IDI – Usability Testing

Usability Testing. Outline

- Concepts
- Usability testing
- Usability laboratories
- Test roles
- Test planning

3

Usability. 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.

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

4

Usability. Concepts

- Usability:
  - 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!**

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

5

Usability. Concepts

- How to test?
  - Ease of use is inversely proportional to the number and severity of difficulties people have in using software.
  - Let's examine the difficulties!!!

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

6

Usability Testing. Outline

- *Concepts*
- Usability testing
- Usability laboratories
- Test roles
- Test planning

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

7

Usability. Usability testing

- Two major families by goals:
  - Determine usability problems (*i.e.text editor*):
    - Discovery, prioritization, and resolution of usability problems
    - May be informal, iterative testing
  - Measure task performance (*i.e. 3D selection*). Include two fundamental tasks:
    - The development of the usability objectives.
    - Iterative testing to determine if the product under test has met the objectives

ViRVIG Visualització, Realitat Virtual i Interacció Gràfica

8

Usability. Usability testing

- Great variety of usability tests:
  - Can be very informal or very formal
  - Observer might sit next to the participant, watch through a one-way glass, or watch the on-screen behaviour of a participant who is performing specified tasks.
  - Often use think-aloud (TA)
  - Observers might watch one or two participants at a time
  - Evaluated software can be varied:
    - Prototypes, under development, competitive products…

ViRVIG Visualització, Realitat Virtual i Interacció Gràfica

9

Usability. Usability testing

- Related Techniques (1): Think-Aloud:
  - Participants must talk about what they are doing as they do it
    - Prompt participants to resume if they stop talking
    - What users say during tasks is more reliable than posterior interviews
      - In interviews users are inclined to answer what they think you would like them to
      - When people verbalize after the experiment, they only note what they remember
      - People tries to rationalize their behaviour (giving reasons why they did not see a button…)

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

10

Usability. Usability testing

- Related techniques (1): Think Aloud:
  - Can be apply to almost any usability testing method
  - Seem to work better with pairs of participants
  - Seem to be best suited than silent participation in problem discovery
  - Better for problem discovery than measurement

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

11

Usability. Usability testing

- Related Techniques (2): Remote testing:
  - Lets participants with special needs, from other countries… to participate
  - May introduce familiar environments
  - May be difficult to have enough visual feedback from the participant
  - May lead to compromised security of unpublished products

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

12

Usability Testing. Outline

- *Concepts*
- *Usability testing*
- Usability laboratories
- Test roles
- Test planning

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

13

Planning usability tests. Environment

- Formal usability tests require a controlled environment
  - Inside a room, outside…
    - Illumination conditions (useful for perception studies)
  - Devices used (e.g. computer with Internet connection and a browser, or a mobile…)
  - Other conditions (e.g. connection quality…)

  Usability lab ☺

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

14

Planning usability tests. Environment

- 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…

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

15

Usability Testing. Outline

- *Concepts*
- *Usability testing*
- *Usability laboratories*
- Test roles
  - Administrator, briefer, camera operator, data recorder, help desk operator, product expert, statistician
- Test planning

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

16

Usability. Test roles

- Test administrator: Designs the usability study
  - Specifies the initial conditions for the test session and the code to use for data logging.
  - Conducts reviews with the rest of the test team
  - Leads the data analysis
  - Puts together the final presentation or report

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

17

Usability. Test roles

- Briefer: Interacts with participants
  - Briefs participants at the start of the test
  - Communicates with them as required during the test
  - Debriefs participants at the end of the test sessions
  - In a think-aloud study, the briefer has the responsibility to keep the participant talking
  - Needs to be familiar enough with the product in order to decide what to answer to the participants

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

18

Usability. Test roles

- Camera Operator: Responsible for running the audio-visual equipment during the test
- Data Recorder: Writes notes during a test session.
  - Is the primary data used for the usability study
  - The camera may not catch the important action at every moment of a usability study
  - Usually uses data-logging software
  - It is a very demanding skill

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

19

Usability. Test roles

- Help Desk Operator: Replaces a real help desk operator
  - Required when the participant experiences enough difficulty to place a call
  - Must behave as a call-center person
- Product Expert: Maintains the product and offers technical guidance during the test
  - Must recover if there is a product failures
  - Helps the other team members understand the system's actions during the test
- Statistician: Extracts the maximum amount of information from the data gathered during a formal test
  - Rarely required for informal tests

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

20

Usability Testing. Outline

- *Concepts*
- *Usability testing*
- *Usability laboratories*
- *Test roles*
- Test planning & reporting results
  - Design (product understanding, test purpose, measures, test goals, participants, task scenarios) pilot test, reporting

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

21

Planning usability tests. Design

- **Before Starting**, the administrator must:
  1. Understand the purpose of the product (p.e.google maps)
  2. Decide which parts of the product are ready for testing
  3. Determine the types of people who will use the product
  4. Determine the use given to the product
  5. Determine the conditions of usage of the product

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

22

Planning usability tests. Design

- **Determining the purpose of the test**
  - Measurement vs Usability problems identification
  - Product or techniques comparison
    - Within-subjects or between-subjects
    - Statistical analyses for these two types of test may be different

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

23

Planning usability tests. Design

- For problem discovery:
  - Prioritize problems
    - Include frequency of occurrence
    - Likelihood of of usage of the occurrence in normal usage
    - Magnitude of impact on the participants
    - Ease of correction
  - Pre-planned number of iterations
  - Number of participants small, but multiple iterations,…

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

24

Planning usability tests. Design

- For measurement tests: *Fundamental global Measures*
  - Successful task completion rates
  - Mean task completion times
  - Mean participant satisfaction ratings (on a task-by-task basis)
    - There are standardized questionnaires for this
  - Other measurements could be:
    - Knowledge acquisition indicators (learnability and learning rate)
    - Number of tasks completed within a specified time limit, number of wrong menu choices, number of user errors, number of repeated errors (same user)

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

25

Planning usability tests. Participants

- **Participants**
  - User profile must be determined (administrator)
    - Sometimes available from the marketing group
  - Must define the characteristics of the target population
    - They are difficult to define:
      - May involve previous experience, education level, age, sex…
  - Can be obtained from employment agencies, internal sources, market research firms, existing customers…

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

26

Planning usability tests. Participants

- Participants. Factors to consider:
  - Demographic locations
  - Age ranges
  - Levels of experience
  - Levels of gender
- Number of users:
  - Will depend on many factors
    - Money and time
    - Type of study: Single-shot (larger) vs iterative (smaller)

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

27

Planning usability tests. Participants

- Sample size estimation:
  - [Virzi, 1992] found that 80% of known usability problems could be surfaced with 5 testers, and 3 that testers would reveal the most severe problems
  - [Nielsen & Landauer, 1993] say that the best benefits are usually obtained testing no more than 5 users and running as many small tests you can afford

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

28

Planning usability tests. Participants

- Sample size estimation:
  - There is a *law of diminishing returns* [Nielsen, 2000]
    - The third tester will do many things that you have already observed with the first or second user
      - Will generate a small amount of new data
    - After the fifth user you are wasting your time by observing the same findings repeatedly but not learning much new

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

29

Planning usability tests. Participants

- Safe values:
  - 3-4 users to find main problems (≈70-80%)
  - 5-6 users to find most problems

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

30

Planning usability tests. Participants

- Sample size estimation:
  - Quality of tests may seriously affect the number of detected problems [Faulkner, 2003]
  - Quality of testers also has an impact on the number of usability problems revealed [Faulkner, 2003]

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

31

Planning usability tests. Implementation

- **Test task scenarios**:
  - Must be representative
    - Core tasks: Features that everybody uses (write a text)
    - Peripheral tasks: Features used less often (table insertion)
  - Once the tasks are defined, scenarios of use must be created
    - Define initial conditions
    - Description of the scenario: what to do and why
      - Some action must be taken on finish
      - Should not provide step-by-step instructions but should include details
      - Not all users must be provided with the same scenarios (may depend on the user profile)

ViRVIG Visualització, Realitat Virtual i Interacció Gràfica

32

Planning usability tests. Implementation

- **Procedure:**

1. Introduction: Purpose of the test, confidentiality…

2.1 Task performance:

- Complete preliminary questions and forms (background questionnaire, informed consent form, confidential disclosure form…)
- Complete training (if required)
- Perform the tasks

ViRVIG Visualització, Realitat Virtual i Interacció Gràfica

33

Planning usability tests. Implementation

- **Procedure:**

  2.2. Task performance:
  - Usually, no help is provided:
    - Refer the users to the documentation
    - If required, provide help, but score the task as failed
    - Try to avoid direct answers to questions
    - If asking questions, try to avoid biasing the participant's response

  3. Give a satisfaction questionnaire at the end of each scenario.

  4. After the scenarios, final questionnaire
  - There are stantardized versions

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

34

Planning usability tests. Implementation

- **Pilot testing:**
  - Usability test must be tested
  - Commonly, a member of the usability team can do the testing.

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

35

Planning usability tests. Reporting

- **Reporting Results:**
  - Describe & prioritize the usability problems
  - Present quantitative measurements

Should lead to a recommendation

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

36

Planning usability tests. Reporting

Problem 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

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

37

Planning usability tests. Reporting

- Reporting. *Usability problems*:
  - 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

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

38

Planning usability tests. Reporting

- Rating the severity of usability problems:
  - Some thoughts on severity and frequency
  - Local evaluation: Jeff Rubin, Jakob Nielsen…
  - Global evaluation: Dumas and Redish

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

40

Planning usability tests. Reporting

- Problem evaluation**. Dumas and Redish:**
  - **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

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

41

Planning usability tests. Reporting

Problem evaluation. **Jeff Rubin:**

- **4: Unusable:** The user is not able to or will not want to use a particular part of the product because of the way that the product has been designed and implemented.
- **3: Severe**: The user will probably use or attempt to use the product here, but will be severely limited in his or her ability to do so.
- **2: Moderate**: The user will be able to use the product in most cases, but will have to undertake some moderate effort in getting around the problem.
- **1: Irritant**: The problem occurs only intermittently, can be circumvented easily, or is dependent on a standard that is outside the product's boundaries. Could also be a cosmetic problem.

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

43

Planning usability tests. Reporting

- Reporting. *Recommendations*:
  - Create a problem grid: frequency/impact
  - Global changes (prevent task completion) first
  - Must be checked:
    - 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

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica

45

Planning usability tests. Reporting

- Problem evaluation. Conclusions
  - 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

ViRVIG
Visualització, Realitat Virtual i Interacció Gràfica