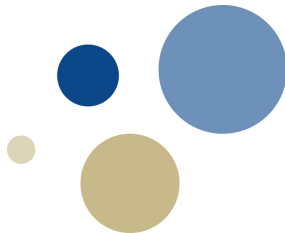# Human Computer Interaction

Design & Evaluation in HCI

Anna Xambó

Department of Music, NTNU

24 October 2019

Learning Outcomes

— Identify key literature in HCI design and ergonomics.
— Understand the trends in the approach to evaluation in the CHI conference / HCI discipline.
— Establish connections between quantitative, qualitative and mixed research methods applied to HCI.
— Conceptualize an approach to evaluation of a self-built musical instrument based on human-centered design.
— Analyze the research methods and evaluation elements in a successful CHI paper submission.

Preparation: Reading

— Send a summary (1 page max.) of the following article:
  - From Mice to Men - 24 years of Evaluation in CHI [1]
    `http://barkhu.us/barkhuus-altchi2007.pdf`

  The summary should include: the main topic of the paper, the approach used, the main findings, and the main contributions.

Class Structure

— 10.15-10.30 Warm-up activity, HCI design & ergonomics.
— 10.30-10.45 Research methods in HCI.
— 10.45-11.00 Group discussion about the prep. reading.
— 11.00-11.15 Human-centered design.
— 11.15-11.30 Team working: (1) What is your research question? (2) How to evaluate your prototype borrowing HCI methods? (3) What makes a good CHI paper?
— 11.30-12.00 Final group discussion & closing.
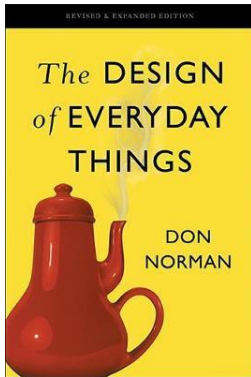
Warm-up Activity

TODO: Similarities / differences between the papers

Similarities — Selected CHI 2019 papers — Differences

MindNode
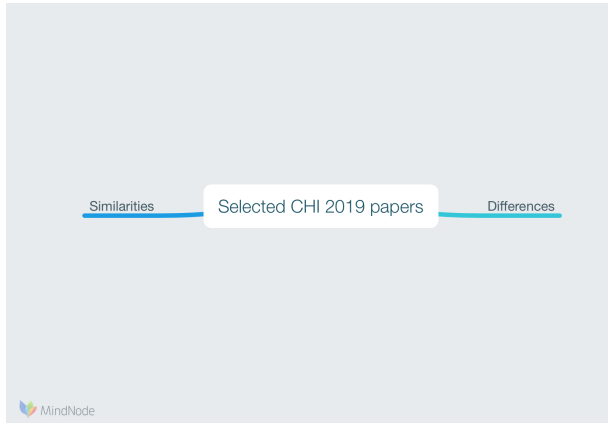
HCI Design & Ergonomics

The Design of Everyday Things

Norman, D. A. (2013/1988). *The Design of Everyday Things*. Basic books. [2]
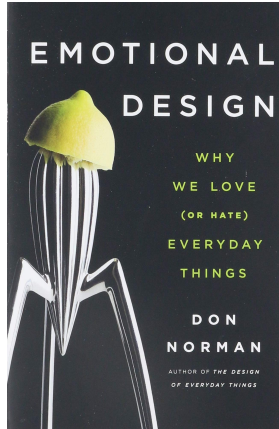
The Design of Everyday Things

FIGURE 1.2. Problem Doors: Signifiers Are Needed. Door hardware can signal whether to push or pull without signs, but the hardware of the two doors in the upper photo, A, are identical even though one should be pushed, the other pulled. The flat, ribbed horizontal bar has the obvious perceived affordance of pushing, but as the signs indicate, the door on the left is to be pulled, the one on the right is to be pushed. In the bottom pair of photos, B and C, there are no visible signifiers or affordances. How does one know which side to push? Trial and error. When external signifiers—signs—have to be added to something as simple as a door, it indicates bad design. (Photographs by the author.)

The Design of Everyday Things

— *Rules*: Make things visible, exploit natural relationships that couple function and control, and make intelligent use of constraints.
— *Goal*: Guide the user effortlessly to the right action on the right control at the right time.

Emotional Design: Why We Love (or Hate) Everyday Things

Norman, D. A. (2004). *Emotional Design: Why We Love (or Hate) Everyday Things*.
Basic Civitas Books. [3]

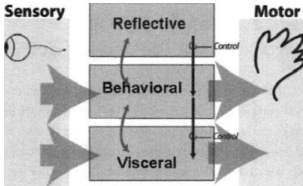# Emotional Design: Why We Love (or Hate) Everyday Things

FIGURE 1.1

Three levels of processing: Visceral, Behavioral, and Reflective.
The visceral level is fast: it makes rapid judgments of what is good or bad, safe or dangerous, and sends appropriate signals to the muscles (the motor system) and alerts the rest of the brain. This is the start of affective processing. These are biologically determined and can be inhibited or enhanced through control signals from above. The behavioral level is the site of most human behavior. Its actions can be enhanced or inhibited by the reflective layer and, in turn, it can enhance or inhibit the visceral layer. The highest layer is that of reflective thought. Note that it does not have direct access either to sensory input or to the control of behavior. Instead it watches over, reflects upon, and tries to bias the behavioral level.

(Modified from a figure by Daniel Russell for Norman, Ortony, & Russell, 2003.)

Emotional Design: Why We Love (or Hate) Everyday Things

The three levels of processing can be mapped to product characteristics [3, p.39]

— *Visceral design*: Appearance, form with powerful emotional signs (e.g. children's toys, clothes, furniture)

— *Behavioral design*: The pleasure of effectiveness of use, emphasis on the use of objects, performance and function matter (e.g. a shower)

— *Reflective design*: Self-image, personal satisfaction, memories, it is about message and culture and the meaning of the product or its use, prestige, perceived rarity and exclusiveness (e.g. a souvenir monument, a smartwatch)

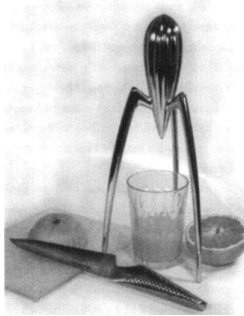# Emotional Design: Why We Love (or Hate) Everyday Things

FIGURE 4.6
Two items of seduction.
Philippe Starck's "Juicy Salif" citrus juicer alongside my Global kitchen knife. Rotate the orange half on the ribbed top of the juicer and the juice flows down the sides and drips from the point into the glass. Except this gold-plated version will be damaged by the acidic fluid. As Starck is rumored to have said, *"My juicer is not meant to squeeze lemons; it is meant to start conversations." (Author's collection.)*

Emotional Design: Why We Love (or Hate) Everyday Things

Characteristics of Philippe Starck's Juicy Salif:
— Entices by diverting attention.
— Delivers surprising novelty.
— Goes beyond obvious needs and expectations.
— Creates an instinctive response.
— Espouses values or connections to personal goals.
— Promises to fulfill these goals.
— Lends the casual viewer to discover something deeper about the juicing experience.
— Fulfills these promises.

Apple Human Interface Guidelines: Accessibility & Inclusive Design
Impairments and Accommodations

— Approximately one in seven people worldwide have a disability or impairment that affects the way they interact with the world and their devices. People can experience impairments at any age, for any duration, and at varying levels of severity. Situational impairments—temporary conditions such as driving a car, hiking on a bright day, or studying in a quiet library—can affect the way almost everyone interacts with their devices at various times.

https://developer.apple.com/design/human-interface-guidelines/accessibility/overview/introduction/

**Apple Human Interface Guidelines: Accessibility & Inclusive Design Impairments and Accommodations**

— Begin designing your app to be inclusive and accessible to everyone by reviewing the four main categories of impairments and the accessibility features that address them: *Vision*, *Hearing*, *Physical and Motor*, *Literacy and Learning*

https://developer.apple.com/design/human-interface-guidelines/accessibility/overview/introduction/

Handbook of Human Factors and Ergonomics
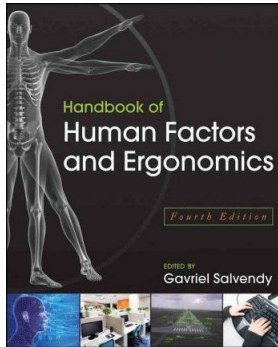
Salvendy, G. (Ed.). (2012). Handbook of Human Factors and Ergonomics. John Wiley & Sons. [4]
Table of contents:
https://onlinelibrary.wiley.com/doi/book/10.1002/9781118131350

Research Methods in HCI

Research Methods in HCI

Choosing which method to use is a highly context-dependent issue related to a variety of factors including the primary purpose of the study, time constraints, funding, the participant pool, and the researchers' experience. [5, p.25] Lazar,

J., Feng, J. H. and Hochheiser, H. (2017). Research Methods in Human-Computer Interaction. Morgan Kaufmann.

Research Methods in HCI

Lazar, J., Feng, J. H. and Hochheiser, H. (2017). Research Methods in Human-Computer Interaction. Morgan Kaufmann. [5]

Descriptive Research, Relational Research, and Experimental Research

Table 2.1 Relationship Between Descriptive Research, Relational Research, and Experimental Research

Type of Research	Focus	General Claims	Typical Methods
Descriptive	Describe a situation or a set of events	X is happening	Observations, field studies, focus groups, interviews
Relational	Identify relations between multiple variables	X is related to Y	Observations, field studies, surveys
Experimental	Identify causes of a situation or a set of events	X is responsible for Y	Controlled experiments

Lazar, J., Feng, J. H. and Hochheiser, H. (2017). Research Methods in Human-Computer Interaction. Morgan Kaufmann. [5, p.27]

Quantitative Methods: "How often?" or "How long?" questions

— **Experimental design** (chapter 2): development of research hypotheses and testing the validity. Null hypothesis vs alternative hypothesis, independent vs dependent variables, randomization. Between-group design (between groups) vs within-group design (repeated measures).

  • Significance tests allow us to determine how confident we are that the results observed from the sampling population can be generalized to the entire population. For example, a test that is significant at $P < 0.05$ suggests that we are confident that 95% of the time the test result correctly applies to the entire population.
  • P-value: Probability value. The lower the value, the more unlikely that the null hypothesis is true. Depending on the result, we accept or reject the null hypothesis (determined by the significance level or threshold).

— **Surveys**: a set of questions to which an individual is asked to respond. They are frequently used to describe populations and to explain behaviors. Considered one of the easiest methods. Typically broad but not deep.

Qualitative Methods: "Why?" questions

— **Diaries**: A document created by an individual who maintains regular recordings about events in their life, at the time that those events occur. See cultural probes by William Gaver [6].

— **Case studies**: An in-depth study of a specific instance (or a small number of instances) within a specific real-life context. Close examination of individual cases can be used to build understanding, generate theories and hypotheses, present evidence for the existence of certain behavior, or to provide insight that would otherwise be difficult to gather.

— **Interviews** (individuals) and **focus groups** (multiple users at once): Direct feedback from interested individuals. Deep but not broad.

— **Ethnography**: A combination of observation, interviews, and participation. Ethnographic research projects use deep immersion and participation in a specific research context to develop an understanding that would not be achievable with other, more limited research approaches.

Mixed Methods

— Mixing of qualitative and quantitative data, methods, methodologies, and/or paradigms in a research study or set of related studies.

- *Quantitatively driven approaches/designs* in which the research study is, at its core, a quantitative study with qualitative data/method added to supplement and improve the quantitative study.
- *Qualitatively driven approaches/designs* in which the research study is, at its core, a qualitative study with quantitative data/method added to supplement and improve the qualitative study.
- *Interactive or equal status designs* in which the research study equally emphasizes (interactively and through integration) quantitative and qualitative data, methods, methodologies, and paradigms.

Creswell, J. W. and Creswell, J. D. (2017). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Sage Publications. [7]

Data collection vs data analysis

— Traditional data collection (users' replies, users' actions, audio recordings, video recordings, field notes).
— Automated data collection indirectly from humans (e.g. key logging and web site logs).
— Data collection directly from humans through sensors focused on the body (e.g. facial EMG and eye-tracking).
— Online data collection (e.g. crowdsourcing and big data).
— Qualitative analysis (transcription, manual annotation) vs quantitative data analysis (statistical analysis).

Working with Human Participants / Subjects

— Working with human subjects involves many challenges.
— Finding the right subjects is often difficult and time consuming, especially for evaluation of systems designed for specific populations or situations.
— Research ethics require that participants must be treated fairly and with respect.
  - *Informed consent*: a mechanism to inform participants about the nature of the study so they can make a meaningful decision as to whether or not they really want to be involved.
  - Check regulations at your institution / country where you are conducting the research (e.g. consent forms, data protection acts, etc).
  - Low- vs high-risk studies.
  - Plan ahead your study.

Group discussion: "From Mice to Men - 24 years of Evaluation in CHI"

— Any surprises?
— Any limitations?
— Is the title accurate?

Human-Centered Design

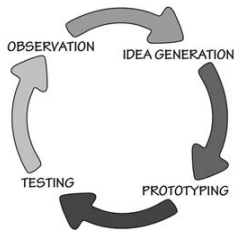# The Human-Centered Design Process

FIGURE 6.2. The Iterative Cycle of Human-Centered Design. Make observations on the intended target population, generate ideas, produce prototypes and test them. Repeat until satisfied. This is often called the *spiral method* (rather than the circle depicted here), to emphasize that each iteration through the stages makes progress.

Norman, D. A. (2013/1988). *The Design of Everyday Things*. Basic books. [2]

The Human-Centered Design Process

— **Observation**: The initial research to understand the nature of the problem itself. Design requirements are determined.

— **Idea generation**: Generation of potential solutions. Generation of numerous ideas. Being creative without regard of constraints. Question everything.

— **Prototyping**: Test the idea. Building of a quick prototype or mock-up of each potential solution. One popular technique is "Wizard of Oz" (mimic a powerful system), which is useful at early stages of development.

— **Testing**: Gather a small group of people (similar to potential target population) and have them use the prototypes as nearly as possible to the way it is intended. Five people is generally a good number to start with (Jakob Nielsen).

— **Iteration**: It enables continual refinement and enhancement. The goal is rapid prototyping and testing. "Fail frequently, fail fast" (David Kelly).

Norman, D. A. (2013/1988). *The Design of Everyday Things*. Basic books. [2]

The Human-Centered Design Process

Failures are to be encouraged—actually, they shouldn't be called failures: they should be thought of as learning experiences. If everything works perfectly, little is learned. Learning occurs when there are difficulties. [2, p.229]

Norman, D. A. (2013/1988). The Design of Everyday Things. Basic books.

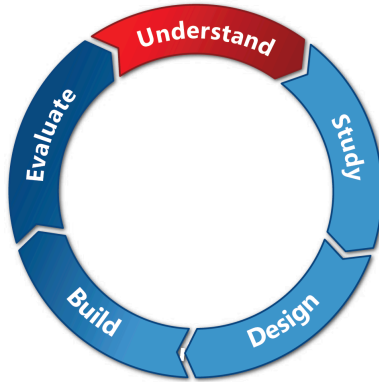# The research and design cycle

Figure: Extended user-centred, five-stage design/research model [8]

Group discussion: How to evaluate your prototype borrowing HCI methods?

What is your research question? What are potential ways of evaluating the music prototypes from the mini-hackathon of the Physical Computing Workshop considering quantitative, qualitative, and mixed research methods...

— Who would be your users?
— What are the challenges of your approach?
— What are the limitations of your approach?

The teams summarize to the group their group discussion about the suitability of the research methods.
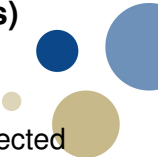
Closing: What makes a good CHI paper?

Discussion on Canvas:
https://uio.instructure.com/courses/22318/discussion_topics/58400

Human-Computer Interaction Day 2 - Group Assignment (post-class)

— Send a summary (1 page max.) of the *research methods* used in the selected
   article discussed in group during class on Tuesday 22 October 2019, from the CHI
   2019 Best Papers (`https:`
   `//chi2019.acm.org/2019/03/15/chi-2019-best-papers-honourable-mentions/`)
   before Friday 25 October 2019 17:00.
   This time you should focus on explaining the research methods used in terms of:
   - whether they are quantitative, qualitative, or mixed methods?
   - how are they related with the research question?
   - what are the limitations of this approach?
   - how do the authors address the limitations?
   - what other research methods could have been used instead and why?

References

[1] Louise Barkhuus and Jennifer A Rode. "From Mice to Men – 24 Years of Evaluation in CHI". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2007, pp. 1–16.

[2] Donald A Norman. *The Psychology of Everyday Things*. Basic Books, 1988.

[3] Donald A Norman. *Emotional Design: Why We Love (or Hate) Everyday Things*. Basic Civitas Books, 2004.

[4] Gavriel Salvendy. *Handbook of Human Factors and Ergonomics*. John Wiley & Sons, 2012.

[5] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. *Research Methods in Human-Computer Interaction*. Morgan Kaufmann, 2017.

[6] William Gaver et al. "Cultural Probes and The Value of Uncertainty". In: *interactions-Funology* 11.5 (2004), pp. 53–56.

[7] John W Creswell and J David Creswell. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage Publications, 2017.

[8] Richard Harper et al. *Being Human: Human-Computer Interaction in the Year 2020*. Microsoft Research Ltd, 2008.