Improving Your Skills in Usability Testing

CHI 2002 Tutorial

Rolf Molich
DialogDesign
Skovkrogen 3,
3660 Stenlose
Denmark

Molich@DialogDesign.dk

April 2002

Table of Contents

Agenda

9.00 – 9.30   **Introduction**
- Overview of tutorial.
- What is quality in usability testing.
- The Comparative Usability Evaluations (CUE-1 and CUE-2).
- Survey of related work.
- The usability lifecycle (taxonomy).

9.30 – 10.00   **Exercise 1: Evaluation Methods**
- Students will review the study designs used in CUE-2 and identify which approaches most suitably meet the requirements of the client scenario.

10.00 – 10.30   **Designing a Usability Study**
- Methods used in CUE: Inspection, Inquiry and Testing
- The Methodological Effect: How results varied based on methodology used in CUE.

10.30 – 11.00   **Break**

11.00 – 11.45   **Exercise 2: Review Usability Test Tasks**
- Each group will review some of the usability test scenarios used in CUE-2 for testing www.Hotmail.com.

11.45 – 12.30   **Tasks and Scenarios**
- The Scenario Effect: How usability test scenarios impact the outcome of a usability study.
- Scenarios used by CUE teams.
- What makes a good usability test scenario?

12.30 – 14.00   **Lunch Break**

14.00 – 14.30   **Identifying and Describing Usability Problems**
- What is a usability problem?

14.30 – 15.00   **Exercise 3: Usability Problem Description**
- Review problem descriptions in the CUE reports. Identify desirable and questionable elements of a usability problem description.

15.00 – 15.30   **Communicating Results**
- Various means of communicating results (other than reports).
- The Report Effect: How the usability report can determine the impact your findings have on directing change in the product development cycle.
- Checklists: Required characteristics of a usability report.

15.30 – 16.00   **Break**

16.00 – 17.00   **Tips and Tricks**
Students will select 2 of the following topics for an in-depth discussion at the tutorial:
- Finding good test participants
- Performing tests.
- Testing intermediate and advanced users.
- International usability testing.
- Assessing the quality of a usability consultancy.

17.00 – 17.20   **Conclusion**
- Critical comments on CUE.
- Reactions from participating CUE test teams to the comparative evaluation of their work ("The psychology of usability testing").
- General lessons learned from CUE.
- Prevention is better than cure.

17.20 – 17.30   **Closing comments, Tutorial Evaluation Forms**

Abstract

Most usability professionals learned their skills by reading books on usability testing or by watching someone else prepare and conduct tests. But how do we know that the particular way we learned to do usability testing is still the best way? Until the [Comparative Usability Evaluation](#) (CUE) project, there was little opportunity for usability professionals to objectively compare their various approaches. As it turns out, there are some striking differences in effectiveness.

This seminar gives rare insights into the practical doings of usability professionals, normally shrouded behind walls of confidentiality. This seminar is not based on the personal opinions of one or two instructors but on the practical accumulated experience - both good and bad - of the thirteen professional usability labs who participated in the CUE-1 and CUE-2 projects, conducting controlled usability tests in realistic, industrial settings.

Insight from these professional lab studies will enable you to assess and improve your abilities in usability test planning, scenario design, usability reporting and usability problem communication.

Objectives of the Course

You will learn

- What constitutes quality in usability testing

- Options available in planning a usability test, and how these choices may affect the outcome

- What makes a good usability test scenario, and how test scenarios impact the quality of a usability study

- What a usability problem is, and how usability problems are identified

- How many test participants are required to find for example 85% of the usability problems in a product

- Required characteristics of a usability report to assure its communicative value

- Novel communication techniques that are vastly superior to traditional usability problem communication through paper reports and video tapes

Instructor Biography

Rolf Molich owns and manages DialogDesign (www.DialogDesign.dk), a small Danish usability consultancy. Rolf conceived and coordinated the comparative usability evaluation studies CUE-1 and CUE-2 where 13 professional usability teams tested the same application. Rolf was a principal investigator in the Nielsen Norman Group's recent large scale usability test of 20 US E-Commerce websites involving more than 60 users. Rolf has worked with usability since 1984; he is the co-inventor of the heuristic inspection method (with Jakob Nielsen) and the author of the best-selling Danish book User Friendly Computer Systems of which 25,000 copies have been sold.

Introduction

Rolf Molich
www.DialogDesign.dk

Overview of this Tutorial

- Focus on four critical areas within usability testing:
    - Usability test planning and methodology.
    - Task and scenario design.
    - Identifying usability problems.
    - Usability problem communication.
- Conclusion.

Facts - Not Opinions

This tutorial

- Gives you a rare insight in the practical doings of real world usability professionals.

- Is mostly based on the practical accumulated experience – both good and bad – of professional usability labs.

- Is not based on the personal opinions of one or two instructors.

Beware of Opinions Disguised as Facts

We have our mouths full of users,
but our practice is full of ourselves.

- Marijke Rijsberman
in a CHI-WEB posting

CUE Overview

Take a website.

Take N professional usability teams.

Let each team usability test the website.

Are the results similar?

CUE Overview

● Investigate the reproducibility of
   usability test results.

● Get an impression of REAL professional
   website usability testing procedures -
   not beautified descriptions.

"Who Checks the Checker?"

- When did YOU last have an objective check of your usability testing skills?

- Who would you trust as an evaluator of your usability testing skills?

Quality in Usability Testing

In my opinion

- A successful usability test sharply reduces the number of disasters experienced by typical users in a user interface.

This means:

- Your success as a usability tester also depends on how much designers and programmers listen to your advice.

CUE-1: Comparative Usability Evaluation 1

- Four professional teams usability tested the same Windows calendar program (Task Timer for Windows).

- Two US teams (Sun, Rockwell), one English (NPL) and
one Irish (HFRG, Univ. Cork).

- See paper in section D of the tutorial documentation.

- Results published in a panel and a paper at UPA98.

- Main conclusions similar to CUE-2.

CUE-2: Comparative Usability Evaluation 2

- Nine teams have usability tested the same website, www.hotmail.com:
  - Seven professional teams.
  - Two student teams.

- Four European, five US teams.

- See paper in section D of the tutorial documentation.

NON Purposes of CUE

- To pick a winner.
- To make a profit.

Usability Test Procedure

- Website address (www.hotmail.com) disclosed at start of three week test period.

- Client scenario.
  (written by Erika Kindlund and Meeta Arcuri)

- Access to client through intermediary. (Erika Kindlund)

- Three weeks to carry out test using preferred standard approach.

- Deliver anonymized usability test report.

Test Reports Freely Available

Some of the anonymous test reports are reproduced in
the tutorial documentation.

You can find the rest at
http://www.dialogdesign.dk

Not .com

On the home page follow the link "In English."

Problems Found

	CUE-2		CUE-1	
Total number of problems	310		141	
Found by seven teams	1		-	
six teams	1		-	
five teams	4		-	
four teams	5		1	
three teams	17		1	
two teams	50		11	
Found only by one team	232	(75%)	128	(91%)

Resources Used in Real-Life Testing

Team	A	B	C	D	E	F	G	H	J
Person hours used for test	136	123	84	(16)	130	50	107	45	218
# Usability professionals	2	1	1	1	3	1	1	3	6
Number of tests	7	6	5	(50)	9	5	11	4	6

The Usability Test Lifecycle - Taxonomy

This is how most teams structured their tests:

● Prepare test.

● Conduct test.

● Communicate results.

This structure applies not only for "Think Aloud" testing but for many other forms of testing like Focus Groups, Interviews, Observations, etc.

Taxonomy: Prepare Test

- Appoint test facilitator.

- Select test tasks and scenarios (together with developers).

- Define test participant profile (together with developers).

- Screen and recruit test participants.

- Inform test participants.

- Prepare test setup.

Taxonomy: Conduct Test

Test participants one by one:

- Greet test participant.

- Questionnaire: Background.

- Ask for expectations.

- Carry out test tasks one by one. Facilitator takes notes.

- Debrief.

- Hand over gift. Goodbye to test participant.

Taxonomy: Communicate Results

- Consensus building with developers.
- Communicate results.
- Follow up on results.

Recommended Readings

Standard textbooks:

- Joseph S. Dumas and Janice ("Ginny") Redish
  A Practical Guide to Usability Testing.
  Intellectbooks 1999, www.intellectbooks.com
  ISBN 1-84150-020-8
  (US distributor: ISBS, Portland, OR, 1-800-944-6190).

- Jeff Rubin
  Handbook of Usability Testing,
  John Wiley, New York NY, 1994
  ISBN 0-471-59403-2.

Other Resources

CHI-WEB listserver:

● http://www.acm.org/sigchi/web/chi-web.html

● 5-50 postings per day.

● Sample, recent discussion subjects:
   – ADA (Standards for web accessibility)
   – Forms evaluation
   – Good uses for Flash

The-usability-list-that-must-not-be-named.
   – Except in a face-to-face conversation.