (8-15 or so) to get good feedback
- Recruit users with the following characteristics:
 - Availability
 - Responsiveness
 - Objectivity
 - Diversity – background, experience, responsibility, ...
 - Represent primary user roles

Test Plan - Design the Test

- Who (cont): Tester roles
 - Test project **leader**, expert
 - **Moderator** – interacts with the participant during the test
 - Data logger / **Note taker**
 - [Technician] – operational responsibility
- Optional observers:
 - Other development team members not involved in the test
 - Other stakeholders

Test Plan –Design the Test

- Define the measurements – reflect usability goals
- Quantitative - objective, measurable
 - Performance data - times, error rates, etc.
 - Subjective ratings, from post test surveys
- Qualitative: subjective
 - Participant comments, survey answers
 - Test team comments, observations
 - Background participant data from user profiles, surveys, questionnaires

UX Target Table

| Work Role: User Class | UX Goal | UX Measure | Measuring Instrument | UX Metric | Baseline Level | Target Level | Observed Results |
|--------------------------|---------|------------|----------------------|-----------|----------------|--------------|------------------|
| | | | | | | | |

- Work role – user category
- UX goal – high level UX design objectives
- UX Measure - usage attribute or characteristic to be measured (e.g., initial performance)
- Measuring instrument – the benchmark task(s) or survey to generate test data
- UX Metric – test measurement values to be collected; e.g., response times
- Baseline level – performance of current system if relevant
- Target level – minimum value for success
- Observed results – measured values

Example UX Target Table

Ticket Kiosk System

| Work Role: User Class | UX Goal | UX Measure | Measuring Instrument | UX Metric | Baseline Level | Target Level | Observed Results | Meet Target? |
|--|----------------------------------|--------------------------|---------------------------------------|--|--------------------------------|--------------|------------------|--------------|
| Ticket buyer: Casual new user, for occasional personal use | Walk-up ease of use | Initial user performance | BT1: Buy special event ticket | Average time on task | 3 min as measured at the kiosk | 2.5 min | 3.5 min | No |
| Ticket buyer: Casual new user, for occasional personal use | Walk-up ease of use for new user | Initial user performance | BT2: Buy movie ticket | Average number of errors | <1 | <1 | 2 | No |
| Ticket buyer: Casual new user, for occasional personal use | Initial customer satisfaction | First impression | Questions Q1–Q10 in questionnaire XYZ | Average rating across users and across questions | 7.5/10 | 8/10 | 7.5 | No |

Prepare for the Test

- Write **test scripts** – to avoid bias due to inconsistent moderator-participant interaction
 - Greet the participant – introductions, set the stage
 - Preliminary interview – warm-up questions
 - Provide instructions
 - Monitor the test – record observations, capture participant's impressions and comments
 - Debrief the participant – wrap-up discussion

Prepare for the Test

- **Script** test and task execution details
 - Length and order
 - Breaks to minimize user fatigue
 - Intervals between tests
 - Flexibility for the unexpected
- Run a pilot test to **rehearse**
 - Be organized
 - Be presentable for a good first impression

Perform the Test

- Pre-Test
 - Greet the participant
 - Have the participant sign the informed consent form
 - Have the participant fill out any pre-test questionnaire
 - Proceed with scripts
- During the test
 - Maintain a log or observation check list for each task
 - Create a problem list to capture anything that is not covered by the check list
 - Note any ideas or theories that occur to you about the problems

Perform the Test

- During the test (cont.)
 - Usability measurements
 - Critical incident observation – emotional impact
- Post-Test
 - Debrief the participant
 - Post-test questionnaire
 - Verbal interview
 - Thank the participant and [provide compensation]
 - Process test data
 - Prepare for the next participant

Process the Data

- Activities performed on the day of the test
 - Collect data
 - Summarize data
 - Organize the material
- Follow-up activities
 - Categorize data – top-down, bottom-up (harder)
 - Analyze data
 - Quantitative data – statistical analysis
 - Qualitative data – summarize, consolidate, correlate to quantitative data

Process the Data

- Identify problems (known and/or suspected)
 - Severity
 - Frequency
 - Errors of omission
 - Errors of commission
- Prioritize problems
- Theorize reasons and solutions
- Identify successes and areas of uncertainty

Test Plan Strawman

- Using the 5W+H heuristic, outline a first version of your project test plan
 - Why - purpose
 - What ...
 - Concerns and goals
 - Tasks
 - Scenarios
 - Measurements
 - UX target table