

# HCI Design & Evaluation

Lecture 5 – Hi-fi prototyping, Tools, Usability testing

Monday 13-12-2021

Randy Klaassen

# About last/this week...

- Do you have nice lo-fi prototypes?
- Exam
  - No dictionary, closed book, no other materials allowed

Friday, 17 December 2021

08:45 - 10:15

---

**Location(s)**

Sportcentrum SC1

**Staff member(s)**

MB van Riemsdijk

R Klaassen

**Student set(s)**

AM/CS MOD06 Group 01 (double degree) CS/BIT MOD06 Group 01 CS/BIT MOD06 Group 02 CS/BIT MOD06 Group 03

CS/BIT MOD06 Group 04 CS/BIT MOD06 Group 05 CS/BIT MOD06 Group 06

**Note**

Partial exam 1

# The course

Week	Lecture (Location in <a href="#">Rooster</a> )	on	Tutorial	Deadlines	Readings
1	Introduction HCI Course overview, topic and project	15.11	Ideation activity * Literature * Brainstorm * top 3 ideas	Proposal report (Fri 19-11)	Klemmer, Lecture 1. Available: <a href="#">YouTube</a> Preece, Chapter 2 <a href="#">Link</a> Lazar, Chapter 1 <a href="#">Link</a>
2	Concepts, Ideation, Value Sensitive Design	22.11	Concept and Ideation * Scenario's, persona's, Values * Final #1 top idea * Video shotlist of above	Concept video, persona and scenario report (Fri 26-11)	Scenario-Based Design - Mary Beth <a href="#">LINK</a> Value Sensitive Design and Information Systems - Batya Friedman <a href="#">LINK</a> CHI - <a href="#">Guide to a Successful Video Submission</a>
3	User confrontation	29.11	User confrontation * Interview script * Interview with peer groups * Interview results and conclusions	Interview report (Fri 03-12)	Lazar, Chapter 5 Surveys <a href="#">Link</a> Lazar, Chapter 8 Interviews <a href="#">Link</a> Lazar, Chapter 11 Qualitative Data <a href="#">Link</a>
4	Lo-fi prototyping	06.12	Prototyping part I * (digital) prototyping * Setup user evaluation	-	Preece, Chapter 12 <a href="#">Link</a> Klemmer, Lecture 2, The Power of Prototyping <a href="#">Link</a> Lazar, Chapter 15 <a href="#">Link</a> Lazar, Chapter 16 <a href="#">Link</a>
5	Hi-Fi Prototyping	13.12	Prototyping part II * (digital) prototyping * User evaluation * Conclusions and results	Lo-Fi prototype report (Fri 17-12) <b>Exam part I (Fri 17-12)</b>	Lazar, Chapter 10 <a href="#">Link</a>
6	Experiment Design, Research methods and Data Analysis	20.12	Prototyping part III * Hi-fi prototyping	-	Lazar Chapter 2 <a href="#">Link</a> Lazar, Chapter 3.1, 3.2, 3.3 <a href="#">Link</a> Lazar, Chapter 4 <a href="#">Link</a> Usability.gov <a href="#">Link</a> ; Quant. Spec. <a href="#">Link</a>
Holiday					
7	Ethics	10.01	Prototyping part III * Hi-fi prototyping * Experiment design + ethics	Hi-Fi Prototype report (Fri 14-01) Controlled Experiment report (Sun 16-01)	<i>provided during lecture</i>
8	Guest lecture	14.01	User testing ( <b>Resits required</b> )	<b>Exam part II (Fri 21-01)</b>	-
9	n/a		n/a	Final project Report (Sun 23-01)	
10	n/a		n/a	Oral Reflection (TBD) (option) resubmission Final project Report (Sun 06-02)	

## Readings List:

Rosson, M. B. & Carroll, J. M. (2002) Scenario-Based Design. [LINK](#)  
 Friedman, B., Kahn Jr., P. H. , & Borning, A. (2013) Value Sensitive Design and Information Systems [LINK](#)  
 Preece, Sharp & Rogers (2019). *Interaction Design: beyond human-computer interaction*. 5th edition. John Wiley. [Link](#)  
 Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research methods in human-computer interaction*. Morgan Kaufmann. [Link](#)

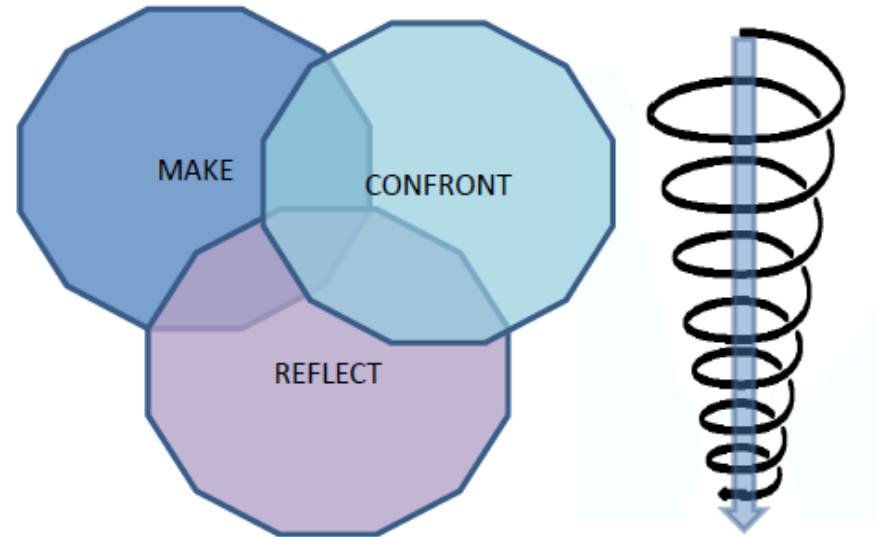


# Today

- Hi-fi prototyping
- Tools – TensorFlow lite, IBM Watson, Dialogflow, ...
- Usability testing

# results after Lo-Fi prototyping

- Challenges?
- Open questions?
- Lessons learned?

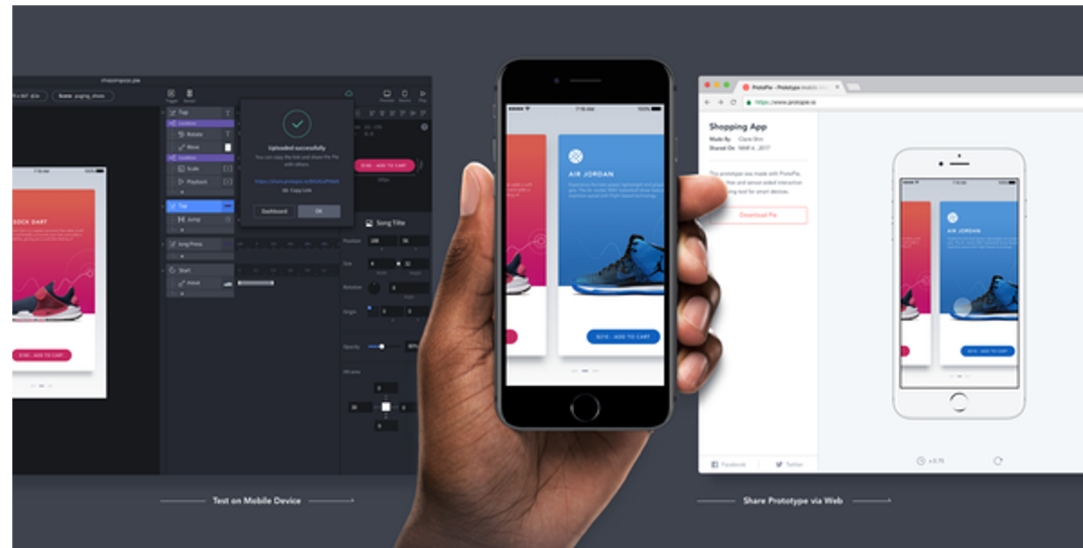


*Figure 1: Activities and phases in the iterative method underlying our design work*

- Today's lessons will bring you a step further towards the final stage of your product.

# Going towards Hi-Fi prototyping

- A high-fidelity (high-fi or hi-fi, sometimes in industry digital or native) prototype is a **computer based** or **physical interactive representation** of (a part of) the concept in its closest similarity to the final design in terms of details and functionality.



# Going towards Hi-Fi prototyping

- Look and feel + interaction /UX + (semi) autonomy / programmed(hard coded): Example Native Prototyping



# Going towards Hi-Fi prototyping

- Look and feel + interaction /UX + (semi) autonomy / programmed: VR prototyping



# Remember! → Prototypes

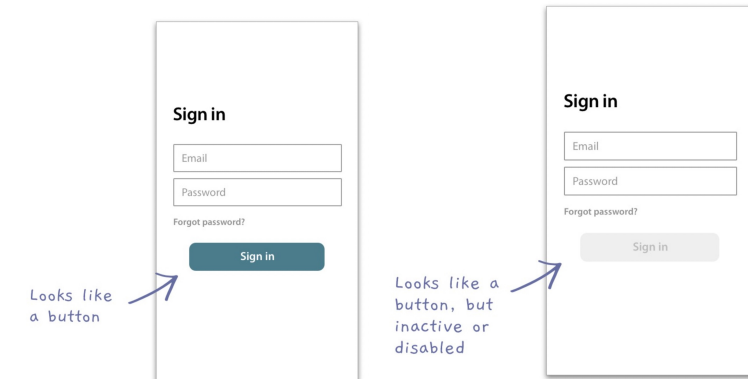
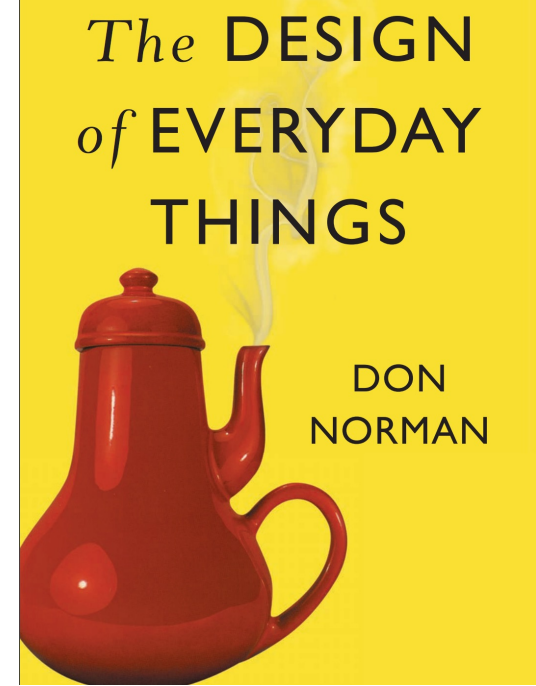
- Helps communicate the story of your app/technology/system for feedback
- Stops you from building a polished product with glaring flaws
  - Spaghetti code, “quick and dirty”, hacking
  - Explore the tech possibilities
- Prevents the need for fixing an error after development
  - Save time and money



# Prototyping – intermezzo

Prototyping can happen on four dimensions:

- **Breadth** (Number of functions/features included)
  - Interaction + UX + Tech
- **Depth** (how well functions/features are worked out in the backend)
  - Tech + Affordances
- **Look** (appearance: fonts, colors, graphics, affordances)
  - Affordances + Signifiers
- **Interaction** (how is input and output handled)
  - UX + Affordances + Signifiers



# Visual Design

## Gestalt Psychology Principles

- The Gestalt theory is guided by 5 principles
  - Proximity
  - Similarity
  - Continuation
  - Closure
  - Figure/Ground.

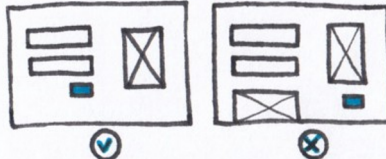


# Proximity

**PROXIMITY**

THE OBJECTS THAT ARE CLOSER TO EACH OTHER ARE PERCEIVED AS MORE RELATED THAN THE ONES THAT ARE NOT POSITIONED NEAR THEM.

ALSO, THE ELEMENTS THAT ARE PLACED CLOSER ARE SEEN AS PART OF THE SAME GROUP OR CATEGORY.



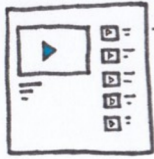
PLACE CONNECTED CONTROLS NEAR TO EACH OTHER!

# Similarity

**SIMILARITY**

IF TWO OBJECTS HAVE SIMILAR CHARACTERISTICS, THESE OBJECTS ARE PERCEIVED AS MORE RELATED THAN THE ONES THAT DON'T SHARE THESE QUALITIES.

WE ALSO ASSUME THAT SIMILAR OBJECTS HAVE SAME FUNCTION/ROLE.



RELATED VIDEOS ON YOUTUBE

SIMILAR CHARACTERISTICS

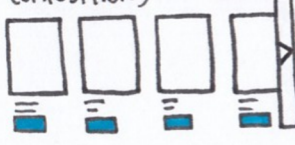
- COLOR
- SIZE
- SHAPE
- FONTS
- TEXTURE ETC.

# Continuation

**CONTINUATION**

THE ELEMENTS THAT ARE PLACED ON A LINE OR A CURVE ARE PERCEIVED AS MORE RELATED THEN THE ONES THAT ARE NOT ALIGNED THAT WAY.

ALSO, OUR PERCEPTION OF SHAPES IS CONTINUED IN THE IMPLIED DIRECTION (EVEN BEYOND THE LAST PART OF THE COMPOSITION)



RECOMMENDED PRODUCTS ON THE SAME HORIZONTAL LINE

# Closure

**CLOSURE** E.G. THIS IS HOW NEGATIVE SPACE WORKS


THE BRAIN FILLS THE MISSING PARTS TO GET A COMPLETE FIGURE; IT CREATES A COMBINATION OF THE PARTS TO ESTABLISH A WHOLE.

→ THERE NEEDS TO BE ENOUGH INFORMATION

→ WE BUILD ON PRIOR KNOWLEDGE & EXPERIENCE

→ CONDITIONS FOR SUCCESSFUL RECOGNITION

E.G. NOTIFICATIONS:



SOME PARTS ARE COVERED, BUT OUR BRAINS HAVE ENOUGH INFO

Not part of the exam material!

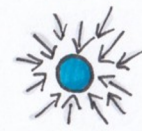
THE WHOLE IS OTHER THAN THE SUM OF THE PARTS.  
(KURT KOFFKA)

# PART 1 GESTALT PRINCIPLES

UX Knowledge Base Sketch #46



**HUMAN PERCEPTION!**  
VISION OPTIMIZATION™:  
WE'RE LOOKING FOR PATTERNS, STRUCTURE, FORM & LOGIC, WE PERCEIVE THE WORLD IN A WAY OF CONSTRUCTING MEANINGFUL THINGS BY COMBINING THE INDIVIDUAL PARTS.



SO MANY INPUTS ⇒ OUR BRAINS START TO ORGANIZE THESE INTO GROUPS.  
E.G.: YOU DON'T ONLY SEE YOUR FRIEND'S HAIR, EYES OR NOSE, YOU RECOGNIZE THE FACE.

TRANSLATED TO UX DESIGN:

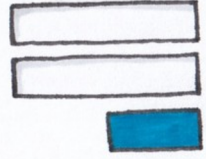
THE ELEMENTS OF A USER INTERFACE AS A WHOLE ARE PERCEIVED DIFFERENTLY THAN THE PARTS ON THEIR OWN.

THIS "WHOLE" DETERMINES HOW A USER INTERACTS WITH THE SYSTEM.

KNOWING THESE PRINCIPLES HELP YOU DESIGN BETTER PRODUCTS!

## SIMPLICITY BASIC LAW OF GESTALT

HUMANS HAVE LIMITED CAPACITY FOR PROCESSING THE INPUT ⇒ OUR MIND IS SEARCHING FOR THE SIMPLEST, YET MOST RELEVANT WAY TO INTERPRET WHAT WE SEE.



WHAT IS THIS? — A DESIGNER'S ANSWER'D BE:  
2 WHITE AND 1-SMALLER-BLUE RECTANGLE?  
A LOG-IN FORM!

## CLOSURE E.G. THIS IS HOW NEGATIVE SPACE WORKS

THE BRAIN FILLS THE MISSING PARTS TO GET A COMPLETE FIGURE; IT CREATES A COMBINATION OF THE PARTS TO ESTABLISH A WHOLE.

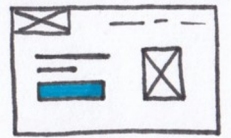
- THERE NEEDS TO BE ENOUGH INFORMATION
- WE BUILD ON PRIOR KNOWLEDGE & EXPERIENCE
- CONDITIONS FOR SUCCESSFUL RECOGNITION



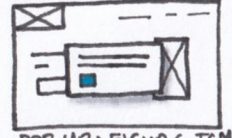
E.G. NOTIFICATIONS:  
SOME PARTS ARE COVERED, BUT OUR BRAINS HAVE ENOUGH INFO

## FIGURE-GROUND

WE TRY TO SEPARATE THE FIGURE FROM THE GROUND: THE FIGURE IS THE ELEMENT WE FOCUS ON, AND THE BACKGROUND IS THE FIGURE'S SURROUNDING.



WE FOCUS ON THE CTA. THE BACKGROUND'S FUNCTION: CONTEXT, COMMUNICATING THE BRAND, MOOD; ADDITIONAL INFORMATION...



POP-UP: FIGURE TEMPORARILY BECOMES THE GROUND

## PROXIMITY

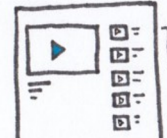
THE OBJECTS THAT ARE CLOSER TO EACH OTHER ARE PERCEIVED AS MORE RELATED THAN THE ONES THAT ARE NOT POSITIONED NEAR THEM.  
ALSO, THE ELEMENTS THAT ARE PLACED CLOSER ARE SEEN AS PART OF THE SAME GROUP OR CATEGORY.



PLACE CONNECTED CONTROLS NEAR TO EACH OTHER!

## SIMILARITY

IF TWO OBJECTS HAVE SIMILAR CHARACTERISTICS, THESE OBJECTS ARE PERCEIVED AS MORE RELATED THAN THE ONES THAT DON'T SHARE THESE QUALITIES.  
WE ALSO ASSUME THAT SIMILAR OBJECTS HAVE SAME FUNCTION/ROLE.



RELATED VIDEOS ON YOUTUBE

- SIMILAR CHARACTERISTICS
- COLOR
  - SIZE
  - SHAPE
  - FONTS

## CONTINUATION

THE ELEMENTS THAT ARE PLACED ON A LINE OR A CURVE ARE PERCEIVED AS MORE RELATED THEN THE ONES THAT ARE NOT ALIGNED THAT WAY.  
ALSO, OUR PERCEPTION OF SHAPES IS CONTINUED IN THE IMPLIED DIRECTION (EVEN BEYOND THE LAST PART OF THE COMPOSITION)



RECOMMENDED PRODUCTS ON THE SAME HORIZONTAL LINE

## ENCLOSURE / COMMON REGION

OBJECTS THAT HAVE A BORDER AROUND THEM ARE PERCEIVED AS MORE RELATED, OR AS A GROUP.  
INSIDE THE BORDER: BELONGING TOGETHER  
OUTSIDE THE BORDER: SEPARATED

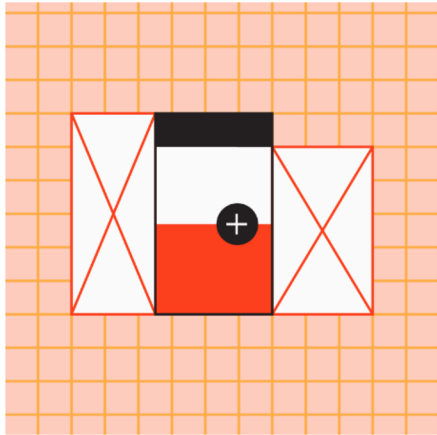


CARD-BASED UI

Not part of the exam material!

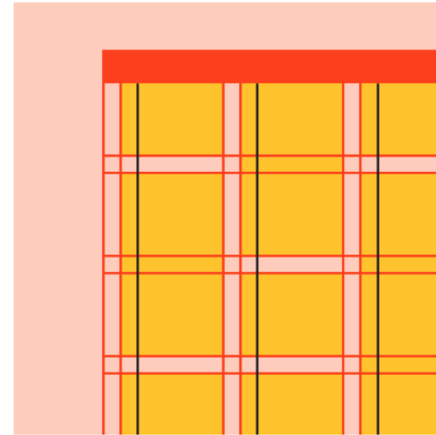
# Real world - Android

## Principles



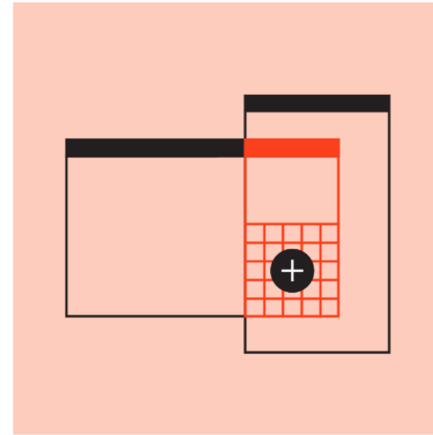
### Predictable

UIs should use intuitive and predictable layouts, with consistent UI regions and spatial organization.



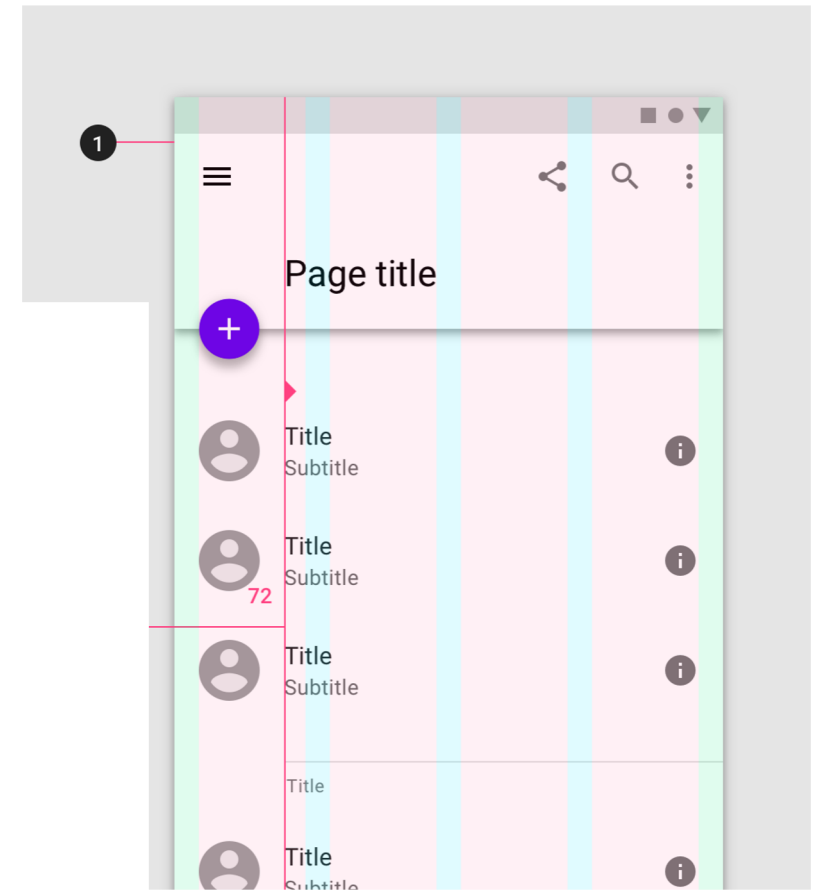
### Consistent

Layouts should use a consistent grid, keylines, and padding.



### Responsive

Layouts are adaptive and react to input from the user, device, and screen elements.



Not part of the exam material!

# Real world - Hololens



Extension above the upper region (looking up) beyond the highlighted area is uncomfortable and limited due to eye and neck morphology.

**Upper region**  
0 to 10 degrees above horizon

**Optimal region**  
0 to 35 degrees below horizon

**Lower region**  
35 to 60\* degrees below horizon

\*determined by maximum neck flexion



Not part of the exam material!

# So...

- Hi-fi prototypes are almost (parts of) the final product that you want to test with users
- Many ways to do it depending on your product
- The 4 dimensions of prototyping need to be taken into account: breadth, depth, look and interaction
- To achieve this, we need to keep into account how humans make sense of objects to interact with them
- Knowledge of human perception, like Gestalt Psychology is very useful to guarantee good design (UI/UX)
















































Tools, frameworks, SDKs

# Tensorflow lite

- <https://www.tensorflow.org/lite/>

## TensorFlow Lite example apps

Explore pre-trained TensorFlow Lite models and learn how to use them in sample apps for a variety of ML applications.

 <b>Image classification</b> Identify hundreds of objects, including people, activities, animals, plants, and places. <b>Model overview</b> → <a href="#">Try it on Android</a>  <a href="#">Try it on iOS</a>  <a href="#">Try it on Raspberry Pi</a> 	 <b>Object detection</b> Detect multiple objects with bounding boxes. Yes, dogs and cats too. <b>Model overview</b> → <a href="#">Try it on Android</a>  <a href="#">Try it on iOS</a>  <a href="#">Try it on Raspberry Pi</a> 	 <b>Pose estimation</b> Estimate poses for single or multiple people. Imagine the possibilities, including stick figure dance parties. <b>Model overview</b> → <a href="#">Try it on Android</a>  <a href="#">Try it on iOS</a>  <a href="#">Try it on Raspberry Pi</a> 	 <b>Digit classifier</b> Classify handwritten digits.  <a href="#">Try it on Android</a>  <a href="#">Try it on iOS</a> 	 <b>Style transfer</b> Apply any styles on an input image to create a new artistic image. <b>Model overview</b> → <a href="#">Try it on Android</a>  <a href="#">Try it on iOS</a> 	 <b>Smart reply</b> Generate reply suggestions to input conversational chat messages.  <b>Model overview</b> → <a href="#">Try it on Android</a> 
 <b>Speech recognition</b> Identify speech commands by recognizing keywords.  <a href="#">Try it on Android</a>  <a href="#">Try it on iOS</a> 	 <b>Gesture recognition</b> Recognize gestures using your webcam.  <a href="#">Try it on Android</a>  <a href="#">Try it on iOS</a> 	 <b>Segmentation</b> Pinpoint the shape of objects with strict localization accuracy and semantic labels. Trained with people, places, animals, and more. <b>Model overview</b> → <a href="#">Try it on Android</a>  <a href="#">Try it on iOS</a> 	 <b>Super resolution</b> Generate a super resolution image from a low resolution image.  <b>Model overview</b> → <a href="#">Try it on Android</a> 	 <b>Audio classification</b> Classify audio using your microphone.  <b>Model overview</b> → <a href="#">Try it on Android</a>  <a href="#">Try it on iOS</a> 	 <b>Reinforcement learning</b> Train a game agent using reinforcement learning and build an Android game using TensorFlow Lite.  <b>Model overview</b> → <a href="#">Try it on Android</a> 
 <b>Text classification</b> Categorize free text into predefined groups. Potential applications include abusive content moderation, tone detection, and more.  <b>Model overview</b> → <a href="#">Try it on Android</a> 	 <b>On-device recommendation</b> Provide personalized on-device recommendations based on events selected by users.  <b>Model overview</b> → <a href="#">Try it on Android</a> 	 <b>Natural language question answering</b> Answer questions based on the content of a given passage of text with BERT.  <b>Model overview</b> → <a href="#">Try it on Android</a>  <a href="#">Try it on iOS</a> 	 <b>Optical character recognition</b> Extract texts from images using Optical Character Recognition with TensorFlow Lite.  <b>Model overview</b> → <a href="#">Try it on Android</a> 	 <b>On-device training</b> Train a TensorFlow Lite model on-device.  <b>Model overview</b> → <a href="#">Try it on Android</a> 	

# Tensorflow lite – pose estimation

- [https://www.tensorflow.org/lite/examples/pose\\_estimation/overview](https://www.tensorflow.org/lite/examples/pose_estimation/overview)

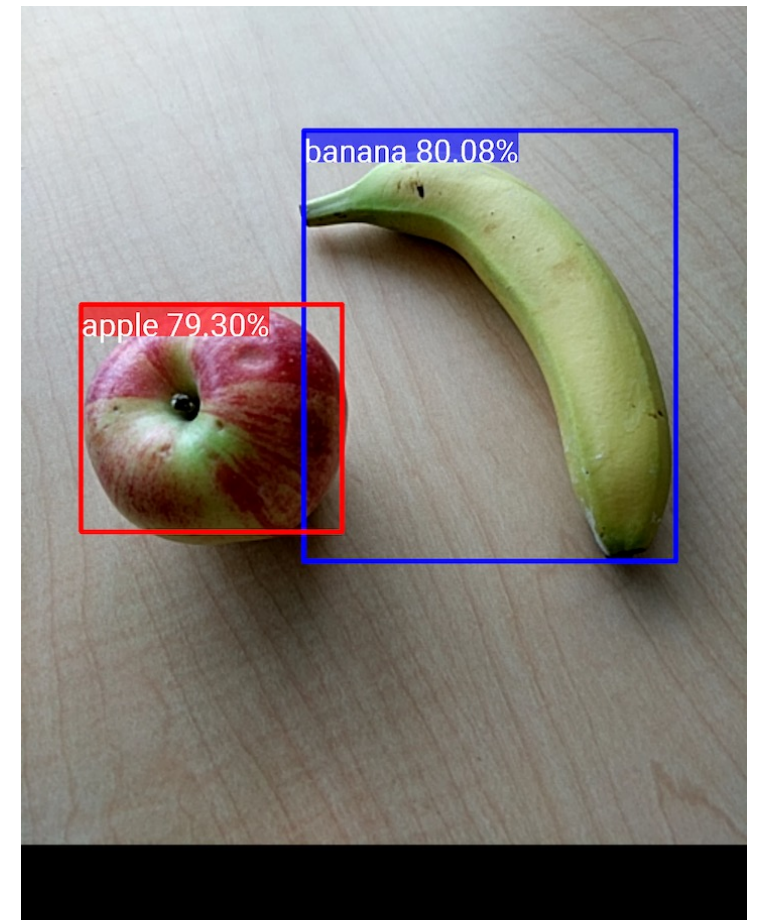


Id	Part
0	nose
1	leftEye
2	rightEye
3	leftEar
4	rightEar
5	leftShoulder
6	rightShoulder
7	leftElbow
8	rightElbow
9	leftWrist
10	rightWrist
11	leftHip
12	rightHip
13	leftKnee
14	rightKnee
15	leftAnkle
16	rightAnkle

# Tensorflow lite – object detection

- [https://www.tensorflow.org/lite/examples/object\\_detection/overview](https://www.tensorflow.org/lite/examples/object_detection/overview)

Class	Score	Location
Apple	0.92	[18, 21, 57, 63]
Banana	0.88	[100, 30, 180, 150]
Strawberry	0.87	[7, 82, 89, 163]
Banana	0.23	[42, 66, 57, 83]
Apple	0.11	[6, 42, 31, 58]



# Tensorflow lite

You can generate a TensorFlow Lite model in the following ways:

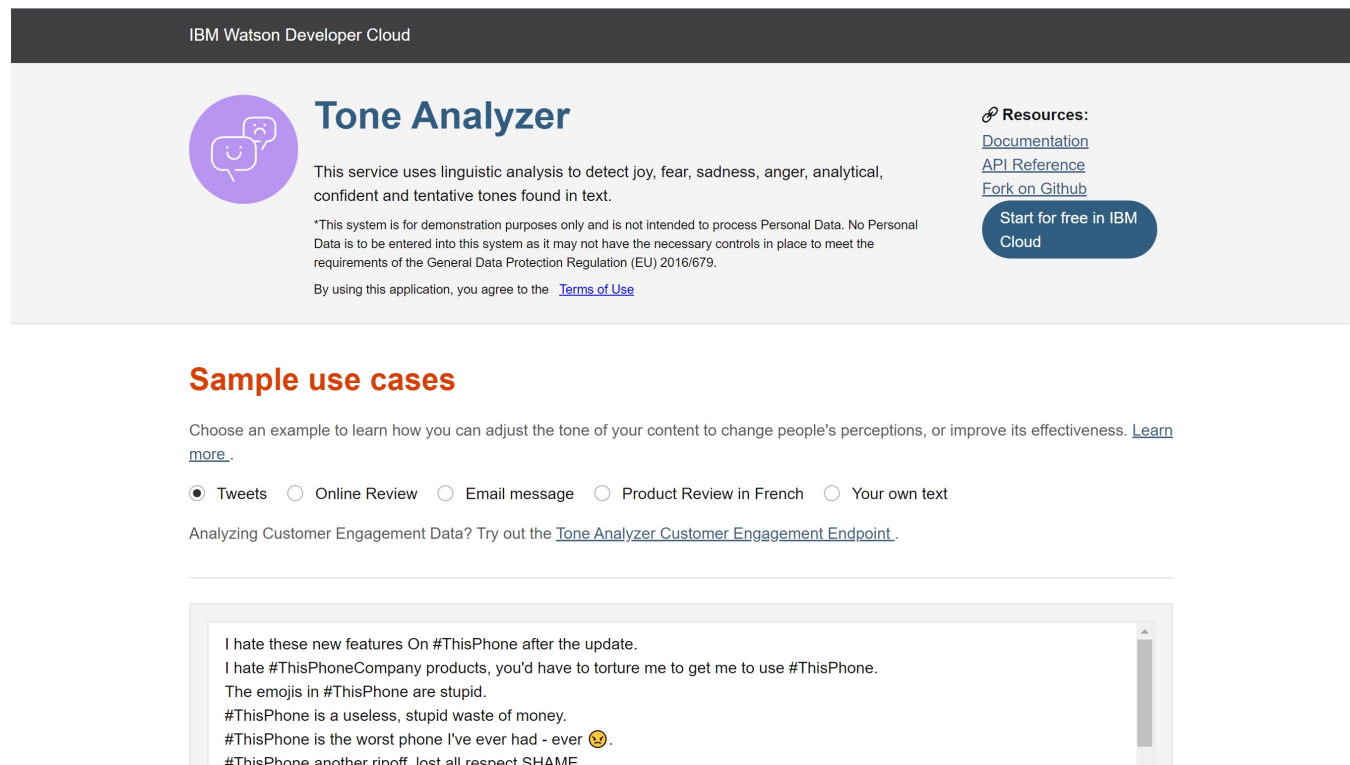
- **Use an existing TensorFlow Lite model:** Refer to [TensorFlow Lite Examples](#) to pick an existing model. *Models may or may not contain metadata.*
- **Create a TensorFlow Lite model:** Use the [TensorFlow Lite Model Maker](#) to create a model with your own custom dataset. *By default, all models contain metadata.*
- **Convert a TensorFlow model into a TensorFlow Lite model:** Use the [TensorFlow Lite Converter](#) to convert a TensorFlow model into a TensorFlow Lite model. During conversion, you can apply [optimizations](#) such as [quantization](#) to reduce model size and latency with minimal or no loss in accuracy. *By default, all models don't contain metadata.*

<https://www.tensorflow.org/lite/guide>

[https://www.tensorflow.org/lite/tutorials/model\\_maker\\_object\\_detection](https://www.tensorflow.org/lite/tutorials/model_maker_object_detection)

# Tools - IBM Watson

- Create a free account: <https://cloud.ibm.com>
- Browse the available services: <https://cloud.ibm.com/catalog>



The screenshot shows the IBM Watson Developer Cloud interface for the Tone Analyzer service. At the top, it says "IBM Watson Developer Cloud". The main heading is "Tone Analyzer" with a purple circular icon containing two speech bubbles, one with a smile and one with a frown. Below the icon, the text reads: "This service uses linguistic analysis to detect joy, fear, sadness, anger, analytical, confident and tentative tones found in text." To the right, under "Resources:", there are links for "Documentation", "API Reference", and "Fork on Github". A blue button says "Start for free in IBM Cloud". A disclaimer states: "\*This system is for demonstration purposes only and is not intended to process Personal Data. No Personal Data is to be entered into this system as it may not have the necessary controls in place to meet the requirements of the General Data Protection Regulation (EU) 2016/679." Below this, it says "By using this application, you agree to the [Terms of Use](#)".

**Sample use cases**

Choose an example to learn how you can adjust the tone of your content to change people's perceptions, or improve its effectiveness. [Learn more.](#)

Tweets  Online Review  Email message  Product Review in French  Your own text

Analyzing Customer Engagement Data? Try out the [Tone Analyzer Customer Engagement Endpoint](#).

The screenshot also shows a text area with the following content:

```
I hate these new features On #ThisPhone after the update.  
I hate #ThisPhoneCompany products, you'd have to torture me to get me to use #ThisPhone.  
The emojis in #ThisPhone are stupid.  
#ThisPhone is a useless, stupid waste of money.  
#ThisPhone is the worst phone I've ever had - ever 😞.  
#ThisPhone another rioff. lost all respect SHAME.
```

<https://tone-analyzer-demo.ng.bluemix.net/>

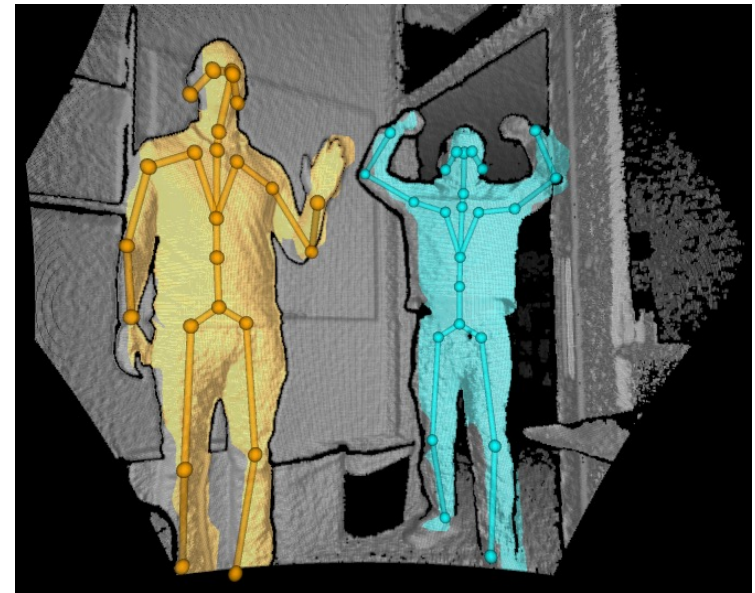
# Google – Dialogflow ES



<https://cloud.google.com/dialogflow/es/docs/video>

# 'Out-of-the-Box AI'

- E.g. Microsoft Azure Kinect
- We (HMI) do have a lot of equipment, discuss with your tutorial leader or ask our lab manager Daniel Davison ([d.p.davison@utwente.nl](mailto:d.p.davison@utwente.nl))
- Decision trees AI Practice?



# Tools

- Options (there is more)
  - Tone Analyzer
  - Visual Recognition
  - Chatbots (Dialogflow)
  - AI course
  - ...
- But remember, we are looking for the interactions with YOUR concept. It depends on your project what you can use and how to use/apply it in your project.

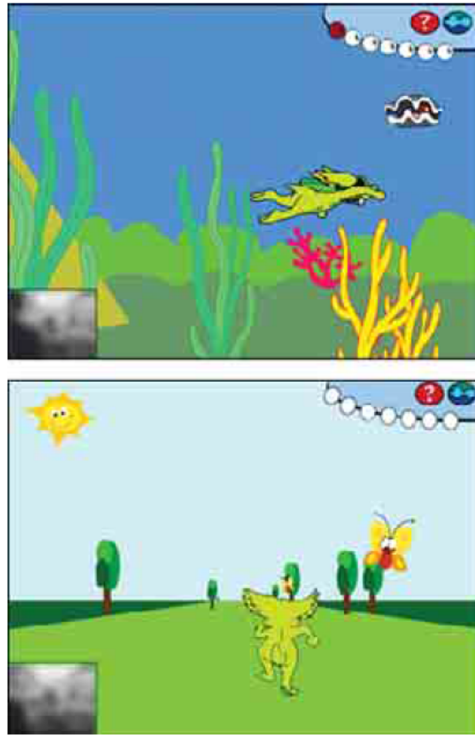
# Hi-fi prototype & Tools

- Again, remember; the goal is not to fully implement a product or system
- We create a prototype to evaluate with users
- It is ok that
  - Only some tasks are implemented
  - Only some use cases are supported
  - WoZ / hardcoded
- Based on the results of the lo-fi prototype test, decide what the interaction scenario with the hi-fi will be and start designing the hi-fi prototypes (next week)



# Wizard of Oz Interaction

- Gesture based games: what are good control gestures?



# Wiz

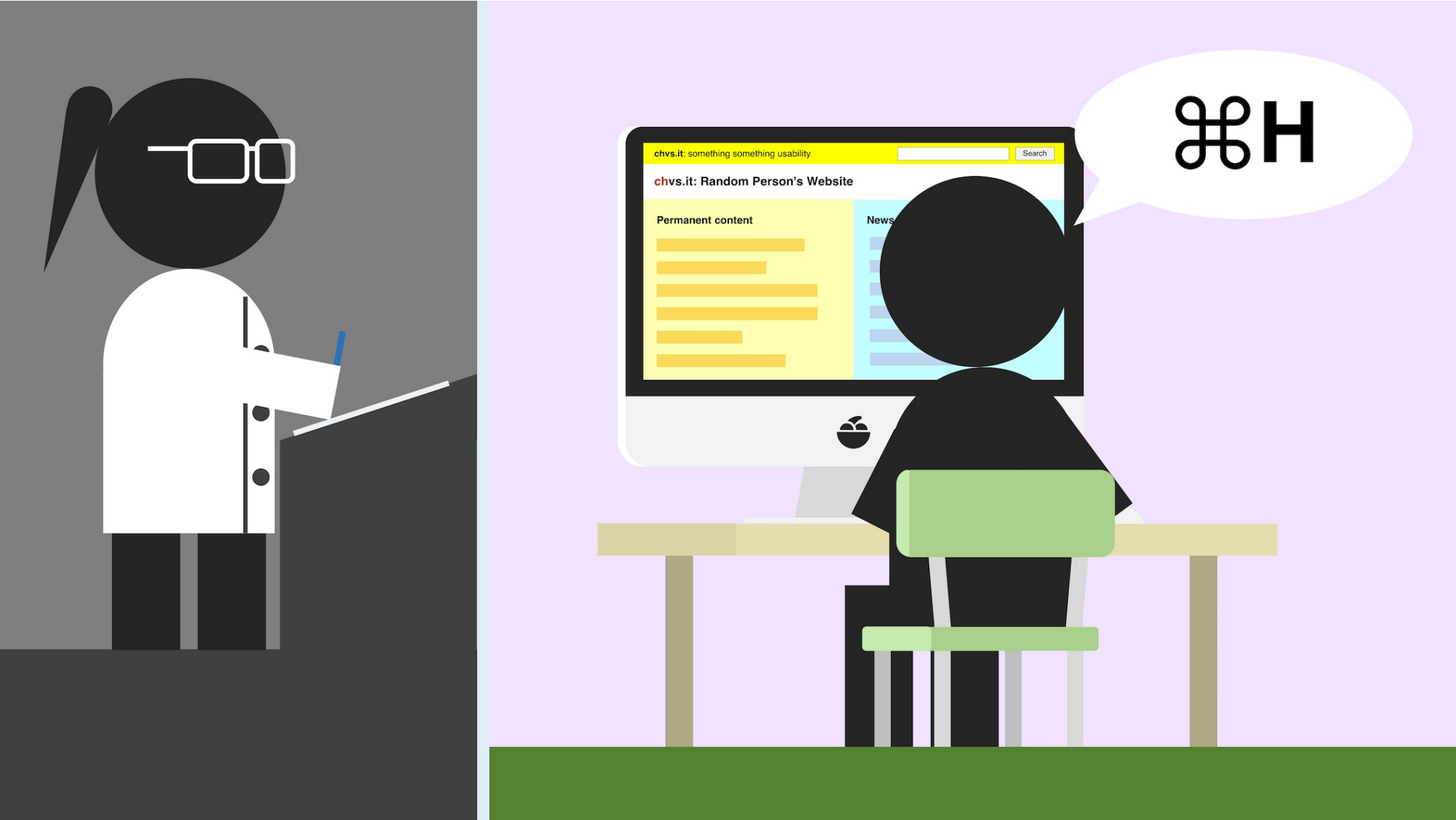
- Ges

	1. swimming	2. diving	3. diving up	4. running	5. jumping
a					
b					
c					
					



<https://userinyerface.com/>

# Think aloud



# Think aloud method

- “think aloud” as they are going through the interface and interaction (known as the “thinking aloud” protocol)
- Influences performance
  - They more they talk they more they are influenced (and the slower they are)
- Provides very useful feedback
  - “I want to get back, hmm there is no back button”  
(although there might be a too small arrow in the corner)
- Reflection/interpretation/retrospective session:
  - Let users perform task, and look at it (e.g. raw video) together afterwards with discussion

# Usability testing

# Usability testing

- What is usability testing?
- *A research method (utilizing representative users and representative tasks) can be used to learn more about how people interact with interfaces, even when the goal is not fixing the interface, but instead learning more about users and interactions (Lazar, p 263)*
- *... involves representative users attempting representative tasks in representative environments, on early prototypes or working versions of computer interfaces (Lazar, p263)*

Us

- W
- A
- ta
- in
- le
- ...
- re
- of

1819 - Bestanden - SURFdrive x | Itemgegevens voor Research n x | Itemgegevens voor Research n x | User Inyerface - A worst-practi x +

← → ↻ 🔒 userinyerface.com

# UI

## User Inyerface

a bagaar frustration

Hi and welcome to User Inyerface,  
a challenging exploration of  
user interactions and design patterns.

To play the game, simply fill in the form  
as fast and accurate as possible.

**NO**

Please [click HERE](#) to GO to the next page

**B** BAGAAR

ve  
ad  
ons

<https://userinyerface.com/>

Card Number

1234 2568 2365 25

Expiry Date

MM /

CVV / CVV2

Name

CARD

### Try Again

There was an error with your information.

Please enter correct:

- a. Card Number
- b. Name on Card
- c. Expiry Date
- d. CVV / CVV2

OK

CONFUSING ERROR MESSAGE

Card Number

1234 2568 2365 25

Must be a 16 digit number

Expiry Date

MM / YYYY

CVV / CVV2

\*\*\*

Name on Card

CARD HOLDER NAME

Cannot be left blank

CLEAR ERROR MESSAGE

# Usability testing

- Types of usability testing or usability engineering

- Expert tests

- E.g. Heuristic review, consistency inspection, cognitive walkthrough

- Automated test, testing based on guidelines

- E.g. accessibility norms, such as captions

- User-based test

- with representative users, representative tasks

**Usability inspection**

**Usability testing**

# Expert-based testing



# Expert-based testing

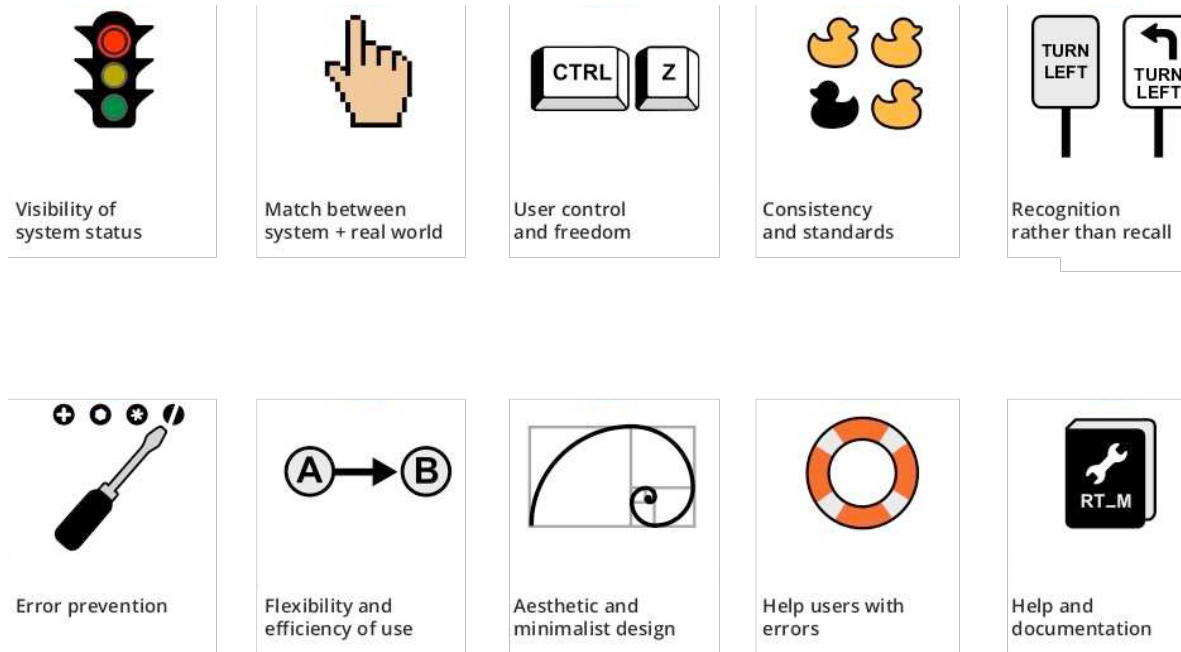
- Done by experts
- Not the developer
  - Too invested
  - [Too self-critical]
  - Experimental point of view [- some bias]
- People should be unfamiliar with the interface
- Representative users are experts in the task
- Usability experts, experts in the interface/usability design

# Expert-based testing: Heuristic review

- Structured inspection
  - Confused wording
  - Inconsistent or misleading layout
  - Color inconsistency
- Before user-based testing occurs
  - Representative users: deeper and task-related interface flaws (Lazar, 2006).

# Expert-based: Heuristic review

- Usually  $\leq 10$  interface rules
  - Schneiderman's 8 golden rules
  - Jakob's Nielsen 10 Usability Heuristics for UI design (not in the book!)



# Heuristic design evaluation

- Proposed method by Nielsen and Molich
- 10 usability criteria are defined (= heuristics)
- Can be used for early design, prototype, storyboard and fully functioning system evaluations
  
- 75% of overall usability problems are discovered with only 5 evaluators



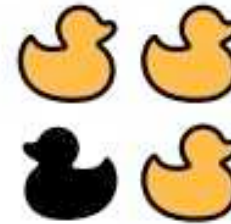
Visibility of system status



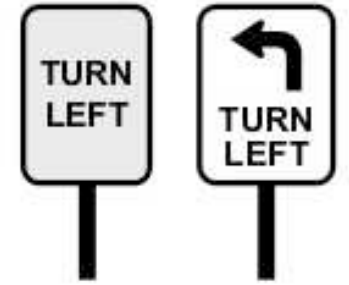
Match between system + real world



User control and freedom



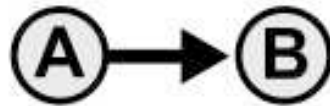
Consistency and standards



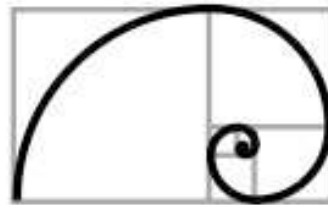
Recognition rather than recall



Error prevention



Flexibility and efficiency of use



Aesthetic and minimalist design



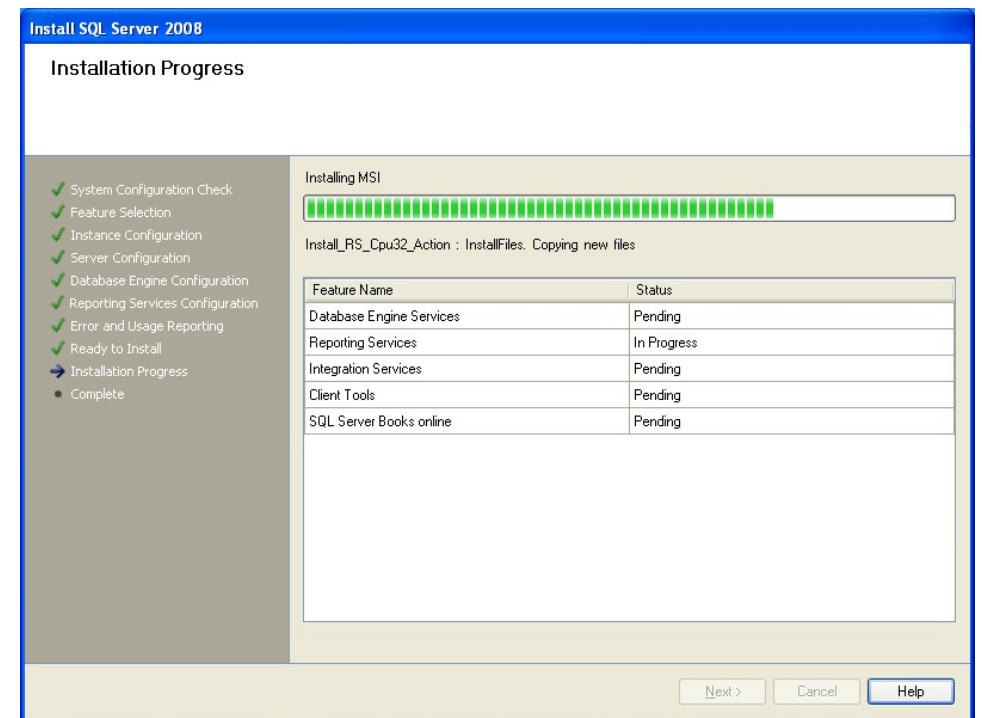
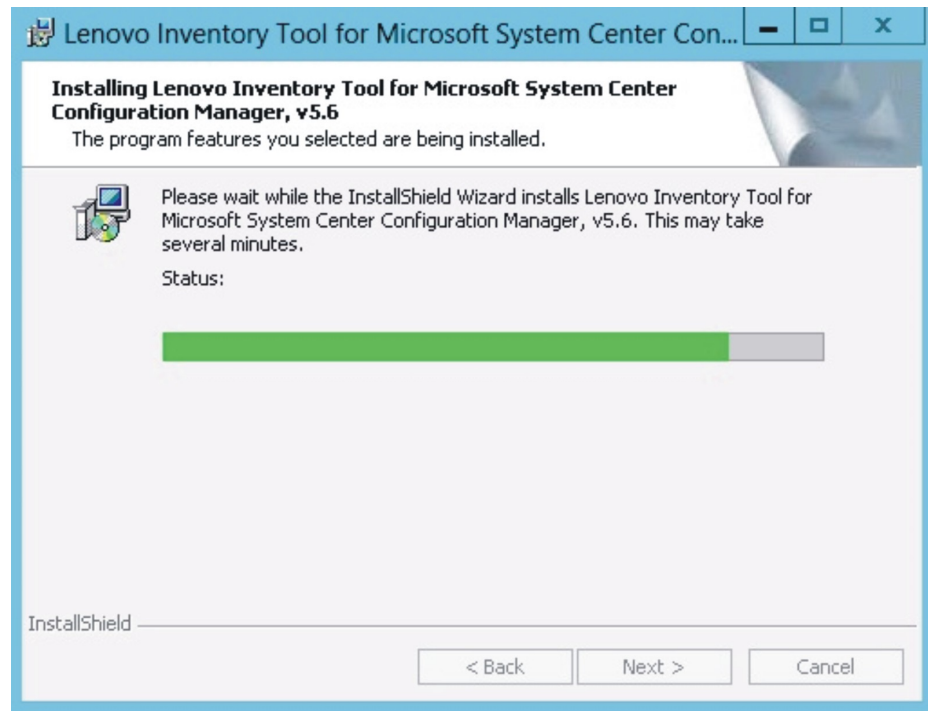
Help users with errors



Help and documentation

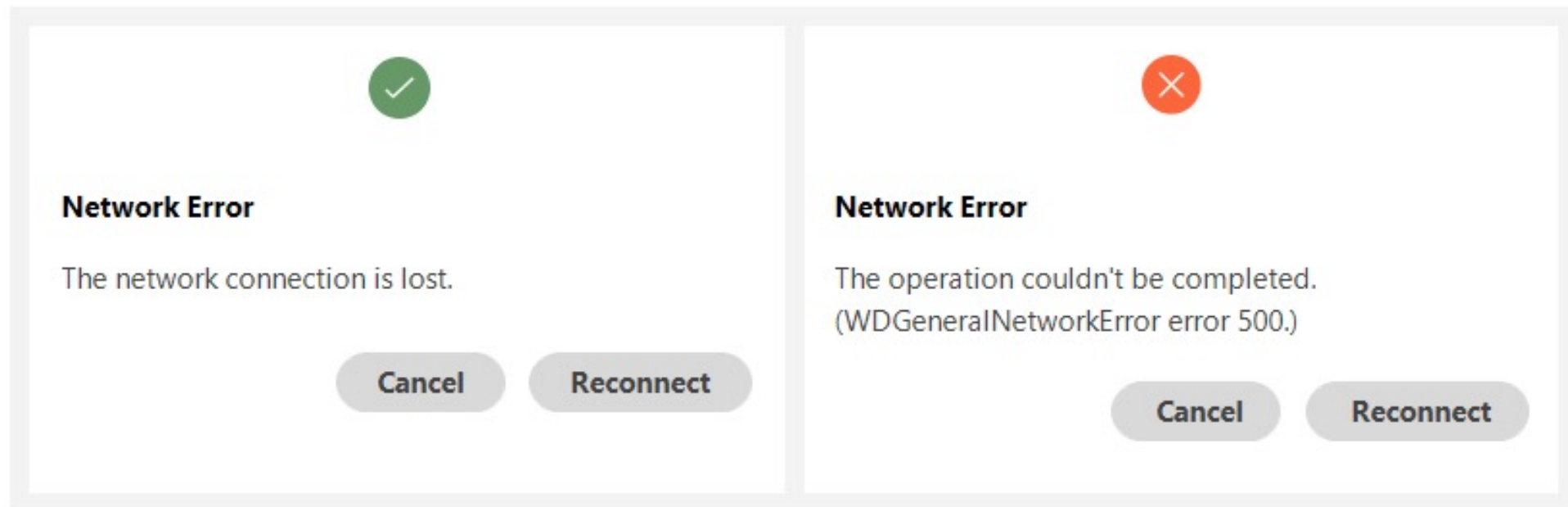
# Visibility of system status

The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.



# Match between system and the real world

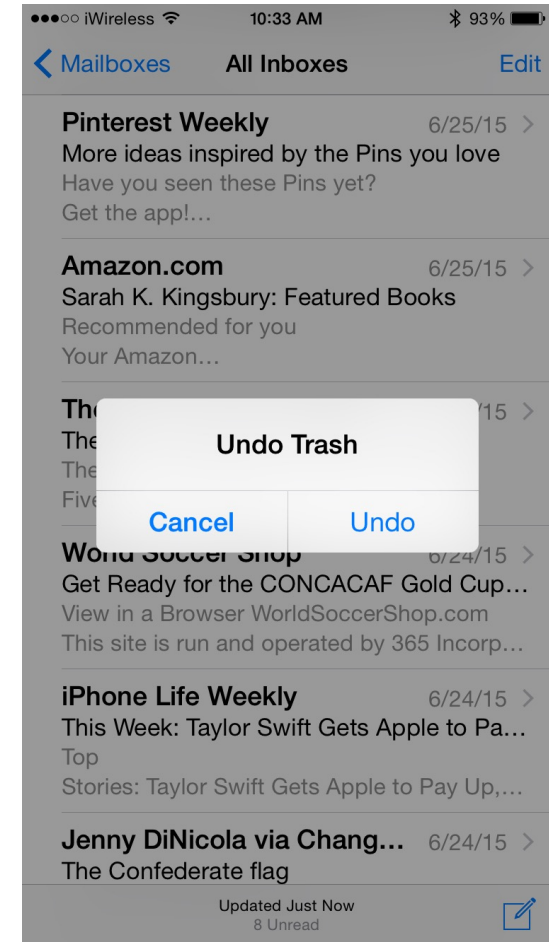
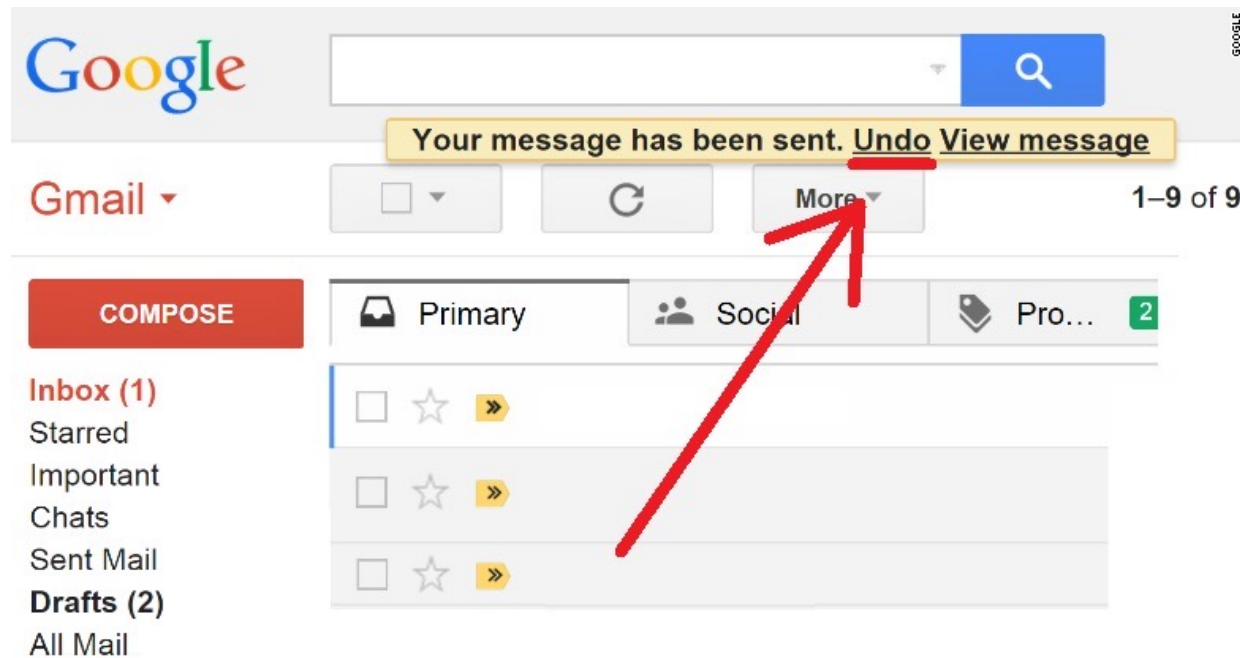
The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.





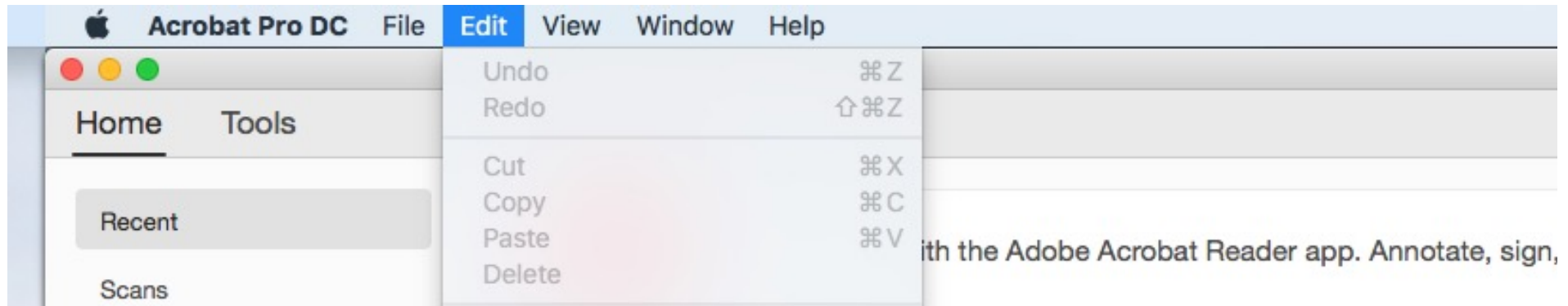
# User control and freedom

Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.



# Consistency and standards

- Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.



# Error prevention

Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.

# Error prevention

The screenshot shows the NS website's journey planning interface. At the top, there is a navigation bar with the NS logo, a menu icon, and the text "Menu". On the right side of the navigation bar, there is a search icon, the text "Search", a user profile icon, the text "Mijn NS", and a flag icon with the text "Nederlands".

Below the navigation bar is a large banner image of a canal in a city. In the top left corner of the banner, it says "Only € 21 for a 2-day ticket". In the top left corner of the banner, there is a "Read more" link with a right-pointing arrow.

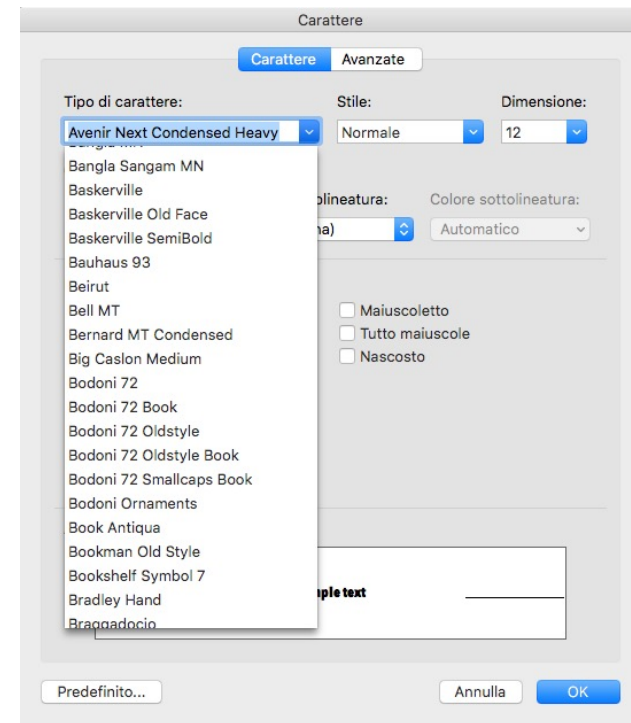
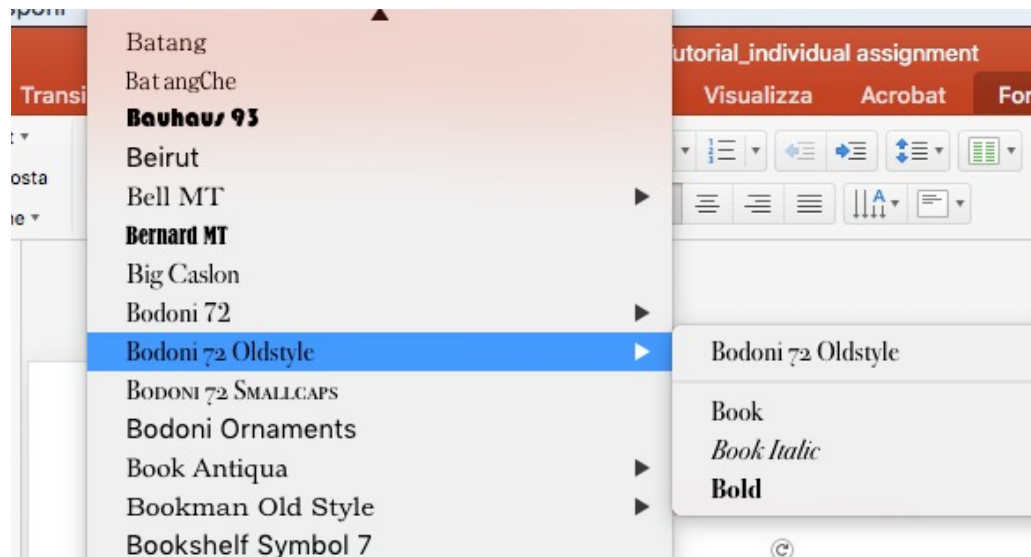
The main content area has a yellow background. It features a "Plan your journey" heading. Below the heading is a search bar with two input fields: "from Hengelo" and "to Tilburg". Each field has a close button (X) and a double-headed arrow between them. Below the search bar, there are several controls: "Departure" and "Arrival" tabs, a "Today" tab with a calendar icon, a time selector showing "12:12" with left and right arrows, and a "Now" tab with a refresh icon. To the right of these controls are two buttons: "More options" and "Plan" with a right-pointing arrow.

At the bottom of the yellow area, there is a warning icon (triangle with exclamation mark) and the text "Disruptions and maintenance work (2)", followed by a right-pointing arrow and the text "More travel information" with a right-pointing arrow.

Below the yellow area, there are three white boxes with icons and text: "Tickets & supplements" with a hand pointing to a screen icon, "Most popular" with a group of three people icon, and "Customer service" with a person and question mark icon.

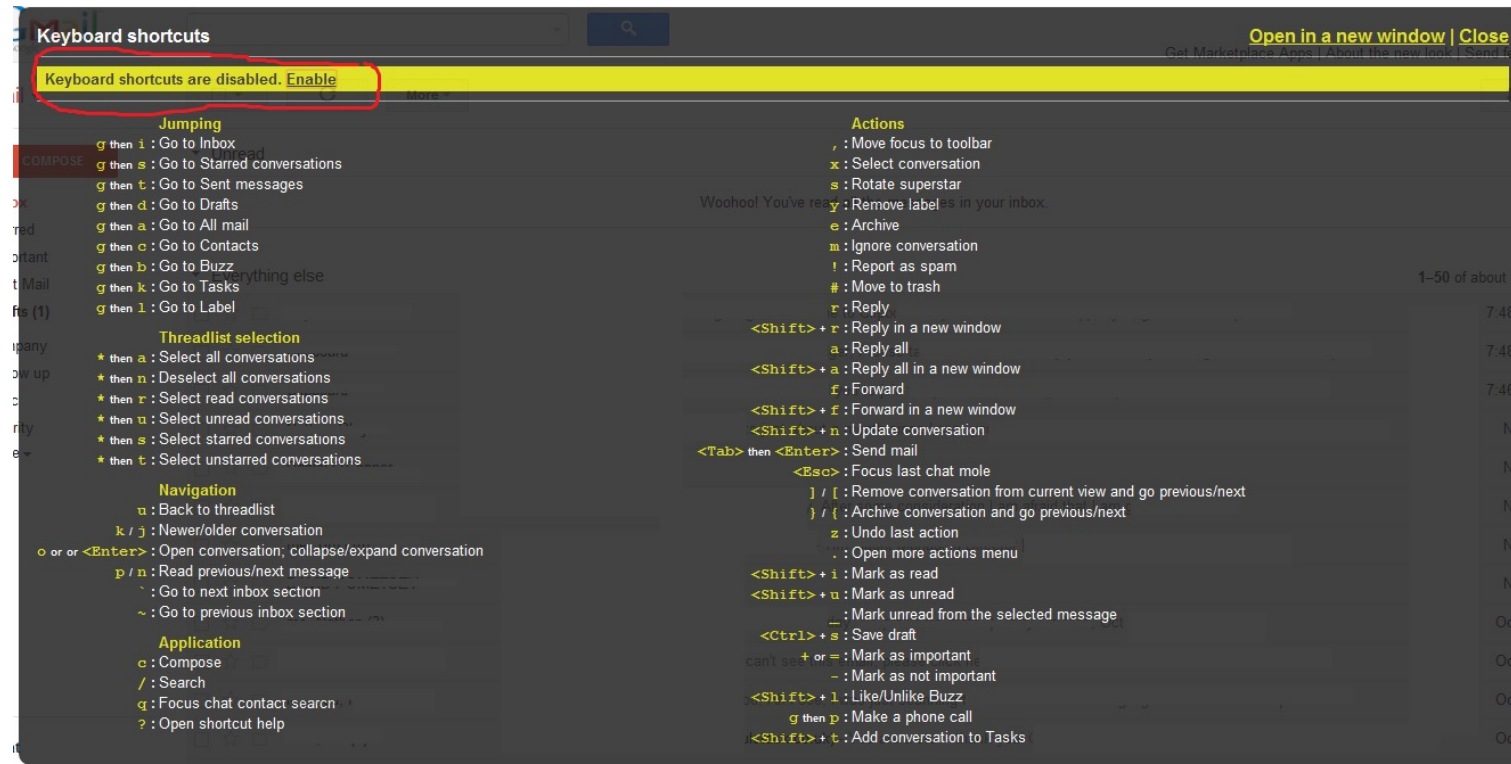
# Recognition rather than recall

Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.



# Flexibility and efficiency of use

Accelerators — unseen by the novice user — may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.



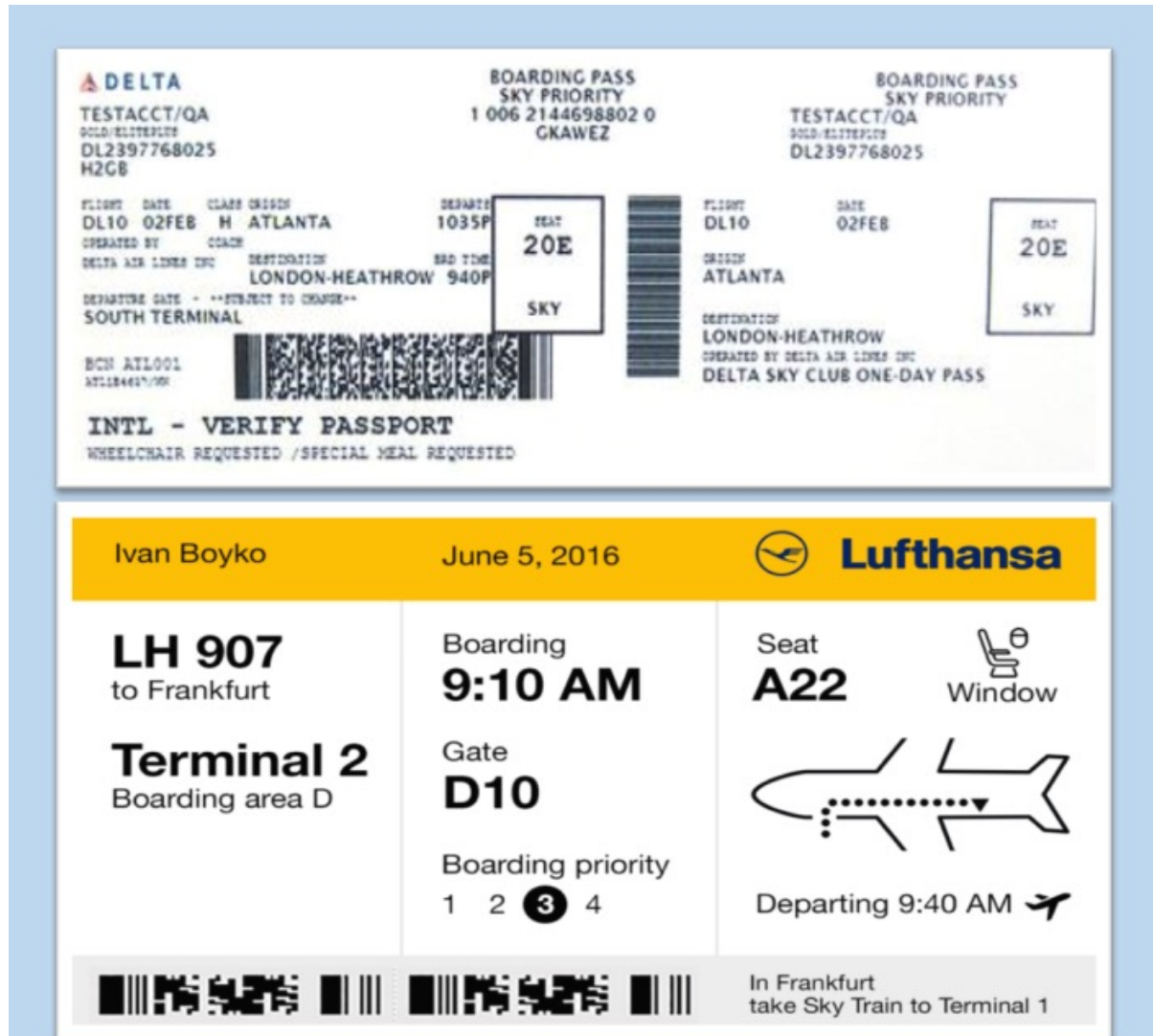
# Aesthetic and minimalist design

Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

# Aesthetic and minimalist design



# Aesthetic and minimalist design



# Help users recognize, diagnose, and recover from errors

Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.



# Help users recognize, diagnose, and recover from errors



Your PC ran into a problem that it couldn't handle, and now it needs to restart.

You can search for the error online: HAL\_INITIALIZATION\_FAILED

# Help users recognize, diagnose, and recover from errors



Your PC ran into a problem and needs to restart. We're just collecting some error info, and then we'll restart for you.

25% complete



For more information about this issue and possible fixes, visit

<http://windows.com/stopcode>

If you call a support person, give them this info:  
Stop code: `CRITICAL_PROCESS_DIED`

# Help and documentation

Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

# Help and documentation



English (US) ▾

[Submit a request](#)

[Sign in](#)

🔍 Search



[Guides](#)



[Troubleshooter](#)



[FAQ](#)



[Announcements](#)

# Expert-based: Heuristic review

Table 10.2 Shneiderman's 8 Golden Rules of Interface Design

1. Strive for consistency
2. Cater to universal usability
3. Offer informative feedback
4. Design dialogs to yield closure
5. Prevent errors
6. Permit easy reversal of actions
7. Support internal locus of control
8. Reduce short-term memory load

(Shneiderman et al., 2017)

See good examples on (e.g. support internal locus of control):

<https://www.interaction-design.org/literature/article/shneiderman-s-eight-golden-rules-will-help-you-design-better-interfaces>

# Expert-based: Consistency Inspection

- Series of screens/web pages to be compared
- Issues of consistency in layout, color, terminology, or language
- “Sometimes, an organization has a specific set of style guidelines (for colors and typefaces) and a consistency inspection can check for overall consistency with those style guidelines.” Organization (e.g. style guides [University of Twente](#))

**UNIVERSITY  
OF TWENTE.**



# Expert-based : guidelines review (mix of consistency + heuristic)

more rules than heuristic (e.g. 10-100)

- OS-based (e.g. [Microsoft Windows apps](#), Android/iOS guidelines)
- Accessibility (webstandards e.g. [WCAG](#)), can be overwhelming!
- Related: Tog's first principles (not for exam, but very useful)
  - <https://asktog.com/atc/principles-of-interaction-design/>

# Cognitive walkthrough

- “walking through” a series of tasks as if they are a user
  - Understanding of first use
  - Exploratory nature
  - Crucial tasks (error recovery)
  - High-frequency task
- Task-based not rule-based

# Beyond scope

- Usability inspection
- Pluralistic walkthrough
  
- Read the book (Lazar, Ch 10)

# Automated usability testing

- Analyzing code
- Often web-based
- Based on interface guidelines
  - Very well-known, alternative code for a graphic:
  - (important for accessibility, and a requirement under the WCAG 2.0),
  - existence of an <alt> attribute in an <img> tag.
  - Screen readers vs. open in new tab of browsers
- Pro's: statistics
  - # of different fonts
  - Average font size
  - Size of buttons
  - Depth of menus
  - Loading time
  - Etc.



# User-based testing

- Representative users
- Representative tasks
- Throughout process
  
- Fixing mistakes early on reduces costs!
  
- WHY?
  - Developers are not users
  - Users often can't clearly identify their (interface) needs
  - [believe me: Users do stupid things]

# Formative usability testing

- Early on in process and tends to be exploratory
- Also known as formative testing
- Lo-fi - easier to criticize
- More communication between moderators and participant
- Formative usability tests tend to focus more on qualitative feedback, moderator observation, and problem discovery

# Summative usability testing

- Example from education?
- More formal prototypes hi-fi (high-fidelity) prototypes
- “evaluate the effectiveness of specific design choices”
- Focus more on task-level measurements, metrics, and quantitative measurements
- Validation tests
  - Benchmark related (comparing to a state-of-the art)
  - E.g. requirement related: “90% of users can complete each task within 1 minute”
  - At the end of the process, right before the public release
  - Less common

# Formative and Summative testing

- Sometimes mixed summative and formative
- Also mixing stages, summative paper prototypes are possible
  - [personally I don't recommend it!]
- Warning
  - Methods such as thinking aloud influence results
    - Cognitive flow interrupted, so for example task takes longer

→ **DEPENDS ON YOUR PROJECT!**

- Choose what most likely leads to actual changes

# Don't re-invent the wheel, especially post-session rating scales:

Also see Tullis & Albert for scales

e.g. System Usability Scale (SUS), CH6

Well known also:

Game Experience Questionnaire

Table 5.1 Survey Tools in HCI

Tool	Citations
Computer System Usability Questionnaire (CSUQ)	Lewis (1995)
Interface Consistency Testing Questionnaire (ICTQ)	Ozok and Salvendy (2001)
Perdue Usability Testing Questionnaire (PUTQ)	Lin et al. (1997)
Questionnaire for User Interaction Satisfaction (QUIS)	Chin et al. (1988) Slaughter et al. (1994) <a href="http://www.lap.umd.edu/quis/">http://www.lap.umd.edu/quis/</a>
Software Usability Measurement Inventory (SUMI)	<a href="http://sumi.uxp.ie/">http://sumi.uxp.ie/</a>
Website Analysis and MeasureMent Inventory (WAMMI)	<a href="http://wammi.uxp.ie/">http://wammi.uxp.ie/</a>

*For more information about existing surveys for usability evaluation, the reader is encouraged to visit <http://garyperلمان.com/quest/>.*

# Stages of usability testing

**Table 10.3** Stages of Usability Testing From Different Authors

Stages of Usability Testing	
(Rubin and Chisnell, 2008)	(Lazar, 2006)
Develop the test plan	Select representative users
Setup the test environment	Select the setting
Find and select participants	Decide what tasks users should perform
Prepare test materials	Decide what type of data to collect
Conduct the test sessions	Before the test session (informed consent, etc.)
Debrief the participants	During the test session
Analyze data and observations	Debriefing after the session
Report findings and recommendations	Summarize results and suggest improvements

# Reminder - “how many users do I need to have?”

- Strict experimental answer:
  - depends on the effect-size , do a power analysis.
- “Lazy” overused excuse HCI answer:
  - “five users will find approximately 80% of usability problems in an interface (Virzi, 1992)”
- Book/my realistic answer: You don’t know in advance
  - You don’t know the number of mistakes
  - [some mistakes might be hidden behind other, requiring multiple tests]

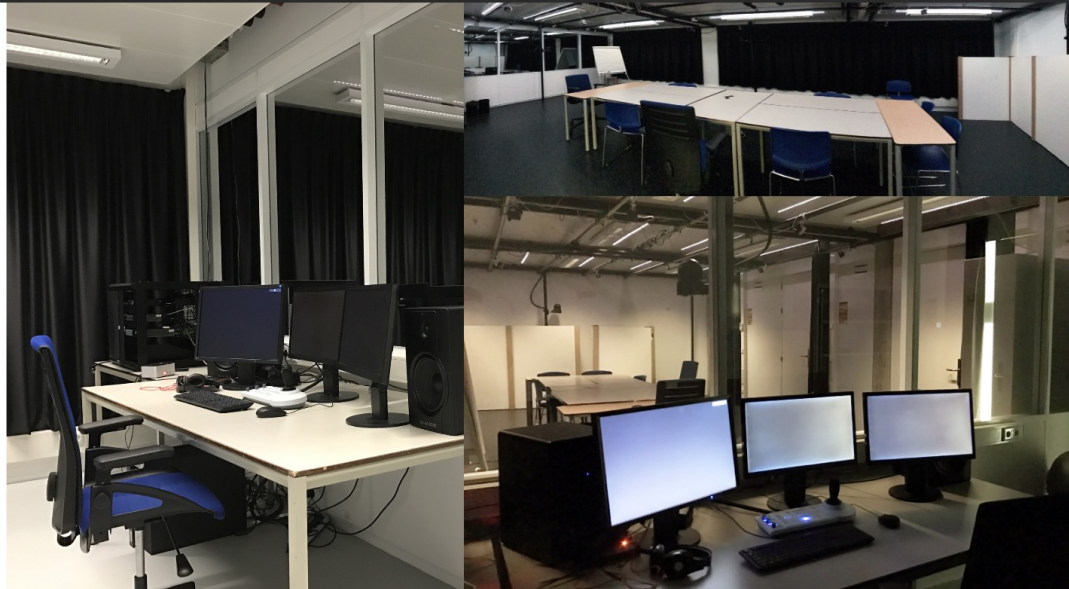
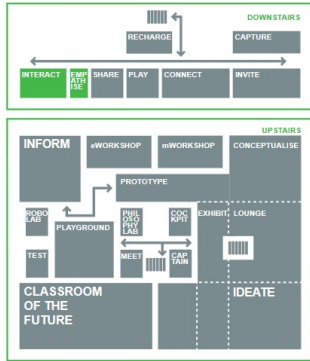
# Reminder - “how many users do I need to have?” :

- Accuracy needed (Lewis 2006)
- Depends on discovery goals (Lewis 2006)
  - [project dependent, generalization needed?]
- How many participants are available? (Lewis 2006)
- Optimal 10+-2 (suggestion Hwang and Salvendy (2010))

In the book :

- “how many users can we afford?”
- “how many users can we get?”
- “how many users do we have time for?”

# Locations for usability testing



VectorStock®

VectorStock.com/23261939



# Take your research on the road

The BMS lab has the ability to host and support a wide range of research projects. However, sometimes the location can be the problem. As a result of its mobile lab facilities, the lab can bring almost any research project to wherever it is needed. The fully equipped "Experivan" can serve a wide range of purposes. It allows researchers to bring lab like environments and equipment to those that are hard to bring to the lab. For example, the elderly or visitors of events. More importantly, the Experivan mobile lab is flexible and can be adjusted to fit a study's needs. The mobile lab includes:

- Ability to run off the grid and on batteries
- Dedicated seating area
- Server, audio, computer and video hardware
- Roofing (tent like dome) for outside the van.
- Access to the internet



[Learn about the mobile lab](#)

# Task lists

- [even for paper prototypes relevant]
- Background scenario to set context
- Clear, accurate and unambiguous tasks
- Task relate to key features of interface
- Questions should require the interaction/interface
- Don't use private data
  - Separate it from the task
  - skip/prepare fake name, age etc;
  - consider cases, takes effort regarding disposal/permission etc.

# Task lists

- Must participants go through tasks in the order listed?
- Can they skip tasks or start at the last task?
- Do certain tasks require that other tasks be completed first?
- How should participants respond if they get stuck and are unable to complete one task?
- Should they skip that task and move onto the next task?
- Is there a time limit per task? (or a minimum time?)
- Is there an overall time limit for the usability testing session?
- Interventions by the moderator?

# Measurement

- Task performance
    - “how many tasks were correctly completed”
    - + “attempted”
  - Time performance
    - How long on completed tasks
    - + How long on unsuccessful
  - User satisfaction (using surveys)
    - Using standardized, validated survey tools
- (sometimes other measures depending on product)

# Making sense of the data

- Similar to analyzing data from any other type of research
- Fewer participants, simple descriptive statistics
- “Rubin and Chisnell (2008) suggest splitting the report into three sections:
  - why you did usability testing and how you prepared;
  - what happened during the testing; and
  - the implications and recommendations. ”
- Never include names or identifying information from participants in your result (this happens every year in the project reports)!

# Thank you!

