Ecological Validity and Behavioural Measures in the Usability Testing of New Applications: A Workshop in Reality Testing

Gitte Lindgaard[1], Bruce Tsuji[2], Shamima Khan[3]
[123]Human Oriented Technology Lab, Carleton University, Ottawa, Ontario, Canada K1S 5B6
[1]gitte_lindgaard@carleton.ca; [2] btsuji@connect.carleton.ca; [3]skhana@connect.carleton.ca

With a focus on ecological validity and behavioural measures, this workshop will address the usability testing of applications based on new or emerging technologies. In many cases we are no longer dealing with the simple one-user-one-desktop-computer situation that was amenable to laboratory testing in the early days. Furthermore, standard usability measures of efficiency and effectiveness are not always relevant to these new situations and environments. It is challenging then, to balance the desire to isolate the effect of individual variables against the business demands of providing meaningful (ecologically valid) and timely feedback to development groups. Another significant issue in usability testing is identifying behavioural measures that capture the right phenomena and yield the information required. We will discuss usability testing best practices with examples from novel wireless and broadband TV applications. Participants and facilitators will work together to determine ways to achieve ecological validity and apply suitable behavioural measures in participants' own research without violating industrial confidentiality.

Workshop participants, please navigate to: http://hot.carleton.ca/projects/hci2005-reality/ or email gitte_lindgaard@carleton.ca

Keywords: ecological validity; usability testing; behavioural measures; confidentiality; industrial secrets; reality testing; new applications; new products

1. TIMELINESS & IMPORTANCE

One of the key barriers associated with HCI and usability professionals learning from each other is the problem of confidentiality. The industries and organizations who employ HCI and usability professionals are reluctant to see trade secrets and potentially lucrative new product ideas shared without significant protections in place. This workshop acknowledges this problem and provides means to protect the ideas while still learning from each other.

2. PROCEDURE

This workshop is highly interactive. Participants will work with and learn from each other under the guidance of the facilitators. First we will lay the groundwork by identifying usability testing challenges across a range of new applications. Secondly we will present participant problems in the form of anonymous case studies to protect organizational confidentiality. We will then engage the entire group in brainstorming, discussion, and evaluation of different strategies and approaches to specific challenges within the case studies. This will provide workshop participants with a set of tools and workable solutions that they will be able to apply to the particular problems they face.

3. DESCRIPTION

Ecological validity refers to the real world meaningfulness of data gathering activities. The term "ecological validity" was originally coined by Egon Brunswik [1]. Brunswik was committed to analyses of the impact of the environment on human judgment. The problem of ecological validity is particularly acute in experimentation with new or

emerging applications and services. New technological devices no longer fit the traditional "one user, one computer" scenario that was once central to HCI research and practice. People use mobile or wireless technologies in different situations. For example, they may listen to music while jogging, talk on the phone while driving, and work on a spreadsheet on the patio. Each of these situations has unique characteristics that affect the task demands and their interaction with the technology. It is difficult, if not impossible, to re-create these kinds of situations in a laboratory with sufficient face validity to yield the kinds of user behaviours one may expect in "real life". In addition, the usual behavioural measures of efficiency and effectiveness as stipulated in the ISO 9241-11 standard [2] are not meaningful in contexts other than work situations in which the primary motive is to increase worker productivity. User satisfaction, the third usability measure in the ISO standard, is also likely to rest on entirely different criteria in the examples described.

The need to identify suitable behavioural measures is another challenge in this new context. How, for example, do you assess usability and acceptance of interactive TV in a typical laboratory? What behaviours will demonstrate adequacy of the technology under scrutiny? Will the expected behaviour even occur at all during the session? Novel technologies and applications thus pose distinctive challenges for achieving ecological validity and choosing appropriate behavioural measures during usability testing.

4. WORKSHOP SCHEDULE

The workshop is one day in duration.

09:00-09:30 Welcome and introductions
09:30-10:30 Workshop scope and process
10:30-10:45 Break
10:45-12:30 Participants' case studies (problems, what worked, what didn't)
12:30-13:30 Lunch
13:30-14:30 Brainstorming solutions
14:30-14:45 Break
14:45-16:30 Discussing/refining results
15:15-16:30 Participants' case studies (problems, what worked) continued
16:30-17:30 Summary of solutions, next steps, wrap up
19:30-23:00 Group dinner

5. CO-ORDINATORS

The three workshop co-ordinators have between them almost 60 years of relevant experience in wired and wireless telecommunications, business intelligence, data mining, medical/health applications, CRM, data security, e-commerce, consumer electronics, streaming media, and other domains. They have substantial university and industrial experience and have worked as HCI practitioners, instructors, consultants, managers, as well as clients of HCI services.

6. REFERENCES

[1] Brunswik, E. (1955). Representative design and probabilistic theory, Psychological Review, 62, 193-217.

[2] ISO 1997. ISO/DIS 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs): Guidance on usability..