HCI: Empirical Research Methods

Learning Objective

- In the previous lecture, we discussed the conceptual framework of empirical research
  - Which is observation-based investigation
  - In the context of HCI, it refers to investigation about the usability of a system with end users
- The investigation is based on testable research questions

  - We saw how to formulate appropriate questions (the validity of questions, the trade-off and how to resolve the trade-off)
- We also discussed the process of observation and measurements

Learning Objective

- In this lecture, we shall discuss the other aspect of empirical research
- In particular, we shall discuss the design of experiments

Themes of Empirical Research

- To recap, the three themes of empirical research are
  - Answer and raise Questions (testable research questions)
  - Observe and Measure (ratio scale of measurement is preferable)
  - User Studies

Themes of Empirical Research

- In the previous lecture, we already discussed the first two themes, namely: (1) Answer and raise Questions, (2) Observe and Measure
- The third theme (User Studies) is equally important since we perform observations in such studies

User Study

- A user study, in the context of HCI, is a scientific way of collecting and analyzing observational data from end users on an interactive system
- Collection of data involve experiments and design of experiments

Experiment Design

- Experiment design is a general term referring to the organization of variables, procedures, etc., in an experiment
- The process of designing an experiment is the process of deciding on which variables to use, what procedure to use, how many participant to use, how to solicit them etc.

Terminology

- Terms to know
  - Participant
  - Independent Variable (Test Conditions)
  - Dependent Variable
  - Control Variable

Terminology

- Terms to know
  - Random Variable
  - Confounding Variable
  - Within Subjects vs. Between Subjects
  - Counterbalancing and Latin Square

Participant

- The people participating in an experiment are referred to as participants
  - When referring specifically to the experiment, use the term participants (e.g., “all participants exhibited a high error rate…”)
  - General comments on the problem or conclusions drawn from the results may use other terms (e.g., “these results suggest that users are less likely to…”

Independent Variable

- An independent variable is a variable that is selected or controlled through the design of the experiment
  - Examples include device, feedback mode, button layout, visual layout, gender, age, expertise, etc.
- The terms independent variable and factor are synonymous

Test Conditions

- The levels, values, or settings for an independent variable are the test conditions
- Provide names for both an independent variable (factor) and the test conditions (levels) for the controlled variable (see the following examples)

Factor	Levels (Test Conditions)
Device	Mouse, Trackball, Joystick
Feedback Mode	Audio, Tactile, Visual
Task	Pointing, Dragging
Visualization	2D, 3D, Animated

Dependent Variable

- A variable representing the measurements or observations on a independent variable
- It is required to provide a name for both the dependent variable and its unit
  - Examples: Task completion time (ms), speed (WPM (word per minute)), selections per minute, etc.), error rate (%), throughput (bits/s (bps))

Control Variable

- Circumstances or factors that might influence a dependent variable, but are not under investigation need to be accommodated in some manner
- One way is to control them or to treat them as control variables (e.g., room lighting, background noise, temperature)

Random Variable

- Instead of controlling all circumstances or factors, some might be allowed to vary randomly
- Such circumstances are random variables

Confounding Variable

- Any variable that varies systematically with an independent variable is a confounding variable
  - For example, if three devices are always administered in the same order, participant performance might improve due to practice; i.e., from the 1st to the 2nd to the 3rd condition; thus "practice" is a confounding variable (because it varies systematically with "device")

Within Subjects, Between Subjects

- The administering of levels of a factor is either within subjects or between subjects
  - If each participant is tested on each level, the factor is within subjects
  - If each participant is tested on only one level, the factor is between subjects. In this case a separate group of participants is used for each level

Within Subjects, Between Subjects

- The terms *repeated measures* and within subjects are synonymous
- A relevant question is, which of the two approaches (within subject and between subject) should be chosen in designing an experiment

Within Subjects, Between Subjects

- Answer: It depends!
  - Sometimes a factor must be between subjects (e.g., gender, age)
  - Sometimes a factor must be within subjects (e.g., session, block)
  - Sometimes there is a choice. In this case there is a trade-off

Within Subjects, Between Subjects

- The advantage of within subject design is, the variance due to participants' pre-dispositions should be the same across test conditions
- Between subjects design, on the other hand, has the advantage of avoiding interference effects (e.g., the practice effect while typing on two different layouts of keyboards)

Counterbalancing

- For repeated measures designs, participants' performance may tend to improve with practice as they progress from one level to the next
  - Thus, participants may perform better on the second level simply because they benefited from practice on the first (this is undesirable)
- To compensate, the order of presenting conditions is counterbalanced

Latin Square

- Participants are divided into groups, and a different order of administration is used for each group
- The order is best governed by a Latin Square (The defining characteristic of a Latin Square is that each condition occurs only once in each row and column)

Latin Square

- Example: suppose we want to administer 4 levels (denoted by A, B, C and D) of a factor to 4 participants (represented by P1, P2, P3 and P4)
  - We can construct a 4×4 Latin square arrangement to depict the order of administering the levels to each participant

P1	A	B	C	D
P2	B	C	D	A
P3	C	D	A	B
P4	D	A	B	C

Latin Square

- In a *balanced* Latin Square, each condition both precedes and follows each other condition an equal number of times
  - We can construct a balanced 4×4 Latin square arrangement for the previous example

P1	A	B	C	D
P2	B	D	A	C
P3	D	C	B	A
P4	C	A	D	B

Expressing Experiment Design

- Consider the statement “3 x 2 repeated-measures design”
  - It refers to an experiment with two factors, having three levels on the first, and two levels on the second. There are six test conditions in total. Both factors are repeated measures, meaning all participants were tested on all test conditions
- Any type of experiment is expressed similarly

Summary

- In this and the previous lectures, we discussed the fundamental ideas associated with empirical research (namely: question formulation, observation and measurement and experiment design)
- In the next lecture, we shall see another important aspect, namely the case for statistical analysis of empirical data