CUE-3 Ideas

CUE-3 is a successor to the successful Comparative Usability Evaluation studies CUE-1 and CUE-2 (www.dialogdesign.dk/cue.html). The focus in CUE-3, however, is on web interface usability inspection as opposed to user-involved usability evaluation.

1. Main Goals

1. Set a benchmark against which other usability professionals can measure their usability evaluation skills.

2. Provide a survey of the state-of-the art within professional usability evaluation of websites.

3. Show participating usability professionals their strengths and weaknesses in usability evaluation, which is one of the core processes of the usability profession.

4. Answer a number of key questions about inspection methods:
   a. Why do evaluators produce different results when usability inspecting an interface? Answering this question will provide a major insight into teaching inspection methods and refining the actual inspection methods.
   b. Does a discussion session between usability evaluators improve the quality of the evaluation results sufficiently to justify the extra resources spent?
      – Are less important or unreasonable problems filtered out?
      – Are important, reasonable problems met with acclaim?
      – Do the discussions cause important but yet unreported problems to be discovered (group effect, “the whole is larger than the sum of the components.”).
      – Does the discussion focus on usability-in-the-large rather than micro-usability?
      – Why are certain important problems not detected?
   c. What are the key strategies used to reach consensus during the discussion sessions?
   d. How do evaluation results compare to usability test results? (Requires that at least two independent teams run a usability test of the website)
   e. Does previous experience influence the quality of the results?

5. Provide the basis for an interesting and entertaining panel discussion at CHI 2002 or another major conference on usability.

6. Collect experiences suitable for a subsequent CUE-4 study. Ideas for CUE-4:
   - Top qualified participants from all over the world.
   - Consensus building will be by email discussion, chat forum or videoconferencing.

Main goals 4, 5 and 6 will not be known by the participants before the study is completed.

2. Organization

I (Rolf Molich) will carry out the day-to-day administration of CUE-3.

I will be assisted by an Advisory Committee, and by an observation team.

Advisory Committee members are chosen based on their practical experience and insight into correct methodological approaches to studies like this one. Members of the Advisory Committee will be expected to comment on the high level methods and goals of the study. They are not expected to do detailed data analysis, but comments on the data analysis strategies are most welcome. All communication within the Advisory Committee will be by email. Face-to-face meetings are not planned.

Confirmed members of the Advisory Committee:
- Erika Kindlund, Intraspect
- Jakob Nielsen, NN Group
- John Morkes, Trilogy
- Marie Tahir, NN Group
- Niels Ebbe Jacobsen, Nokia
- Robin Jeffries, Sun

The observation team will consist of four experienced usability professionals who will act as observers during the discussion sessions.

Confirmed members of the observation team:
- Lene Byskov, DialogDesign
- Niels Ebbe Jacobsen, Nokia
- Rolf Molich, DialogDesign

3. Proposed step-by-step procedure

1. **Select an appropriate meeting place and date**
   Suggestion: End of August in Copenhagen.

2. **Invite approximately 15 experienced usability professionals to participate**
   - Participation will be mainly by invitation to members of the Danish usability community. At the moment I do not think it will be necessary to advertise the study in public to get the required number of participants.
   - At most two people from any given company can participate.
   - Minimum qualifications to participate: Must have at least one year's experience with professional usability work. Must have conducted at least two professional usability tests.
   - Published reports will be anonymous unless participants explicitly agree to the opposite.
   - Each participating organization will cover all of its own expenses in connection with CUE-3. Participants must pay for their own travel and lodging (if required) but they will not be charged for organizing CUE-3 and meeting expenses.

3. Select a suitable website for evaluation.

- The language of the website must be English.
- Participants will be asked to evaluate a full website or a substantial part of a website, for example the registration section of Hotmail.
- The website must be highly interactive.
- The website must not require any special domain expertise.
- It must be possible to specify a reasonable evaluation scenario that makes the website relevant for Danes.

4. Seven working days before the meeting date: Start of evaluation.

Participants will receive

- The evaluation scenario.
- Detailed reporting requirements.
- A questionnaire about their personal qualifications.

The scenario includes the URL of the website. Participants will not be able to contact the website project team during the study. Any questions for the project team that might arise during the evaluation should be directed to the administrator who will substitute for the website project team.

5. Detailed reporting requirements.

The evaluation report format will be strictly prescribed since the topic of differences in reporting format was covered adequately by CUE-1 and CUE-2. Evaluation reports should be written in English; participants who feel uncomfortable expressing themselves in English may submit a report in Danish. A detailed report template and examples will be provided.

A report should contain:

a. One page title, author(s) and affiliation.

b. One page executive summary.
   - 2-3 most important problems
   - 2-3 most important positive findings
   - High-level managerial recommendations.

c. Comments on the website.
   - Each comment must be classified as a problem description or a positive comment.
   - Each problem must be rated on three scales:
     - Frequency (how often does the problem occur),
     - Impact (how serious is the problem when it occurs),
     - Persistence (will users learn how to get around the problem).

   Estimated frequency and impact must be shown in a table with the following format:

	% users who will not be affected by this problem	% of users who will suffer minor problems (brief delays)	% of users who will experience serious problems (significant delays; users eventually complete their task)	% of users who will experience disasters (unable to solve task; voices strong irritation; solves task incorrectly)	Persistence
Problem 1	70	20	10	0	1
Problem 2	0	0	20	80	3
…					

Persistance must be estimated according to the following scale:

1. Users quickly learn how to get around the problem,
2. Users only learn to get around the problem after encountering it several times,
3. Users never learn how to get around the problem.

The gross way to determine whether a problem should be fixed is to add (or multiply) the three numeric values and see whether the problem gets a high score.
Include as many comments as you would consider appropriate in an industrial setting.

6. Questionnaire about qualifications.

Each participant must answer the following questions:

a. Name
b. Job title
c. Experience with usability work (number of years). Please specify both experience from university and from professional work.
d. Approximate number of usability tests conducted (with users)
e. Approximate number of usability evaluations conducted (without users)
f. How much time do you normally spend on a usability evaluation?

The questionnaire is used by the observation team to divide the participants into teams.

7. Two working days before the meeting date: Receive evaluation reports from each participant.

Each participant must submit an individual evaluation report and a questionnaire about his/her qualifications. A participant who does not submit a proper report in time will not be admitted to the meeting.

8. Two working days before the meeting date: Send out methodology questionnaire.

Each participant will be asked to answer the following questions about the actual evaluation methodology that he or she used:

a. How much time did you spend on
   - Examining the interface?
   - Reporting the results?
   - Other activities that are part of the evaluation? Please provide a short description of each of these activities.
b. Please describe briefly how you carried out the evaluation.
c. If you used heuristics, please attach a list of the heuristics you used
   - Did the heuristics help you in discovering usability problems?
d. If you used cognitive inspection, please attach a list of the user tasks you used.
   - Did the user tasks help you in discovering usability problems?
e. Please describe the differences, if any, between this evaluation and previous evaluations you have carried out.

9. Meeting.

The observation team will divide the participants into four teams. Each team will consist of 3-4 participants. Participants from the same company can not be on the same team. Participants will not get copies of each other's reports before the meeting; copies will be available at the start of the discussion session but they will be distributed only if the team asks for them.

Meeting agenda:

10.00 – 10.20 Welcome and introduction.

10.20 – 12.30 Discussion session.

During the discussion sessions each team will try to reach consensus on a list of comments on the website. At the end of this session the team must provide a list of comments

Discussions will be in Danish. Each team will have access to a pc with internet access attached to a projector.

Each team will be observed by a member of the observation team; these observers will not interact with the teams. We will consider videotaping each of the four discussions.

12.30 – 13.30 Lunch.

13.30 – 15.30 Plenum discussion:

- Problems and positive comments on the website.
- Evaluation methodologies.
  - How useful do you consider the method? Why?
  - How and where did you learn the method?
  - How many others have you taught the method to?
  - What do you think is the best ways to teach/learn it?
  - How do you prefer to rate or prioritize the issues you find?
- Were the discussion sessions worthwhile?
- Social aspects of usability evaluation.
- CUE-3 process and validity.
- Was CUE-3 a worthwhile experience for the participants?
- How to improve CUE-3.

10. Checklist for observation team.

In addition to the main goals specified in section 1.4, observation team members should pay particular attention to the following points while observing the discussion session:

a. How does the team reach consensus?
   Basic methods:
   – Quid pro quo
   – Merging lists
   – Sensible discussion
   – Other – specify.
b. How much time is spent on coming to consesus about the findings vs. consensus about inspection method heuristics and techniques?
c. Is there a competitive element in the discussion? How does it show? Is the competitive element constructive or destructive?
d. What kinds of "face-to-face" strategies do team members use to convince their colleagues that a problem is important enough to prioritize above others?

11. Capture relevant pages from the website.

12. Collect information from the evaluation reports and observations and make conclusions.

4. Goals of Each Component of the Study

The goal of the Reports (submitted by participants and analysed by advisory team) is to compare results and methodologies.

The goal of the Discussion Session is to identify consensus-building strategies.

The Plenum Discussion is a de-briefing, high-level conclusions session.

Usability Evaluation Scenario *

In preparation for a major revision of the website www.Avis.com, Avis wants your assessment of the usability for Scandinavian users of the current website.

Avis is a leading worldwide supplier of rental cars. Avis wants its website to have a similar leading position on the World Wide Web.

Avis has recently noticed a growing interest in its website from international customers, mostly from Scandinavia (Denmark, Norway and Sweden). At the moment, approximately 20% of Scandinavian customers make their reservations on-line. Avis has also noticed, however, that many customers start to make a reservation but abandon the website before the reservation is complete. Avis's webteam has not been able to recognize any pattern in where people abandon the website. Further details about these statistics are not available for this evaluation.

Below is a list of user tasks that Marketing and Development have identified as benefiting from user feedback. The tasks are listed in order of importance to Marketing with the most important tasks listed first.

1. Find out what it costs to rent a certain car.
2. Make a reservation.
3. Get an overview of the kind of cars that Avis offers.
4. Find out where you can rent a car.
5. Determine if Avis has special deals.

Avis's management wants you to focus on the core tasks listed above, and not the other services now offered from the site. They are not interested in the following areas:

1. Rates, reservations, etc., for corporate users, travel agents, travel partners, etc.
2. Rates, reservations, etc., for vehicles that require special driving licenses like trucks, buses, motorcycles, etc.
3. Programs and services.
4. Maps and Weather.
5. Information about the company Avis.
6. The usability of Avis's non-US websites like www.Avis.dk.

User Profile

Most customers of the Avis website are between 18 and 50 years old. Approximately 70% are male and 30% are female.

Avis's goal is that the website should be usable for anyone who is reasonably familiar with the web. Previous experience with on-line reservation systems should not be required.

Reporting Requirements

Your evaluation report must be submitted to molich@dialogdesign.dk not later than Wednesday, August 29, 2001, at 17.00 hours.

Evaluation reports should be written in English; participants who feel uncomfortable expressing themselves in English may submit a report in Danish.

The following report formats are accepted:
- Microsoft Word version 97 or earlier.
- Adobe pdf.

Your evaluation report should contain:

a. One page title, author and affiliation.
You may decide to anonymize your report. If you decide to do so, then you should make sure that your author name and affiliation does not appear anywhere in your report.
Each evaluator will have a code letter assigned to him or her. The code letter should be displayed in the header of each page regardless of whether the report is anonymous or not.

b. One page executive summary.
- 3 most important problems
- 3 most important positive findings
- Up to three high-level managerial recommendations.

c. Comments on the website.
Each comment must be classified as a problem description or a positive feature.
Report one usability problem or one positive feature per comment. Avoid comments that are conglomerates of several problems or positive features.
Include as many comments as you would consider appropriate in an industrial setting.
Number your comments consecutively, starting from 1.

d. Estimated problem frequency, impact and persistence.
Each problem must be rated on three scales:
- Frequency (how often does the problem occur),
- Impact (how serious is the problem when it occurs),
- Persistence (will users learn how to get around the problem).

Estimated frequency, impact and persistence must be reported in a table with the following format:

	% users who will not be affected by this problem	% of users who will suffer minor problems (brief delays)	% of users who will experience serious problems (significant delays; users eventually complete their task)	% of users who will experience disasters (unable to solve task; voices strong irritation; solves task incorrectly)	Persistence
Problem 1	70	20	10	0	1
Problem 2	0	0	20	80	3
…					

Persistence must be estimated according to the following scale:
1. Users quickly learn how to get around the problem,
2. Users only learn to get around the problem after encountering it several times,
3. Users never learn how to get around the problem.

A report template with examples will be provided separately.

Addendum

In addition to the evaluation report, please also submit a separate addendum with the following information:

a. **Deviations from your standard usability evaluation procedure.**

b. **Resources used for the evaluation (person hours).**
   Provide a day-by-day timesheet for the period where you carried out your evaluation.

c. **Comments on how realistic the evaluation has been.**

d. **Other comments (if applicable).**

There are no specific formatting requirements for this report except that your code letter should appear in the header of each page. The addendum may be anonymous, just like the evaluation report.

Questions

Please direct any questions that might arise during your evaluation for Avis's management or website project team to Rolf Molich (molich@dialogdesign.dk).

* Important note.

The above usability evaluation scenario was written by DialogDesign for the explicit purpose of doing a sample usability evaluation of a state-of-the-art website.

DialogDesign has chosen the www.Avis.com website for this usability evaluation because we think that this website presents a number of interesting usability challenges and solutions, whose real usability would be of interest to a wider audience.

DialogDesign has written this requirements specification based on our general experience with such specifications. Although the scenario presented in this scenario is entirely reasonable according to our experience, it is also completely fictitious. Neither Avis's management nor any Avis employee has approved or even seen this scenario.