Experiments III: Threats to Validity

Prof. Narges Mahyar
Advanced Methods in HCI
CHECK IN & FEEDBACK

• Presentations and your submissions grades and feedback are posted

• Need help for your project? Come see us
  – No office hours this week

• Evaluation of teammate's work
  – “It happens often in team projects that at least one person will not do nearly enough work and other members will need to carry the load”.
READING DISCUSSION

• Ephemeral adaptation
• Methodology matters
EPHEMERAL ADAPTATION

APPROACH

Abrupt onset of predicted items

Gradual onset of non-predicted items

DESIGN BENEFITS

Temporary adaptive support

Maintains spatial consistency

Based on literature in visual attention

[Findlater, Moffatt, McGrenere, and Dawson, CHI 2009]
WHAT IS EPHEMERAL ADAPTATION?

- an adaptive method of highlighting menu items that reduces visual search time while maintaining spatial consistency
HOW IS AN EXPERIMENT DESIGN REPORTED?

- how easy/difficult was this paper to read?

- what were the elements that made it
  - easy?
  - difficult?
VALUE OF PILOTING AND 2 STUDIES

• what was the benefit of piloting and having two separate studies (study 1 and study 2)?
  (i.e., why not just do one BIG study???)
CAREFUL DETERMINATION OF VARIABLES

• Too much to test in one study (likelihood of success – learning something meaningful – would have been very low)

• At each stage (piloting, Study 1, Study 2) they were able to clarify which variables were important and at which values (i.e., determine factors and factor levels)
THREATS TO VALIDITY

• What are the threats to validity?
Validity

• Are you measuring what you say you are measuring?

• Methodological Soundness
  – Anticipate potential threats to validity
  – Create procedures to eliminate or reduce threats
RELIABILITY

• The degree to which a test or measure will produce the same results when applied in the same circumstances
**Type of Validity**

1. **Internal**: demonstration of causality
2. **Statistical**: accuracy of the conclusion drawn from a statistical test
3. **Construct**: finding support for the theory or construct
4. **External**: generalizability
I. INTERNAL VALIDITY

Demonstration of causality (did A cause B?)

• is there a causal relation between independent & dependent variables?

• e.g., Hawthorne effect – subjects change their behavior because they know they are being studied
2. STATISTICAL VALIDITY

Accuracy of the conclusion drawn from a statistical test

- Were the statistical tests used appropriate?
- Are the measures used to measure Dependent Variables reliable?
3. CONSTRUCT VALIDITY

How well the results support the theory or construct?

• are we measuring what we think we are measuring?

• e.g., create a questionnaire to assess early “adopted-ness”, but in fact it assesses financial ability to buy new technology instead
4. EXTERNAL VALIDITY

Do the results generalize?

• e.g., sample not representative of true population
• e.g., insufficient description of experiment protocol
ON DECK...

- Thursday
  - Final prototype review
  - Milestone Test 2