CPP 524: Program Evaluation II

Prof. Jesse Lecy

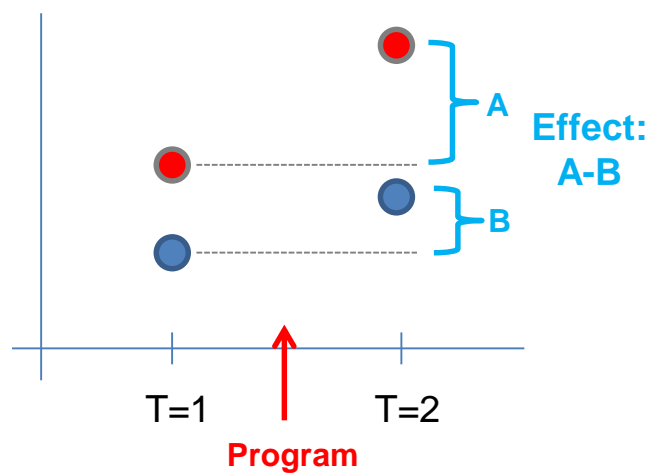**Campbell Score SOLUTIONS**

CH5: Pretest-Posttest Control Group Design

Improving Cognitive Ability in Chronically Deprived Children

Pre-Post With Control

Omitted Variable Bias

(1) **Selection**: +1

Group assignment was randomized which should control for unobservables and selection. But, candidates for the study had to meet specific criteria for the study. Candidates had to represent the most malnourished, underserved children in the population, and these criteria were based on the judgment of individual medical staff.

The groups that collective comprised the {T1 to T4} represent both the treatment and the control groups. So, for example, at time=1 the treatment group is T1 and the control group is {T1,T2,T4}. At time=2 the treatment group is {T1,T2} and the control group is {T3,T4}, etc. So to test whether randomization was "happy" we need to make sure these group do not differ significantly. The authors state in Table 2 on page 66 that "*There are no statistically significant differences among groups T1-T4.*"

Also, neither HS nor T0 are a true control group. They are used to represent upper and lower bounds of expected outcomes, and in this way to put the results (program effects) in context. Due to the nature of the design we need not be too concerned that there are statistical differences between the combined group {T1 to T4} and T0 since the individuals in T0 were already distinguished from the test subjects by researchers during the selection/assignment process.

(2) **Non-Random Attrition**: +1

Attrition seemed to be a risk to the study from the beginning. Out of a group of 333 at the beginning, there were 53 children that discontinued treatment before its end, mostly due to emigration. The study states that the individuals that discontinued did not differ on selection covariates to a statistically significant degree from the remaining participants (pg. 63), but statistics are not presented. The attrition is considered to be random since the measured covariates of those that left and those that stayed do not to differ significantly.

Trends in the Data

(3) **Maturation**: +1

Maturation in a study on child/youth development is inevitable. Use of the control group allowed for comparing treatment and maturation trends across the treatment and control groups. Figure 2 (pg.69) demonstrates clearly that the treatment was having an effect through the intervention gains of the experimental group(s) as compared to the maturation gains of the control group(s). One caution might be to consider that throughout study the experimental

control was constantly shrinking as control participants moved to the treatment group, and for the final observation period there was no true control group to compare the results.

(4) Secular Trends: +1

Much like maturation, the use of control group allows for the trends between secular trends and effect size to be differenced out. Caution as to whether secular trends have been controlled in this study might be justified for the same reasons as maturation.

(5) Seasonality: +1

No apparent seasonality in child development for preschool children.

(6) Testing: +1

Because the study tests for cognitive ability and tests are applied multiple times, we should be concerned about testing effects. It is possible that the sampling measures and the various observation points could have led to some sort of testing bias. In this particular case, both the control and the treatment groups receive the same tests and as a result if there is a gain they will be differenced out when the effects are calculated. We would worry most about testing if one group received more tests than another, or if we are using a pre-post design without a comparison group.

(7) Regression to the Mean: +1

The study may be susceptible to regression to the mean, because the subjects represent children in the population most at risk of hindered cognitive development due to malnutrition, and any intervention (effective or not) may appear to be effective. But since the treatment and the comparison groups both come from the same pool at the low end of the distribution, any regression to the mean experienced by one group should also be experienced by the other. As a result, when you calculate the effect score you will difference this out.

Study Calibration and Measurement

(8) Time-Frame of Study: +1

The time frame was specifically chosen based on child development literature and research. The intervention was meant to occur until the children were old enough for primary school. But, there were still significant effects measured at the end of the study.

(9) **Measurement Error**: +0

The researchers took exceptional precautions at ensuring inter-tester reliability (pg. 66). Testers were extensively trained, and they practiced recording data before every observation point.

The outcome (cognitive ability) is a latent construct, however, so there WILL be some level of measurement error in the study.

It is good to see that efforts were made to reduce it, but if we end up with a lack of statistical significance it may be because measurement error in the DV inflated standard errors, making our results less significant.

Note that measurement error in the DV will not bias the estimate, however. If measurement error is in the treatment variable then it causes attenuation bias and makes program impacts look smaller than they actually are.

(10) **Intervening Events**: +0

There was a fire that terminated the study early, before the start of primary school, and reduced the number of treatment days (pg. 64). It may have affected the treatment intensity, and thus the effect-size during that period. Specifically, the intervening event in this case ONLY affected the treatment and not the control group. As a result, in attempting to generalize the study the true program effects might be less than could be expected if there was no interruption.

Was it a big enough interruption to significantly bias our results? Probably not in this case (it was late in the study, schools adapted). But we cannot eliminate it as a possible cause for changes (or lack of changes) we observe in the data. Conservatively we can say that an intervening event exists.

Campbell Score = 8/10