

HCI Evaluation & Usability Testing

11/1/18

Professor Nathan J. McNeese

The aims:

- Explain how to do usability testing
- Outline the basics of experimental design
- Describe how to do field studies

Usability testing

- Involves recording performance of typical users doing typical tasks.
- Controlled settings.
- Users are observed and timed.
- Data is recorded on video & key presses are logged.
- The data is used to calculate performance times, and to identify & explain errors.
- User satisfaction is evaluated using questionnaires & interviews.
- Field observations may be used to provide contextual understanding.

Checklist for planning usability testing

- <https://www.nngroup.com/articles/usability-test-checklist/>

Experiments & usability testing

- Experiments test hypotheses to discover new knowledge by investigating the relationship between two or more variables.
- Usability testing is applied experimentation.
- Developers check that the system is usable by the intended user population for their tasks.

Usability testing & research

Usability testing

Improve products

Few participants

Results inform design

Usually not completely replicable

Conditions controlled as much as possible

Procedure planned

Results reported to developers

Experiments for research

Discover knowledge

Many participants

Results validated statistically

Must be replicable

Strongly controlled conditions

Experimental design

Scientific report to scientific community

Usability Testing conditions

- Usability lab or other controlled space.
- Emphasis on:
 - selecting representative users;
 - developing representative tasks.
- 5-10 users typically selected.
- Tasks usually around 30 minutes
- Test conditions are the same for every participant.
- Informed consent form explains procedures and deals with ethical issues.

Usability types of data

- Time to complete a task.
- Time to complete a task after a specified time away from the product.
- Number and type of errors per task.
- Number of errors per unit of time.
- Number of times online help and manuals accessed.
- Number of users making an error.
- Number of users successfully completing a task.

How many participants is enough for user testing?

- The number is a practical issue.
- Depends on:
 - schedule for testing;
 - availability of participants;
 - cost of running tests.
- Typically 5-10 participants.
- Some experts argue that testing should continue until no new insights are gained.

Usability lab with observers watching a user & assistant



Figure 14.1 A usability laboratory in which evaluators watch participants on a monitor and through a one-way mirror

Portable equipment for use in the field



Figure 14.3 The Tracksys lab-in-a-box system, which comprises components that pack into a heavy duty padded flight case plus a PC system

Source: Courtesy of Harry Brignull.

Portable equipment for use in the field



Figure 14.4 The Tracksys system being used with a mobile device camera that attaches to a flexible arm, which mounts on a mobile device, and is tethered to the lab

Source: Courtesy of Harry Brignull.

Mobile head-mounted eye tracker



Figure 14.5 The mobile head-mounted eye-tracker

Source: Picture courtesy of SensoMotoric Instruments (SMI), copyright 2010.

Usability testing the iPad

7 participants with 3+ months experience with iPhones

Signed an informed consent form explaining:

- what the participant would be asked to do;
- the length of time needed for the study;
- the compensation that would be offered for participating;
- participants' right to withdraw from the study at any time;
- a promise that the person's identity would not be disclosed; and
- an agreement that the data collected would be confidential and would be available to only the evaluators

Then they were asked to explore the iPad

Next they were asked to perform randomly assigned specified tasks

Examples of the tasks

App or website	Task
iBook	Download a free copy of <i>Alice's Adventures in Wonderland</i> and read through the first few pages.
Craigslist	Find some free mulch for your garden.
eBay	You want to buy a new iPad on eBay. Find one that you could buy from a reputable seller.
<i>Time</i> Magazine	Browse through the magazine and find the best pictures of the week.
Epicurious	You want to make an apple pie for tonight. Find a recipe and see what you need to buy in order to prepare it.
Kayak	You are planning a trip to Death Valley in May this year. Find a hotel located in the park or close to the park.

Table 14.1 Examples of some of the tests used in the iPad evaluation (adapted from Budiu and Nielsen, 2010).

Source: Copyright Nielsen Norman Group, from report available at <http://www.nngroup.com/reports/>.

Example of the equipment



Figure 14.6 The setup used in the Chicago usability testing sessions

Source: Copyright Nielsen Norman Group, from report available at <http://www.nngroup.com/reports/>.

Problems and actions

Problems detected:

- Accessing the Web was difficult
- Lack of affordance and feedback
- Getting lost
- Knowing where to tap

Actions by evaluators:

- Reported to developers
- Made available to public on nngroup.com

Accessibility for all users important

Guerilla Testing

<https://www.youtube.com/watch?v=0YL0xoSmyZI>

Usability Activity- User Satisfaction:

- Pair up in dyads.
- Each of you pick a mobile app that you have never used before.
- Spend 5 minutes exploring the application.
- Ask each other the following questions and take notes (10 minutes each):
 - What kind of phone did you use?
 - What apps do you use most on your phone?
 - General thoughts regarding the homescreen?
 - General thoughts regarding aesthetics of the app?
 - General thoughts regarding navigation within the app?
 - How would you describe the app you just saw? What is it for?
 - Is this application something you personally want to install on your phone? (yes/no)
 - I would/wouldn't install this app on my phone because:
 - Any additional questions for me?

Experiments

- Test hypothesis
- Predict the relationship between two or more variables.
- Independent variable is manipulated by the researcher.
- Dependent variable influenced by the independent variable.
- Typical experimental designs have one or two independent variables.
- Validated statistically & replicable.

Experimental designs

- Different participants - single group of participants is allocated randomly to the experimental conditions.
- Same participants - all participants appear in both conditions.
- Matched participants - participants are matched in pairs, e.g., based on expertise, gender, etc.

Different, same, matched participant design

Design	Advantages	Disadvantages
Different	No order effects	Many subjects & individual differences a problem
Same	Few individuals, no individual differences	Counter-balancing needed because of ordering effects
Matched	Same as different participants but individual differences reduced	Cannot be sure of perfect matching on all differences

Field studies

- Field studies are done in natural settings.
- “In the wild” is a term for prototypes being used freely in natural settings.
- Aim to understand what users do naturally and how technology impacts them.
- Field studies are used in product design to:
 - identify opportunities for new technology;
 - determine design requirements;
 - decide how best to introduce new technology;
 - evaluate technology in use.

Data collection & analysis

Observation & interviews

- Notes, pictures, recordings
- Video
- Logging

Analyzes

- Categorized
- Categories can be provided by theory
 - Grounded theory
 - Activity theory

Data presentation

The aim is to show how the products are being appropriated and integrated into their surroundings.

Typical presentation forms include:

- Vignettes,
- Excerpts,
- Critical incidents,
- Patterns, and narratives.

Key points

- Usability testing takes place in controlled usability labs or temporary labs.
- Usability testing focuses on performance measures, eg. how long and how many errors are made when completing a set of predefined tasks. Indirect observation (video and keystroke logging), user satisfaction questionnaires and interviews are also collected.
- Affordable, remote testing systems are more portable than usability labs. Many also contain mobile eye-tracking and other devices.
- Experiments test a hypothesis by manipulating certain variables while keeping others constant.

Key points

- The experimenter controls independent variable(s) in order to measure dependent variable(s).
- Field studies are evaluation studies that are carried out in natural settings to discover how people interact with technology in the real world.
- Field studies that involve the deployment of prototypes or technologies in natural settings may also be referred to as ‘in the wild’.
- Sometimes the findings of a field study are unexpected, especially for in the wild studies in which explore how novel technologies are used by participants in their own homes, places of work, or outside.