Name ________________________________

Laboratory for Repeated Measures Designs

• Modeling Individual Differences and Order (Position) Effects

People differ. We can take that to the bank. But why do they differ? Lots of reasons, but for now let's consider individual differences as due to a limited number of factors. Specifically, let's consider Gender, IQ, and Motivation as three factors that would influence performance in an experiment. In fact, let's model these differences as follows:

Gender	Male = +2	Female = +3
IQ	Low = -1	High = +4
Motivation	Low = -1	High = +1

Now, let's consider a group of participants in a simple repeated measures experiment with two levels. For the following participants, compute their "baseline" scores:

Participant	Characteristics	Score
A	Male, HiIQ, HiMotiv	
B	Male, LoIQ, LoMotiv	
C	Male, LoIQ, HiMotiv	
D	Male, HiIQ, HiMotiv	
E	Female, LoIQ, LoMotiv	
F	Female, HiIQ, LoMotiv	
G	Female, HiIQ, HiMotiv	
H	Female, HiIQ, HiMotiv	

Those are the scores that these participants would receive without any treatment, just baseline measures. Now, let's assume that our experiment will produce a practice effect. We could model such an effect in different ways, but let's assume that because of a lack of familiarity with the experimental setting, on the first treatment, participants will uniformly get -1. Then, on the second treatment, because they are more comfortable, participants will uniformly get +3.

[For simplicity's sake, we won't model random effects, but you could imagine randomly adding +2 to one or two participants.]

Name ________________________________

OK, we can now look at the scores of these participants in a study. First, let's assume no counterbalancing, so all participants first receive the Control condition and then receive the Experimental condition. When you consider the participants' scores (as above), but now modified by the practice effect, what would they look like?

Participant	Control	Experimental
A		
B		
C		
D		
E		
F		
G		
H		
Mean		
Variance		

Compute the mean and variance of each set of scores. Note that the two means would now differ, even though we've not introduced any treatment effect. And that's not good! Thus, we need to counterbalance the orders in which these participants receive the treatments. But first, what's true of the two variances?

With only two conditions, we would use complete counterbalancing. Thus, half of our participants would receive the Control condition first and the other half would receive the Experimental condition first. Let's assume that Participants A, B, E, and F receive the order Control -> Experimental (i.e., are unchanged from above) and Participants C, D, G, and H receive the order Experimental -> Control. Record their scores below and then compute the mean and variance.

Participant	Control	Experimental
A		
B		
C		
D		
E		
F		
G		
H		
Mean		
Variance		

What's now true of the two means and the two variances? Is that news uniformly good?

Name ________________________________

• Carry-over effects and Order effects

Whether we're dealing with order effects (as above) or carry-over effects, we need to be concerned and take protective action—counterbalancing.

• Complete Counterbalancing

Complete counterbalancing is ideal, but produces increasingly large numbers of orders as the number of conditions increases. With *k* conditions, there would be *k*! orders. Thus, with two conditions (A and B), there would be two orders (A -> B, B -> A). With three conditions (A, B, and C) there would be six orders:

Order #	Position 1	Position 2	Position 3
1	A	B	C
2	A	C	B
3	B	A	C
4	B	C	A
5	C	A	B
6	C	B	A

With four conditions (A, B, C, and D) there would be 24 orders. To ensure that you understand the implications of complete counterbalancing, complete the following (partial) list of orders:

Order #	Position 1	Position 2	Position 3	Position 4
1	A			
2	A			
3	A			
4	A			
5	A			
6	A			
7	B			
8	B			
9	B			
10	B			
11	B			
12	B			

To be sure that you understand the implications of complete counterbalancing, provide the number of orders generated for each of the number of conditions shown below:

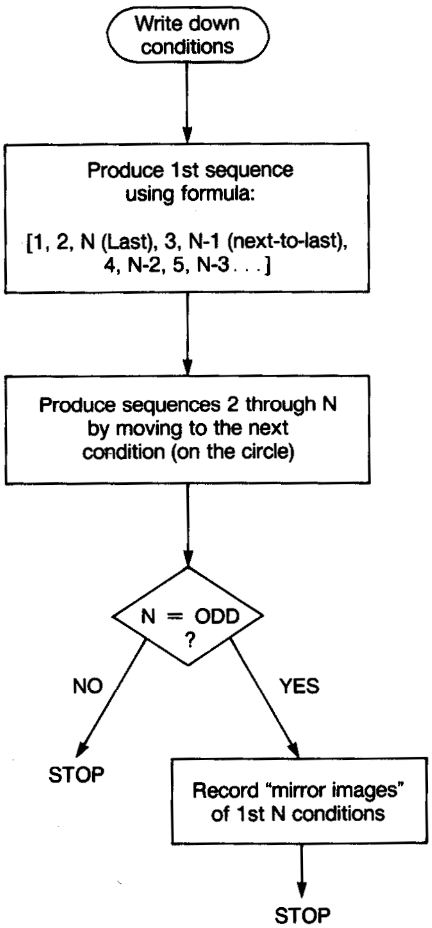
Number of Conditions	Number of Orders
5	
6	
7	

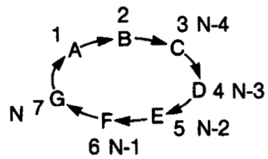
Name ______________________________

• Incomplete counterbalancing (digram balanced)

Obviously, as the number of conditions gets above five, the impact of complete counterbalancing is that the number of necessary orders gets prohibitively large for most researchers. Thus, there is a need for an alternative. Below, you see an example of how you could generate incomplete counterbalancing using the digram balanced approach.

Illustration of algorithm to produce an incomplete counterbalancing scheme for N conditions. The results next to the algorithm are for $N = 7$, with labels for the conditions: A B C D E F G.





1st	A	B	G	C	F	D	E
2nd	B	C	A	D	G	E	F
	C	D	B	E	A	F	G
	D	E	C	F	B	G	A
	E	F	D	G	C	A	B
	F	G	E	A	D	B	C
	G	A	F	B	E	C	D

E	D	F	C	G	B	A
F	E	G	D	A	C	B
G	F	A	E	B	D	C
A	G	B	F	C	E	D
B	A	C	G	D	F	E
C	B	D	A	E	G	F
D	C	E	B	F	A	G

Name ________________________________

To be sure that you understand the principle, generate incomplete counterbalancing orders for five (A, B, C, D, and E) and six (A, B, C, D, E, and F) conditions:

Order #	Position 1	Position 2	Position 3	Position 4	Position 5
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

Order #	Position 1	Position 2	Position 3	Position 4	Position 5	Position 6
1						
2						
3						
4						
5						
6						

To be sure that you understand the implications of number of conditions on number of orders needed with incomplete counterbalancing, complete the table below:

Number of Conditions	Number of Orders
7	
8	
9	

Name ________________________________

To be sure that you understand the implications of counterbalancing, answer the following questions:

What happens to order and carry-over effects as a result of counterbalancing?	
What impact does counterbalancing have on the likely error term (MS_{Error}) in the ANOVA?	

• Comparing Independent Groups and Repeated Measures ANOVAs

Dr. Alucard is interested in the auditory discrimination abilities of bats. First, he trains 54 bats to fly from a perch at one end of the experimental chamber to the food box at the other end of the room. He then constructs three different mazes of vertical wires (with wires that go from floor to ceiling with a small opening to allow the bat to move through one wall of wire to the next) between the starting perch and the food box. In one maze the wires are Thick, in one maze the wires are Intermediate, and in the final maze the wires are Thin. Each maze is different (openings are in different locations). The room in which the mazes are constructed is made completely dark before running each bat. The dependent variable he decides to use is the number of times a bat hits the wires of the maze. First, consider this experiment as an independent groups design.

Research Question: How sensitive are bats to stimuli using only auditory information?

Statistical Hypotheses: H_0: $\mu_{Thin} = \mu_{Intermediate} = \mu_{Thick}$
H_1: Not H_0

Decision Rule: If $F_{Obt} \geq F_{Crit}$, Reject H_0.

Data:

	Thin	**Intermediate**	**Thick**
	1	0	0
	5	3	1
	3	1	0
	4	4	1
	3	1	0
	5	2	0
	4	2	0
	3	2	1
	2	2	1
	3	2	1
	2	1	0
	1	1	0
	3	1	0
	5	2	0
	3	2	0
	3	1	0
	2	0	0
	3	2	0
M	3.06	1.61	0.28
T (ΣX)	55	29	5
ΣX^2	193	63	5
SS	24.94	16.28	3.44

Name ________________________________

Source Table:

Source	***SS* Formula**	***SS***	***df***	***MS***	***F***
Treatment	$\sum \frac{T^2}{n} - \frac{G^2}{N}$				
Error	ΣSS per group				
Total	$\sum X^2 - \frac{G^2}{N}$				

F_{Max}:

Decision:

Post Hoc Test:

Interpretation:

Effect Size:

How do the results of the two analyses compare?

Name ________________________________

Now, lets assume that we have the exact same data, but from a repeated measures design. The first change to note is that the exact same data would come from 18 bats, not 54. That's one indication of the efficiency of the repeated measures design.

Research Question: How sensitive are bats to stimuli using only auditory information?

Statistical Hypotheses: H_0: $\mu_{Thin} = \mu_{Intermediate} = \mu_{Thick}$
H_1: Not H_0

Decision Rule: If $F_{Obt} \geq F_{Crit}$, Reject H_0.

Data:

Bat	Thin	Intermediate	Thick	P
1	1	0	0	1
2	5	3	1	9
3	3	1	0	4
4	4	4	1	9
5	3	1	0	4
6	5	2	0	7
7	4	2	0	6
8	3	2	1	6
9	2	2	1	4
10	3	2	1	6
11	2	1	0	3
12	1	1	0	2
13	3	1	0	4
14	5	2	0	7
15	3	2	0	5
16	3	1	0	4
17	2	0	0	2
18	3	2	0	5
M	3.06	1.61	0.28	
T (ΣX)	55	29	5	
ΣX^2	193	63	5	
SS	24.94	16.28	3.44	

Name ________________________________

Source Table:

Source	***SS* Formula**	***SS***	***df***	***MS***	***F***
Treatment	$\sum \frac{T^2}{n} - \frac{G^2}{N}$				
Within Groups	ΣSS per group				
Subject	$\sum \frac{P^2}{k} - \frac{G^2}{N}$				
Error	SS_{Within} - $SS_{Subject}$				
Total	$\sum X^2 - \frac{G^2}{N}$				

Decision:

Post Hoc Test:

Interpretation:

Effect Size: