

Human Computer Interaction

Usability Evaluation

Dr Pradipta Biswas, PhD (Cantab)
Assistant Professor
Indian Institute of Science
<http://cambum.net>

Why we need it



Why we need it

Pradipta Biswas x Hermes Webmail Service x Social TV - Select Session x

stv.rtbi.in/SocialTV/session.php

Social TV

Welcome Pradipta







Home Upload Schedule Feedback Help Notification(0) Logout

Home > Entertainment

Select the Video:

Select the Language English

Search stream: Search

 CricketWorldCupWinningMoment Duration : 00:02:11	 YesPrimeMinisterBBCcomedy Duration : 00:02:52	 EaglesHotelCaliforniaLive Duration : 00:06:34
 Failure_is_the_pillar_of_success Duration : 00:01:16	 NAMMA_OORU_final Duration : 00:03:09	 Motivational_Video Duration : 00:03:25

Home Profile

RTBii
IITM's Rural Technology and Business Incubator

IU-ATC
India-UK Advanced Technology Centre of Excellence in Next Generation Networks, Systems and Services

12:19
05/09/2012

Why we need it

The screenshot displays the 'INDIAN RAILWAYS PASSENGER RESERVATION ENQUIRY' website. The main search form is titled 'Trains/Fare/Accommodation Availability between Important Stations'. It includes the following fields and options:

- Source Station:** NEW DELHI - NDLS
- Destination Station:** MUMBAI - BCT
- Class:** Sleeper Class
- Journey Date:** 07 Dec

A red circle highlights the 'Get It' button, and a red arrow points to the 'Clear Data' button. To the right of the search form, there is a box for 'For e-Payment of' which lists: Corporate Tax, Wealth Tax, Income Tax, Gift Tax, and TDS & FBT.

At the bottom of the page, there is a blue banner for '24 x 7 Online Tax Payment Facility' and a footer with copyright information: 'Copyright © 2006, Centre For Railway Information Systems, Designed and Hosted by CRIS | Disclaimer: Best viewed at 1024 x 768 resolution with Internet Explorer 5.0 or Netscape Navigator 5.0 and higher.'

Strategies

- Heuristic evaluation
- Guidelines
- Cognitive walkthrough
- Think aloud protocol
- Cognitive dimensions of notation
- Simulation
- Survey
- Controlled experiment

Heuristic evaluation

- Nielsen's Usability Heuristics
 - Visibility of system status
 - Match between system and the real world
 - User control and freedom
 - Consistency and standards
 - Error prevention
 - Recognition rather than recall
 - Flexibility and efficiency of use
 - Aesthetic and minimalist design
 - Help users recognize, diagnose, and recover from errors
 - Help and documentation

Guidelines

The screenshot shows a web browser window displaying the W3C website. The address bar shows the URL www.w3.org/TR/tr-technology-stds. The page features a blue header with the W3C logo and navigation links for STANDARDS, PARTICIPATE, MEMBERSHIP, and ABOUT W3C. A search bar is located in the top right corner. The main content area is titled "STANDARDS ONLY (SORTED BY TECHNOLOGY)" and includes a filter menu set to "Standards Only" sorted by "Technology". A "Show View" button is also present. Below the filter, there are radio buttons for "Show details" and "Hide details". The main list of standards includes:

- ▶ Accessibility (All)
- ▶ Authoring Tool Accessibility Guidelines (ATAG)
- ▶ Best Practices for Authoring HTML
- ▶ CC/PP
- ▶ CSS
- ▶ Declarative Web Applications
- ▶ Device Description Repository
- ▶ DOM
- ▶ DOM events
- ▶ Efficient XML Interchange
- ▶ GRDDL
- ▶ HTML

The browser's taskbar at the bottom shows the date and time as 09:04 18/10/2012.

Issues

- Easy to use
- Needs multiple evaluators
- Tools available for automatic checking (<http://www.w3.org/WAI/RC/tools/complete>) , but not works for all
 - E.g.: Checking usability / accessibility of dynamic web content

Cognitive walkthrough

- Experts simulate users' interaction
- Walkthrough high frequency to low frequency tasks
- Good for exploratory interfaces
- Can quickly identify errors/ wrong assumptions in structure /sequence of interfaces

Think aloud protocol

- Users undertake task while 'thinking aloud'
- Provides rapid, high-quality, qualitative user feedback
- Allows meaningful, direct dialogue
 - Designer understands users' way of thinking and can clarify
- Can be video recorded for later analysis

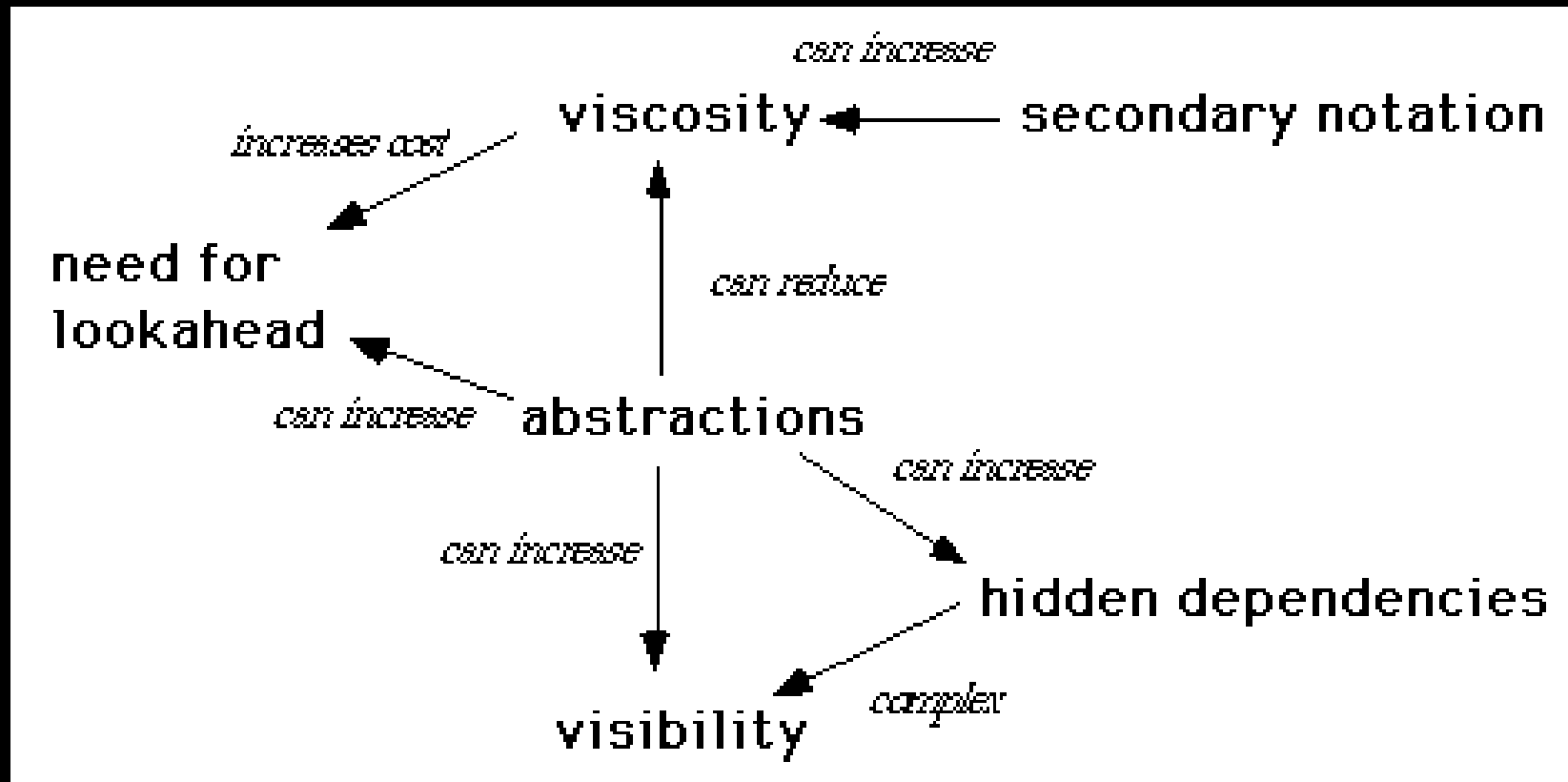
Cognitive Dimensions of Notations

- A 'lightweight' approach to evaluation
 - Easy to apply
- Define a common set of vocabulary to discuss design
- Define design trade-offs in terms of the vocabulary

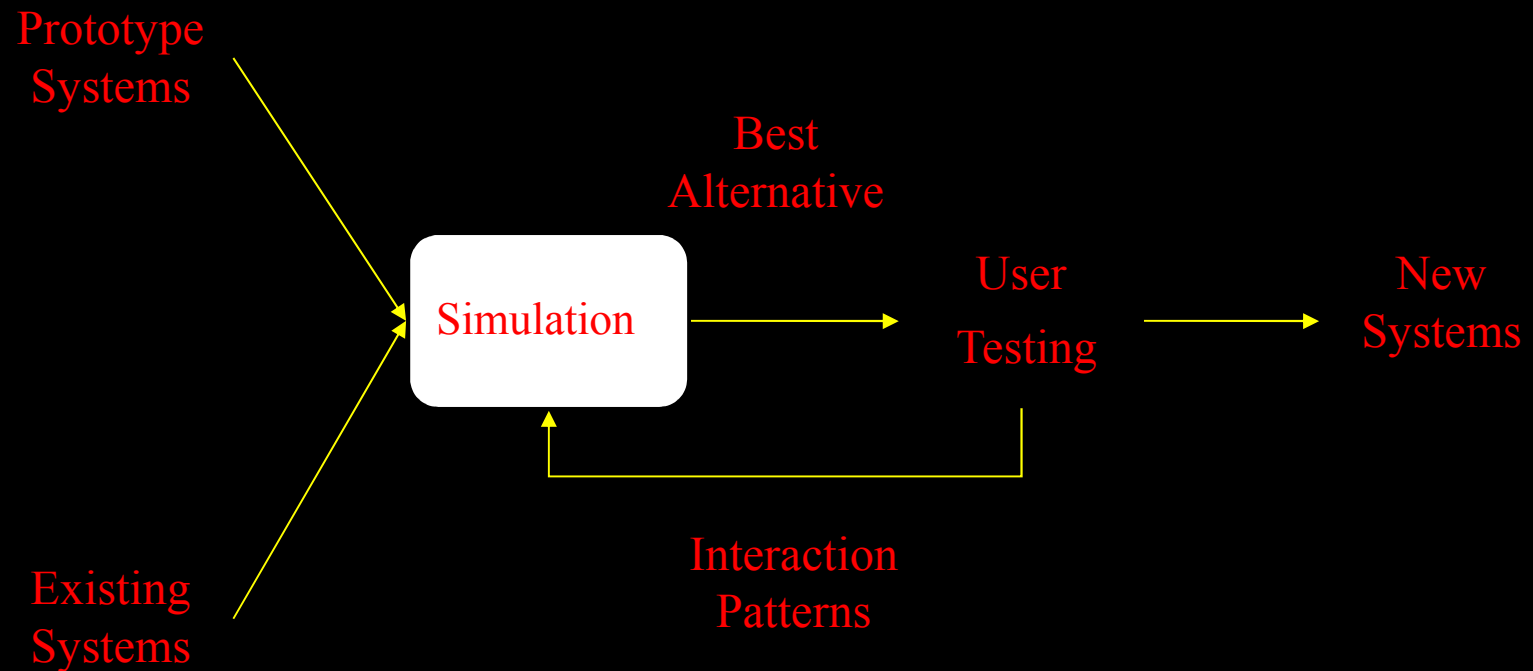
Cognitive Dimensions of Notations

<i>dimension</i>	<i>thumbnail description</i>
Viscosity	resistance to change
Hidden Dependencies	important links between entities are not visible
Visibility and Juxtaposibility	ability to view components easily
Imposed Lookahead	Constraints on order of doing things
Secondary Notation	extra information in means other than program syntax
Closeness of Mapping	representation maps to domain
Progressive Evaluation	ability to check while incomplete
Hard Mental Operations	operations that tax working memory
Diffuseness/Terseness	succinctness of language
Abstraction Gradient	amount of abstraction required, amount possible
Role-expressiveness	purpose of a component is readily inferred
Error-proneness	syntax provokes slips
Perceptual mapping	important meanings conveyed by position, size, colour etc
Consistency	Similar semantics expressed in similar syntax

Cognitive Dimensions of Notations



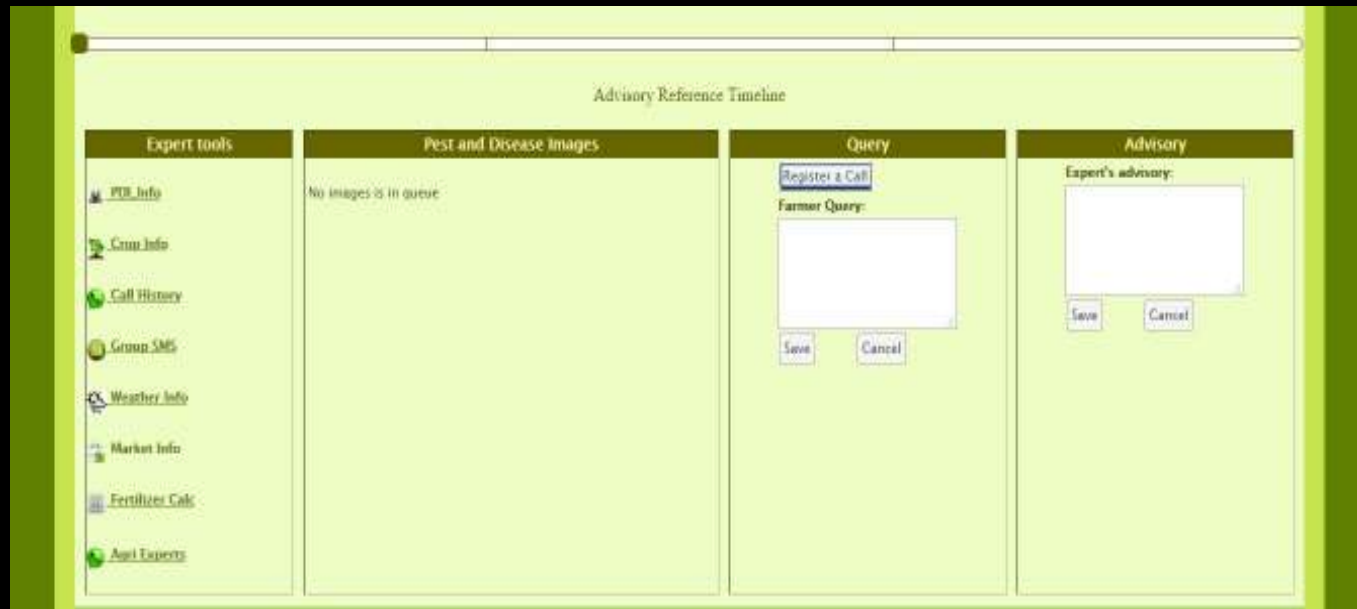
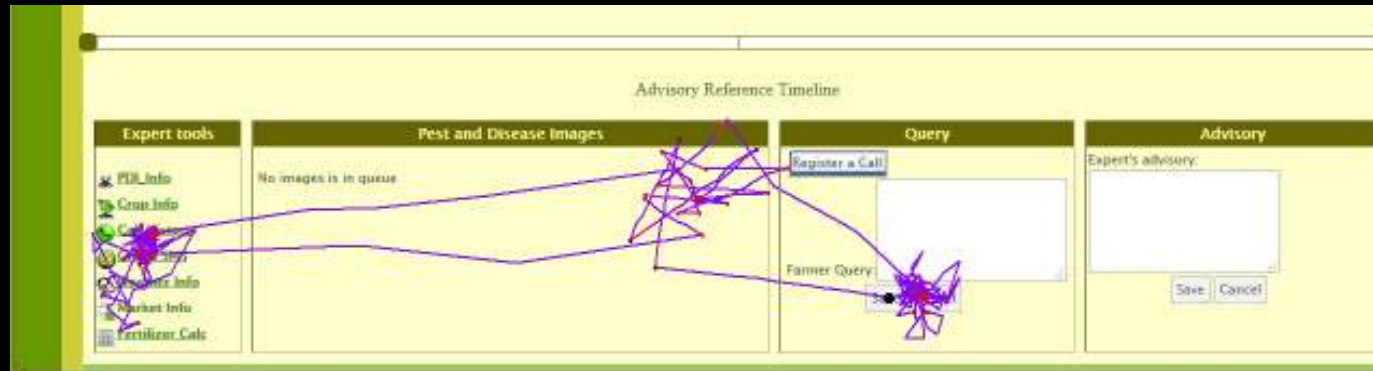
Evaluation through simulation



Simulation



Simulation



Simulation



Surveys

- IBM User Satisfaction Questionnaire
 - Usefulness
 - Information quality
 - Interface quality
- Shneiderman's Questionnaire for User Interaction Satisfaction
 - System Experience
 - User reaction
 - Screen design
 - Learning
 - On line tutorial and so on

Issues

- Easy to collect a lot of data
- Needs careful consideration in
 - Questionnaire design
 - Avoid negative or double question
 - Experimenter's bias
- Easy to crowd-source
- People tend to pretend 'good', which may not be true in reality

Cognitive Load Measurement

- NASA TLX
 - Average performance
 - Peak Performance
- System Usability Scale
- Bedford Workload Scale (BWS)

SUS

I think that I would like to use this system frequently.

I found the system unnecessarily complex.

I thought the system was easy to use.

I think that I would need the support of a technical person to be able to use this system.

I found the various functions in this system were well integrated.

I thought there was too much inconsistency in this system.

I would imagine that most people would learn to use this system very quickly.

I found the system very cumbersome to use.

I felt very confident using the system.

I needed to learn a lot of things before I could get going with this system.

NASA TLX

Workload

Mental Demand:

How mentally demanding was the task?



Physical Demand:

How physically demanding was the task?



Temporal Demand:

How hurried or rushed was the pace of the task?



Effort:

How hard did you have to work to accomplish your level of performance?



Frustration:

How insecure, discouraged, irritated, stressed, or annoyed were you?



Performance: Please note that the following scale is a measure of how well you think you did on the task.

Performance:

How successful were you in accomplishing the task?



BWS

BRS

Expt ID:

Participant ID:

Was it possible to complete the task?

Was workload tolerable for the task?

Was workload satisfactory without reduction?

Workload Insignificant

Workload Low

Enough spare capacity for all desirable tasks

© Pradipta Biswas, University of Cambridge

BRS

Expt ID:

Participant ID:

Was it possible to complete the task?

Was workload tolerable for the task?

Was workload satisfactory without reduction?

Insufficient spare capacity for easy performance of additional tasks

Reduce spare capacity. Additional tasks cannot be given desired attention

Little spare capacity. Level of effort allows little attention to additional tasks

© Pradipta Biswas, University of Cambridge

BRS

Expt ID:

Participant ID:

Was it possible to complete the task?

Was workload tolerable for the task?

Very little spare capacity. Maintenance of effort in the primary tasks not in question

Very high workload with almost no spare capacity. Difficulty in maintaining level of effort

Extremely high workload. No spare capacity. Serious doubts about the ability to maintain level of effort

© Pradipta Biswas, University of Cambridge

Controlled experiments

- Psychological experiment in controlled laboratory setting
- Results suitable for formal statistical analysis
- Good to measure effect of an interface or interaction on users' performance
- Not suitable for exploratory analysis or at design phase, should be used as a confirmatory test
- Will be discussed in detail in the next lecture.

Other techniques

- **Pluralistic walkthrough**
 - uses group meetings where users, developers, and human factors people step through a scenario, discussing each dialogue element.
- **Feature inspection**
 - lists sequences of features used to accomplish typical tasks, checks for long sequences, cumbersome steps, steps that would not be natural for users to try, and steps that require extensive knowledge/experience in order to assess a proposed feature set.
- **Consistency inspection**
 - designers who represent multiple other projects inspect an interface to see whether it does things in the same way as their own designs.
- **Standards inspection**
 - an expert on an interface standard inspect the interface for compliance.
- **Formal inspection**
 - Experts hold courtroom style meeting with designers

Usability metrics

Short term –based on representative task / application

- Task completion time
- Number of errors / correct selection
- Cognitive load
 - NASA TLX score

Usability metrics

Long term metrics

- increased enjoyment of using technology
- feel immersed in experiences
- feel closer to / more socially connected with collaborators
- feel less overwhelmed by information overload -- feel that they are getting the info they need, not being distracted by the irrelevant
- increased connection to others
- enhanced productivity
- enables new experiences
- enables participation by the disabled in aspects of life that were formerly closed to them
- less air travel (due to increasing effectiveness of remote collaboration)

Take away points

- Introduction to different usability evaluation techniques
- Can be used in different phases of system development life cycle, for example
 - Early design: Simulation, CDN
 - Late design: Heuristic evaluation, Simulation
 - Early prototype: Guidelines, Surveys
 - Advanced prototype: Controlled experiment
- Techniques should be used in conjunction to standard software testing routines.