SIM 17/18 – T5.2
Discount Evaluation Methods

Miguel Tavares Coimbra

Summary

- Discount Evaluation Methods
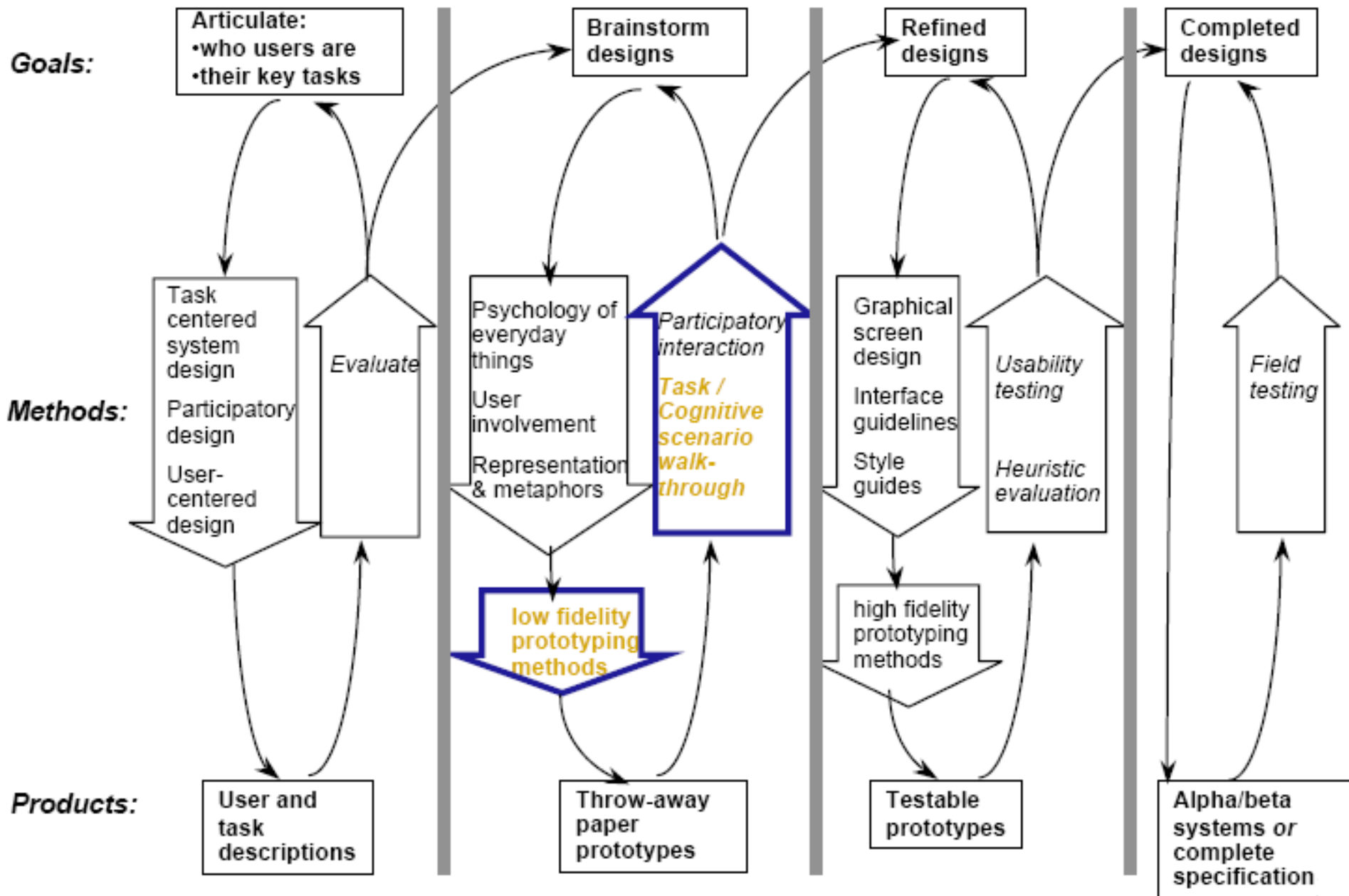
- Cognitive Walkthrough

- Heuristic Evaluation

Discount usability engineering

- Cheap (thus 'discount')
  - No special labs or equipment needed
  - Doesn't need to involve users *directly*
  - the more careful (and informed by users) you are, the better it gets
- Fast
  - On order of 1 day to apply
  - Standard usability testing may take a week
- Easy to use
  - Can be taught in 2-4 hours

Cognitive walkthrough

- Evaluate **"mental model"**
  - Assesses "exploratory learning stage"
    - New users, occasional users
  - *What mental model does the system image facilitate?*
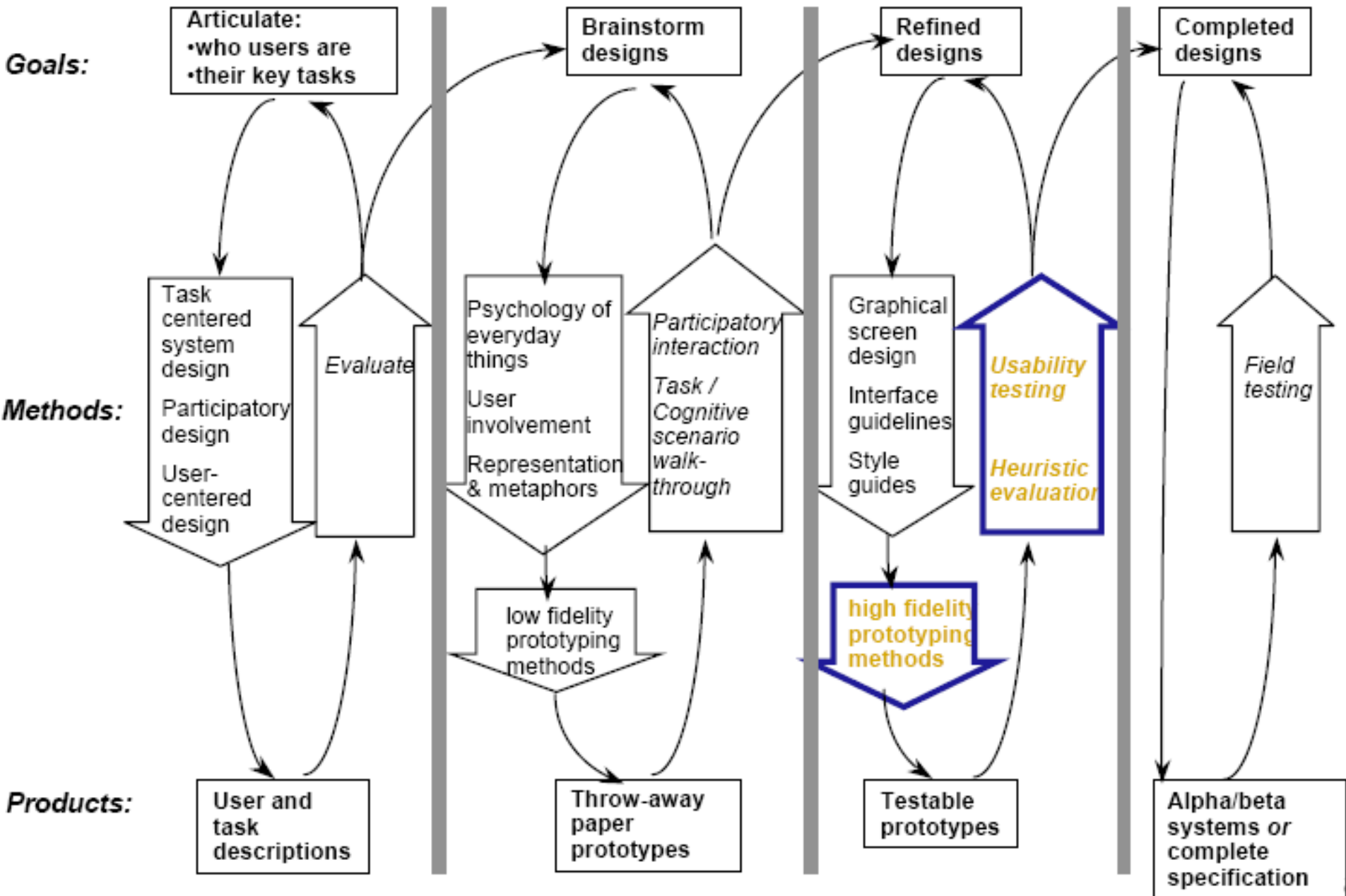 - Done by non-experts and/or domain experts

Interface Design and Usability Engineering

Goals:

| Articulate: • who users are • their key tasks | Brainstorm designs | Refined designs | Completed designs |

Methods:

Task centered system design

Participatory design

User-centered design

Evaluate

Psychology of everyday things

User involvement

Representation & metaphors

Participatory interaction

Task / Cognitive scenario walk-through

low fidelity prototyping methods

Graphical screen design

Interface guidelines

Style guides

Usability testing

Heuristic evaluation

high fidelity prototyping methods

Field testing

Products:

| User and task descriptions | Throw-away paper prototypes | Testable prototypes | Alpha/beta systems *or* complete specification |

Heuristic Evaluation

- **"Fine tune"** system image
  - Targets broader use range
    - New, occasional and expert users
  - Fine-tunes the interface
  - HCI professionals apply a list of heuristics while simulating task execution

Interface Design and Usability Engineering

Goals:

| Articulate:
•who users are
•their key tasks | Brainstorm designs | Refined designs | Completed designs |

Methods:

Task centered system design

Participatory design

User-centered design

Evaluate

Psychology of everyday things

User involvement

Representation & metaphors

Participatory interaction

Task / Cognitive scenario walk-through

Graphical screen design

Interface guidelines

Style guides

Usability testing

Heuristic evaluation

Field testing

low fidelity prototyping methods

high fidelity prototyping methods

Products:

| User and task descriptions | Throw-away paper prototypes | Testable prototypes | Alpha/beta systems *or* complete specification |

Cognitive Walkthrough

Cognitive walkthrough

- **What for**: assessing how well a **new user** will be able to figure out the interface
- **Not for**: assessing performance at highly skilled, frequently performed tasks; or finding radically new approaches
- **Additional advantages**: helps work out task sequence models through observation
- **Disadvantages**: limited utility for frequent-use interfaces, narrow focus, relatively time consuming & laborious (compared to HE)
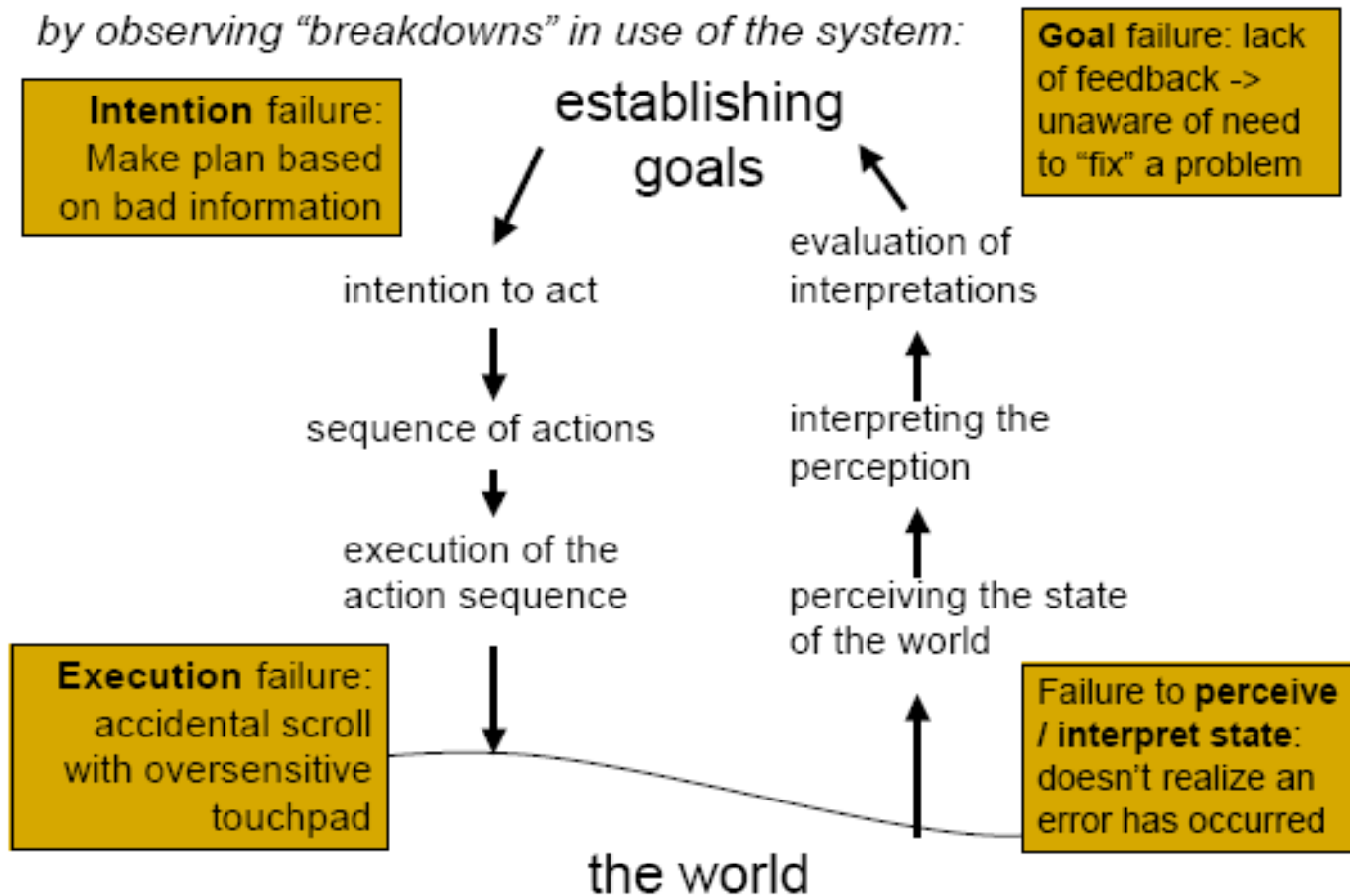
Cognitive walkthrough

- Possible outputs:
  - Loci & sources of confusion, errors, dead ends
  - Estimates of success rates, error recovery; performance speed less evident
  - Helps to figure out what activity sequences could or should be
- What's required: complete interface description
  - Horizontal prototype (paper prototype, Balsamiq, etc)
- Who does it:
  - anyone – different benefits will accrue from using design team members, naïve users or expert outside analysts. **More distance = better!**

How? Roughly:

- Start with a scenario (a **design-specific task)**

- Ask these questions at each step as relevant:
  - Q1: Will the correct action be evident?
  - Q2: Will the user recognize the correct action?
  - Q3: Will the user interpret the result correctly?
  - Q4: Will the user be able to progress towards goal?

Suggestion: Follow Norman's diagram for each task

by observing "breakdowns" in use of the system:

Intention failure: Make plan based on bad information

Goal failure: lack of feedback -> unaware of need to "fix" a problem

establishing goals

intention to act

evaluation of interpretations

sequence of actions

interpreting the perception

execution of the action sequence

perceiving the state of the world

Execution failure: accidental scroll with oversensitive touchpad

Failure to **perceive / interpret state**: doesn't realize an error has occurred

the world

Simpler approach

	User 1	User 2	User 3
Task 1	🟩	🟩	🟩
Task 2	🟧	🟧	🟩
Task 3	🟥	🟥	🟥

1 – User completed the task successfully
2 – User completed the task with difficulty
3 – User did not complete the task

U.PORTO FC

Heuristic Evaluation

Heuristic evaluation

- **What for:** identifying (listing & describing) problems with existing prototypes (any kind of interface)

- **Not for:** coming up with radically new solutions

- **Additional advantages**: contributes valuable insights from objective observers

- **Disadvantages:**
  - Reinforces existing design - better solutions might exist
  - Not very repeatable

Heuristic evaluation

- What's required:
  - A good model of the proposed interface (e.g., at least a paper prototype)
  - A list of design heuristics to be applied
  - A scenario (task example + design prototype)
- Who does it:
  - Team of 3 to 5 experienced, objective people ("**experts")** who **aren't on the design team.**
- General idea:
  - Independently check compliance with usability principles ("heuristics")

  step 1: each evaluator works with interface *alone* (different evaluators will find different problems)
  step 2: evaluators aggregate findings afterwards

One list of heuristics (Nielson, '93)

- H2-1: visibility of system status
- H2-2: match between system & the real world
- H2-3: user control & freedom
- H2-4: consistency and standards
- H2-5: error prevention
- H2-6: recognition rather than recall
- H2-7: flexibility and efficiency of use
- H2-8: aesthetic and minimalist design
- H2-9: help users recognize, diagnose & recover f/ errors
- H2-10: help and documentation

U.PORTO FC

Step 1: Individual evaluation

- **At least two passes for each evaluator**
  - First to get feel for flow and scope of system
  - Second to focus on specific elements
- **Each evaluator produces list of problems**
  - Explain problem w/reference to heuristic or other info
  - Be specific and **list each problem separately**
  - Assign rating of **severity** to each violation

Severity ratings

- Each violation is assigned a **severity rating**
- Combination of:
  - Frequency
  - Impact
  - Persistence (one time or repeating)
- Used to:
  - Allocate resources to fix problems
  - Estimate need for more usability efforts
- Done independently by all evaluators

Severity & extent scales

- **One severity scale (others possible):**
  - 0 - don't agree that this is a usability problem
  - 1 - cosmetic problem
  - 2 - minor usability problem
  - 3 - major usability problem; important to fix
  - 4 - usability catastrophe; imperative to fix
- **One extent scale:**
  - 1 = single case
  - 2 = several places
  - 3 = widespread

U.PORTO FC

Step 2: aggregating results & making recommendations

- **Evaluation team** meets and compares results

- Through discussion and consensus, each violation is documented and categorized in terms of severity, extent

- Violations are ordered in terms of severity

→ **combined report goes back to design team.**

Heuristic evaluation

- Advantages
  - The "minimalist" approach
    - General guidelines can correct for majority of usability problems
    - Easily remembered, easily applied with modest effort
    - **"black box": systematic technique that is reproducible with care.**
  - *Discount usability engineering*
    - Cheap and fast way to inspect a system
    - Can be done by usability experts and end users
- Problems:
  - Principles must be applied intuitively and carefully
    - Can't be treated as a simple checklist
    - Subtleties involved in their use
  - Doesn't necessarily predict users/customers' overall satisfaction
  - May not have same "credibility" as user test data
    - A solution: include design team & developers in usability evaluation

Resources

1. Kellogg S. Booth, Introduction to HCI Methods, University of British Columbia, Canada
   http://www.ugrad.cs.ubc.ca/~cs344/current-term/