



WP4

EVALUATING USABILITY ASSESSMENT METHODS FOR WEB BASED CULTURAL HERITAGE APPLICATIONS

Davide Bolchini (2), Nicoletta Di Blas (1), Franca Garzotto (1), Paolo Paolini (1), Elisa Rubegni(2)

HOC-LAB Politecnico di Milano, IT
TEC-LAB USI, University of Lugano, CH

Outline

- Premise
- Goals
- CH-MiLE+
- Empirical study:
 - Experiment 1: „*quick inspection*“
 - Experiment 2: „*usability project*“
- Results and Discussion

Premise

- A proliferation of general Usability Evaluation Methods (UEMs)
 - Different philosophies - conception of „quality“, „usability“ and their interrelationships -, thus approaches and techniques
- Very few empirical studies on the quality of UEMs
- Limited support to domain specific UEMs

3

CH-MiLE+ : Milano-Lugano Evaluation for Cultural Heritage

- **CH-MiLE+** is a usability inspection method
 - Evolution of two previous usability methods:
 - SUE (Systematic Usability Evaluation)
 - MiLE
 - Borrow general concepts from mainstream usability inspection approaches
 - Partially developed withing EPOCH – WP4
 - Promotion of systematic, structured approach to the analysis, yet aimed at being particularly suitable to novice evaluators
 - Key feature:
 - provision of a domain specific usability heuristics and indicators, workflow of activities, inspection tasks
 - A very large based of use (in educational and cultural institution settings)

4

Goals of our work

- Evaluating CH-MiLE+ in a systematic and reliable way
- Support arguments with empirical data

5

The Quality Space for Method Evaluation

- Quality attributes concerning the **output** of the evaluation
 - Effectiveness
 - Number of usability problems discovered ...
 - Reliability (consistency of results across different inspectors)...
 - Thoroughness (found problems vs existing problems) ...
 - Validity (correcting predicting user's behaviour, no or minimized false positives...)
 - Productivity...
 - Scope...
- Quality attributes concerning the **acceptability and adoption**
 - Learnability
 - Applicability and Compatibility in current practice
 - Verticalization on domains
 - Reusability
 - Cost-effectiveness

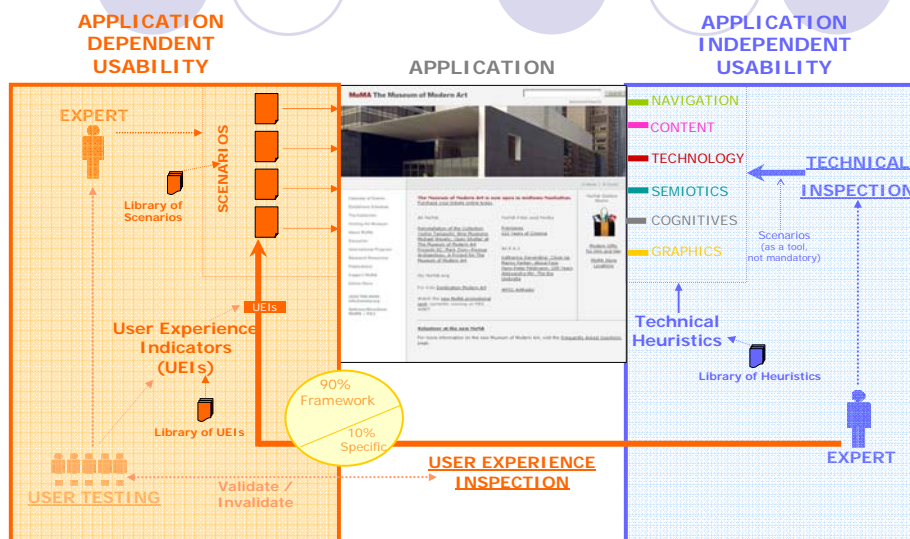
6

Our focus for CH-MiLE+ Evaluation

- Focus on few key attributes that we could measure in a **realistic CH setting** and to support in **effective adoption**:
 - Performance
 - Efficiency
 - Cost-effectiveness
 - Learnability

7

CH-MiLE+ Activities Framework



8

CH-MiLE+ technical heuristics (Example)

- **(82) Technical Heuristics**, coupled by a set of operational guidelines that suggest the inspection tasks to undertake in order to measure the various heuristics.
- Organized by design dimensions
 - **Navigation: (36)** heuristics addressing the website's navigational structure
 - **Content: (8)** heuristics addressing the information provided by the application
 - **Technology/Performance: (7)** heuristics addressing technology-driven features of the application
 - **Interface Design: (31)** heuristics that address the semiotics of the interface, the graphical layout, and the "cognitive" aspects (i.e., what the user understands about the application and its content or functionality)

9

Operationalized Attributes to measure - 1

- **Performance:**
 - Performance indicates the degree at which a method supports the detection of all existing usability problems for an application.
 - It is operationalized as the average rate of the number of different problems found by an inspector (P_i) in given inspection conditions (e.g. time at disposal) against the total number of existing problems (P_{tot})
- **Performance = avrg (P_i)/ P_{tot}**

10

Operationalized Attributes to measure - 2

● Efficiency:

- Efficiency indicates the degree at which a method supports a “fast” detection of usability problems.
- It is operationalized as the rate of the number of different problems identified by an inspector in relation to the time spent, and then calculating the mean among a set of inspectors:

$$Efficiency = avrg\left(\frac{P_i}{t_i}\right)$$

- P_i is the number of problems detected by the i -th inspector in a time period t_i .

11

Operationalized Attributes to measure - 3

● Cost-effectiveness:

- Cost-effectiveness denotes the effort - measured in terms of person-hours - needed by an evaluator to:
 - carry out a complete evaluation of a significantly complex web application
 - produce an evaluation documentation that meets professional standards, i.e., a report that can be proficiently used by a (re)design team to address the usability problems.

12

Operationalized Attributes to measure - 4

● Learnability:

- Learnability denotes the ease of learning a method.
- We operationalize it by means of the following factors:
 - the effort, in terms of person-hours, needed by a novice to become “reasonably expert” and to be able to carry on an inspection activity with a reasonable level of performance
 - the novice’s perceived difficulty of learning, i.e., of moving from “knowing nothing” to “feeling reasonably comfortable” with the method and “ready to undertake an evaluation”
 - the novice’s perceived difficulty of applying application, i.e., of using the method in a real case.

13

Empirical study: general conditions

- The overall study involved 42 participants
 - Students from HCI course, Politecnico di Milano (Milano and Como campus)
 - „novice inspectors“
 - Preconditions:
 - No previous exposure to usability
 - Basic background in web development
 - Etherogeneous profile in terms of age and technical background
 - Preparatory conditions:
 - 5 hours classroom training on usability and MiLE+
 - Assignment of learning material to study (MiLE+ overview, technical heuristics library, 2 real-life case studies, excerpts from an online course on usability and MiLE+)

14

Exp.1: „quick inspection“

- Inspectors
 - 16 graduate students (Como)
- Purpose
 - measure *efficiency* and *performance*
 - *Learnability* hypothesis: study effort to become proficient ≤ 2 days
- Assigned Inspection Goals:
 - Inspect a museum website (Cleveland Museum of Modern Art) with CH-MiLE+ technical inspection
- Setting:
 - Concurrent individual inspection
 - 3 hour time
 - Limited inspection scope (2 main sections, around 300 page instances)
 - One week after MiLE+ classes
- Output produced:
 - inspection notes including, for each usability problem, name, design dimension, description (max 3 lines), page URL

15

Key Results: from zero to hero...

Experiment 1

- Avg number of problems discovered: **14.8**
- Hourly efficiency: avg **4.9** problems per hour
- Existing usability problems (team of experts): 41
- Performance: **36%**

- **After 6 hours of training and a maximum of 15 hours of study, a novice can become able to detect more than one third of the existing usability problems.**

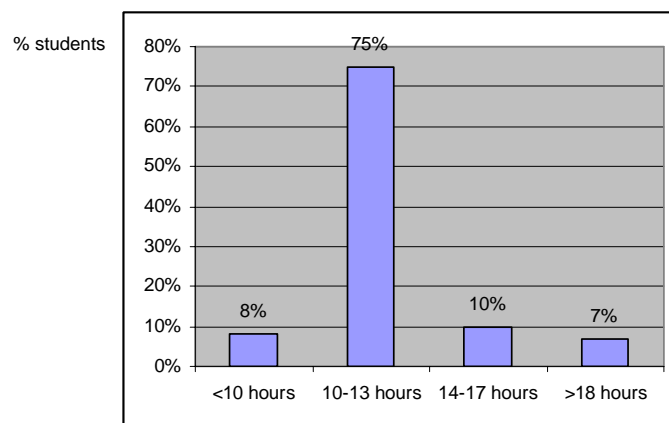
16

Exp.2: „usability project“

- Inspectors
 - 26 graduate students (Milano)
- Purpose
 - measure *perceived difficulty* in *learning* and *using* MiLE+ and *effort* needed to produce a professional evaluation report
- Assigned Inspection Goals:
 - Inspect the full museum website (Cleveland Museum of Modern Art)
- Setting:
 - Asynchronous, team inspection
 - Two months period
 - One week after MiLE+ classes
- Output produced:
 - Complete usability evaluation report

17

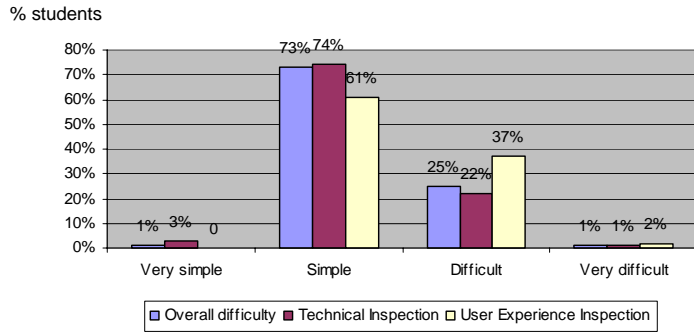
Key Results: Learning Effort



Time invested in the study of MiLE+ before and during the project

18

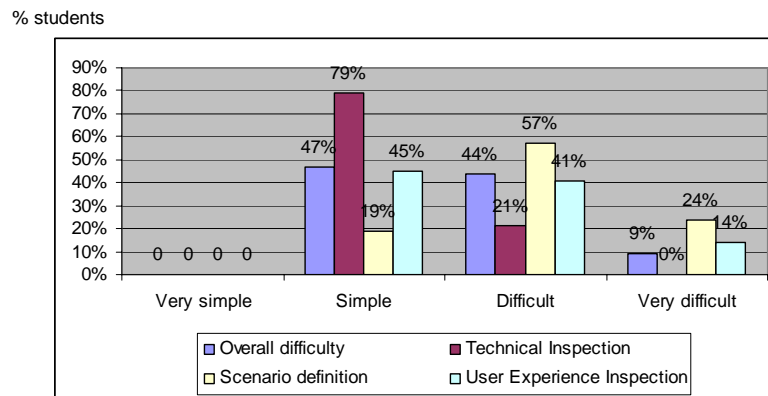
Key Results: Learning Difficulty



Perceived difficulty in learning the various MiLE+ tasks

19

Key Results: Difficulty in Using

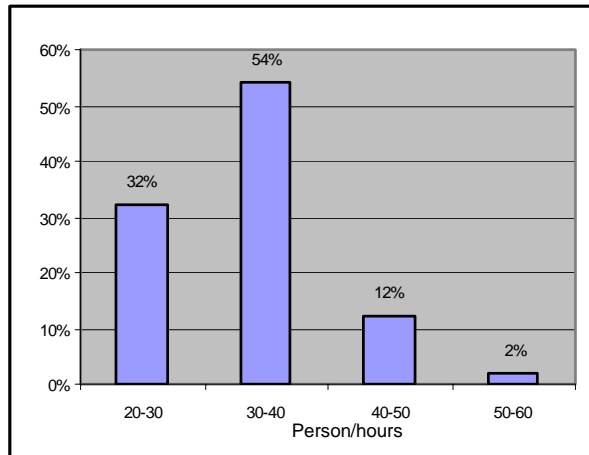


Perceived difficulty in using the various MiLE+ tasks

20

Key Results: Individual Effort

% students

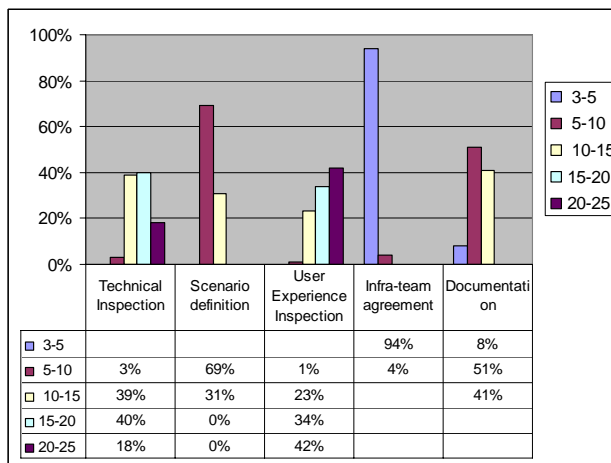


Estimated effort for the entire evaluation process

21

Key Results: Individual Effort per Task

% students



Estimated effort for the entire evaluation process BY task

22

Conclusions

- Promote usability evaluation methods for **adoption in CH settings**
 - Fostering learnability and cost effectiveness
- We have empirically substantiated the adoption suitability of CH-MILE+, with encouraging results
 - **Performant, efficient, cost-effective, easy to learn and use**
- Ongoing diffusion of the method among CH institutions:
 - Courses for professionals

23