**Usability testing**

**Concepts**

Usability: ease of use and acceptability of a system or product for a particular class of users carrying out specific tasks in a specific environment.

* Where “ease of use” affects user performance (efficacy, efficiency), satisfaction (comfort).
* And “acceptability” affects whether or not the product is used.

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.

To be useful, usability has to be specific. It must refer to particular tasks, particular environments and particular users.

SO HAS TO BE ITS TESTING!

How to test?

* Ease of use is inversely proportional to the number and severity of difficulties people have in using software.

**Usability testing**

Two major families by goals:

* Determine usability problems:
  + Discovery, prioritization, and resolution of usability problems.
  + Iterative testing.
* Measure task performance, include two fundamental tasks:
  + The development of the usability objectives.
  + Iterative testing to determine if the product under test has met the objectives.

Great variety of usability tests:

* Can be very informal or very formal.
* Often use think – aloud (TA):
  + More reliable than posterior interviews.
  + Doesn’t affect efficiency.
  + Better for problem discovery than measurement.
* Remote or local.
* Evaluated software can be varied:
  + Prototypes, under development, competitive products…

Testing techniques:

* “Formal” usability tests.
* Remote testing.
* Guerilla usability testing.
* Steve Krug’s “usability testing on 10 cents a day”.
* Heuristic / expert evaluation.

**Formal usability tests**

Environment:

“Formal” usability tests require a controlled environment:

* Inside a room, outside… (illumination conditions: useful for perception studies).
* Device used (e.g., computer with Internet connection and a browser, or a mobile…).
* Other conditions (e.g., connection quality…).
* Usability lab.

Set of soundproofed rooms:

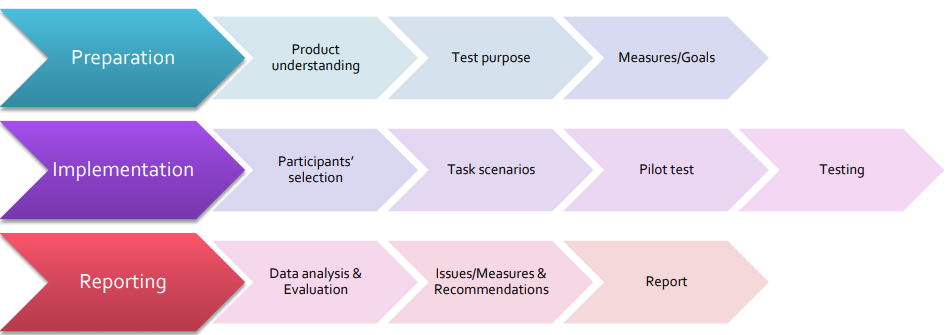
* Proper recording and avoiding distractions to participants.

Different areas and equipment:

* Participant area (where the experiment is carried out).
* Observer area with one-way glass.
* Executive viewing area behind the primary observer area.
* Video cameras and microphones, telephone…

Tasks and roles:

Usability test workflow:



Usability test roles:

* A: Test administrator.
* B: Briefer.
* CO: Camera Operator.
* DR: Data Recorder.
* HD: Help Desk Operator.
* PE: Product Expert.
* S: Statistician.

***PREPARATION (A)***

*Product understanding (A + PE):*

1. Understanding the purpose of the product.
2. Parts of the product are ready for testing.
3. Types of people who will use the product.
4. Determine the use given to the product.
5. Conditions of usage of the product.

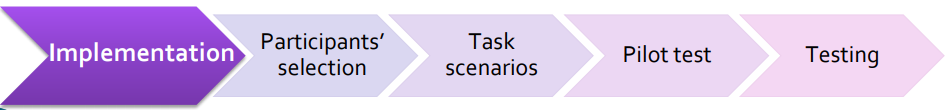
*Test purpose (A + S):*

* Product comparison, within / between subjects.

*Measures / Goals (A + S):*

* Measures / Goals: number of iterations, counting mistakes / errors, timings…

***IMPLEMENTATION***



*Participants’ selection:*

* It’s complicated.
* Should be representative: people that could be real users (no other managers).
* No-show rates above 10%.
* Should be paid. Factors to consider:
  + Time needed, qualification of the participant.
  + Non-monetary for internal participants that are already being paid for their time.
  + Always offer food and beverages.
* For statistical significance: 10 – 12 participants.
* Less formal: 4 – 5 participants per user group.
* Ensure recruiting criteria reflects user characteristics. (e.g., for a website, ensure participants have prior experience browsing).

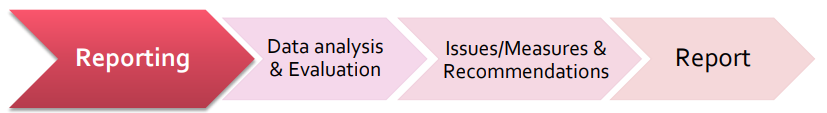
*Test task & scenarios*

* Task must be representative:
  + Core tasks: features that everybody uses (write a text).
  + Peripheral tasks: features used less often (table insertion).
* Scenarios must be determined:
  + Define initial conditions.
  + Description of the scenario: what to do and why.
    - Should not provide step-by-step instructions but should include details.
    - Some action must be taken on finish.
    - Not all users must be provided with the same scenarios (may depend on the user profile).

*Pilot test & testing:*

* Brief participants: B.
* Initial questionnaire: B.
* Develop tasks: B, CO, DR, [HD], A.
* Debrief: B.
* Final questionnaires: B.

***REPORTING***



*Data analysis & evaluation:*

* Frequency: number of users that find a problem divided by the number of users testing the app or web. (easy objective to evaluate).
* Severity: importance of the problem.
  + Might be completely catastrophic or simply cosmetic.
  + Difficult (more subjective) to evaluate.
* Should indicate the importance: severity, can be classified:
  + Mistakes: errors due to incorrect intention.
  + Slips: errors due to appropriate intention but incorrect action.
* Expertise does not affect on the number of errors, but affects how fast they are handled.
* Problems evaluation. Dumas and Reddish:
  + Level 1: prevents task completion.
  + Level 2: creates significant delay and frustration.
  + Level 3: problems have a minor effect on usability.
  + Level 4: subtle and possible enhancements / suggestions.

*Issues / Measures & recommendations:*

* Create a problem grid: frequency / impact.
* Global changes (prevent task completion) first: a missing help may be a global problem or something related with a concrete UI.
* Try to give at least one recommendation for each problem: present the different trade-offs clearly.

*Problem evaluation. Conclusions of the report:*

* Do not use a large number of categories: do not get obsessed by the number of categories either.
* Different evaluators may disagree on some problems’ severity.
* Treat frequency separately from severity.
* Do not forget to point out positive findings.

**Simplified usability tests**

Testing just a single person early is much better than 50 near the end. The point of testing is to inform your judgment.

***GUERRILLA USABILITY TESTING***

* Take someone in a coffee or public space and ask her to use a website for a couple of minutes.
* Observe users:
  + Ask open-ended questions such as “What would you do here?”
* Get to know them a bit: offer coffee or bagels.
* Analyze captured data: considering your audience.

***USABILITY TESTING ON 10 CENTS A DAY***

* Prepare some tasks to evaluate.
* Grab somebody from the company as user.
* Gather stakeholders in an observing room.
* Let the user do a set of tasks.
* Capture gestures, mouse, record…
* Discuss over lunch (order a pizza for everybody).
* Report.

***REMOTE TESTING***

* Like traditional tests but participants and facilitator are in different physical locations:
  + Participants can do the test at home.
  + Facilitator watches remotely.

Advantages:

* Cheaper and easier test setup.
* Usually faster (in terms of allocating / securing facilities travel…).
* Can get geographically dispersed users.

Disadvantages:

* Cannot read body languages.
* Difficult to decide when to talk / interact.
* Variability in participants’ motivation.
* No-show rates higher than in-person studies.

Two types:

Unmoderated remote testing:

* Users do the task completely alone.
* Users don’t have real-time support.
* Don’t get any clue on how the session went.
* No opportunity to ask detailed questions: sometimes the software allows to have some of them predefined.
* Preferable to work only on a few specific elements than a broad view of a product.
* Good for tight timeframes.

Moderated remote testing:

* Users have access to a facilitator.
* Facilitator can change or reorder task as needed.
* Facilitator can ask follow-up questions or clarifications.
* Participant is less likely to spend time on tasks not related to the test.
* Test sessions can be longer (usually about an hour).
* Can perform more in-depth tests.
* The team can watch the test and discuss afterwards.

***HEURISTIC EVALUATION***

* 3 – 5 usability experts evaluate an app or UI.
* Use pre-defined principles (heuristics).
* Can highlight usability issues before user testing.

Advantages:

* Can be quick and cost effective: if we have internal resources.
* Can be used early in the design process.
* Can give a comprehensive usability status of a product’s UI.
* Is compatible with other usability testing methods.

Process:

* Collect the UI.
* Understand the business and users’ needs.
* Understand user motivations and tasks to accomplish.
* Define the heuristics.
* Use a minimum of 3 experts.
* Set up a consistent evaluation system.
* Highlight problem(s) and its ratings.
* Compare and analyze the results of multiple experts.