Post-lecture Questions IV.2 – Pre-test/Post-test Designs

Study Questions

Why can't you test the efficacy (effectiveness) of a treatment using a simple (one-group) within-subjects design?

What problem would make you avoid running a simple (one-measure) between-subjects design to test a treatment?

In general, why do you need a (no-treatment) control group when testing whether a treatment or therapy is effective?

What is the standard form of analysis that is applied to the data from a two-group, pre-test/post-test design? (By "standard" I mean what has been used most often in the past.) How is the output from the analysis interpreted?

What is the fancier form of analysis that should be applied to the data? How is the output from this analysis interpreted?

1. The main reason that standard experimental designs cannot be used to test the efficacy of a treatment or intervention is because _____ .
   (A) treatments and interventions rarely work
   (B) treatments and interventions always have some effect, but it often isn't helpful
   (C) you cannot counter-balance the order of the conditions
   (D) you are always forced to use a non-equivalent-groups design

Answers to Study Questions

A properly-conducted within-subject design involves the counter-balancing of the order of conditions. This can't be done when you are testing the efficacy of a treatment because treatments are supposed to have permanent effects, so you cannot ever run the "no-treatment" condition second – it must be before vs after for no-treatment vs treatment – so you can't counter-balance order. In other words, a simple (one-group) within-subjects design would have a perfect (and therefore awful) confound between time/order and with vs without treatment.

What's the biggest worry (from an internal validity point-of-view) when it comes to between-subject designs? Answer: the two groups might not have been the same to start with (i.e., a failure of random assignment). Given that many treatment-efficacy experiments are relatively small (in terms of the number of subjects) and people needing treatment are quite variable, a simple between-subjects design would be quite susceptible to this threat. A second problem with simple, two-group, post-test-only designs is that they don't tell you if one group got better or the other group got worse (or a bit of both).

A control group is needed to get an estimate of how much better people would get just due to the passage of time. Even if the treatment is completely ineffective, scores on various measures of interest (e.g., depression or anxiety) will often go down between the time of the pre-test and the time of the post-test due to various other things. The control group is used to get an estimate of this change, which can then be "subtracted out" using a form of fourth type of logic for dealing with confounds. (Note: the confound in this case is between *before* vs *after* treatment and *earlier* vs *later* in time.)

The standard form of analysis is to compare the change scores between the two groups. If the treatment group changes more (in the "good" direction) than the control group, then the therapy was effective.

The proper analysis uses the pre-scores as a covariate, since removes the problems caused by failures of random assignment via the third level of confound-control logic. You can then either compare the post-test scores between the two group or compare the change scores between the two groups; statistically, it makes no difference, so using the change scores is the best approach because it allows you to see whether the treatment group got better and/or whether the control group got worse. Either way, if the treatment group differs from the control group (in a "good" way), then the therapy was effective.

1. The main reason that standard experimental designs cannot be used to test the efficacy of a treatment or intervention is because (C) you cannot counter-balance the order of the conditions.