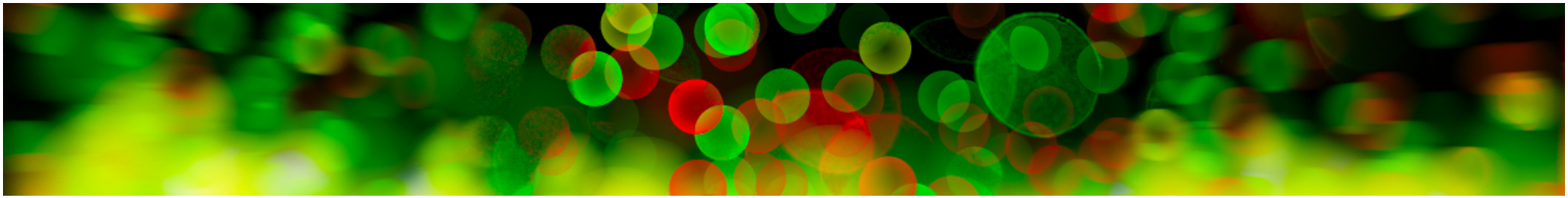
User eXperience Evaluation

Which Method to Choose?



Instructors

Virpi Roto

Aalto University
School of Art and Design

Arnold Vermeeren

TU Delft Delft University of Technology

Effie Law

University of Leicester

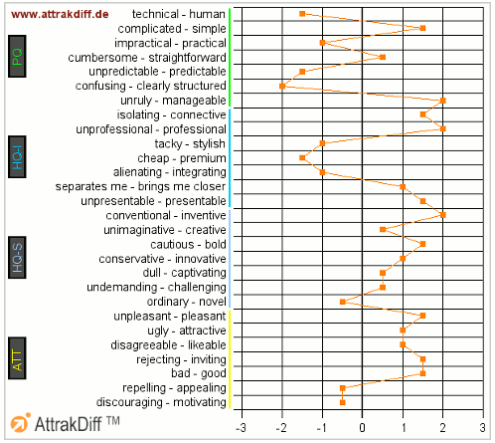
Kaisa Väänänen-Vainio-Mattila

TAMPERE UNIVERSITY OF TECHNOLOGY
Human-Centered Technology

Marianna Obrist

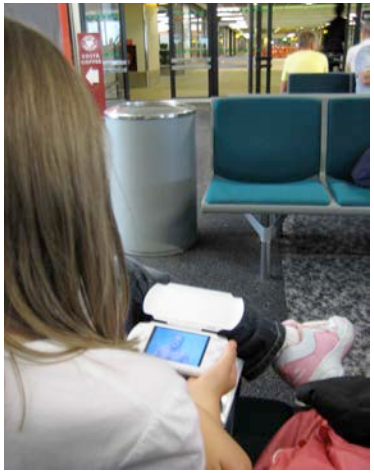
Newcastle University

How to evaluate user experience?

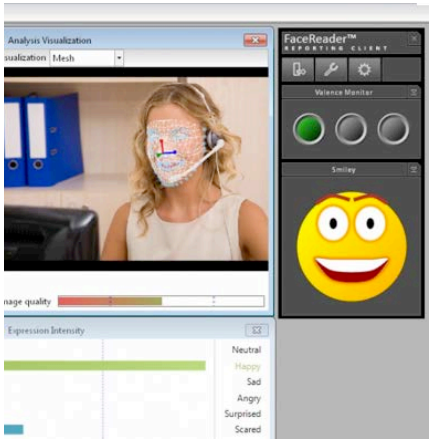


questionnaires

observing



…etc…



measuring in the lab

Picture from http://www.noldus.com/

self-reporting during field studies

Scope of this tutorial

UX = user experience
UXEM = UX evaluation method

User experience evaluation

– methods to reveal how user feels about your design

- Overview of the kinds of UX evaluation methods available
- How to choose the right UX evaluation method for the purpose
- The basics of the most popular UXEMs of different types
- Where to find more information on those methods

Programme

Part 1: Overview of different types of UX evaluation methods

Introduction to UX evaluation (+ an exercise)

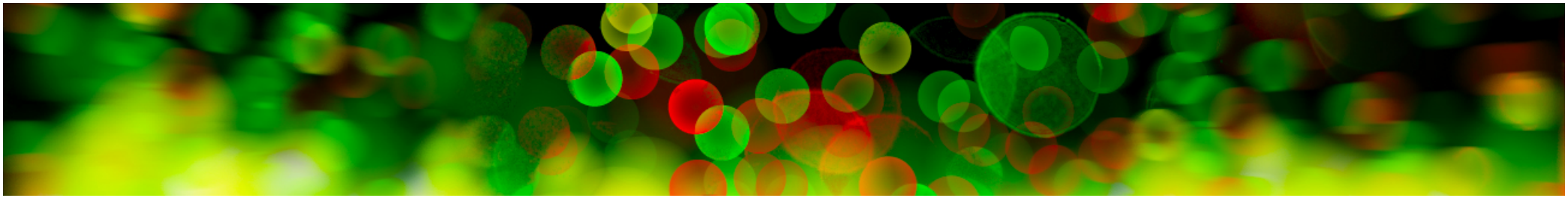
Overview of the types of UXEMs available (+ an exercise)

Part 2: Method introductions

Examples of different types of UXEMs (+ an exercise)

- Experience sampling method
- Day Reconstruction Method
- iScale
- AttrakDiff questionnaire
- Expert evaluation with UX heuristics
- Psychophysiological measurements
- Facereader
- PrEmo
- Sensual Evaluation Instrument

Introduction to UX evaluation



Why evaluate?

- To determine the quality of UX (summative)
  - Measuring UX, e.g. to get out a score
  - Requirements fulfilled? Targets achieved?
  - Comparison of several designs

- To find out improvement areas (formative)
  - Why does certain UX emerge?
  - What you cannot evaluate, you cannot improve

User Experience Evaluation Methods

UX evaluation is different from usability evaluation

User experience:

- How a person *feels* about using a system
- Beyond pragmatic utility and usability
- Subjective, holistic, emotional, long-term

User Experience White Paper – Clarifying the concept of user experience

www.allaboutux.org/uxwhitepaper

User Experience Evaluation Methods

Evaluation:

- Some kind of material being evaluated
- Not only understanding user behavior/needs
- Focus is NOT on inspirational / co-creation methods
- Part of the iterative design/development process, resulting in input for the next phase

User Experience Evaluation Methods

Method:

- Procedure of conducting the evaluation
  - Also techniques that help to collect experiential data
- Not (plain) tools
- Not (plain) measures of UX

Different methods for different purposes

- No one killer method exists for all purposes
- Select the right method for your evaluation case
  - To convince management that solution A is better than solution B
  - To experiment the physical reactions while playing a fun game
- Combinations of methods recommended (method triangulation)
  - Moment-to-moment measurement, and a post-test questionnaire
  - Observation, log of interaction, and interview

UX evaluation in academia vs. industry

Academia	Industry
Understand the phenomenon	Improve product UX
Validate(d) methods	Cost-efficient methods
Big, well-planned study	Frequent and small evaluations
Controlled variables	Real life contexts
Project team of UX experts	UX people a minority in the project
Skilled researchers	UX evaluation often outsourced

So, evaluation situations differ

Before choosing a specific method, decide whether you want to

- get out a score or improvement suggestions (summative or formative)
- conduct the study in a lab, on the field, or on the Web
- run the test in one or more locations
- run the test with lots of users or a handful of them
- …

Such decisions are similar to those in usability evaluations

- Guidance for making these decisions exists elsewhere
- We assume you know your evaluation needs on this level
- Once you know the needs, we help you to find a method

Example case: Personal health service

MA thesis work by
Iris Tomaszewski,
Aalto University

Product	Mobile app (+ iPad app at pharmacy)
Design stage	Concept design
Product representation	Storyboard
Purpose of evaluation	Find best concept idea for personal health service
Study location	Meeting room
Participants	4 young pharmacy customers (mothers, young adults)
Time restrictions	3 weeks (from assignment to results report)
Equipment, tools	Audio recorder, assistant for note taking
Skills of researchers	New media student

Exercise 1

Describe a situation in which you are in need of a UX evaluation method

Situation description:

Attribute	Examples
• Product	Smart phone, E-commerce site, Lamp
• Design stage	Concept design, Prototyping, On market
• Product representation	Storyboard, Flash animation, Participant's own device
• Purpose of evaluation	Convince management, Find best design alternative
• Study location	Usability lab, Participants' daily environments, Online study
• Participants	UX experts, All visitors of a web site, Kids (with a parent)
• Time restrictions	3 months, 1 week, Lunch hour (from assignment to results)
• Equipment, tools	Special sensors, Remote video & logging tool, None
• Skills of researchers	Anthropologist, Statistician, Summer trainees, Students

Scenarios of evaluation situations

Case 1

- UX consultancy project in industry, test a prototype

Case 2

- Research project in academia, study existing product on market

Case 3

- Joint research project, compare UX of early design ideas

Case 1

- UX consultancy project in industry, test a prototype
- Anna is a usability specialist in a UX consultancy company of 30 employees. Their client has asked for offers to evaluate an application prototype on vitamin information for use in a specific pharmacy. The client wants to **test the usability of the vitamin info application and its potential to improve the service experience** of typical pharmacy customers during rush hours. The client gives one month to complete the study.
- Anna has now one day to formulate an offer for their client, describing the evaluation process on high level. She is excited to learn about UX evaluation (which their company would like to do more) and wants to see which UX evaluation method(s) could be used to find out answers to the service experience request of the client. She has to make a competitive offer for the client that is most interested in improvement proposals for the prototype.

Case 2

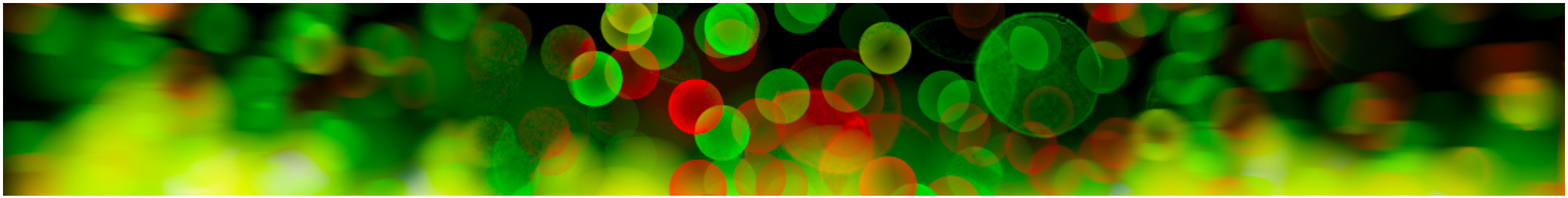
- Research project in academia, study existing product on market
- A group of researchers in a university wants to know **how game play experience changes over time**. They are planning a study to identify the differences between Chinese, British, and North American culture in the long-term UX. They plan to study Angry Birds players over 4 months with a sample of 100 players per country.

Case 3

- Joint research project, compare UX of early design ideas
- A joint research project between academia and industry has ideated some new concepts for families to save energy at home. Six families with 8-15-years-old children living near the university have participated in the ideation phase and they have agreed to be involved in the studies throughout the project. The team now has storyboards (cartoons) of 4 different ideas, and they should select which of the ideas to take further for prototyping. One of the companies involved has the responsibility to build the prototype, but they have resources for implementing one of the ideas only. The project will select one of the ideas based on four criteria: potential impact in energy saving; technical feasibility; cost of the system; and user experience.

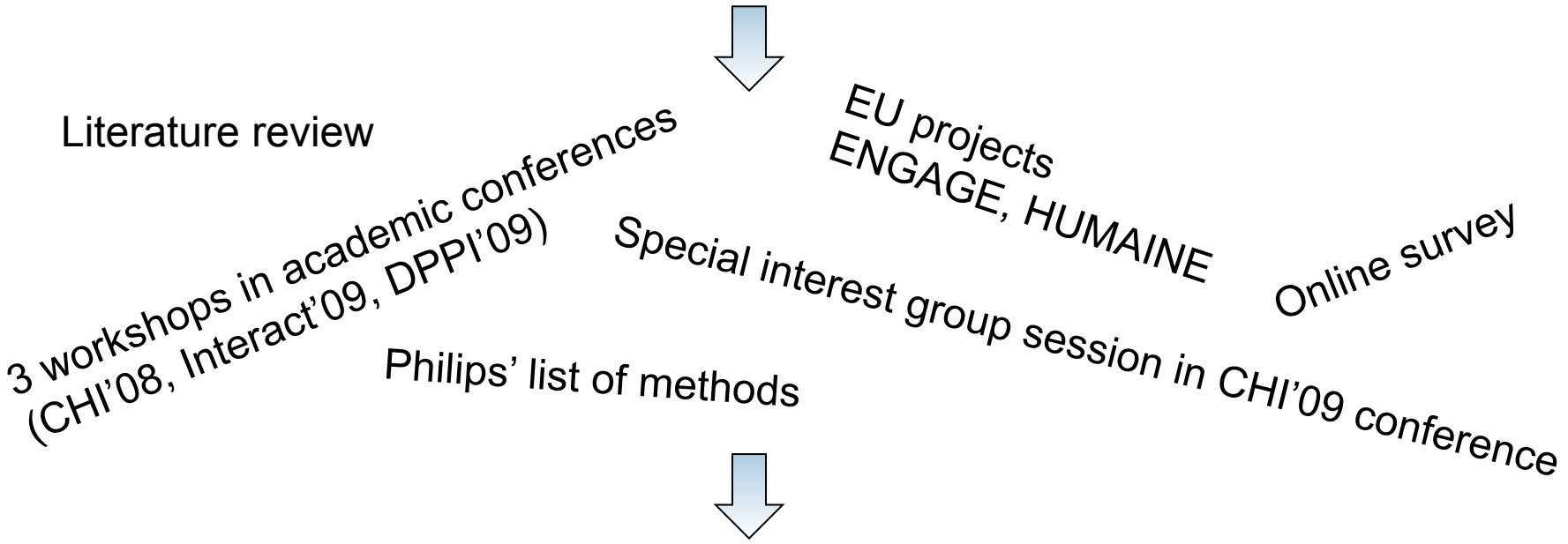
- Tom, the only user experience researcher in the project, has now one week time to estimate **which of the ideas would generate the most excited emotional response** among the family members. The team just wants him to deliver a “UX score” for each idea, so that implementation can start as soon as possible.

Overview of the types of UXEMs available



Collecting UXEMs

Goal: Collect experiential methods from academia and industry



Outcome: Descriptions of 101 UX evaluation methods

Published: 80 UX evaluation methods at allaboutux.org

Method categories

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

See the next 2 slides

Methods for different periods of experience

Moment	Episode	Long-term
e.g. during game play	e.g. reflections after playing the game	e.g. 1 year after starting to play the game

Evaluating emotions	Evaluating an episode	Evaluating long-term UX
Observation	**Observation**	**Self-reporting**
Facial, body, vocal expressions (e.g. smile, lean back, sigh)	Experience think aloud	Questionnaires, Laddering, iScale, Repertory Grid Technique
Psychophysiological measurements	**Self-Reporting**	
Muscle, pupil, heart, skin reactions detected with sensors	Experience sampling, AttrakDiff, Interviews, Day Reconstruction	
Self-reporting		
Verbal: PANAS, Affect Grid Non-verbal: EmotionSlider, EmoCards, PrEmo		

Methods for different phases of development

Conceptual design	Low-fi prototypes	Functional prototypes
e.g. Cartoons about the idea	e.g. Flash demo; 3 core features working	e.g. (Almost) all parts in place; products on market

Design concept	Low-fi prototype	Functional prototype
Visual design	**Visual design**	**Lab test**
Emotional expressions,reactions	Emotional expressions,reactions	Emotional expressions,reactions AttrakDiff
Idea description	**Interaction**	**Field study**
Expert evaluation, Role play: Perspective-Based Inspection	Experience think aloud	Experience sampling, Diary, Day Reconstruction Method
		Market feedback
		Questionnaires, UX Curve / iScale

Experience think aloud:
Traditional think aloud method used in usability testing, but special focus on experiential/ emotional comments (not on allaboutux.org)

Example case: Personal Health service

MA thesis work by
Iris Tomaszewski

Method type

- Field studies
- **➢ Lab studies**
- Online studies
- Questionnaires / Scales

Data

- **➢ Qualitative**
- Quantitative

Studied period of experience

- Before usage
- Snapshots during interaction
- An experience (of a task or activity)
- **➢ Long-term UX**

Development phase,
i.e. material available

- **➢ Scenarios, sketches; i.e. concepts**
- Early prototypes
- Functional prototypes
- Products on market

Evaluator / Information provider

- UX experts
- **➢ One user at a time**
- Groups of users
- **➢ Pairs of users**

Exercise 2

Which method categories would be valid for your sample situation?

Method type
- Field studies
- Lab studies
- Online studies
- Questionnaires / Scales

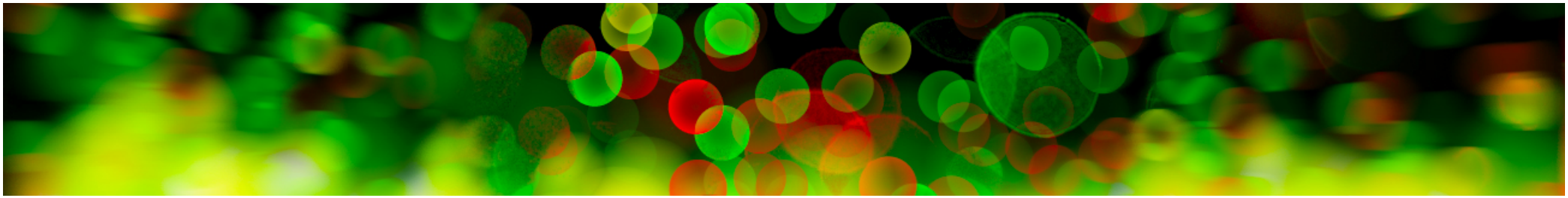
Data
- Qualitative
- Quantitative

Studied period of experience
- Before usage
- Snapshots during interaction
- An experience (of a task or activity)
- Long-term UX

Development phase,
i.e. material available
- Scenarios, sketches; i.e. concepts
- Early prototypes
- Functional prototypes
- Products on market

Evaluator / Information provider
- UX experts
- One user at a time
- Groups of users
- Pairs of users

Examples of different types of UXEMs



…Examples of different types of UXEMs…

Experience Sampling Method

Source: http://www.trackyourhappiness.org/

Experience Sampling Method (ESM)

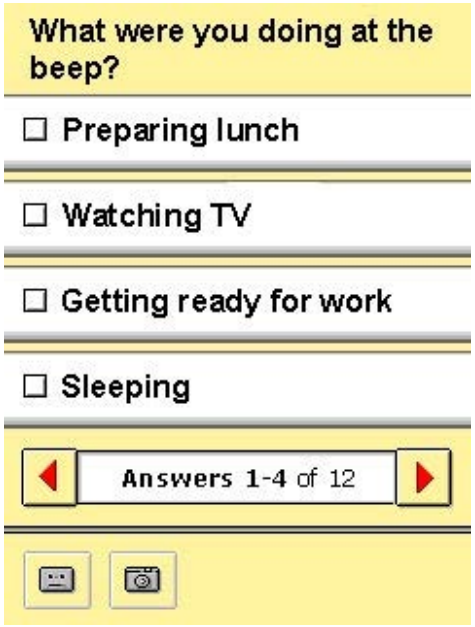
- ESM is a set of methods that is designed to repeatedly request people to document and report their thoughts, feelings and actions outside the laboratory, within the context of their everyday life.
  - Sending requests can be timed, random, triggered by context variables or triggered by user actions
  - Requests and answers can be sent through a variety of devices (internet, web browser, text messaging on mobile phone, etc.)

- Strengths:
  - Does not depend on recall (no memory effects)
  - Provides insights into variability over time
- Weakness:
  - Requires high compliance, high effort (interferes with activity)

Experience Sampling Method (ESM)

Data collection options:

multiple choice question screen



audio note sample

sampling by taking a picture

Source: S. S. Intille, J. Rondoni, C. Kukla, I. Anacona, and L. Bao, "A context-aware experience sampling tool," in *Proceedings of CHI '03 Extended Abstracts on Human Factors in Computing Systems*. New York, NY: ACM Press, 2003, pp. 972-973.

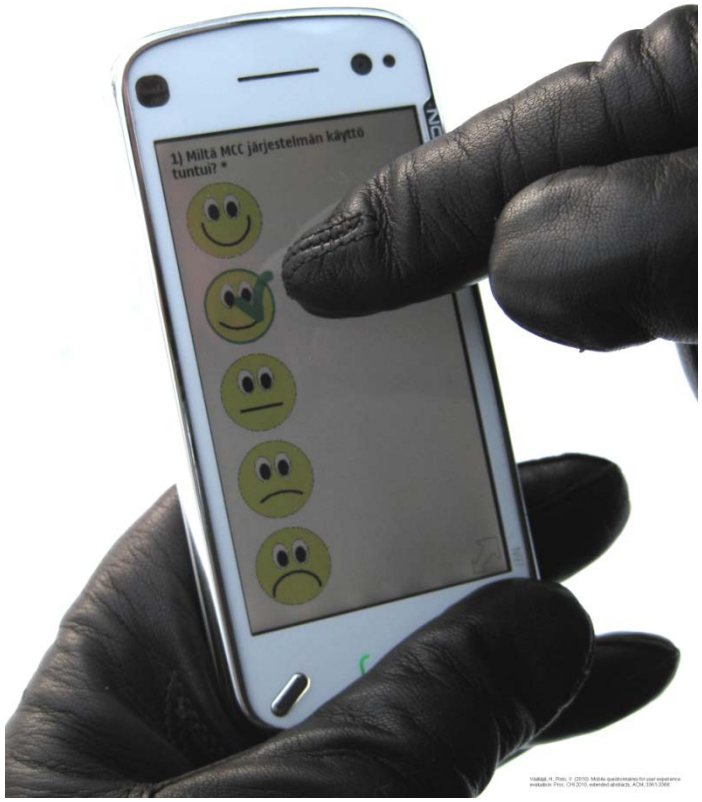
Example case

Case: UX evaluation of a mobile journalism application

Journalists used the mobile jounalism application on their mobile devices on the field

Right after delivering the story, journalists received an SMS with a link to an online questionnaire:

- 5 point scale for emotion/valence
- 5 point scale for emotion/arousal
- Modified *Attrakdiff* questionnaire: hedonic and pragmatic quality perception and judgment of appeal



Heli Väätäjä, Virpi Roto (2010): Mobile questionnaires for user experience evaluation. In CHI 2010 extended abstracts. p 3361-3366.

Applicability

Method type

- ✓ Field studies
- Lab studies
- Online studies
- Questionnaire

Development phase

- Concepts
- Early prototypes
- ✓ Functional prototypes
- ✓ Products on market

Studied period of experience

- Before usage
- ✓ Snapshots during interaction
- An experience (of a task or activity)
- ✓ Long-term UX

Evaluator / Information provider

- UX experts
- ✓ One user at a time
- Groups of users
- Pairs of users

Data

- ✓ Qualitative
- ✓ Quantitative

Applications

- ✓ Web services
- ✓ PC software
- ✓ Mobile software
- ✓ Hardware designs

Requirements

- ✓ Trained researcher
- ✓ Special equipment **(tools available)**

ESM Tools

- Timed ESM, e.g., TrackYourHappiness.org
  - The traditional way: prompts come at certain time
- Context-aware ESM, e.g., MyExperience[1], SocioXensor[2], Maestro[3]
  - ESM installed in the examined system
  - Tracks user actions and prompts the user after interesting activity

References:

1. Froehlich, J., et al., “MyExperience: A System for In Situ Tracing and Capturing of User Feedback on Mobile Phones,” Proc of MobiSys ‘07, (Jun 2007), pp.57-70.
2. Ter Hofte, G.H. Xensible interruptions from your mobile phone In: Proceedings of the ACM MobileHCI 2007, September 9-12, 2007, Singapore, pp. 176-179
3. Meschtscherjakov, A., Reitberger, W., Tscheligi, M. MAESTRO: Orchestrating User Behavior Driven and Context Triggered Experience Sampling. Measuring Behavior 2010, Eindhoven, The Netherlands.

The original ESM

References

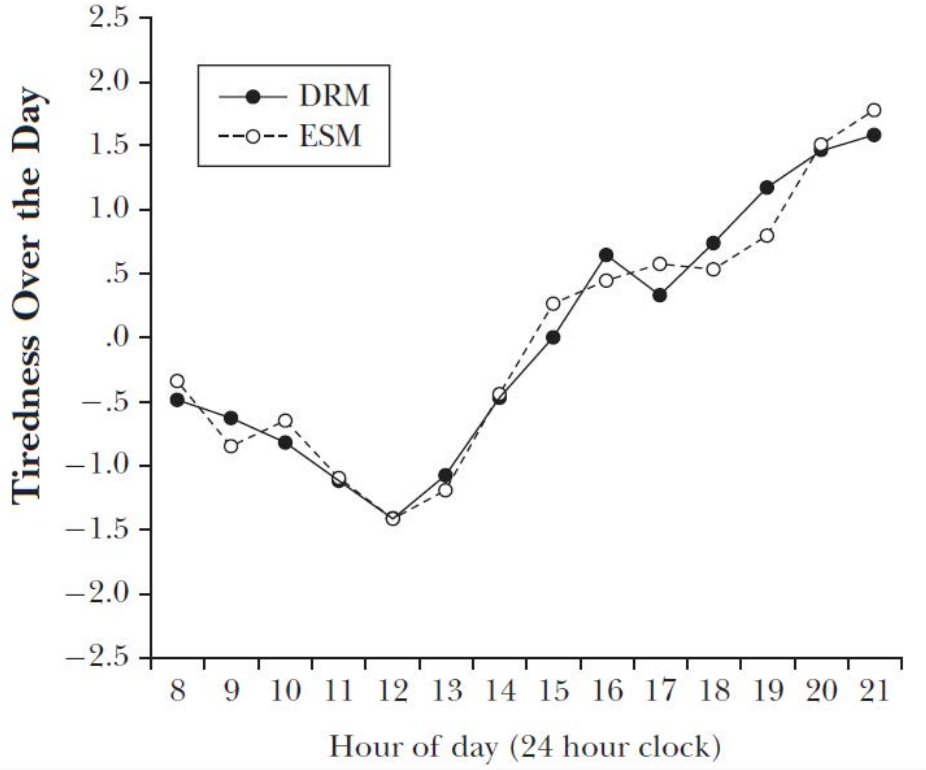
- Larson, R., & Csikszentmihalyi, M. 1983. The experience sampling method. In H. T. Reis (Ed.),
  Naturalistic Approaches to Studying Social Interaction. New Directions for Methodology of Social and Behavioral Science, Vol. 15: 41-56. San Francisco: Jossey-Bass.

- Csikszentmihalyi M, Larson R. 1992. Validity and reliability of the Experience Sampling Method.

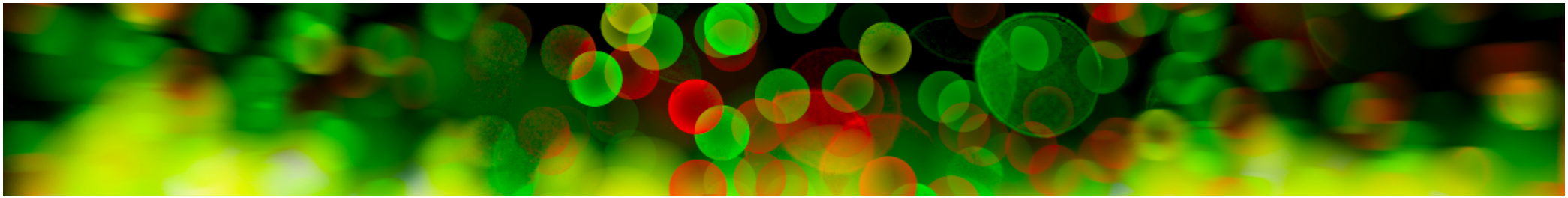
- ScollonCN, Kim-Prieto C, Diener E. Experience sampling:promise and pitfalls, strengths and weaknesses. J Happiness Studies 2003;4:5-34.

…Examples of different types of UXEMs…

Day Reconstruction Method



Kahneman D, Krueger AB, Schkade DA, Schwarz N, Stone AA: A survey method for characterizing daily life experience: the day reconstruction method. Science. 2004 Dec 3;306(5702):1776-80.



Day Reconstruction Method (DRM)

- A retrospective (offline) diary method (Kahneman et al. 2004)
- DRM respondents first revive memories of the previous day by constructing a diary consisting of a **sequence of episodes**
- Then they describe each episode by answering questions about the situation and about the feelings that they experienced, as in experience sampling
- The goal is to provide an accurate picture of the experience associated with activities (e.g., commuting) and circumstances (e.g., a job with time pressure)
- Each experience is recalled in relation to preceding and succeeding ones
- As a result, participants reconstruct the emotional experience on the basis of sufficient episodic information
- Increasing contextual cues, thus avoiding retrospective biases

- DRM may achieve an accuracy close to that of online reporting, e.g. Experience Sampling Method

Example case

- 4-week field study on taking iPhone into use
- 6 participants captured their daily experiences at the end of each day by
  - Day reconstruction: participants listed all activities of the day that somehow related to their iPhone. A brief name and an estimation of time spent for each activity.
  - Experience narration: participants were asked to pick the three *most impactful,* either satisfying or dissatisfying, experiences of that day. For each of the three experiences, participants were asked to write a story that describes in detail the situation, their feelings and their momentary perceptions of the product.
- Data analysis: Categorization of 482 narratives
- Results: Experiences changed over time. Phases of orientation, incorporation, and identification were recognized.

Karapanos, E., Zimmerman, J., Forlizzi, J., Martens, J-B: User Experience Over Time: An Initial Framework. CHI 2009, Boston, USA.

Applicability

Method type

- ✓ Field studies
- Lab studies
- ✓ Online studies
- Questionnaire

Development phase

- Concepts
- Early prototypes
- ✓ Functional prototypes
- ✓ Products on market

Studied period of experience

- Before usage
- Snapshots during interaction
- ✓ An experience (of a task or activity)
- ✓ Long-term UX

Evaluator / Information provider

- UX experts
- ✓ One user at a time
- Groups of users
- Pairs of users

Data

- ✓ Qualitative
- Quantitative

Applications

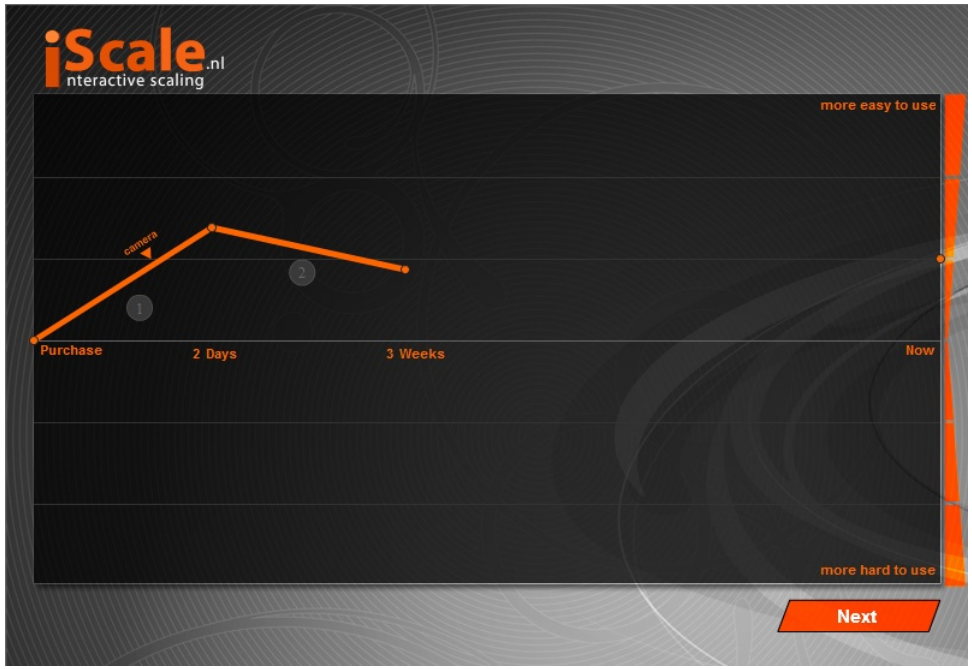
- ✓ Web services
- ✓ PC software
- ✓ Mobile software
- ✓ Hardware designs

Requirements

- ✓ Trained researcher
- Special equipment

…Examples of different types of UXEMs…

iScale



http://ekarapanos.com/

iScale tool

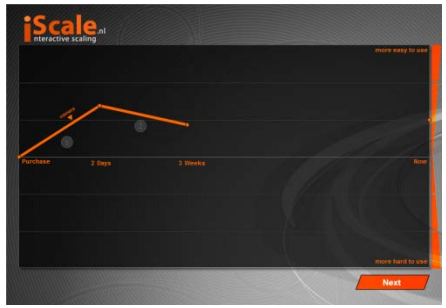
- An online tool for the **retrospective elicitation of longitudinal UX data**
  - Concept partly based on Day Reconstruction Method (DRM, Kahneman et al. 2004)
- iScale uses sketching in the reconstruction of one's experiences with the aim to **minimize retrospection bias**
- **Chronological order helps** the reconstruction of experiences
  - more contextual details surrounding the experienced events
  - the felt emotion is constructed on the basis of the recalled contextual details
  - increases the amount, the richness, and reliability of recalled information
- The results provide support for retrospective techniques as cost-effective alternatives to longitudinal studies

  Video: http://ekarapanos.com/iscale-final/iscale-final.html

Karapanos, E., Martens, J.-B., Hassenzahl, M. (2010) On the Retrospective Assessment of Users' Experiences Over Time: Memory or Actuality?**.** *CHI'10 extended abstracts.*

iScale tool

A validation study of iScale



- Comparison of **two forms of sketching** to help recall:
  - Chronological sketching - users recall step-by-step in forward chronological order
  - Top-down sketching - users first assess the overall change, and split this into periods
  - Chronological sketching was found result to an increase in the amount, the richness, and the test–retest consistency of recalled information. Top-down sketching not.
- Comparison of **iScale tool with free-hand sketching** (pen and paper):
  - Free-hand sketching is more expressive: less rules for the graph and annotations
  - iScale allows modifications of the graph, which helps in reconstructing experiences
  - Free-hand sketching is an alternative to the iScale tool (see also: UX Curve method)
- iScale helps **understanding the long-term UX** of products
- **Sketching assists in the reconstruction** of the context of the experiences

Karapanos, E., Martens, J.-B., Hassenzahl, M. (2012) Reconstructing Experiences with iScale. International Journal of Human-Computer Studies, 70 (11), pp. 849-865

Applicability

Method type

- ✓ Field studies
- ✓ Lab studies
- ✓ Online studies
- Questionnaire

Development phase

- Concepts
- Early prototypes
- ✓ Functional prototypes
- ✓ Products on market

Studied period of experience

- Before usage
- Snapshots during interaction
- An experience (of a task or activity)
- ✓ Long-term UX

Evaluator / Information provider

- UX experts
- ✓ One user at a time
- Groups of users
- Pairs of users

Data

- ✓ Qualitative
- ✓ Quantitative

Applications

- ✓ Web services
- ✓ PC software
- ✓ Mobile software
- Hardware designs

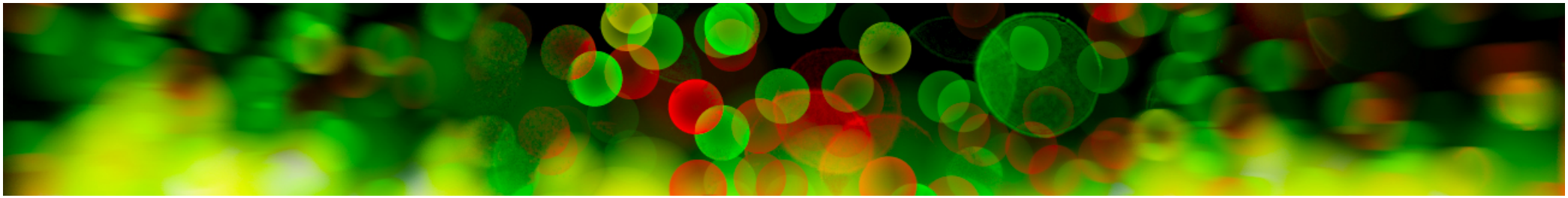
Requirements

- ✓ Trained researcher
- ✓ Special equipment: iScale software

…Examples of different types of UXEMs…

AttrakDiff™ questionnaire

interesting	_ _ _ _ _ _ _	boring
extravagant	_ _ _ _ _ _ _	cheap
exciting	_ _ _ _ _ _ _	dull
exclusive	_ _ _ _ _ _ _	standard
impressive	_ _ _ _ _ _ _	nondescript
original	_ _ _ _ _ _ _	ordinary
innovative	_ _ _ _ _ _ _	conservative



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, e.g. in terms of usability and appearance
- AttrakDiff2 consists of 21 word-pairs (semantic differentials)
  - Seven-step items whose poles are opposites (e.g. "confusing - clear", "good - bad")
- Covers the following components of user experience:
  - pragmatic quality (PQ)
  - hedonic quality: identification (HQ-I)
  - hedonic quality: stimulation (HQ-S)

Hassenzahl, M., Brumester, M. and Koller, F. (2003) 'Attrakdiff: Ein Fragebogen Zur Messung Wahrgenommener Hedonischer Und Pragmatischer Qualität', in Z.J. Bewegung, G. Szwillus and B.G. Teubner (Eds) Interaktion. Mensch & Computer,187–196. Stuttgart, Leipzig: B.G.Teubner

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

Short version of AttrakDiff

- 10-item version of AttrakDiff2 was used by Hassenzahl & Monk 2010
  - Why shorter questionnaire: evaluate a large number of web sites
- In this case, the purpose was to clarify the relation between perceived usability and beauty (may influence the selection of items to the set of 10)

Marc Hassenzahl & Andrew Monk (2010): The Inference of Perceived Usability From Beauty, Human–Computer Interaction, 25:3, 235-260

Pragmatic quality (PQ)

- confusing – structured
- impractical – practical
- unpredictable – predictable
- complicated – simple

Hedonic quality (HQ)

- dull – captivating
- tacky – stylish
- cheap – premium
- unimaginative – creative

Goodness and beauty

- good – bad
- beautiful – ugly

Applicability

Method type

- ✓ Field studies
- ✓ Lab studies
- ✓ Online studies
- ✓ Questionnaire

Development phase

- Concepts
- Early prototypes
- ✓ Functional prototypes
- ✓ Products on market

Studied period of experience

- Before usage
- Snapshots during interaction
- ✓ An experience (of a task or activity)
- ✓ Long-term UX

Evaluator / Information provider

- UX experts
- ✓ One user at a time
- Groups of users
- Pairs of users

Data

- ✓ Qualitative
- ✓ Quantitative

Applications

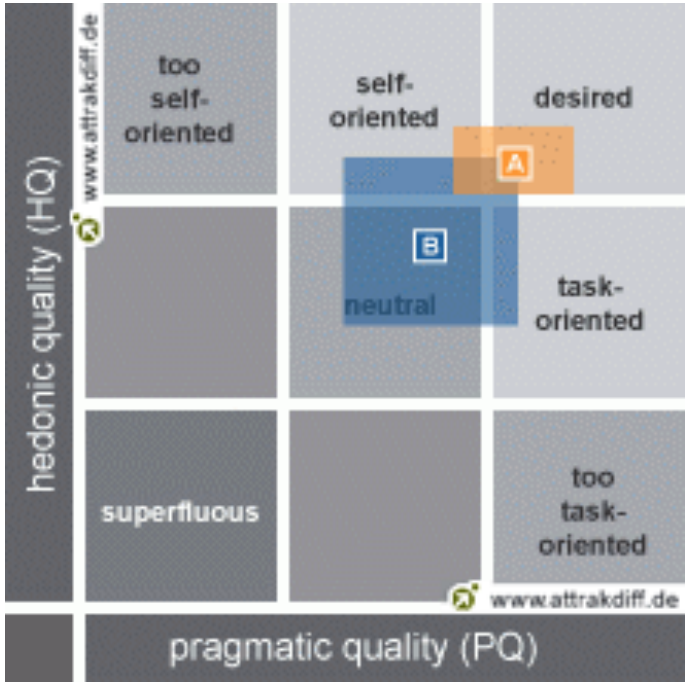
- ✓ Web services
- ✓ PC software
- ✓ Mobile software
- ✓ Hardware designs

Requirements

- Trained researcher
- Special equipment (**tool available**)

Tool: AttrakDiff Basic

- AttrakDiff™ Basic (free) enables you to have a product evaluated online by users
  - Single Evaluation, or
  - Comparison before – after, or
  - Comparison of product A with product B
- Graphic summary report is produced by the tool →
- Available in English and in German
- (The Premium version is licenced)

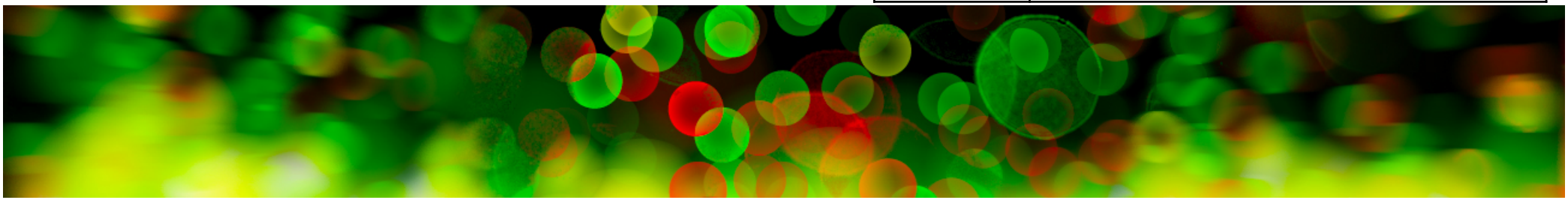


- Adds 7 items on 'Attractiveness' to AttrakDiff2
- In total 28 items

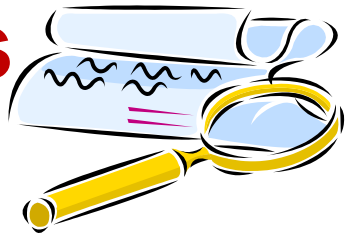
http://www.attrakdiff.de/en/Services/AttrakDiff-Basic/

UX Expert evaluation with UX heuristics

	Name of the heuristic
H1	*Usage and creation of composite services*
H2	*Cross-platform service access*
H3	*Social interaction and navigation*
H4	*Dynamic service features*
H5	*Context-aware services and contextually enriched content*
H6	*General UX-related issues*



Expert evaluation with UX heuristics



- UX experts evaluate a service using a predefined set of heuristics
- Heuristics indicate both pragmatic and hedonic instances of each heuristic (Väänänen-Vainio-Mattila & Wäljas, 2009)
- Experts mark down both positive and negative findings (experiences)
- Resulting in a broader understanding of UX of a product, when compared to ”standard” heuristic evaluation

	Name of the heuristic
H1	*Usage and creation of composite services*
H2	*Cross-platform service access*
H3	*Social interaction and navigation*
H4	*Dynamic service features*
H5	*Context-aware services and contextually enriched content*
H6	*General UX-related issues*

H4: Dynamic service features

Pragmatic aspects of the heuristic:

- When users enter the service, it is possible to gain an overview of the recent changes in the service
- While using the service users can easily find the new, interesting content

Hedonic aspects of the heuristic:

- The service feels like a lively place where it is enjoyable to spend time
- The service satisfies users’ curiosity/seeking of knowledge by frequently offering interesting content

Case: Cross-platform Web services

- Development of a set of heuristics for Web2.0 types of cross-platform (multi-device) services (see Väänänen-Vainio-Mattila & Wäljas 2009, 2010)
- Evaluation of three Web2.0 services (Facebook, Nokia Sports Tracker, TripAdvisor travel service) by three experts for each service,
- Approximately two days evaluation time over 1-2 weeks
- Of the total number of 255 evaluation findings, 93 (36%) were labelled positive and 162 (64%) negative, 191 (74%) were seen pragmatic and 64 (26%) hedonic
- Rich picture of service UX

References:

Väänänen-Vainio-Mattila, K., Wäljas, M. Developing an Expert Evaluation Method for User eXperience of Cross-Platform Web Services. Proceedings of Mindtrek'09, ACM 2009.

Väänänen-Vainio-Mattila, K., Wäljas, M. Evaluating User Experience of Cross-Platform Web Services with a Heuristic Evaluation Method, accepted for Publication in Int. Journal of Art and Technology (IJART), 2010.

Applicability

Method type

- Field studies
- ✓ Lab studies
- ✓ Online studies
- ✓ Questionnaire

Development phase

- ✓ Concepts
- ✓ Early prototypes
- ✓ Functional prototypes
- ✓ Products on market

Studied period of experience

- ✓ Before usage
- ✓ Snapshots during interaction
- ✓ An experience (of a task or activity)
  Long-term UX

Evaluator / Information provider

- ✓ UX experts
- One user at a time
- Groups of users
- Pairs of users

Data

- ✓ Qualitative
- ✓ Quantitative

Applications

- ✓ Web services
- ✓ PC software
- ✓ Mobile software
- ✓ Hardware designs

Requirements

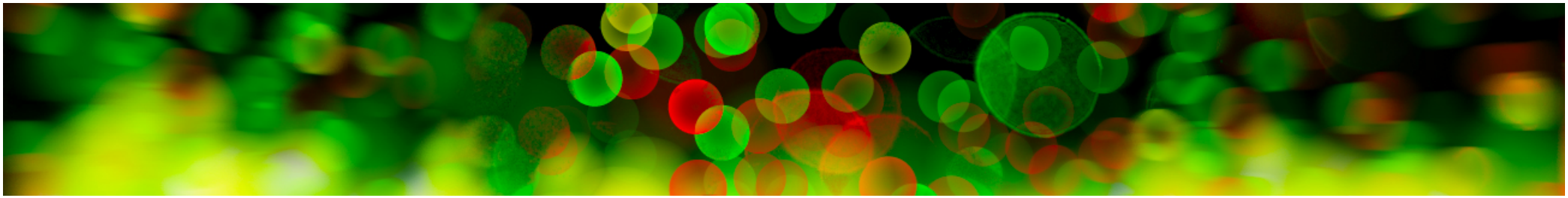
- ✓ Trained researcher
- Special equipment

More expert evaluation methods in our pool

- Immersion
- Perspective-based inspection
- Playability heuristics
- Property checklists

…Examples of different types of UXEMs…

Psycho-physiological measurements



Psycho-physiological measurements

- Brain, Body, and Bytes CHI 2010 Workshop Presentations
  http://www.eecs.tufts.edu/~agirou01/workshop/

Psycho-physiological measurements

- 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

Example Case

- Game play against a friend vs. a computer
- Participants experienced higher physiological values when playing against a friend than against a computer

- Reference: Octavia, J. et al: Exploring Psycho-physiological Measures for the Design and Behavior of Socially-Aware Avatars in Ubicomp Environments. Psychophysiological User Interaction Workshop at CHI 2010 http://www.eecs.tufts.edu/~agirou01/workshop/

A well-known reference to psychophysiological measures for UX:
Mandryk, R.L., Inkpen, K., and Calvert, T.W.
Using psychophysiological techniques to measure user experience with entertainment technologies.
Behaviour and Information Technology (Special Issue on User Experience), 25,2, (2006), 141-158.

Applicability

Method type

- Field studies
- ✓ Lab studies
- Online studies
- Questionnaire

Development phase

- ✓ Concepts
- Early prototypes
- ✓ Functional prototypes
- ✓ Products on market

Studied period of experience

- ✓ Before usage
- ✓ Snapshots during interaction
- An experience (of a task or activity)
- Long-term UX

Evaluator / Information provider

- UX experts
- ✓ One user at a time
- Groups of users
- Pairs of users

Data

- Qualitative
- ✓ Quantitative

Applications

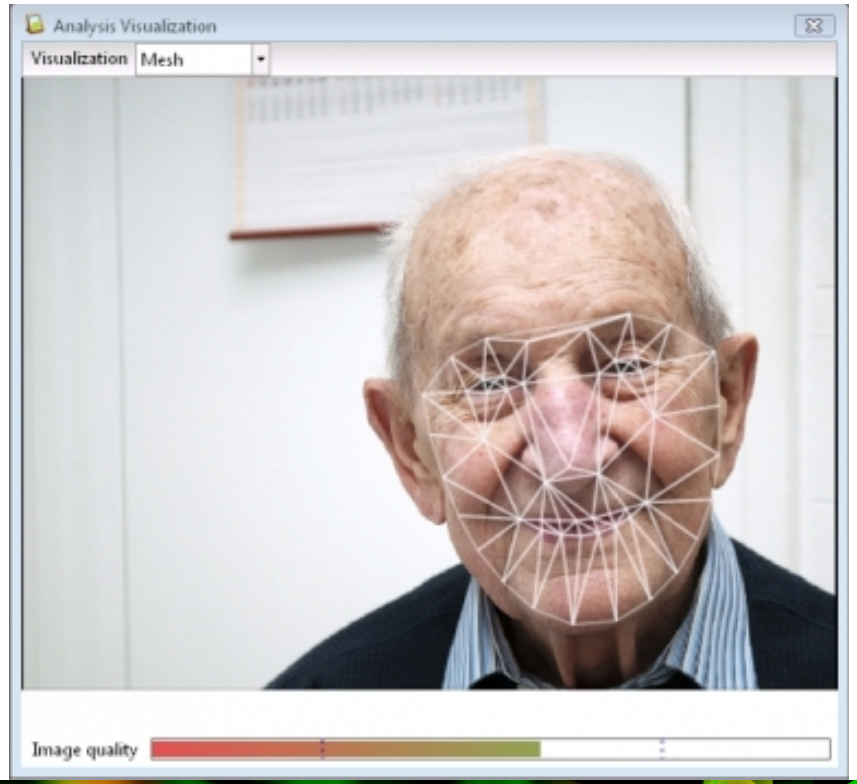
- ✓ Web services
- ✓ PC software
- Mobile software
- Hardware designs

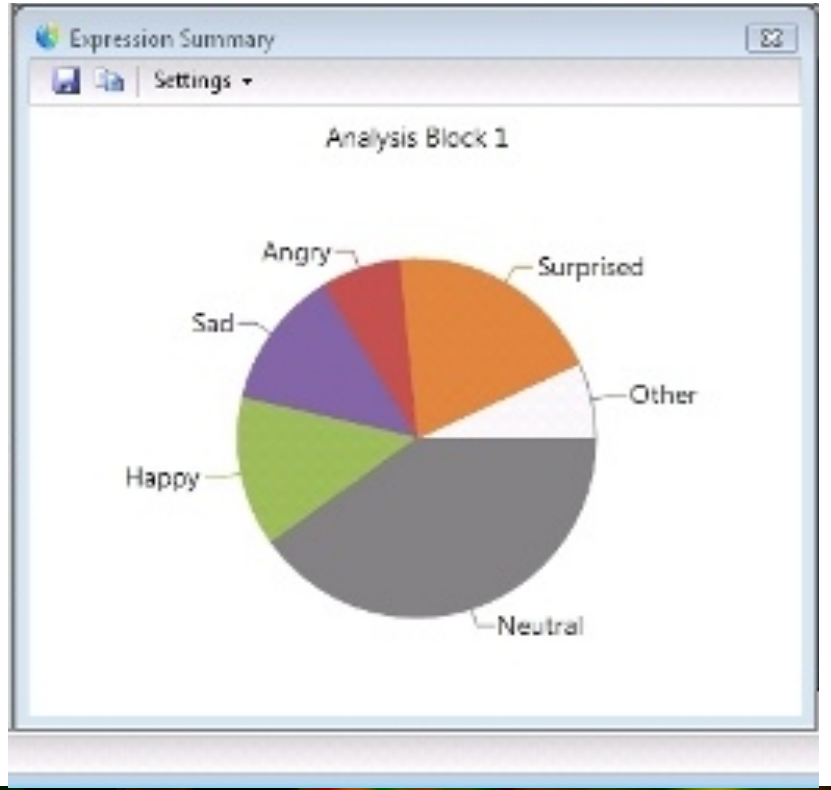
Requirements

- ✓ Trained researcher
- ✓ Special equipment

…Examples of different types of UXEMs…

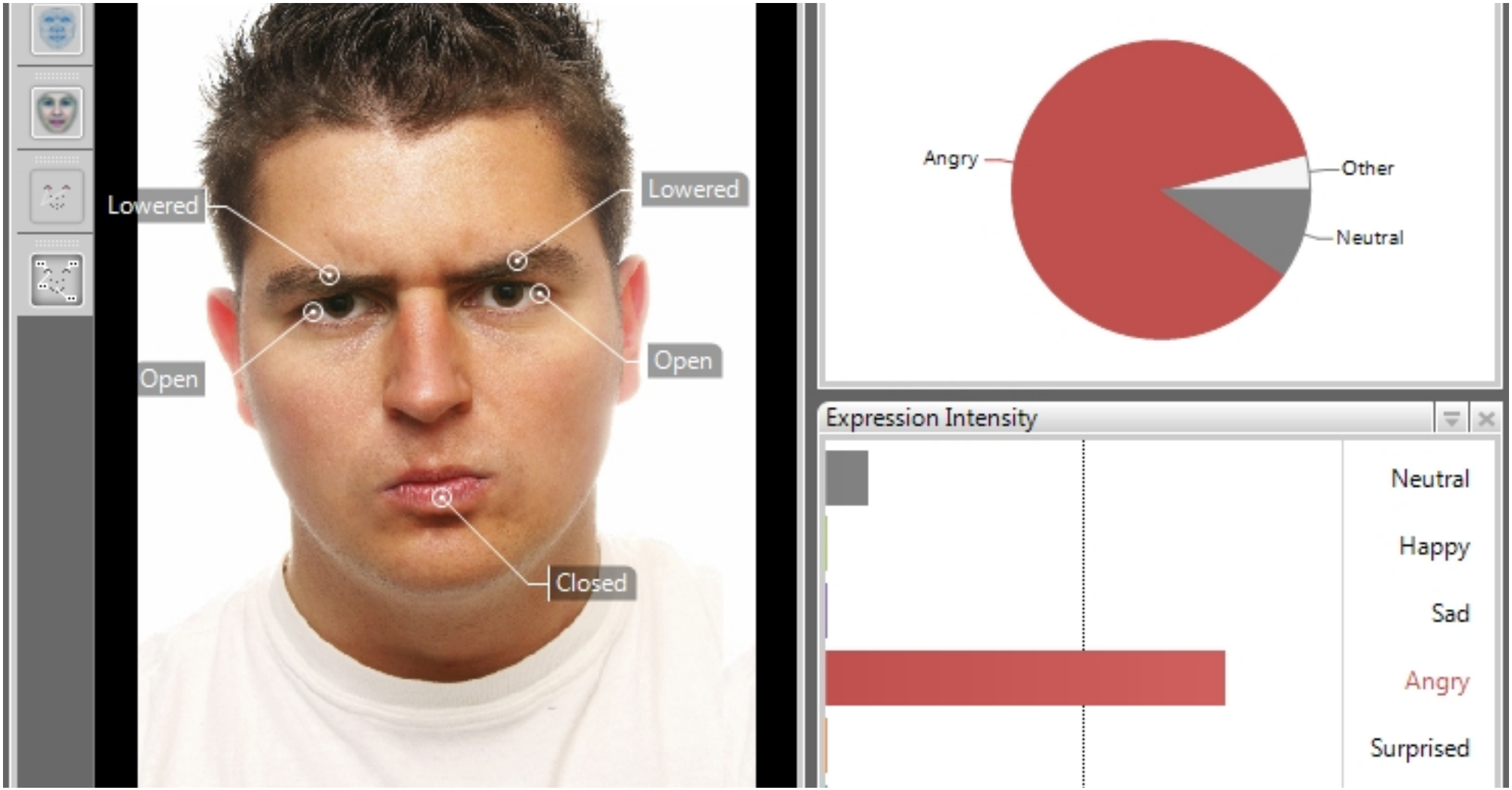
FaceReader





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

Facial expressions communicate affect



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 or live via a webcam.
- FaceReader™ constructs a model of the face from the video and a neural network classifies the facial expressions.
- 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:

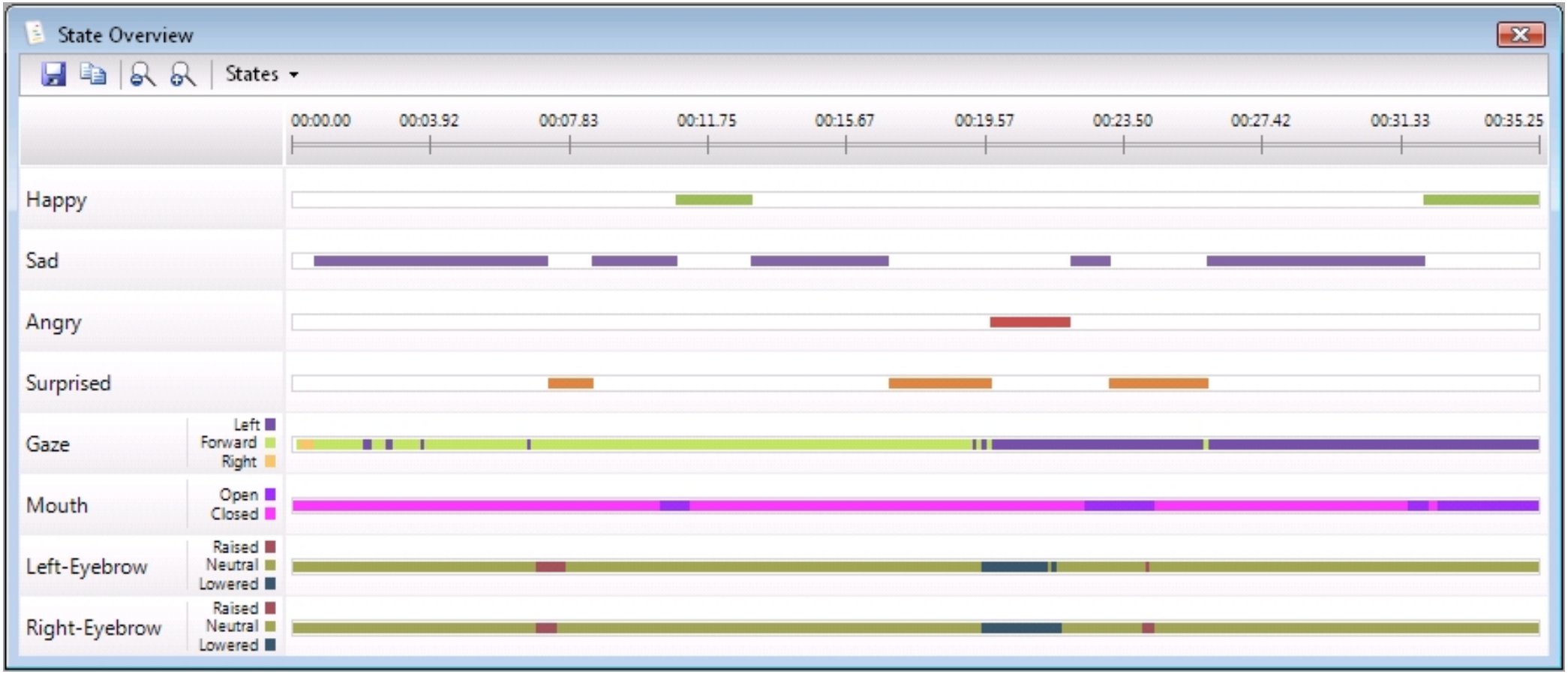
- Objective assessment of a person's emotion

Weakness:

- Data limited to six basic emotions.

FaceReader

Click for Demo Video



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

Multidisciplinary application domain

- Psychology: how do people respond to stimuli, e.g. in fear research?
- Education: observing students' facial expressions can support the development of educational tools.
- **Human-computer interaction**: facial signals can provide valuable information about the user experience.
- Usability testing: emotional expressions can indicate the ease of use and efficiency of user interfaces.
- Market research: how do people respond to a new commercial?
- Consumer behavior: how do participants in a sensory panel react?

Example case

- Automatically evaluate emotions as part of a game

- The “Emotional Flowers” game harnesses the player’s emotions as the primary means for the game interaction. Within the game the player’s facial expression of emotion is used to control the growth of a flower.

Bernhaupt R, Boldt A, Mirlacher T, Wilfinger D, and Tscheligi M (2007). Using emotion in games: emotional flowers. In Proceedings of the international Conference on Advances in Computer Entertainment Technology (Salzburg, Austria, June 13 - 15, 2007). ACE '07, vol. 203. ACM Press, New York, NY, 41-48.

FaceReader

Applicability

Method type

- Field studies
- ✓ Lab studies
- Online studies
- Questionnaire

Development phase

- Concepts
- Early prototypes
- ✓ Functional prototypes
- ✓ Products on market

Studied period of experience

- Before usage
- Snapshots during interaction
- ✓ An experience (of a task or activity)
- Long-term UX

Evaluator / Information provider

- UX experts
- ✓ One user at a time
- Groups of users
- Pairs of users

Data

- Qualitative
- ✓ Quantitative

Applications

- ✓ Web services
- ✓ PC software
- ✓ Mobile software
- Hardware designs

Requirements

- Trained researcher
- ✓ Special equipment

More information

References:

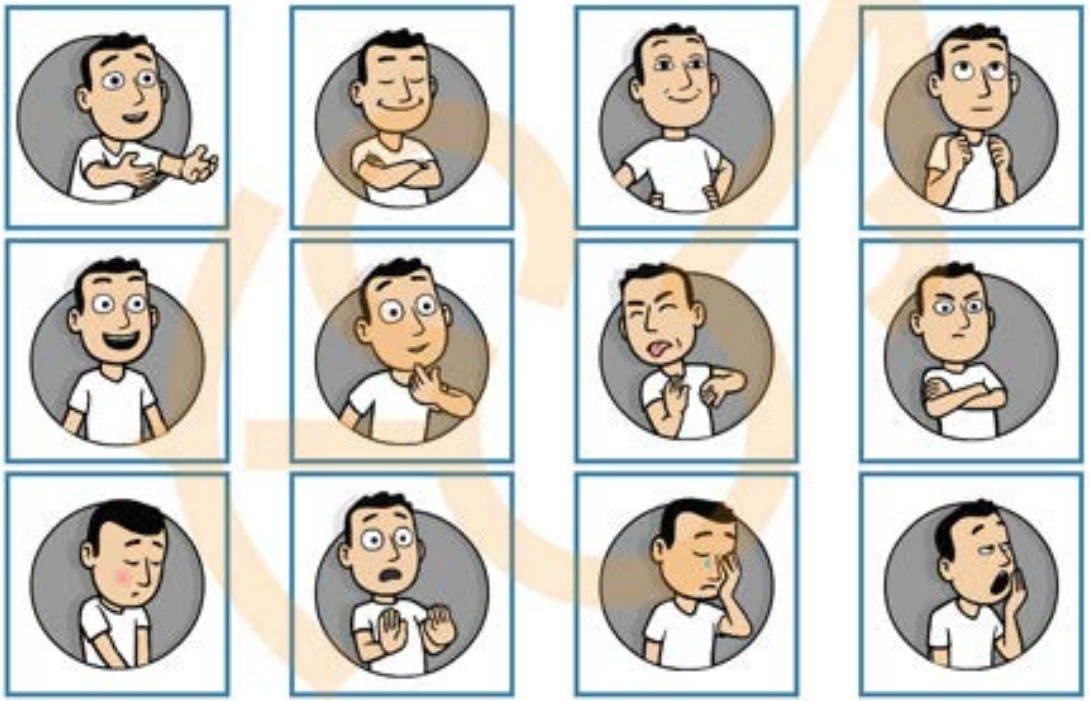
- Uyl, M. den, Kuilenburg, H. van & Lebert, E. (2005). FaceReader: an online facial expression recognition system. Measuring Behavior 2005, 5th International Conference on Methods and Techniques in Behavioral Research, 30 August - 2 September 2005, Wageningen, The Netherlands.

- FaceReader is available from Noldus Information Technology

  www.Noldus.com

…Examples of different types of UXEMs…

PrEmo



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.

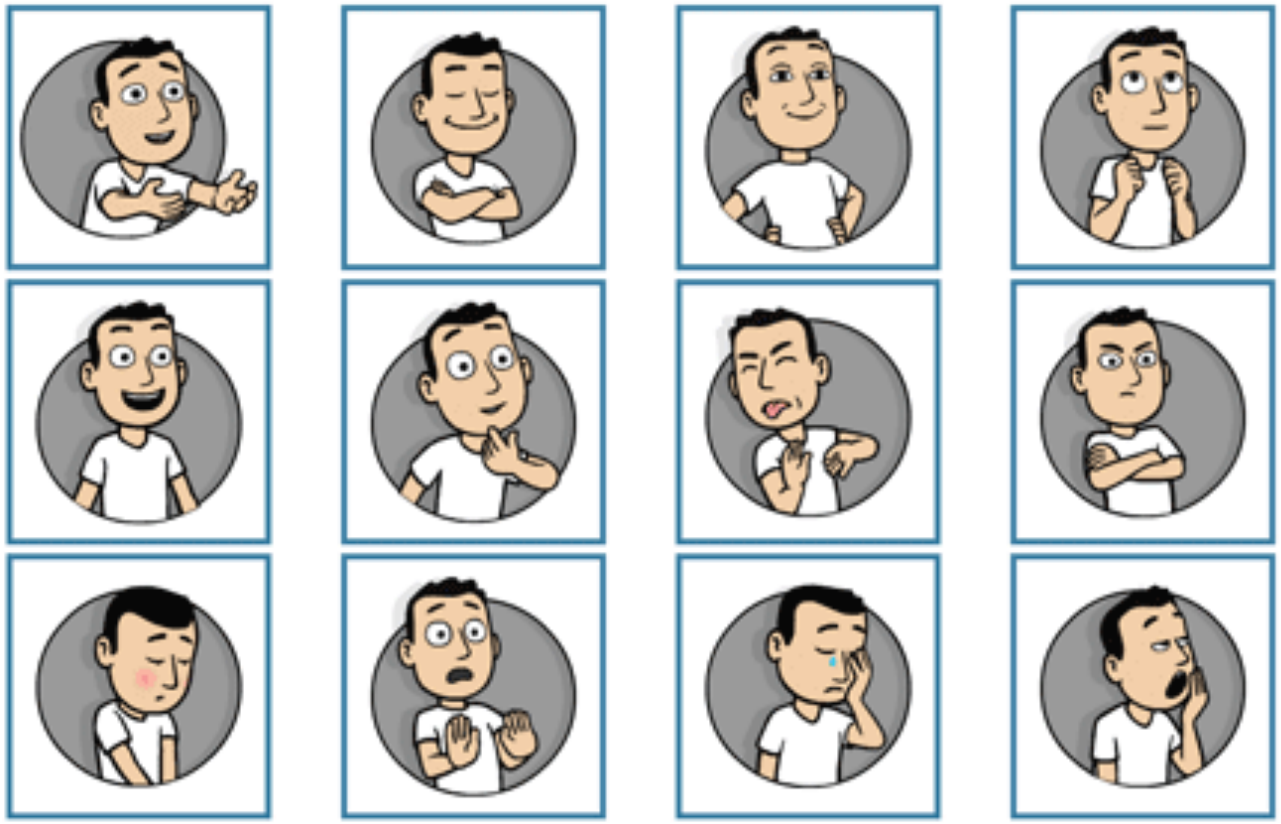
PrEmo

- Procedure:
  - Show product, system, etc. to participant
  - For each of 14 emotions ask participant to indicate how intense the emotion is felt.

    "To which extent do the feelings expressed by the characters correspond with your own feelings towards the stimulus?"

- PrEmo is a licensed commercial product available through Susagroup.com: www.premotool.com

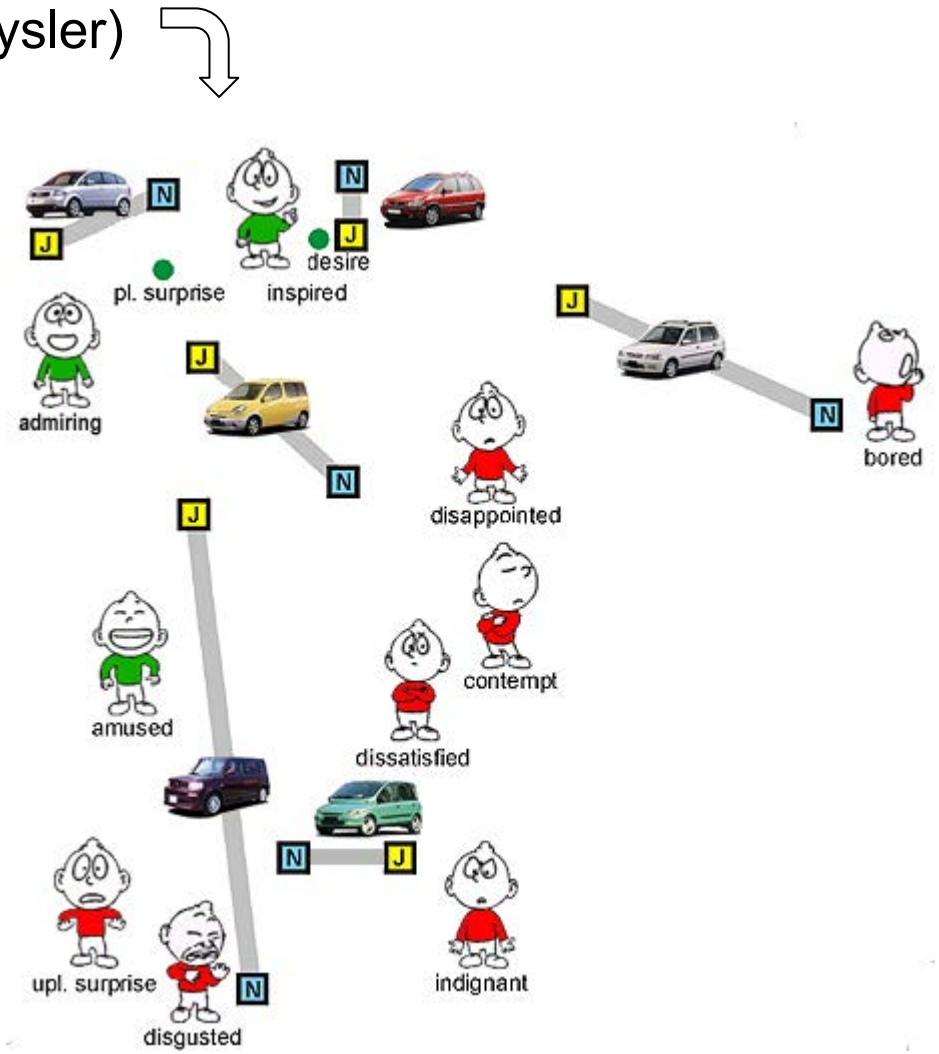
PrEmo (latest version) http://www.premotool.com/about-premo/how-does-premo-work



I feel this strongly 4 3 2 1 I do not feel this 0

Example cases

- automotive design (Mitsubishi; Daimler-Chrysler)
- biscuit packages (Bolletje)
- office chairs (Ahrend)
- mobile telephones (KPN Telecom)
- websites (Starchild foundation)
- gearlevers (Johnson Controls)
- cooking oil packages (Diamant)
- city districts (city of Enschede)
- sport shoes (Nike)
- childrens wheelchairs (Havenkamp)
- fragrances (Procter & Gamble)
- meals and snacks (KLM; Barilla)



More at http://www.premotool.com/cases

PrEmo

Applicability

Method type

- ✓ Field studies
- ✓ Lab studies
- ✓ Online studies
- Questionnaire

Development phase

- Concepts
- Early prototypes
- ✓ Functional prototypes
- ✓ Products on market

Studied period of experience

- ✓ Before usage
- ✓ Snapshots during interaction
- An experience (of a task or activity)
- Long-term UX

Evaluator / Information provider

- UX experts
- ✓ One user at a time
- Groups of users
- Pairs of users

Data

- ✓ Qualitative
- ✓ Quantitative

Applications

- ✓ Web services
- ✓ PC software
- ✓ Mobile software
- ✓ Hardware designs

Requirements

- Trained researcher
- ✓ Special equipment: licenced software

More information

References:

http://studiolab.io.tudelft.nl/desmet/PrEmo

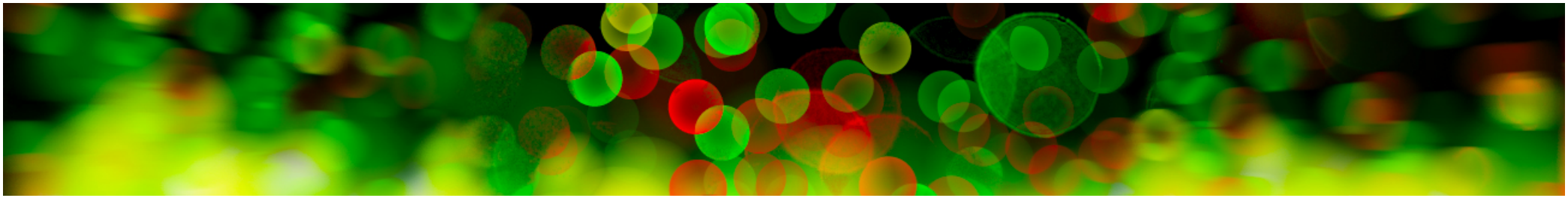
Desmet, P.M.A. (2003). Measuring emotion; development and application of an instrument to measure emotional responses to products. In: M.A. Blythe, A.F. Monk, K. Overbeeke, & P.C. Wright (Eds.), Funology: from usability to enjoyment (pp. 111-123). Dordrecht: Kluwer Academic Publishers.

Available through Susagroup.com:

www.premo-online.com

…Examples of different types of UXEMs…

Sensual Evaluation Instrument



Sensual Evaluation Instrument (SEI)

- Sculpted objects to help participants describe their emotion (i.e. non-verbal)
- Works for stationary tests to investigate feelings

How?

Procedure:

1. Describe participant the idea of indicating affect by objects of different shapes.
2. Let the participant get familiar with the objects by encouraging to handle them.
3. Instruct the participant to pick one or more objects that best describe the emotion during using a system (or after showing other stimulus).
4. Depending on the case, researcher can stay or leave the room.
5. Interview the user afterwards about the emotions they experienced when the user picked an object. A video recording helps them to recall the cases.

Applicability

Method type

- Field studies
- ✓ Lab studies
- Online studies
- Questionnaire

Development phase

- Concepts
- Early prototypes
- ✓ Functional prototypes
- ✓ Products on market

Studied period of experience

- Before usage
- ✓ Snapshots during interaction
- An experience (of a task or activity)
- Long-term UX

Evaluator / Information provider

- UX experts
- ✓ One user at a time
- Groups of users
- Pairs of users

Data

- ✓ Qualitative
- Quantitative

Applications

- ✓ Web services
- ✓ PC software
- Mobile software
- ✓ Hardware designs

Requirements

- ✓ Trained researcher
- ✓ Special equipment: the objects

More information

Strengths:

- Multicultural, since does not rely on words
- Picking an object is less disturbing than filling in a questionnaire or verbal explanations of the emotion
- Fun for participants

Weaknesses:

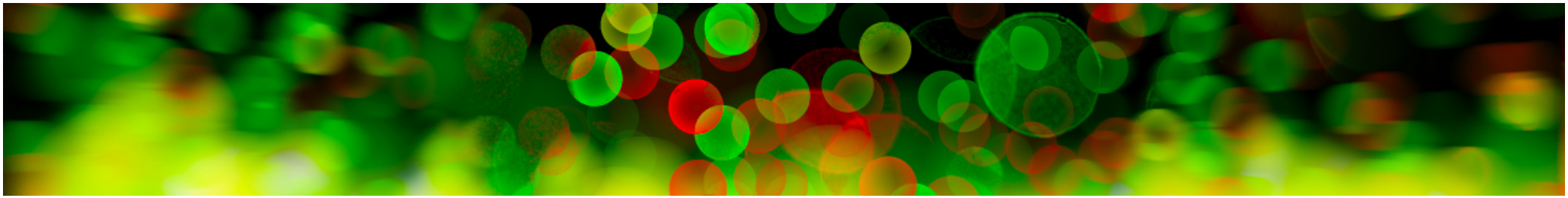
- You need to acquire the objects
- Bound to a stationary place where objects are easily accessible
- Since objects may mean different things to different participants, analysis is hard

References

Isbister, K., Höök, K., Sharp, M., and Laaksolahti, J. 2006. The sensual evaluation instrument: developing an affective evaluation tool. In Proc CHI '06. ACM, New York, NY, 1163-1172

Conclusion

Various methods are available for evaluating how people feel about a design



What's best for You?

No lack of methods, but how to choose the best one for the purpose?

www.allaboutux.org

- Aims to help in choosing the right method
- Launched late 2010
- Now more than 80 UX evaluation methods
  - We published only the ones with a reference to more information
- Method descriptions come from various people
  - Participants in UXEM workshops
  - Participants in a SIG session in CHI conference
  - Other existing method repositories: ENGAGE, HUMAINE

Exercise 3

Selecting one or more specific methods

- Choose one of the collected situations you want to find a method for
  - See exercise 1
- Find out which method categories would fit the chosen situation
  - See exercise 2
- On www.allaboutux.org site, search for a method suitable for the situation
  - You can use the Advanced Search tool
  - You can browse the method lists

Thank you!

We hope this was helpful.

All the best with your evaluation efforts!

Virpi, Arnold, Effie, Kaisa, and Marianna