The George Washington University
Dissertation Directed by
C. Dianne Martin

Development and Validation of a Multimedia User Interface Usability Evaluation Tool in the Context of Educational Web Sites

by Hana Al-Nuaim

Research Problem

- The explosive growth of the WWW
- The need for user involvement in the design process
- Application of HCI and usability principles to the WWW
- Limitations of existing evaluation instrument

Research purpose

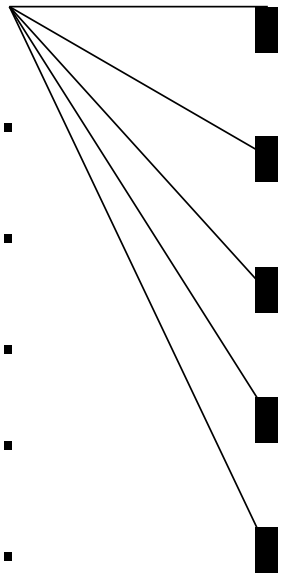
- Including the user in the design process
- Developing an evaluation instrument for them to use that:
  - Addresses traditional user interface issues that can be applied to the WWW
  - Addresses multimedia and the integration of media
  - Reliable and Valid
  - Useful

Hypothesis

- Novices (N) used the MUIU evaluation tool (M) and/or a free format tool (F)
- Hypothesis:
  - **H1 : $N_{M+F} > N_F$ Usability**
  - **H2 : $N_{M+F} > N_F$ Design Defects**
  - **H3 : $N_{M+F} > N_F$ MM and Integration of media**
  - **Null : No difference**

Theoretical Framework

- Traditional GUI principles
- Multimedia
- The WWW Environment
- Cognition and Hypermedia



- Learnability
- Efficiency
- Memorability
- Low error rates
- Subjective satisfaction

Subjects (Experts)

- 4 distinct groups of experts
  - Review of version 1 of the MUIU evaluation tool
  - Selection of important items of version 2
  - Web site selection
  - Use of version 3 of the MUIU evaluation tool to evaluate 2 educational Web sites

Subjects (Novices)

- Novices in user interface, multimedia and Web development
  - Novice review of Version 1
  - Novice evaluation of Web site (Group 1a, 1b, 2a, and 2b)

Development of Version 1

- Identification of evaluation criteria for multimedia, user interface, the WWW, and cognition
- Identification of the usability principles
- The items of the MUIU evaluation tool were chosen from a pool of evaluation criteria items that had an effect on usability

Version 1 of the MUIU Evaluation Tool

- 12 Categories
- 82 items/questions
- Semantic differential 7-point scale (-3 to +3) using bipolar adjectives
- Overall satisfaction scale added to each category

Development of Version 2

- A review of the tool was conducted using user interface and instructional design experts.
- The revision included
  - Item omissions
  - Item modifications
  - Item additions
  - Category modification

Development of Version 3

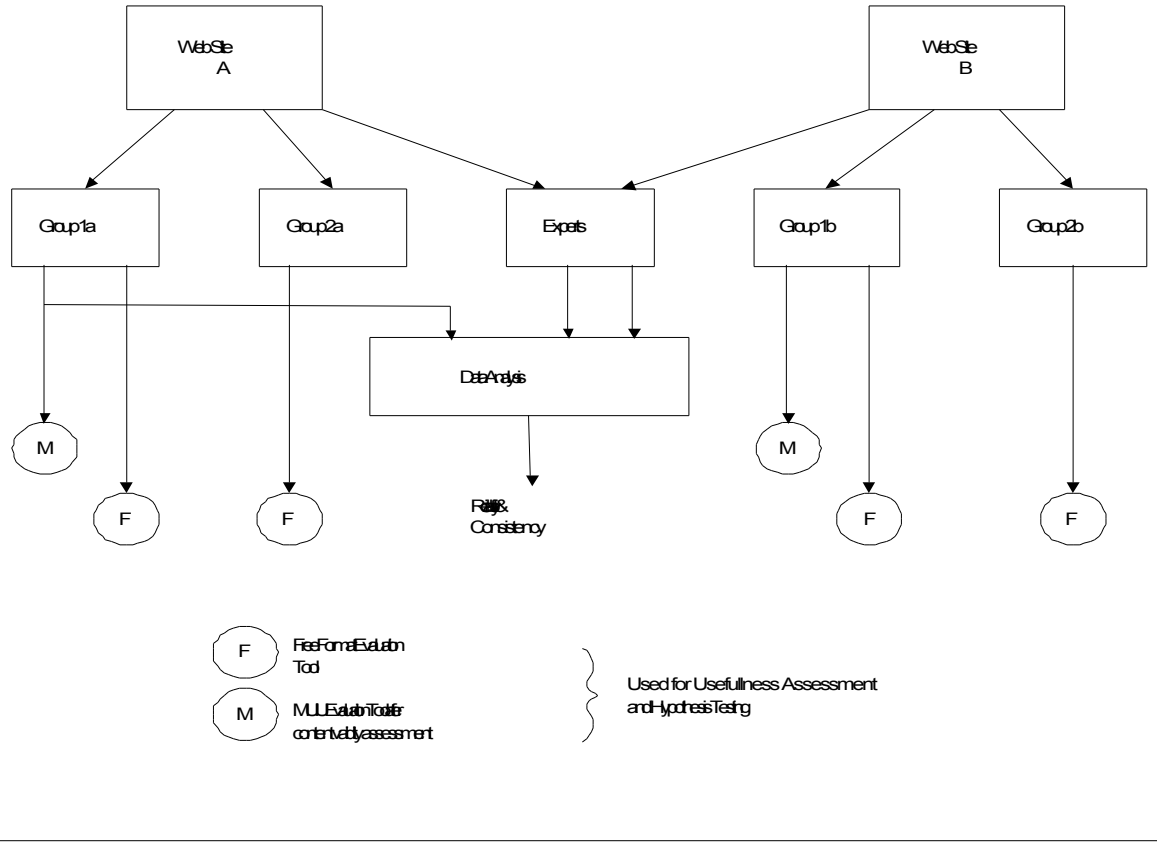
- Experts rated each item on a 5-point scale
- Criteria for item elimination
  - Conceptual: comments and suggestions
  - Numerical: means, standard deviation, and frequency

Version 3 of the MUIU Evaluation Tool

- 36 items
- 8 categories
- 8 overall satisfaction items
- 4 WWW general questions

The Method

Figure 3.1 Methodology Diagram



Reliability

- Experts evaluation of Web site A
  - Overall alpha = 0.87
  - Item analysis : no item when excluded significantly affect overall alpha
- Experts evaluation of Web site B
  - Overall alpha = 0.95
  - Item analysis : no item when excluded significantly affect overall alpha

Content Validity

- The representative of items with respect to the domain
  - The methodology of development of version 3 (expert review and ratings) helped establish its content validity

Expert vs. Novice Evaluation

- t-test on expert and novice evaluation of Web site "A"
- for each item $p > 0.05$ no significant difference

Consistency of Expert Evaluation

- Compare mean of category with mean of satisfaction scale
- Pearson Correlation Coefficient = 0.84
- t-test results:
  - no significant differences in Multimedia and Multimedia Integration categories
  - there exists a difference in the Screen Layout and Navigation categories

Usefulness of the MUIU Evaluation Tool

- Novice evaluation
- Use of free format tool (FF)
- Compare information of Novices $_{\text{MUIU + FF}}$ with Novices $_{\text{FF}}$

Content Analysis of Free Format Tool

- Category definition
- Text collection
- Coding sheet
- Coder training (two coders)
- Inter-rater reliability

  r = .915 (between coders)

Novice Evaluation (Hypothesis Testing)

- Novices $_{\text{MUIU + FF}}$ > Novices $_{\text{FF}}$ (usability)
  - Reject null H1 and accept alternative H1
- Novices $_{\text{MUIU + FF}}$ > Novices $_{\text{FF}}$ (design defects)
  - Reject null H2 and accept alternative H2
- Novices $_{\text{MUIU + FF}}$ > Novices $_{\text{FF}}$ (MM design defects)
  - Reject null H3 and accept alternative H3

Conclusion

- The MUIU evaluation tool:
  - Is a valid and reliable tool
  - Addresses multimedia and the integration of media
  - Addresses traditional user interface issues that can be applied to the WWW
  - Addresses WWW design factors
  - Differentiates between design features
  - Easy to use and administer (paper format)

Implications of the Research

- Identification of design factors that affect the usability of the multimedia user interface for the WWW.
- The MUIU evaluation instrument focuses on the user
- Questionnaire helps user recall specific experiences

Implications of the Research (con't)

- users lack the terminology for expressing design defects and usability problems
- Traditional GUI fundamentals and usability principles still hold true today as they did in the 1980's.

Implications of the Research (con't)

- The MUIU is a valuable addition to the field of UI, Multimedia, WWW and HCI.
  - A validated and reliable instrument for conducting formative and comparative evaluations.
  - A guide in the design or redesign of multimedia Web-based user interfaces
  - A tool for assessing potential areas of improvement for Web-page design.
  - A test instrument to be used in usability labs

Recommendations for Further Research

- Additional validation studies
  - Control of some variables
  - Context for testing
  - Control for equal time
  - Unguided evaluation
- Exposure to usability and multimedia terminology

Recommendations for Further Research (con't)

- Identification if a particular design feature negatively affects overall satisfaction
- Novices' contribution to usability criteria
- Designers' utilization of the information from MUIU
- Remote evaluation with anonymity
- Advancement in technology

Recommendations for Further Research (con't)

- Language and culture differences
- Effects of evaluators' background on evaluation
- The screen layout category
- Online evaluation

Online Evaluation (Advantages)

- Easier to update
- Automatic computation of results
- Users can view results
- Can be accessed directly from anywhere
- Tailoring of tool
- No data encoding errors

Online Evaluation (Disadvantages)

- Technical knowledge
- Maintenance
- Usability evaluation
- Separate window for evaluation form

Discussion

- Q & A