Experimental Methods

COMP2044: Human-Computer Interaction (2024-2025)

Matthew Pike

Overview

Learning Objectives

- Introduce standard scientific experimentation terminology.
- Explore the development, design, and execution of HCI studies.
- Examine subjective and objective measures.

Experimental Evaluation

Experimental Factors

- Participants - sometimes referred to as subjects or users.
  - Who? How many? Are they representative of the target population?
- Variables - things that we can manipulate or measure.
- Hypothesis - What you'd like to show.
  - A statement that predicts the relationship between variables.
- Experimental Design - How you're going to do it.
  - The plan for how the experiment will be conducted.

Research Questions

- Formal questions that guide experimental investigations.
- Define what needs to be measured/compared.
- Should be:
  - Focused and answerable through experimentation.
  - Grounded in existing theory/practice.
  - Structured to enable clear measurement.
- Example HCI Questions:
  - Does gesture input improve task completion time compared to menus?
  - How does error rate vary between novice and expert users?
- Related to Variables (next slide):
  - What is the relationship between IV _____ and DV _____?
  - Does _____ moderate the effect of _____ on _____?

Variables

Independent Variable (IV)

- The variable that is manipulated by the researcher to produce different conditions.
- Examples:
  - Interface style: perhaps text vs. icons.
  - Menu Design: X vs Y Number of items.

Dependent Variable (DV)

- The variable that is measured to determine the effect of the independent variable.
- Examples:
  - Time to complete a task.
  - Number of errors made.
  - User satisfaction.

Hypothesis

- A hypothesis is a statement that predicts an outcome based on the relationship between the independent and dependent variables.
  - Examples: “Error rates will increase as font size decreases.”
- A researcher will also state the null hypothesis (H0), which states that there is no relationship between the variables. The aim of the experiment is to reject the null hypothesis.
  - Example: “Error rates will not increase as font size decreases.”
- Hypotheses are often, but not always, derived from existing theories or previous research.

Experimental Design

Within-Subjects Design

- Each participant completes all conditions.
- It's possible that participants learn from one condition to the next, ultimately biasing the results.
  - Counterbalancing can help mitigate this.
- Less costly (as fewer participants are needed) and less prone to participant variability.

Between-Subjects Design

- Each participant completes only one condition.
- Removes learning effects or "transfer" effects.
- Requires more participants.
- Prone to participant variability, which can bias the results.

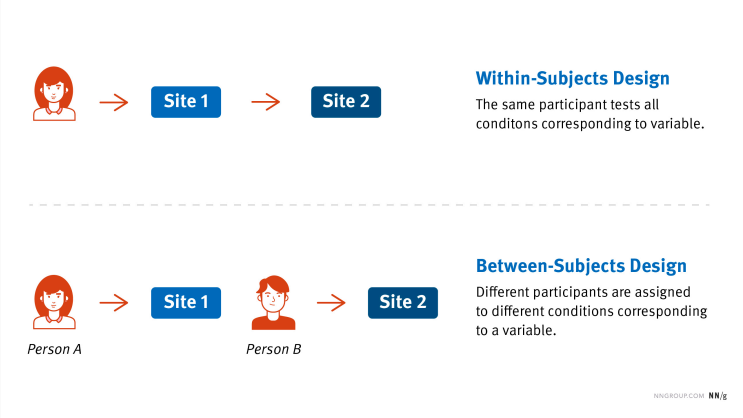

Figure 1: Differences between within-subjects and between-subjects designs - Image Credit

Field vs. Lab Studies

Field Studies

- Pros:
  - Ecological validity - behaviour is observed in a natural setting.
  - Enables longitudinal studies - observing behaviour over time.
- Cons
  - Potential distractions which may affect the results.
  - Uncontrolled settings - may introduce confounding variables.
- Suitable for
  - Ecologically-focused outcomes.

Lab Studies

- Pros:
  - Allows the use of specialised sensors/measures in a controlled environment.
  - Undisturbed environment.
- Cons
  - Limited ecological validity - we wont know if the results will generalise to the real world.
  - Requires access to a suitable space.
- Suitable for
  - Risky or dangerous studies.
  - Studies that require precise control over variables.

Activity: Experimental Design

Touch vs Physical Controls in Cars

You're a researcher at a car company and you're interested in understanding the impact of touch vs physical controls on driver performance. You decide to conduct an experiment to evaluate the impact of touch vs physical controls on driver performance.

Determine the following:

1. Research Question(s);
2. Hypothesis;
3. Independent Variable(s);
4. Dependent Variables;
5. Experimental Design (Within-Subjects or Between-Subjects);
    - What conditions will you have in your experiment?;
6. Experimental Environment (Field or Lab).

Measures

Qualitative Measures

- Non-numerical data capturing subjective experiences, opinions, and behaviors.
- Descriptive and expressed in terms of language rather than numerical values.
- Examples include:
  - User interviews (semi-structured or unstructured).
  - Observational field notes (ethnographic studies).
  - Think-aloud protocols (passive or active).

Pros

- Rich, detailed insights into user experiences.
- Captures context and meaning.
- Flexible data collection methods.
- Allows for exploration of unexpected findings.

Cons

- Subject to researcher bias and interpretation.
- Time-consuming to analyse and interpret.
- Harder to compare across studies.

Quantitative Measures

- Numerical data that can be statistically analyzed.
- Used to measure specific variables and establish relationships.
- Examples include:
  - Task completion time
  - Error rates
  - Likert scale ratings
  - Physiological measures (e.g. eye tracking, heart rate variability)

Pros

- Objective and reliable measurements.
- Easier to analyse statistically.
- Enables direct comparisons between groups.
- Can identify trends and patterns.

Cons

- May miss contextual factors and nuances.
- Limited depth of understanding.
- Requires careful operationalisation of variables.
- May not capture the full user experience.

Eye Tracking (Quantitative, Physiological)

- Eye tracking is a method used to measure eye movements, gaze fixation, and pupil dilation.
- Eye movements are used to infer cognitive processes.
    - e.g. attention, memory, problem-solving, etc.
- “What a person is looking at is assumed to indicate the thought “on top of the stack” of cognitive processes (Just & Carpenter, 1976)” - (Poole & Ball, 2006)

Figure 2: Tobil Pro Glasses 3 are non-intrusive and can be used in a variety of environments and scenarios - Image Credit

Heart Rate Variability (HRV) (Quantitative, Physiological)

- HRV measures the variation in the time interval between heartbeats.
- An indicator of the autonomic nervous system activity.
    - e.g. stress, relaxation, etc.
- High temporal resolution - can be used to detect changes in real-time, but prone to noise.
- Slightly invasive as some sensors are attached to the chest.
- Has been used in HCI to detect stress, cognitive load, and emotional states (Rowe et al., 1998).

Figure 3: The Polar H10 is a highly regarded chest strap that can be used to measure heart rate variability - Image Credit

Brain-based Measures (Quantitative, Physiological)

- A variety of brain-based measures can be used to infer cognitive processes.
  - e.g. EEG (Frey et al., 2016), fNIRS (Solovey et al., 2009), fMRI (Kohrs et al., 2016), etc.
- Each method has its own strengths and limitations.
  - e.g. EEG has high temporal resolution, but low spatial resolution. EEG is also very sensitive to noise.
  - fNIRS has low temporal resolution but better spatial resolution than EEG, and is less sensitive to noise.
- The choice of method depends on the research question and the context.

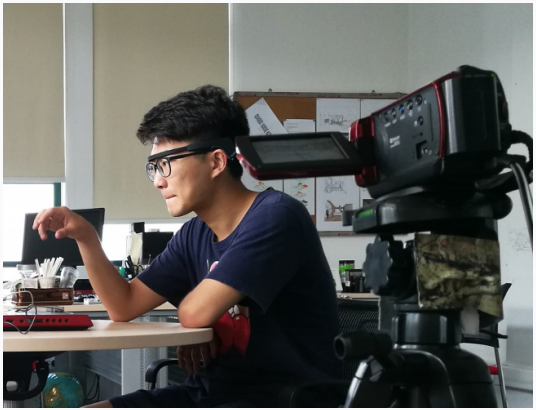

Figure 4: You'll remember the EEG cap from the first lecture - the Muse EEG

Combining Measures

Figure 5: Combining measures may provide additional insights into underlying physiological and cognitive processes Image Credit.

Web Articles

- Between-Subjects vs. Within-Subjects Study Design
  - https://www.nngroup.com/articles/between-within-subjects/
- How to Conduct Usability Studies for Accessibility
  - https://www.nngroup.com/reports/how-to-conduct-usability-studies-accessibility/
- How to Conduct Eyetracking Studies
  - https://www.nngroup.com/reports/how-to-conduct-eyetracking-studies

References i

Frey, J., Daniel, M., Castet, J., Hachet, M., & Lotte, F. (2016). Framework for electroencephalography-based evaluation of user experience. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2283–2294.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*(4), 441–480.

Kohrs, C., Angenstein, N., & Brechmann, A. (2016). Delays in human-computer interaction and their effects on brain activity. *PloS One*, *11*(1), e0146250.

Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. In *Encyclopedia of human computer interaction* (pp. 211–219). IGI global.

References ii

Rowe, D. W., Sibert, J., & Irwin, D. (1998). Heart rate variability: Indicator of user state as an aid to human-computer interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 480–487.

Solovey, E. T., Girouard, A., Chauncey, K., Hirshfield, L. M., Sassaroli, A., Zheng, F., Fantini, S., & Jacob, R. J. (2009). Using fNIRS brain sensing in realistic HCI settings: Experiments and guidelines. *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology*, 157–166.