Lecture 10: Essentials of Experimental Design for Interface Evaluation

Part 02

Research Design

Learning Objectives

To provide an introduction to the topic of research design.

To provide an overview of some of the key concepts used in research design.

To provide an overview of effective experimental design.

To provide an overview of ethics as applied to research.

Learning Outcomes

To be able to describe the key concepts used in research design.

To be able to describe the main types of research design.

To be able to describe the different types of measurement scales, the implementation of experimental and control groups, and randomisation.

To develop an appreciation of the possible confounds that may affect experimental outcomes.

To be able to describe the main components of ethical experimental design.

Below is a sample of the Materials and Methods section taken from a published article: Cassarino, M., Maisto, M., Esposito, Y., Guerrero, D., Chan, J.S. and Setti, A., 2019. Testing attention restoration in a virtual reality driving simulator. *Frontiers in psychology*, *10*, p.250.

Compare understanding pre-and post lecture:

Participants:

"In line with Berto (2005), a total of 38 participants (Mean age = 22.1, SD = 3.43; 44% female) were recruited through convenience sampling among undergraduate and graduate students at University College Cork, Ireland. Participants were randomly assigned to an urban or rural environmental exposure (n = 19 in each group). Half of the participants (n = 19) were fully licensed drivers with an average of 5.5 years of driving experience (SD = 3.24), whereas the other half (n = 19) included individuals with no full license and mean driving experience of 2.3 years (SD = 3.81). All participants read and signed a consent form prior to data collection in accordance with the Declaration of Helsinki. Ethical approval for the study was received by the School of Applied Psychology Ethics Committee, University College Cork. All participants read an information sheet briefing on the aims of the study and all were asked to read and sign a written consent form prior to participation in the study. No vulnerable populations were included in the study."

Below is a sample of the Materials and Methods section taken from a published article: Cassarino, M., Maisto, M., Esposito, Y., Guerrero, D., Chan, J.S. and Setti, A., 2019. Testing attention restoration in a virtual reality driving simulator. *Frontiers in psychology*, *10*, p.250.

Compare understanding pre-and post lecture:

Design:

"*A 2 × 2 mixed between-within design was employed, with the participants' performance at SART, (assessed pre- vs. post- exposure to virtual reality environments in a full vehicle driving simulator) as the within-subjects factor; and Environment type (urban vs. rural) as the between-subjects factor.*"

Overview Of Research Design

Designing an experiment to evaluate your interface, we will be considering the following:

Key concepts and definitions
Types of study
Questions and Hypotheses
Measurement scales
Variables
Groups and Controls
Randomisation
Confounds
Limitations
Ethics

Criteria for Scientific Evidence When Conducting a Study

Study Design: How the study is set up and conducted.

Empirical Research: Evidence learned by observation.

Objectivity: Researcher biases/attitudes/subjective impressions should not influence outcomes.

Systematic: Observations obtained in a methodical manner.

Control: Simplify situations by controlling for any extraneous factors that could influence outcomes.

Types of Research

Case study: Assessment/evaluation of single person's experience.

Questionnaires/Surveys: Dependent on questions asked and response scales used.
Interviews/focus groups: Structured/unstructured. Individual/group.

Controlled Experiments: Often involve a sample of individuals (representative of the population about whom we wish to make a statement).

Types of Research, Claims & Methods

Type of Research	Focus	General Claims	Typical Methods
Descriptive	Describe a situation or a set of events	X is happening	Observations, field studies, focus groups, interviews
Relational	Identify relations between multiple variables	X is related to Y	Observations, field studies, surveys
Experimental	Identify causes of a situation or a set of events	X is responsible for Y	Controlled experiments

10

Hypotheses

A controlled experiment <u>tests</u> a hypothesis: e.g., effects of a designed change on some measurable performance indicator.

A hypothesis should be:

Testable

Falsifiable

Parsimonius

Null Hypothesis

Start with an initial null hypothesis.

<mark>Null hypothesis typically states that there is no difference between experimental treatments.</mark>

E.g. Experiment to test if right-handed users are quicker to respond via a novel interface than left-handed users.
Start with the assumption that there is no difference between right- or left-handed users **(Null hypothesis H_0)**
Then carry out the experiment, analyse the results and decide if to accept the null hypothesis, or reject it and accept the **Alternative Hypothesis (H_1).**

Null Hypothesis and Alternative Hypothesis

Example:
Website developers decide to conduct an experiment to find out which menu design (pull-down or pop-up) will allow the users to navigate the site more effectively (measure time spent locating pages). For this research case, the null and alternative hypotheses can be stated in classical statistical terms as follows:

Null Hypothesis
H_0: There is no difference between the pull-down menu and the pop-up menu in the time spent locating pages.

Alternative Hypothesis
H_1: There is a difference between the pull-down menu and the pop-up menu in the time spent locating pages.

Note: Experiments can examine multiple pairs of null and alternative hypotheses.

Components of an Experiment

After a research hypothesis is identified, design of an experiment consists of three components:

Treatments/Conditions: Different techniques/devices/procedures that we want to compare
Units: Objects to which we apply the experiment treatments. In typical user-based studies this will be human participants.
Assignment Method: Way in which the experimental units (participants) are assigned different treatments.

Example

Experiment: Compare time taken to haptic device 1 and haptic device 2.

Measuring: Time

2 experimental conditions: Haptic device 1 and haptic device 2

20 participants.

20 participants randomly assigned to the two experimental conditions

Haptic device 1   Haptic device 2

10 participants   10 participants

Between-subject design with 2 conditions

Measurement Scales

4 main scales of measurement:

Nominal scale: Numbers refer to categorical labels.

Ordinal scale: Numbers refer to ranks.

Interval scale: Typical numerical scale.

Ratio scale: Typical numerical scale with an absolute zero point.

Variables

A well-defined hypothesis also clearly states the dependent and independent variables of the study.

Independent variable

Dependent variable

Variables

Independent variables: Factors that researchers are interested in studying or the possible "cause" of the change in the dependent variable.

Dependent variables: Outcome or effect that researchers are interested in.

Researcher wants to know: Whether and how, changes in independent variables induce changes in dependent variables.

Variables

Previous Null Hypothesis:

There is no difference between the pull-down menu and the pop-up menu in the time spent locating pages.

Independent variable = Type of menu (pull-down or pop-up).

Dependent variable = Time spent in locating web pages.

Questions (5 mins)

Read the following hypotheses and identify the dependent variables and independent variables in each case.

1. There is no difference in users' reading speed and retention rate when they view content on a desktop computer or a PDA.

2. There is no difference in the target selection speed and error rate between joystick, touch screen, and gesture recognition

3. There is no difference in the technology adoption rate between two speech-based applications with different dialog designs.

4. There is no difference in the reading skills of children who used educational software for 6 months compared to those who have never used the software.