Within-Subjects Design Pt2

Limitations of Counterbalancing

TABLE 9.2

Hypothetical Data Showing How Counterbalancing Distributes Order Effects Evenly between the Treatment Conditions

(a) Original Scores with No Order Effect		(b) Modified Scores with a 5-Point Order Effect		
Treatment I	Treatment II	Treatment I		Treatment II
20	27	20	—— order ——→	32 (27 + 5)
23	29	23	——————→	34 (29 + 5)
25	29	25	——————→	34 (29 + 5)
19	26	19	——————→	31 (26 + 5)
26	31	(26 + 5) 31	←—— order ——	31
17	22	(17 + 5) 22	←——————	22
14	20	(14 + 5) 19	←——————	20
16	24	(16 + 5) 21	←——————	24
Mean = 20	Mean = 26	Mean = 22.5		Mean = 28.5

Counterbalancing & Variance

- while counterbalancing distributes order effects evenly across conditions
  - counterbalancing can still distort group means
- counterbalancing adds order effects to only some participants in each treatment but not others
  - Book: adds extra points to participants in whichever treatment was counterbalanced
  - changes in scores then increase variance within treatments
  - difference between group means may be the same but the large amount of variance within groups decreases the chance of a significant finding

Assymmetrical Order Effects

- we assume that the order effects affecting our findings are symmetrical
  - both groups have the same amount of an order effect
- possible that order effects affect groups differently
- if one group happens to be more fatigued than the other group, the additional order effects experienced may have different effects on participant responses
  - counterbalancing cannot do much about this

Counterbalancing & the Number of Treatments

- to use counterbalancing to address order effects, you need to address every possible sequence of conditions/treatments
  - this is where we have **complete counterbalancing**
- if you have more than two conditions/treatments, you need to account for unique effects occurring in different sequences of the order of conditions/treatments
- two conditions, counterbalancing is easy
  - there are only two possible sequences of the order of your conditions
    - t1 –> t2; t2 –> t1

- more than two treatments, counterbalancing can get complex

$$n! = n(n-1)(n-2)(n-3)$$

- Four treatment groups, results in 24 groups

$$4! = 4(3)(2)(1)$$

ABCD	BACD	CABD	DABC	Note that the sequence ABCD indicates
ABDC	BADC	CADB	DACB	that treatment A is first, B is second, C is
ACBD	BCAD	CBAD	DBAC	third, and D is fourth.
ACDB	BCDA	CBDA	DBCA	
ADBC	BDAC	CDAB	DCAB	
ADCB	BDCA	CDBA	DCBA	

- to completely counterbalance the design, you must divide participants into these 24 equal-sized groups and assign them into one of these groups

- **partial counterbalancing** is a simpler option than complete counterbalancing
  - instead of every possible sequence, partial counterbalancing uses enough different sequences to ensure each condition occurs first in the sequence for one group, occurs second for another group, third for another group, and fourth for another group
    - ABCD, CADB, BDAC, DCBA
    - A shows up in all positions

- there is also **Latin square counterbalancing**
  - uses all four conditions using a 4 x 4 matrix to fill in with the letters A, B, C, and D
  - the matrix has each element which appears once in each column and row

A	B	C	D
D	A	B	C
C	D	A	B
B	C	D	A

- there is also **reverse counterbalancing**
  - where all participants receive all conditions in a specific order and the again in another order
    - two conditions = t1 –> t2 –> t2 –> t1

Advantages of Within-Subjects Designs

- less participants
- removes differences between groups that may appear in between-subjects designs
  - even with randomization, matching, etc.
- each participant serves as their own baseline
  - removes variation caused by individual differences
- in statistics, more within-subjects designs are generally more powerful

Disadvantages of Within-Subjects Designs

- participants go through a series of conditions/treatments
  - with differing times, there is a possibly for order effects or time-related factors
- **participant attrition,** or participants withdrawing from the study before the completion of the study
  - lose interest, forget about the study, pass away
  - best option is to over sample
    - get more participants than you think is necessary for your design knowing you'll lose participants

Choosing Within- or Between-Subjects Designs

- some research questions will require one design or the other making your decision easier
  - some questions can be tailored to either a within-subjects or a between-subjects design, or even a mixed design
- individual differences between your groups can be an issue in between-subjects designs
  - if large individual differences are anticipated, go with a within-subjects design
    - Ex: amount of time on a computer in college students
    - Ex: blood pressure measurement

- time-related factors & order effects
  - these don't affect between-subjects designs because you may only test participants once
    - participants won't have order effects because they only see one condition
- fewer participants are necessary in a within-subjects design because you're comparing treatments/conditions
  - if you can get a large enough sample, then a between-subjects design is still feasible

Matched-Subjects Designs

- a **matched-subjects design** uses separate groups for each condition; however, each individual in a group is matched with a corresponding individual in every other group
  - match participants by demographics or by variables of interest
  - Ex: stressed college students may be matched on GSR (galvanic skin response) level
- attempt to utilize the benefits of both within- and between-subjects designs
  - can match on more than one variable
  - try to make it appear that you are comparing the same people

Applications & Statistical Analysis of Within-Subjects Designs

Two-Treatment Designs

- paired-samples t-test
- repeated measures ANOVA
  - better option
  - can control for other potential confounding variables)
- both work for a continuous outcome, which includes
  - average construct values from scales
  - number of correct on an exam
  - percentage correct
- other tests are necessary when data are measured on an ordinal scale (rank ordered)

Multiple-Treatment Designs

- more than two conditions/treatments can only be done with a repeated measures ANOVA, mixed-effects/hierarchical/multi-level regression, or mediation analyses
- issues with attrition –> missing data
  - different types of missing data
  - **MCAR** (Missing completely at random)
    - can be handled with modern approaches to missing data
  - **MAR** (Missing at random)
    - maybe participants missed certain questions
    - can be handled with modern approaches to missing data
  - **MNAR** (Missing not at random)
    - participants deliberately decided not to answer specific questions
    - techniques can be used to fill in data but are complex
    - traditional techniques cannot be used