Precautions with Within-Subjects Experimental Designs

Although within-subjects (repeated-measures) analyses have a major statistical power advantage over between-subjects analyses, there are some important precautions to be noted. For within-subjects experiments (and often, similarly to repeated measures and longitudinal studies with no within-subjects experimental manipulation), problems can arise that may be methodological confounds for hypothesized results—what Keppel and Wickens (2004) call *incidental factors*.

Practice effects

There may be systematic changes in measures because of multiple testing administrations. A simple example is giving the same ability test twice. Students will be more likely to get questions correct the second time. The consequence is improvement in the performance that is not due to increased ability.

For within-subjects experiments, use of a single-order of administration of the treatments, such as control group then treatment group, may indicate an improvement, in which the change may be a result of practice effects rather than the effect of the treatment.

The usual remedy is to counterbalance the order of the treatments, with half of the participants assigned to treatment first and half assigned to control first. With more than two levels, counterbalancing can be done by using all possible orderings or a Latin square, which balances sequencing but does not require all possible orders (see section 13.6 in Myers et al., 2010 for more detail). The counterbalancing approach on average should help balance out the order effect, but it may create differences for the two order groups.

Practice effects can be examined by conducting a mixed between- by within-subjects ANOVA using Order by Treatment to determine if differences between the experimental groups differ by order group.

Differential carryover effects

Differential carryover effects is a methodological artifact that occurs when counterbalancing does not balance out order effects. One particular ordering, either treatment then control or control then treatment, creates a reaction to the dependent variable for that ordering only. This will be shown in a significant interaction.

Keppel and Wickens provide a nice summary of the course of action when this occurs.

Table 17.3: Steps to test an effect after removing the influence of an incide[ntal] factor from the individual scores.

1. Organize the data by the levels of the nuisance factor P, and calculate the individual means $\bar{Y}_{P_k}$ and the grand mean $\bar{Y}_T$. Use these values to determine the effects $\bar{Y}_{P_k} - \bar{Y}_T$.
2. Subtract the appropriate incidental effect from each of the original scores to equate the levels of the nuisance factor. Verify that the P means are identical and that the means (or sums) for the other factors have not changed.
3. Reorganize the adjusted data by the treatment factor (or factors).
4. Calculate the sums of square for any effects to be tested and for their error terms using the standard formulas.
5. Adjust the degrees of freedom for each error term by subtracting the degrees of freedom for the effect from its original value.
6. Using the adjusted degrees of freedom, calculate means squares and F ratios as usual.

References and Recommended Readings

Keppel, G., & Wickens T. D. (2004). *Design and analysis: A researcher's handbook, 4th Edition*. Englewood Cliffs, NJ: Prentice Hall. Table 17.3.
Kirk, R.E. (2013). *Experimental design: Procedures for the behavioral sciences, fourth edition*. Sage.
Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(3), 785. https://doi.org/10.1037/0278-7393.21.3.785