

Ubiquitous Computing Seminar

Session 5 Usability & Evaluation

João Pedro Sousa, Jens-Peter Kaps

IT 821, SWE 699, ECE 699
George Mason University

outline

- [Scholtz 04] framework for evaluation
- [Iachello 06] "paratype" inquiring technique
- [Chung 04] evaluation of design patterns for Ubicomp
- [Trevor 05] compare personalization methods

[Scholtz 04] framework for evaluation

proposes conceptual measures
for usability evaluation along 9 aspects:

- application adoption
- appeal
- match to user's conceptual models
- trust & privacy
- user attention
- interactions
- invisibility
- side effects
- robustness

adoption & appeal

- market share
 - demographics
 - new users/unit of time
 - number of tasks supported
 - (not/)originally envisioned
 - enjoyment when using/loss when not available
 - how much users are willing to pay for it
 - pride in using the app
 - peer pressure
 - rating of look and feel
- appeal

conceptual models & trust

- match between user's expectations and
 - app features
 - interaction primitives
 - behavior
 - app assumptions about user responsibilities
 - create awareness & offer control mechanisms
 - which data is collected, stored, and disseminated
- trust/privacy

user attention & interaction

- effectiveness in supporting tasks
 - time to completion
 - time on task vs. overhead
 - shifts on focus of attention
 - unnoticed events
 - comparison of diff devices/modalities
 - scalability
 - #tasks per user
 - #users/collaboration groups
 - effectiveness in managing multiple users and conflict resolution
- attention

invisibility, side effects & robustness

- appropriate inference of context and intentions
- explanation of beliefs and disambiguation
- learning and personalization invisibility
- added value relative to alternatives
- impact on indirect stakeholders
- requirements of use (awkward goggles)
- resilience to untrained users robustness
- ease of recovering from mistakes
- self-handling of errors, keeping users informed

instantiate evaluation framework for concrete apps

determine with stakeholders

- which aspects are more important
- select concrete measures
that instantiate the relevant conceptual measures
- establish success criteria (metrics)

aside: experiment types in HCI formative experiments

- paratypes
 - situate end users in real/realistic situations and ask them to "use" and evaluate a *mockup*
 - e.g. Palm Pilot; recording cell-phone
- prototypes
 - develop illustrative functionality, maybe just UIs, stub out major functional components; show to real users
- Wizard of Oz
 - same as prototype, where major functional components are replaced by human operators "behind a curtain"

aside: experiment types in HCI formative experiments

- paratypes
 - situate end users in real/realistic situations and ask them to "use" and evaluate a *mockup*
 - e.g. Palm Pilot; recording cell-phone
 - prototypes
 - develop illustrative functionality, maybe just UIs, stub out major functional components; show to real users
 - Wizard of Oz
 - same as prototype, where major functional components are replaced by human operators "behind a curtain"
-
- early in the life-cycle:
do not require a full product
 - help form the design

aside: experiment types in HCI empirical evaluation

- in the lab
 - design limited but realistic scenarios;
 - bring representative users in to perform the scenarios under controlled conditions
 - typical duration: one day
- in situ
 - researchers observe real users in the field
 - users may be asked to perform scenarios, or left free to act
 - typical duration: a few days, maybe spread out
- in the wild
 - real users are free to use, or not, the product
 - typical duration: weeks to several months
- late in the life-cycle:
require a fully usable product
- validate the design

● evaluate user performance
and perception

aside: experiment types in HCI empirical evaluation

- in the lab
 - design limited but realistic scenarios;
 - bring representative users in to perform the scenarios under controlled conditions
 - typical duration: one day
- in situ
 - researchers observe real users in the field
 - users may be asked to perform scenarios, or left free to act
 - typical duration: a few days, maybe spread out
- in the wild
 - real users are free to use, or not, the product
 - typical duration: weeks to several months