Measurability and Predictability of User Experience (UX)
Who and Where am I?

- PhD in Psychology
- Usability since 2000
  - COST Action 294 MAUSE (2005-2009)
    - "Towards the Maturation of Usability Evaluation"
    - http://www.cost294.org/
- User Experience since 2006
  - Invited talk of Marc Hassenzahl
  - COST Action IC0904 TwinTide (2009-2013)
    - http://www.twintide.org/
    - "Towards the Integration of Trans-sectorial IT Design and Evaluation"
- A series of scientific work:
  - CHI SIG 2008 on UX Definition
  - CHI Paper 2009 on UX Definition
  - INTERACT Workshop 2009 on UX Methods
  - CHI SIG 2011 on UX Theories
  - Special Issue on UX Modelling
  - Dagstuhl Seminar on Demarcating UX 2010
Overview

- UX Definition
- UX Models
- UX Evaluation Methods
- Measurability of UX
- Predictability of UX
- Conclusion & Outlook
Why is there not yet a common definition of UX?

- UX is associated with a broad range of **fuzzy** and **dynamic** concepts, e.g., experience, emotion, affect, and aesthetics
- Unit of analysis for UX is too stretchable (solo vs. social)
- The landscape of UX research is complicated by diverse conceptual models with different foci

No common understanding of what UX is!

Reference:
Why bother to have one?

1. Facilitate scientific discourses; otherwise communication breakdown

2. Enable practical applications of UX, operationalize it into measures, compare across similar products

3. Help teaching UX, its nature and scope

Caveat:
Reaching a common definition is *not* a panacea for resolving a number of problems related to UX, but it serves as crucial step towards an integrated framework of UX
UX Definitions: Selected Examples

- **ISO 9241-210: 2010**: A person’s perceptions and responses that result from the use and/or anticipated use of a product, system or service.

- **Hassenzahl & Tractinsky (2006)**: A consequence of a user’s *internal* state (...), the characteristics of the designed *system* (...), and the *context* within which the interaction occurs.

- **Nielson & Norman Group**: all aspects of the end-user's interaction with the company, its services, and its products. The first requirement for an exemplary user experience is to meet the exact needs of the customer, without fuss or bother. Next comes *simplicity* and elegance that produce products that are a *joy to own*, a joy to use....

- **Wikipedia**: UX highlights the *experiential, affective, meaningful and valuable* aspects of human-computer interaction and product ownership, but it also includes a person’s perceptions of the practical aspects such as *utility, ease of use and efficiency* of the system.
Example of Confusing UX Usage

- Is it a marketing strategy to entitle the book with the phrase “the User Experience” (NB: small font for the subtitle: “… Usability metrics”)

- Do the authors treat UX and Usability as synonyms? …. *Seemingly No!*
  - *We take a very broad view of usability and examine the entire user experience. When we talk about “measuring usability”, we’re really looking at the entire user experience* (p.4)
  - *S.5.4.1 Severity ratings based on the User Experience* (pp.105-106)
  - *S.6.7.1 Assessing Specific Attributes*
    - Visual appeal, ..., Enjoyment ...
    - “Covering in detail the ways you might assess all the specific attributes you are interested in is beyond the scope of this book.” (p.158)
<table>
<thead>
<tr>
<th><strong>USABILITY</strong></th>
<th><strong>UX</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pragmatic quality</td>
<td>Hedonic quality</td>
</tr>
<tr>
<td><em>do-goal</em> (to find an e-book)</td>
<td><em>be-goal</em> (to feel competent)</td>
</tr>
<tr>
<td>Product: performance, task</td>
<td>Experience: emotion, affect</td>
</tr>
<tr>
<td>Reductionist?</td>
<td>Holistic?</td>
</tr>
<tr>
<td>Partly objective</td>
<td>Highly subjective</td>
</tr>
<tr>
<td>Relatively persistent</td>
<td>Inherently dynamic</td>
</tr>
<tr>
<td>Standard usability metrics exist</td>
<td>Standard UX metrics yet to be created (Note 3 of ISO 9241-210: usability criteria can be used to assess aspects of user experience)</td>
</tr>
<tr>
<td>- Efficiency, Effectiveness, User satisfaction</td>
<td></td>
</tr>
</tbody>
</table>

Relationship btw. Usability and UX
Donald Norman, early 1990s, “User Experience Architect Group” at Apple
(Smith et al., 1982, Byte; Roy Nierenberg et al. 1985, InfoWorld)

“I invented the term because I thought Human Interface and usability was too narrow; I wanted to cover all aspects of a person’s experience with a system, including industrial design, graphics, the interface, the physical interaction, and the manual. Since then, the term has spread widely, so much so that it is starting to lose its meaning.” (1998, personal comm. Peter Merholz)

“…today I include much more; user experience is the whole totality, … experience is actually more based upon memory than upon reality…” (2008, the UX Week)

Ref: http://vimeo.com/2963837
Observe real uses in real situations
(basics of design and evaluation)

Replace ‘users’ with ‘people’
(terminology confusion; “people experience”)

Total experience that matters
(trajecotory of experience; which momentary or episodic events carry more weights)

UX designers learn to speak the language of business, using plausible numbers
(reasonable but not necessarily real) to sell our ideas and bringing spreadsheets
to the executives; allying with marketing folks;
(ethical issues, formulae for ROI -> the same conundrum for usability)

Job title confusion → UX consultant should have a clear role in design
(identity crisis and concept chaos in both practice and research)
Norman’s Emotional Design

- **Visceral** (aesthetic): beauty (website visual aesthetics/attractiveness Lavie & Tractinsky, 2004)
- **Behavioural** (pragmatic): usability
- **Reflective** (hedonic):
  - Identification: self-image, other-oriented
  - Stimulation: novelty, surprise
  - Evocation: memories; temporality of UX
**UX Modelling**


**Structural models**: To establish the *(cause-and-effect)* relations between constructs and inform the **design** of a system.

**Measurement models**: To allow constructs to be **measured** and inform the **evaluation** of a system.
Generic UX Models: Practitioner-based

UX Honeycomb, 2005
Peter Morville, Information Architect

Restructured UX Honeycomb, 2007
Magnus Revang, UX Practitioner
Specific Research-based UX Model: Gameplay Experience

Elements of children’s pleasurable gameplay experience, Emri & Mäyrä (2005)

Sensory, Challenge-based Immersion (SCI) Model, Emri & Mäyrä (2005)
# Gameplay Experience: Constructs and Metrics

<table>
<thead>
<tr>
<th>Construct</th>
<th>Metric</th>
<th>Description</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immersion</td>
<td>• Sensory – perceptual impacts</td>
<td>• Being engrossed by audio and graphics</td>
<td>• Brown &amp; Cairns (04)</td>
</tr>
<tr>
<td></td>
<td>• Imaginative-richness of narrative</td>
<td>• Being absorbed (i.e. distorted time perception and awareness of extraneous happenings) by the storyline and game worlds; • Identify empathetically with characters;</td>
<td>• Emri &amp; Mäyrä (03, 05)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Csikszentmihalyi (93)</td>
</tr>
<tr>
<td>Excitement</td>
<td>• Bodily reactions</td>
<td>A range of psycho-physiological measures tension, anxiety, nervousness</td>
<td>• Mandryk et al. (06)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Lennart et al. (10)</td>
</tr>
<tr>
<td>Enjoyment</td>
<td>Perceived hedonic quality</td>
<td>Fun and pleasure derived from gameplay</td>
<td>• Hassenzahl (01)</td>
</tr>
<tr>
<td>Challenge</td>
<td>Balance between tasks and skills</td>
<td>Availability of difficulty levels; Meaningfulness of the game structure and gaming activities in terms of outcomes</td>
<td>• Klimmt et al (07)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Sweetser &amp; Wyeth (05)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Csikszentmihalyi (93)</td>
</tr>
</tbody>
</table>
There are MORE structural models than measurement models!
UX Evaluation Methods (UXEMs)


http://www.allaboutux.org/
Collection of UXEMs

Goal: Collect experiential methods from academia and industry

- Literature review
- EU project ENGAGE
- Online survey
- 3 workshops in academic conferences (CHI'08, Interact'09, DPPI'09)
- Special interest group session in CHI’09 conference
- Philips’ list of methods

Outcome: Descriptions of 101 UX evaluation methods
Categorisation of UXEMs

UX evaluation methods can be categorized in various ways:

Who are the participants: Experts, Users, User groups
Time restrictions: 1 day ... Several months
Expertise required: Special skills of researchers
Place of evaluation: Lab, Field, Online
Data type: Quantitative or Qualitative
Period of experience: Moment, Episode, Overall UX
Product development phase: Concept, Prototype, Ready product
Temporal UX Model

(UX White Paper
http://www.allaboutux.org/
Dagstuhl Seminar)

- Dagstuhl Seminar, Germany
  15-18 September 2010
  http://www.dagstuhl.de/en/program/calendar/semhp/?semnr=10373

- 30 researchers and practitioners from the User Experience (UX) community, including from the US and Japan

  to the period before first use, or any of the three other time spans of UX
For longer time spans, UX can be structured in terms of a lifecycle.

Previous experiences influence a future one, for example, reflecting or recounting after one usage episode will frame anticipations of future ones.

The phases of experiencing overlap and interleave in a variety of orders, there is no fixed sequence from anticipating to recounting.

(Ref: UX White Paper)
### Evaluation Methods per Period
(adapted from UX Tutorial, Roto et al., NordiCHI 2010).

<table>
<thead>
<tr>
<th></th>
<th>Momentary e.g. during the gameplay</th>
<th>Episodic e.g. reflection on the 10-minute gameplay episode</th>
<th>Long-term e.g. 3 months after the initial observation (on and off)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Evaluating emotions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observation</td>
<td>Observation</td>
<td>Self-reporting</td>
<td>Questionnaires, Laddering, UX Curve, Repertory Grid Technique</td>
</tr>
<tr>
<td>Facial, body, vocal expressions (e.g. smile, lean back, sigh)</td>
<td>Experience thinking</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Psychophysiological measurements</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muscle, pupil, skin reactions, measured with sensors</td>
<td>Experience sampling, AttrakDiff, Interviews, Day Reconstruction</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Self-reporting</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal: PANAS, AffectGrid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-verbal: EmotionSlider, EmoCards, PrEmo</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
What are the challenges and obstacles for advancing the research work on measuring UX qualities?

Can they be resolved and overcome?
A **metric** is a way of measuring or evaluating a phenomenon or an object quantitatively based on:

- references to a standard definition;
- derivation from models;
- systematic observations;
- reliable procedure agreed upon by a community of practice;

A **metric** can:

- add structure to the design and evaluation process
- gain insights into findings
- provide information to decision makers, especially ROI;
- verify improvement

"*To measure is to know*"
"*If you cannot measure it, you cannot improve it*"

(Lord Kelvin, n.d.)
Quantitative vs. Qualitative

Model-based, Quantitative
Paul van Schaik
Marc Hassenzahl
Alistair Sutcliffe
Noam Tractinsky

Mixed
Andrew Monk
Effie Law

Design-based, Qualitative
John McCarthy
Peter Wright
Mark Blythe
Kia Höök
Jodi Forlizzi

Migration between Camps
AGAINST

“... the reduction of experience into a number of factors or processes... such approaches may be useful for experimental analysis but they can miss some of the insights available in accounts that resist such reduction ... qualitative data provides a richness and detail that may be absent from quantitative research”

(Swallow, Blythe & Wright, 2005)

FOR

“... rich accounts of experience might require an outstandingly reflective and attentive ‘experientor’.... I suspect experiences with technology (as many other experiences as well) to be far less unique and far less variable as implied by the proponents of the "phenomenological" approach. [McCarthy & Wright, 2004] ... Accounts of according experiences might differ in their quality, the experience itself does not.”

(Hassenzahl, 2008)
Measurability is not an issue, or making a rather bold claim: There is nothing that cannot be measured. However, it is the design of data collection method (including procedure, technique, tool, and expertise as well as experience of people involved in the process) that is of critical importance to determine the meaningfulness of UX measures.
<table>
<thead>
<tr>
<th>UX Dimension</th>
<th>Measurement</th>
<th>Data collection method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic UX</td>
<td>No</td>
<td>Interview, Collage</td>
</tr>
<tr>
<td>Affect/Emotion</td>
<td>Yes</td>
<td>Scales (SAM), Psychophysiology</td>
</tr>
<tr>
<td>Enjoyment/Fun</td>
<td>Yes</td>
<td>Scales, Pictures</td>
</tr>
<tr>
<td>Aesthetic, appeal</td>
<td>Yes</td>
<td>Scales (classic/expressive aesthetics)</td>
</tr>
<tr>
<td>Hedonic quality</td>
<td>Yes</td>
<td>Scales (AttrakDiff)</td>
</tr>
<tr>
<td>Engagement, flow</td>
<td>Yes</td>
<td>Scales, Interview</td>
</tr>
<tr>
<td>Motivation</td>
<td>No</td>
<td>Probes</td>
</tr>
<tr>
<td>Enchantment</td>
<td>No</td>
<td>Interview</td>
</tr>
<tr>
<td>Frustration</td>
<td>No</td>
<td>Interview, Observation</td>
</tr>
<tr>
<td>Other constructs</td>
<td>??</td>
<td>??</td>
</tr>
</tbody>
</table>

Adapted from Bargas-Avila & Hornbaek (2011), Table 2

Methodologies of UX studies are largely “borrowed” from traditional HCI!
Examples of UXEMs with Measurements

- **iScale**: *Generic UX*
- **AttrakDiff**: *Attractiveness, Visual appeal*
- **Psychophysiological measurement**: *Emotional states*
- **ProEmo**: *Emotional states*

Ref:
UX Evaluation Methods Tutorial,
Roto, V., Vermeeren, A., Law, E. et al., NordiCHI 2010,
iScale tool

- A survey tool for the retrospective elicitation of longitudinal UX data
  - Concept partly based on Day Reconstruction Method (DRM, Kahneman et al. 2004)
- iScale uses sketching in the reconstruction of one’s experiences with the aim to minimize retrospection bias
- iScale tool imposes a chronological order in the reconstruction of one’s experiences
  - more contextual details surrounding the experienced events
  - the felt emotion is constructed on the basis of the recalled contextual details
- iScale results in an increase in the amount, the richness, and reliability of recalled information (Karapanos et al. 2010)
- The results provide support retrospective techniques as cost-effective alternatives to longitudinal studies
AttrakDiff™ questionnaire

interesting ---- boring
elegant ------ cheap
exciting ------- dull
exclusive ------ standard
impressive ------ nondescript
original ------- ordinary
innovative ------ conservative

AttrakDiff is owned and managed by UI Design GmbH,
http://www.attrakdiff.de/en/AttrakDiff/What-is-AttrakDiff/
AttrakDiff™ questionnaire

- Evaluates **hedonic** and **pragmatic** quality of interactive products (Hassenzahl et al. 2003)
- The data enables to evaluate how the **attractiveness** of the product is experienced, in terms of usability and appearance
- AttrakDiff™ consists of 23 word-pairs (semantic differentials) representing the extreme opposite
  - Seven-step items whose poles are opposite adjectives (e.g. "confusing - clear", "unusual - ordinary", "good - bad")
  - The middle values of an item group creates a scale value for pragmatic quality (PQ), hedonic Quality (HQ) and attractiveness (ATT)
- Supports the distinction between sub-qualities of hedonic quality, stimulation and identity
Psycho-physiological measurements

- Brain, Body, and Bytes CHI 2010 Workshop Presentations
  http://www.eecs.tufts.edu/~agirou01/workshop/
E.g. heart beat, skin perspiration, and facial muscles tell about the emotional state of a person.

Physiological reactions are recorded with sensors attached to the participant. This objective data can be used in combination with self-report data to find out what the user experienced.

**Strengths**
- Investigates momentary experiences without intervening in the interaction

**Weaknesses**
- Expensive setup
- Momentary emotions are important in some domains only
FaceReader

- FaceReader is a tool to track the user affective state while using products or software without resorting to self-report.
- Real-time analysis of facial expression from a video.
- FaceReader constructs a model of the face from the video and automatically evaluate several elementary facial movements (action units).
- Based on these movements it calculates the likeliness that each of six basic emotions (joy, anger, sadness, surprise, fear and disgust) is felt at any given time.

Strength:
- Objective assessment of a person’s emotion

Weakness:
- Data limited to six basic emotions.
FaceReader

http://www.noldus.com/human-behavior-research/products/facereader
PrEmo

- PrEmo is a non-verbal self-report software instrument that measures 14 emotions that are often elicited by product design.
- Emotional responses difficult to measure because
  - their nature is subtle (low intensity)
  - often mixed (i.e. more than one emotional response at the same time).
- Does not rely on words
- Each of the emotions is portrayed by an animation of dynamic facial, bodily, and vocal expressions.
- For use in internet surveys, formal interviews, and qualitative interviews, e.g.,
  - to identify the concept with the most pleasant emotional impact
  - as a discussion tool in consumer interviews.
PrEmo (interface at work – latest version)

Use the instrument below to reflect what you feel for the product pictured on the previous screen.
Procedure:
- Show product, system, etc. to participant
- For each of 14 emotions ask participant to indicate how intense the emotion is felt.

PrEmo is a licensed commercial product
“Where are the Ionians of User Experience Research?”
(Kari Kuutti, NordiCHI, 2010)

Babylonians – practical empiricists
Ionians – speculative theorists

There are not enough Babylonians either!
James-Lange Theory of Emotion (1880s)

- It hypothesizes the intricate relationships between human perception, action and cognition.
- Emotion arises from our conscious cognitive interpretations of perceptual-sensory responses.
- We see and act before we feel
- UX is a cognitive process that can be modeled and measured
Memory: the mental faculty of retaining and recalling past experience based on the mental processes of learning, retention, recall, and recognition.

Daniel Kahneman on happiness – experiencing self vs. remembering self (three cognitive traps):

“. ... The second trap is confusion between experience and memory: basically it's between being happy in your life and being happy about your life or happy with your life. And those are two very different concepts, and they're both lumped in the notion of happiness ... “
It is particularly promising to shed light onto the understanding of UX, which is essentially psychological construct determined by actors’ motives and needs which are in turn shaped by the socio-cultural context where actors are situated.

Predictability of UX

- Does the trajectory of UX follow any model, pattern, or rule of ‘evolution’? (positive feelings ‘survive’?)
- Is user experience predictable?
- Which UX factors should be included and excluded when predicting UX for a specific artifact in a specific context of use?
- How to address tradeoffs and reciprocal relationships between different UX factors, between different UX qualities and between UX factors and qualities?
  - UX factor is distinct from UX quality. The former influences how the latter will be instantiated – type, intensity, and extensity.
Two levels of prediction

- **UX-factor-quality-loop**: integration as well as interaction of specific UX factors (predictors) allows us to predict which UX qualities (criteria) a user is very likely to experience with an interactive entity of interest.
  - **Fidelity of prototype**: the accuracy of prediction hinges crucially on the extent to which an early prototype or a design concept resembles the fully executable version.

- **UX-behaviour-loop**: a specific set of user experiences (predictors), be they negative or positive, determines the likelihood a user (or a customer) will likely purchase or adopt a system/product/service (criterion).
  - UX accumulated over time may shape cognitive processes and behavioural tendencies.
User Experience is a research field that is still being defined!

Measurability and predictability of UX look promising, thought still more work to be done

Gaps between UX academics and practitioners to be bridged

Sound theoretical frameworks for UX to inform the definition and operationalisation of UX qualities and the development of data collection methods;

Effective algorithms to enable the combinatorial integration of a (large) set of UX factors and qualities with reasonable accuracy and efficiency.

HCI Education
Thank you very much for your attention!

Questions? Comments?