

Experiments . . .

Please review quickly your readings and notes for the past several weeks. For a summary, see Wikipedia, here:

http://en.wikipedia.org/wiki/Usability_testing

In this review, experiments are typically “A/B” and a more nuanced variant - Nielsen’s brand of usability testing; that is, a formal comparison is made on certain dependent behaviours of the tested group (usually representative of the ultimate users in some meaningful way, i.e. a random sample).

So what should your assignment look like. We will take parts of a real study and shrink them to the scope of this assignment . . .



Figure 2. The Nokia 3310 and 5510

Premise and introduction: Different cellphone shapes predispose different usability advantages.

Independent variables: Phone shape and button layout. It’s clear that the left set is a single-hand unit, whilst the right-hand one is for two handed use. It’s larger as well. Screen size is equal

Controlled variables: Screen size, response time, button size/feel, electronic capabilities (i.e. screen graphics, typeface, audio quality), user tasks and environment (all subjects brought to the same test site). Individual differences partly controlled by random selection of test subjects (in reality, this is rarely done!)

Dependent variables: speed and errors of doing two tasks in two use conditions . . .

1. Placing a simple 20 character text message
 - (a) while walking
 - (b) while sitting at a desk
2. Playing a 2 user game involving on-screen evasion
 - (a) while walking
 - (b) while sitting at a desk

So we have four conditions used on two phones. You can probably guess the hypotheses that I predict ahead of actual testing:

H1: The single-hand unit is better for a brief text message while walking

H2: The two-handed unit is better for game interaction while seated at a desk.

But...

H3: The two-handed unit is worse for brief text messaging when seated at a desk.

And...

NULL for H1, H2, H3: The two handsets make *no difference* to task or situation of use (walking or seated).

Verification of hypotheses: Significant difference in average (mean) speed and error rates according to test conditions.

Validity: Speed and error rates are common, industry accepted quantitative usability measures. Question: what could I do to get at enjoyment or frustration with these handsets, going beyond the confines of this study? How would I discuss EXTERNAL validity?

Reliability: I randomly select a second sample of test users and run the experiment again. The results should correlate highly (be similar) between the two test runs for the four (4) test conditions.

Feasibility of study: This could be done at low cost with a small sample (n=30), alternate handsets with the same subjects. I would have to rotate across subjects whether they first did the test with the single-handed or two-handed set (Why?)

Utility: I can demonstrate design potential of two basic handset types and their crossover capabilities for different uses. Weighed against likely cell user need, we can use these data to set manufacturing and promotion priorities.