

# Lecture 10: Essentials of Experimental Design for Interface Evaluation

## Part 01

# Overall Learning Objectives

To provide an introduction to the topic of experimental design and research.

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

To provide an overview of research data analysis and interpretation.

# Overall Learning Outcomes

To develop an appreciation of the steps involved in developing an effective experimental study to evaluate an interface.

To be able to describe the main concepts and considerations in research and design.

To be able to describe the main data analysis and interpretation methods.

# Relevance

You create a new interface, and you may want to:

Test how it performs in comparison to other interfaces.

Test how different groups of individuals use the interface.

Test how individuals interact with the interface over a long period of time.

# Sample Article

Below is a sample of the Participants 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.”*

# Sample Article

Below is a sample of the 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  $2 \times 2$  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.”*

# Sample Article

Below is a sample of the Results 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:

## Results:

*“Participants’ performance at the SART was analyzed in terms of d-prime ( $d'$ : a measure of signal detection sensitivity, calculated as the standardized difference (z-scores) between the proportion of correct responses on non-lures minus the proportion of incorrect responses on lures), overall mean accuracy (proportion of correct responses on lures and non-lures), mean accuracy on non-lures (pressing the bar), accuracy on lures (not pressing the bar when number three appears), reaction times (in milliseconds) of correct responses (related to pressing the bar in the presence of a non-lure), and inverse efficiency, a measure of speed-accuracy trade-off calculated as the ratio of reaction times over accuracy on non-lures (Bruyer and Brysbaert, 2011). Comparisons between the two exposure groups in terms of gender were conducted using Chi-square test and potential differences in age and driving experience were investigated via an independent samples t-test.”*

# Sample Article

Below is a sample of the Results 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:

## Results:

*“A  $2 \times 2$  mixed-design ANOVA was conducted with Environment (rural vs. urban) as the between-subjects factor, and SART (pre- vs. post-drive) as the within-subjects factor to investigate effects of environmental exposure on changes in attentional performance pre- and post-drive. Post hoc comparisons were conducted via t-test statistics. Comparisons between exposure groups in terms of driving behavior were assessed via independent t-test. In addition, potential effects of driving on attention were tested through a  $2$  (SART session)  $\times$   $2$  (environmental exposure)  $\times$   $2$  (driving vs. passenger condition) ANOVA with Driving (driver or passenger) and Environment (urban vs. rural) as the between-subject factors, and SART (pre- vs. post-drive) as the within-subjects factor. We conducted a test of normality on the ANOVA unstandardized residuals as well as the Levene’s test of homogeneity; for measures that did not appear to meet the assumptions of normality, we conducted the analyses using non-parametric tests and found no differences in results.”*



# Sample Article

Below is a sample of the Results 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:

## Results:

*“Environmental Exposure Effects on Attention: The two exposure groups ( $n = 19$  in each group) did not differ significantly in terms of gender ( $\chi^2_{1} = 0.11$ ,  $p = 0.74$ ), age ( $t_{36} = -0.42$ ,  $p = 0.67$ ) or driving experience ( $t_{36} = 0.16$ ,  $p = 0.87$ ).*

*The  $2 \times 2$  mixed-design ANOVA indicated no significant interaction between environmental exposure and SART pre- and post-drive for any of the measures of interest.*

*There was a main effect of environmental exposure for the measure of  $d'$  ( $F_{1,36} = 4.18$ ,  $p = 0.048$ ,  $\mu^2 = 0.11$ ), with participants in the rural exposure group ( $M = 1.26$ ,  $SD = 1.07$ ) showing overall higher sensitivity (i.e., better performance) than the urban exposure group ( $M = 0.62$ ,  $SD = 0.84$ ). There was also a main effect of environmental exposure for the measure of accuracy on lures ( $F_{1,36} = 4.61$ ,  $p = 0.04$ ,  $\mu^2 = 0.11$ ), with participants in the rural group ( $M = 0.64$ ,  $SD = 0.25$ ) being overall more accurate than those in the urban group ( $M = 0.48$ ,  $SD = 0.21$ ). In both cases, however, the size of the effect was small.”*

# Lecture 10 Structure

Part 01: Introduction and sample article text as a work-through example.

Part 02: Research Design

Part 03: Data Analysis and Interpretation

Part 04: Examples