Basic Types of Evaluation

Evaluation

You can evaluate a design, such as a wireframe or a working prototype:

- To determine usability issues
- Identify positive design and development elements
- Should take place as often as possible and in every stage of the product creation
- Can happen in the lab or in the field
- Can happen with or without (yes without!) users

Summative Evaluation

- Usually happens at the end of the creative process
- The results are used to ‘judge’ the creation and for quality control
- Usually has a high focus on quantitative measures such as performance and comparisons

Formative Evaluation

- Happens during the whole process of creation
- The results are used to improve the creation
- Has a mix methods approach with qualitative and quantitative measures

Within / Between-Subjects Experimental Design

Within

- Each participant takes part in each experiment condition
- Commonly use when testing two or more different types of systems
- Individual participant factors are less apparent in results

Between

- Each participant takes part in one condition only
- Useful when we want to negate effects of learnability

Observational Evaluation

- Method:
  - Observe and register the real use of a system by the users.
  - Participants may or may not be allowed to use a ‘think aloud protocol’ while taking part in the evaluation testing which is to explain in real time what they are thinking. Alternatively, they will be asked to relate thoughts after the observation – this not affecting their behavior and performance.
  - The investigator usually does not intervene during the observation.

There are several ways in which we can register and record the observations:

- Pen and paper
- Audio recording
- Visual recording
- Computer Logs and / or screen capture
- Eye – Tracking
- Bio- sensors
- Many many more….

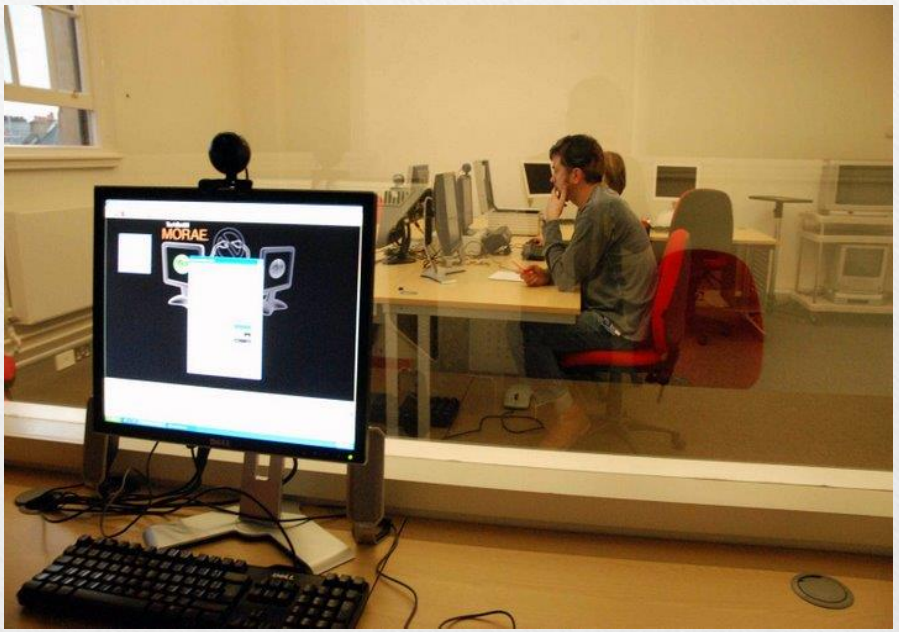
Lab-Based

Usually involves cameras, streaming and microphones. There is also the option for eye-tracking.

Sometimes the investigator is inside and sometimes its outside.

The procedure must be the same for every participant.

There needs to be a clear protocol and enough description so that there can be reproducibility.



https://www.youtube.com/watch?v=0SeXtreV-w8

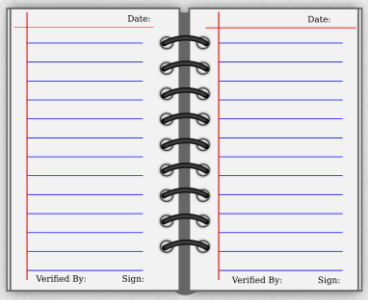
Users in their natural environment

• Without the presence of the investigator

• To detect usability issues in a natural environment without influence, when used in the users' 'standard' way.

• Often requires the users to communicate findings themselves (or remember to record their behaviour

• Diary Studies are a usual methodology for this…



Query-based

See lecture on qualitative and quantitative analysis (questionnaires / interviews etc.)

Amazon Mechanical Turk

Heuristic evaluation

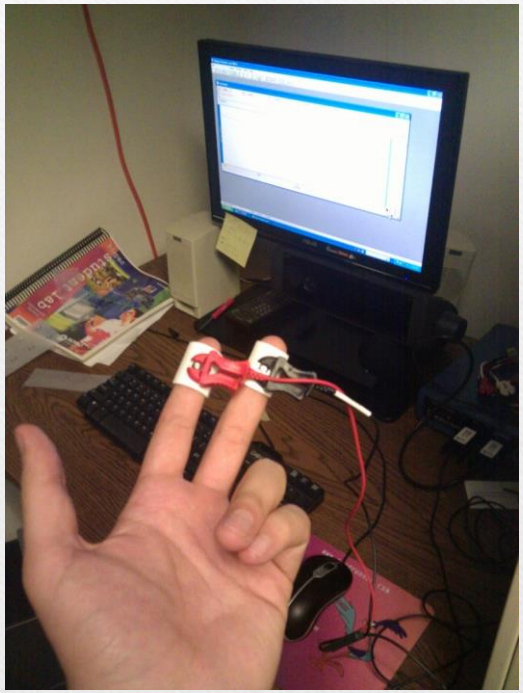
Predetermined and tested guidelines and measures

Can be used for design as well as evaluation.

Cheap (discount usability engineering)

Caution: tends to find false positives (false problems)
(Hollingsed & Novick, 2007)

Other Measures



Galvanic Skin Response

Brain wave patterns (hearing lie testing)

Usability Metrics

Time to Complete

Error Rates

Learnability

Memorability

User Satisfaction

Expert review

Systematic review from experienced individuals to assess the effective design of a system.

Can take place throughout the design process.

Usually 3-5 experts are involved. Sometimes only 1 is hired (consultant).

Give all use cases and all parts of the system.

Initially they (experts) evaluate the system individually. They then meetup and make a common report.

References

• Blandford, A., Cox, A., & Cairns, P. (2008). Controlled Experiments. In P. Cairns & A. Cox, *Research Methods for Human-Computer Interaction (pp. 1-*16). Cambridge University Press.
• Cairns, P., & Cox, A. (2008). Using Statistics in Usability Research. In *Research Methods for Human-Computer Interaction (pp. 112-137). Cambridge*University Press.
• Jacob, R., & Karn, K. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In J. Hyona, R. Radach, & H. Deubel, *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research (pp. 573-605).*
• Kahn, M., & Prail, A. (1994). Formal usability inspections. In J. Nielsen & R. Mack, *Usability inspection methods (pp. 141-171). New York: John Wiley &* Sons.
• Koutsabasis, P., Vlachogiannis, E., & Darzentas, J. S. (2010). Beyond Specifications : Towards a Practical Methodology for Evaluating Web Accessibility. *Journal of Usability Studies, August 2010.*
• Nielsen, J. (1993). *Usability Engineering. San Fransisco: Morgan Kaufmann.*
• Nielsen, J. (1994). Usability inspection methods. In *Conference on Human Factors in Computing Systems (pp. 413-414). Boston, Massachusetts, United* States.
• Nielsen, J. (2005). Formal Usability Reports vs. Quick Findings. *Jacob Nielsen's Alertbox, April 25, 2005. Retrieved from*
http://www.useit.com/alertbox/20050425.html.
• Snyder, C. (2003). *Paper Prototyping. Morgan Kauffman.*
• Tullis, T., & Albert, B. (2008). *Measuring the User Experience: Collecting Analysing and Presenting Usability Metrics. Morgan Kaufmann.*