Universidade de Aveiro
Departamento de Electrónica,
Telecomunicações e Informática

Evaluation in Visualization
Analytic methods

Beatriz Sousa Santos, University of Aveiro, 2019

How can we produce a Visualization?

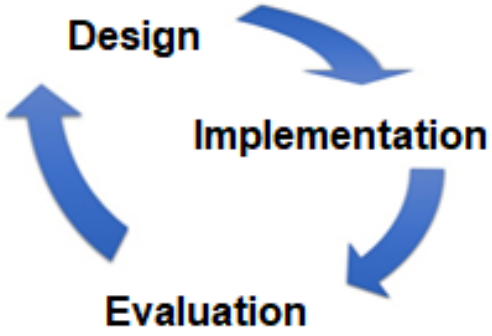
- There are **principles** (derived form human perception and cognition)

  paradigms (examples resulting form past experience)

  and many **methods**

- To obtain **efficacy** it is fundamental:
  - a correct definition of **goal** and **user tasks**
  - apply adequate methods and **evaluate**

  in **several iterations** until the goals are satisfied …



How can we evaluate?

- Many methods can be used to evaluate a Visualization application (some specifically developed, others adapted)

- Evaluation methods from other disciplines may and have been adapted and used to evaluate Visualization applications, as methods from:

  - Human- Computer Interaction
  - Image Processing
  - S/W Engineering

- Applications to visually explore data are interactive and should be usable

- Usability is, according to ISO 9241-11:

“the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use”

- How to measure it??

- We can use methods used in Human-Computer Interaction

Usability evaluation Methods

- Analytical (without users)
  - Heuristic Evaluation
  - Cognitive Walkthrough
  - Model based methods
  - Review methods
  - ...
- Empirical (involving users)
  - Observation > usability tests
  - Query
  - Controlled Experiments
  - ...

Heuristic Evaluation (Nielsen and Molich 1990)

- A "**discount usability engineering method**" for quick, cheap, and easy evaluation of a UI design
- Most popular usability inspection method; yet is **subjective**
- It is a **systematic inspection** of a design for usability
- Meant to find the usability problems in the design so that they can be attended to as part of an iterative design process
- Involves a small set of analysts judging the UI against a list of usability principles ("**heuristics**")

How to perform HE

- Should be performed by several evaluators (one person will never be able to find all the problems)

- Evaluators should work independently:
  - First get a general idea of the UI
  - Then perform a detailed inspection using a set of heuristics
  - Listing usability problems (heuristics not followed and severity degree)

- Findings of all evaluators should be integrated in the same report

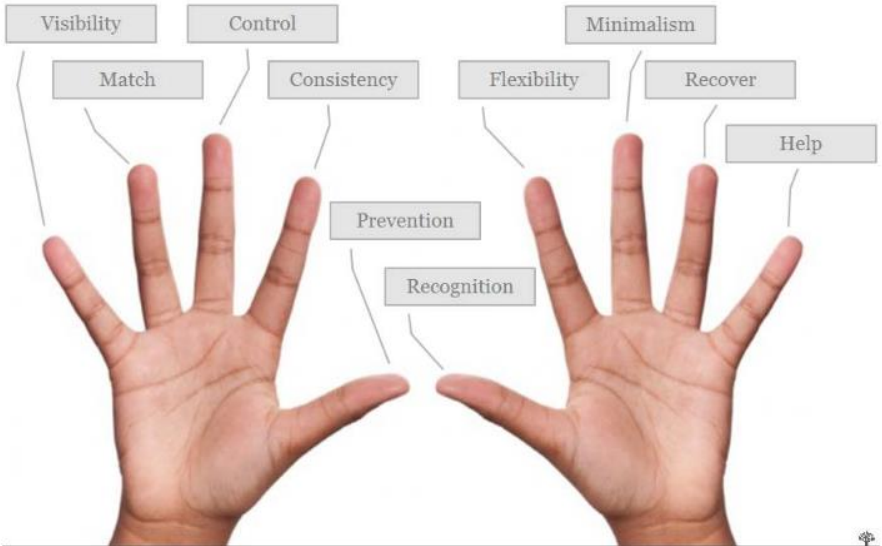
- The report should help the development teem to prioritize problem fixing

- Nielsen proposed **10 general usability heuristics**, yet **there are other sets** (e.g., visualization, web, mobile applications for seniors or children…)

- More details on how to conduct a heuristic evaluation at: http://www.nngroup.com/articles/how-to-conduct-a-heuristic evaluation

- And how to rate severity of the usability problems found: http://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/

- **The list of problems and severity rates should help the development team to priorityse problem fixing**



List of recognized usability principles (“the heuristics”)

1-Visibility of system status

2-Match between system and the real world

3-User control and freedom

4-Consistency and standards

5-Error prevention

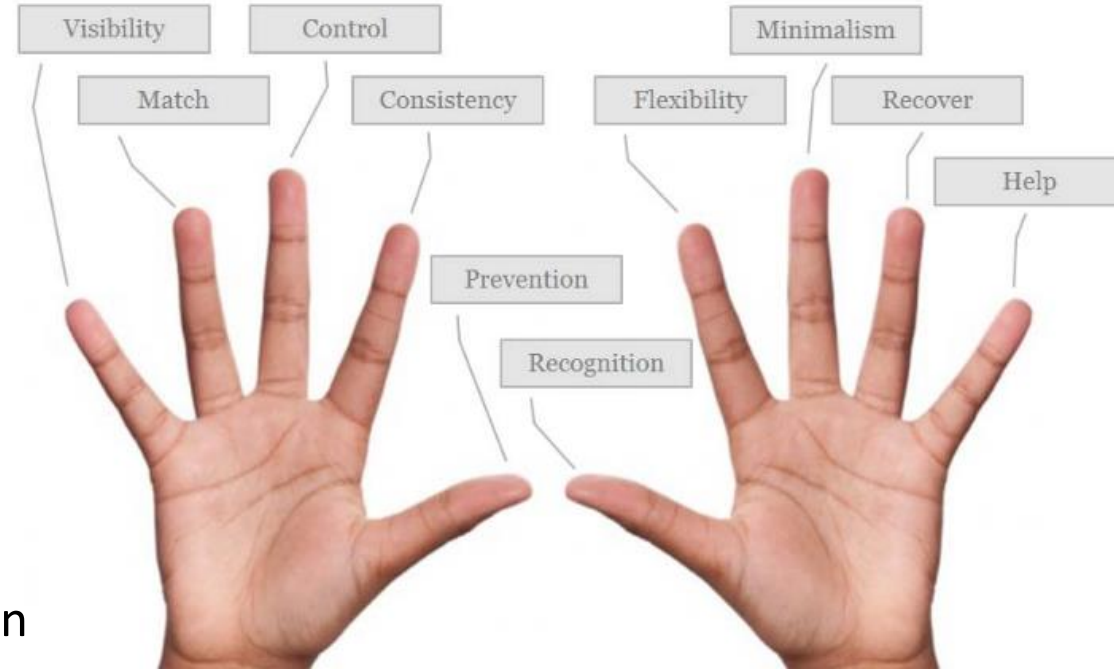
6-Recognition rather than recall

7-Flexibility and efficiency of use

8-Aesthetic and minimalist design

9-Help users recognize, diagnose, and recover from errors

10-Help and documentation



https://www.nngroup.com/articles/ten-usability-heuristics/

Specific Heuristics for Visualization

- Zuk's Perceptual and Cognitive heuristics (Zuk *et al*., 2006)
- Forsell's. heuristic set for evaluation in InfoVis (Forsell and Johanson, 2010)
- Shneiderman's "Visual Information-Seeking Mantra"
- Freitas's *et al*. Ergonomic Criteria for Hierarchical Information Visualization Techniques (Freitas et al., 2009)
- Amar and Stasko's Knowledge and task-based framework
- ...

Zuk and Carpendale's (2006) heuristics

1-Ensure visual variable has sufficient length

2-Don't expect reading order from color

3-Color perception varies with size of colored item

4-Local contrast affects color & gray perception

5-Consider people with color blindness

6-Preattentive benefits increase with field of view

7-Quantitative assessment requires position or size variation

8-Preserve data to graphics dimensionality

9-Put the most data in the least space

10-Remove the extraneous (ink)

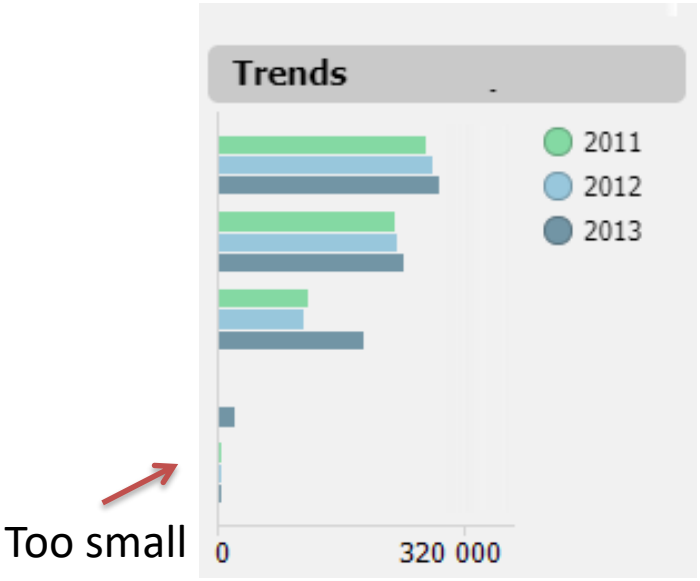
11-Consider Gestalt Laws

12-Provide multiple levels of detail

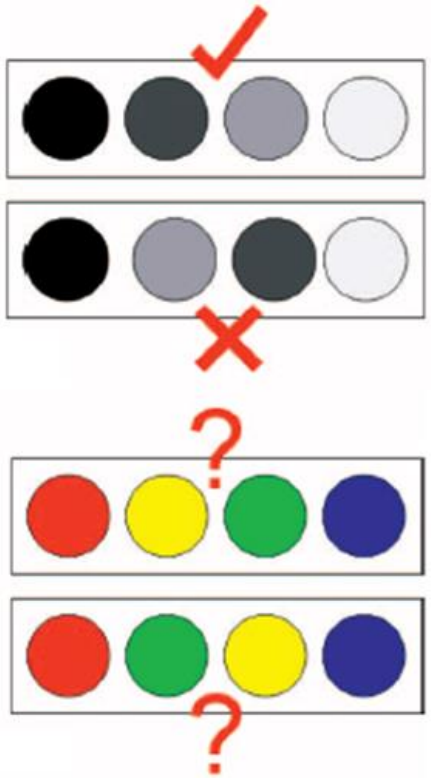
13-Integrate text whenever relevant

Explaining some of the Specific Heuristics for Visualization

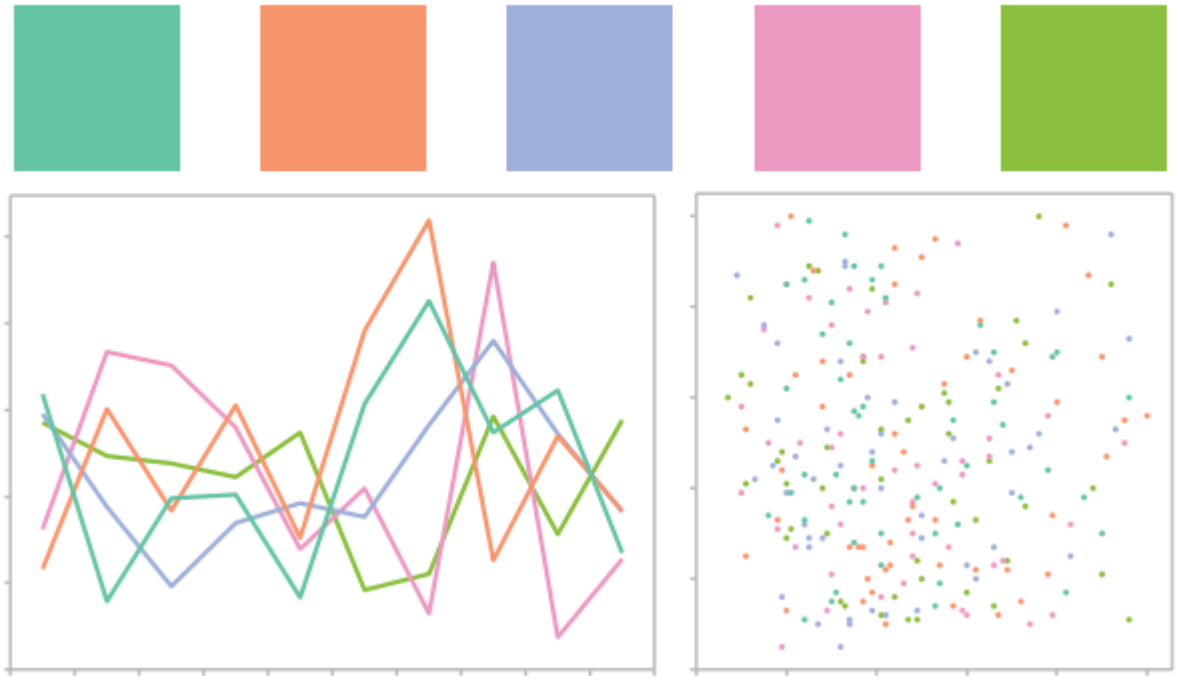
Visual variables must have sufficient length



Do not expect to easily perceive order from color



Color perception varies with size of colored item



A set of colors with different hues but equal luminance (L* = 72). In the squares, these colors are distinct and visually balanced (www.colorbrewer.org).
They are harder to distinguish in smaller items

Stone, M., "In color perception, size matters", IEEE *Computer Graphics & Applications*. 32, 2, 2012, pp. 8-13

Consider people with color blindness

The most common form of color blindness is deuteranopia (“daltonism”)

There are color blindness simulators

Normal vision

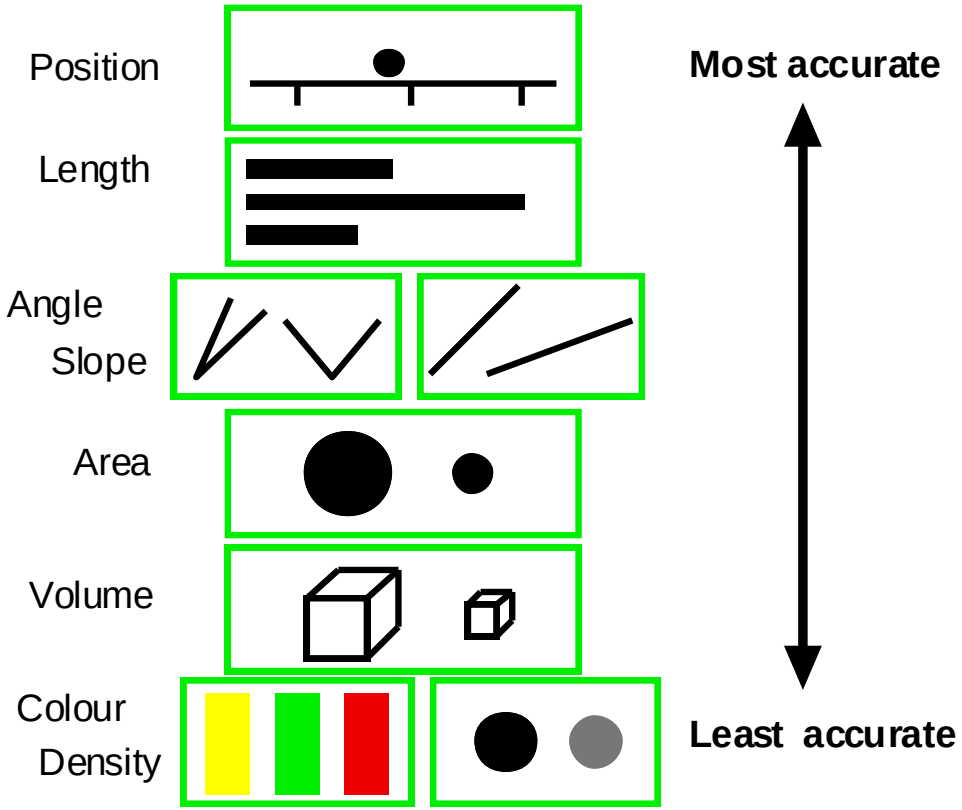
http://www.color-blindness.com/coblis-color-blindness-simulator

Deuteranopia

Tritanopia

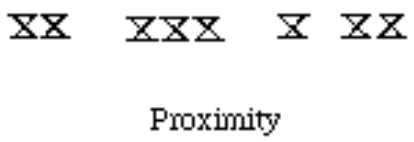
http://www.colourblindawareness.org/

Quantitative assessment requires position or size variation

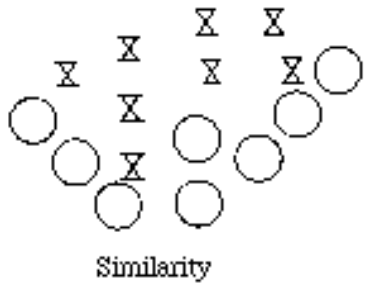


The relative difficulty of assessing quantitative value as a function of encoding mechanism, as established by Cleveland and McGill (Spence, 2007)

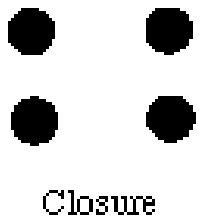
Gestalt Laws



Near stimuli are perceived as a group



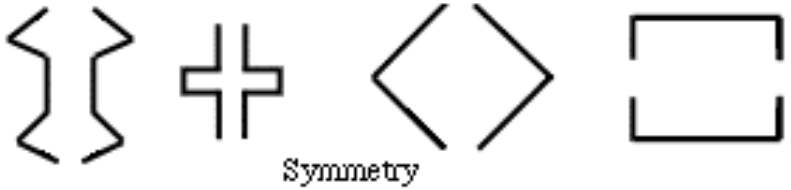
Similar stimuli tend to be grouped
(may override proximity)



Stimuli tend to be grouped
in complete figures

Ambiguous stimuli tend to be resolved
Using the simplest explanation

Stimuli tend to be grouped as to minimize
variations or discontinuities



Regions delimited by symmetric tend
to be perceived as coherent figures

Forsell´s et al. (2010) heuristics

B5. Information coding. Perception of information is directly dependent on the mapping of data elements to visual objects. This should be enhanced by using realistic characteristics/techniques or the use of additional symbols.

E7. Minimal actions. Concerns workload with respect to the number of actions necessary to accomplish a goal or a task.

E11: Flexibility. Flexibility is reflected in the number of possible ways of achieving a given goal. It refers to the means available to customization in order to take into account working strategies, habits and task requirements.

B7: Orientation and help. Functions like support to control levels of details, redo/undo of actions and representing additional information.

B3: Spatial organization. Concerns users' orientation in the information space, the distribution of elements in the layout, precision and legibility, efficiency in space usage and distortion of visual elements.

Uses heuristics from other sets:

B- Freitas et al.

C - Nielsen

D- Zuck and Carpendale

E- Bastien & Scapin

E16: Consistency. Refers to the way design choices are maintained in similar contexts, and are different when applied to different contexts.

C6: Recognition rather than recall. The user should not have to memorize a lot of information to carry out tasks.

E1: Prompting. Refers to all means that help to know all alternatives when several actions are possible depending on the contexts

D10: Remove the extraneous. Concerns whether any extra information can be a distraction and take the eye away from seeing the data or making comparisons.

B9: Data set reduction. Concerns provided features for reducing a data set, their efficiency and ease of use

Cognitive Walkthrough (Wharton, et al., 1992)

- Usability **inspection method** (thus not involving users)
- Based on the fact that users usually prefer to learn a system by using it (e.g., instead of studying a manual)
- Focused on **assessing learnability** (i.e., how easy it is for new users to accomplish tasks with the system)
- May produce results quickly at a low cost
- Applicable at **early phases**, before any coding

How to perform a cognitive walkthrough

1- Task analysis: sequence of steps or actions required by a user to accomplish a task, and the system responses

2- Designers and developers walkthrough as a group, asking themselves a set of questions at each step

3- Data gathering during the walkthrough: answering the questions for each subtask usability problems are detected

4- Report of potential issues

5- UI redesign to address the issues identified

CW Four questions:

- **Will the user try to achieve the effect that the subtask has?**
  (Does the user understand this subtask is needed to reach the goal?)

- **Will the user notice that the correct action is available?**
  (E.g. is the button visible?)

- **Will the user understand that the wanted subtask can be achieved by the action?**
  (E.g. the button is visible but the user doesn't understand the text and will not click on it)

- **Does the user get feedback?**
  Will the user know that they have done the right thing?

Common issues

- The evaluator doesn't know how to perform the task; the method involves the **optimal** sequence of actions

- Involves an extensive analysis and documentation and often too many potential issues are detected, resulting very **time consuming**

Thus:

Lighter variants of Cognitive Walkthrough were proposed to make it **more applicable** in S/W development companies

Streamlined Cognitive Walkthrough (Spencer, 2000)

- A “lighter” version with only two questions:

comprises the 3 first questions of CW

- Will the user know what to do at this step?

- If the user does the right thing, will they know that they did the right thing, and are making progress towards their goal?

- And a set of rules to streamlining the walkthrough and trade-off granularity for coverage

Limitations of Analytical Methods

- Are **subjective**
- Involve several usability experts
- Cannot find all usability problems

Thus, empirical methods (involving users) are needed

observation

query

controlled experiments (scientific approach)

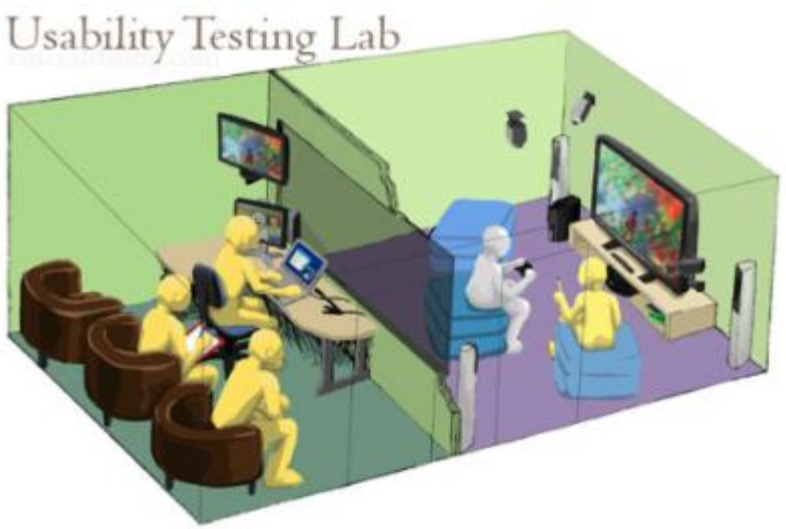
Usability test (engineering approach)

Recommendation: use analytical methods in early phases to debug major usability problems before using empirical methods

Usability tests

- “Engineering approach”
- Involve observation and query
- Main aspects:
  - Participants
  - Tasks
  - Test facilities and systems
  - Experimental design
  - Usability measures
  - Data analysis
- May have a complex logistics

Observation



Has many variants from very simple
to very complex and expensive:

- Direct: observer takes notes
- Indirect: through audio/ video – more complex and time consuming
- Think Aloud: users are asked to explain what they are doing
- Logging: users activity is logged by the system
- Combinations of the previous, etc.

Controlled experiments

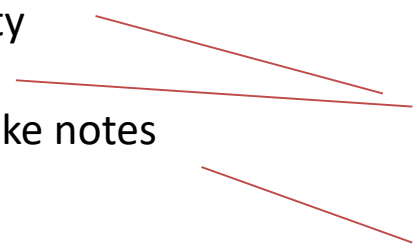
- The **work horse of experimental science** ...

- Important issues to consider:

  - Hypothesis
  - Variables (input or independent; output or dependent)
  - Secondary variables
  - Experimental design (within groups; between groups)
  - Participants (number, profile)
  - Statistics

Controlled experiment

- Define an hypothesis
- Define input (independent), output (dependent) and secondary variables
- Define experimental design (within-groups / between groups)
- Select the participants
- Prepare all the documentation:
  - list of tasks and perceived difficulty
  - final questionnaire
  - list of tasks for the observer to take notes



- Run a pilot test
- Take care of the logistics … and after the experiment analyze data

Participants

Important issues in usability tests and controlled experiments:

- The total number of participants to be tested
  (a valid statistical analysis implies a sufficient number of subjects)
- Segmentation of user groups tested, if more than one
- Key characteristics and capabilities of user group
  (user profile: age, gender, computing experience, product experience, etc.)
- How to select participants
- Differences between the participant sample and the user population
  (e.g. actual users might have training whereas test subjects were untrained)

Tasks

- The task scenarios for testing (or experiments)
- Why these tasks were selected

  (e.g. the most frequent tasks, the most troublesome tasks)
- The source of these tasks

  (e.g. observation of users using similar products, product specifications)
- Any task data given to the participants
- Completion or performance criteria established for each task

  (e.g. n. of clicks < N, time limit)

Test Facilities and equipment

- The setting and type of space in which the evaluation will be done

  (e.g. usability lab, cubicle office, meeting room, home office, home family room, manufacturing floor, etc.)

- Any relevant features or circumstances that can affect the results

  (e.g. video and audio recording equipment, one-way mirrors, or automatic data collection equipment)

- Participant's Computing Environment

  (e.g. computer configuration, including model, OS version, required libraries or settings, browser name and version; relevant plug-in, etc.)

- Display and input devices characteristics

- Any questionnaires to be used

Experimental design

- Procedure/ protocol: the logical design of the test/experiment
- Participant general instructions and task instructions
- The independent variables and control variables
- The usability measures to be used:
  - a) for effectiveness (completeness rate, errors, assists…)
  - b) for efficiency (times)
  - c) for satisfaction

Usability Evaluation Bibliography - Papers

- Spencer, R. The Streamlined Cognitive Walkthrough Method, Working Around Social Constraints Encountered in a Software Development Company, CHI 2000 vol.2 issue 1, 2000, pp353–359

- Wharton, C. Bradford, J. Jeffries, J. Franzke, M. Applying Cognitive Walkthroughs to more Complex User Interfaces: Experiences, Issues and Recommendations CHI '92, 1992, pp381–388

- Gilbert Cockton, Usability Evaluation. In: Soegaard, Mads and Dam, Rikke Friis (eds.), *The Encyclopedia of Human-Computer Interaction*, 2nd Ed, 2013, Aarhus, Denmark: The Interaction Design Foundation, 2013
  http://www.interaction-design.org/encyclopedia/usability_evaluation.html