Exam 2 PS 217, Spring 2003

1. Below is a set of data, followed by a repeated measures ANOVA on the data.

a_1	a_2	a_3
3	6	9
4	5	5
6	7	8
2	5	9
5	7	10
4	5	8

ANOVA Table for A

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Subject	5	15.333	3.067				
Category for A	2	52.333	26.167	18.256	.0005	36.512	.998
Category for A * Subject	10	14.333	1.433				

Means Table for A
Effect: Category for A

	Count	Mean	Std. Dev.	Std. Err.
a1	6	4.000	1.414	.577
a2	6	5.833	.983	.401
a3	6	8.167	1.722	.703

Suppose that these same data had been analyzed using an independent groups ANOVA, instead of the repeated measures ANOVA. What would the source table have looked like, had the analysis been conducted with an independent groups ANOVA? [5 pts]

SOURCE	**SS**	**df**	**MS**	**F**
Between	**52.33**	**2**	**26.167**	**13.22**
Within	**29.67**	**15**	**1.98**	
Total	**82**	**17**		

Note that the *Within* term is a combination of the *Subject* term and the *Error* term (*Category for A * Subject*) — for both SS and df.

2. People with agoraphobia are so filled with anxiety about being in public places that they seldom leave their homes. Knowing this is a difficult disorder to treat, a researcher tries a long-term treatment. A sample of individuals suffering agoraphobia report how often they have ventured out of the house in the past month. Then they receive relaxation training and are introduced to trips away from the house at gradually increasing durations. After two months of treatment, participants report the number of trips out of the house they made in the past 30 days. The data are seen below. Analyze the data to determine the effectiveness of the treatment. [15 pts]

Participant	Before Treatment	After Treatment	P
A	0	4	4
B	0	0	0
C	3	14	17
D	3	23	26
E	2	9	11
F	0	8	8
G	0	6	6
ΣX (T)	8	64	72
ΣX^2	22	922	944
SS	12.86	336.86	349.72
$\bar{X}$	1.14	9.14	

The first step is to recognize the design as a repeated measures design. The computations would then follow.

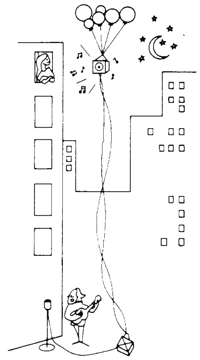
$H_0: \mu_B = \mu_A$ $H_1: \mu_B \neq \mu_A$

SOURCE	SS	df	MS	F
Between	$\frac{4160}{7} - \frac{5184}{14} = 594.3 - 370.3 = 224$	**1**	**224**	**11.3**
Within	**12.86 + 336.86 = 349.72**	**12**		
Subject	$\frac{1202}{2} - \frac{5184}{14} = 601 - 370.3 = 230.7$	**6**		
Error	**349.72 – 230.7 = 119**	**6**	**19.8**	
Total	$944 - \frac{5184}{14} = 573.7$	**13**		

$F_{Crit}(1,6) = 5.99$
Decision: Reject H_0, $F_{Obtained} > F_{Critical}$

Because there are only two groups, a post hoc test is not needed. Thus, we can conclude that the number of trips out of the house was significantly greater after treatment than before treatment.

3. There were five conditions in the Bransford & Johnson (1972) memory study (as compared to the two conditions we used in lab). These conditions are: Complete-context picture before passage; Complete-context picture after passage; Partial-context picture before passage; Hearing passage once without picture; and Hearing passage twice without picture. Below is the Complete-context picture (as a recall cue for you).



Suppose that the results for this study had turned out as seen below. Complete the source table and then analyze the study as completely as you can. [15 pts]

ANOVA Table for Recall

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Condition	4	301.387	75.347	113.344	<.0001	453.375	1.000
Residual	70	46.533	.665				

Means Table for Recall
Effect: Condition

	Count	Mean	Std. Dev.	Std. Err.
Complete-after	15	2.933	.704	.182
Complete-before	15	7.667	1.113	.287
No pic - Hear twice	15	2.400	.737	.190
No pic-Hear once	15	2.267	.704	.182
Partial-before	15	4.133	.743	.192

F_{Max} = 1.24/.50 = 2.48 $F_{MaxCrit}$ = 4.37 No concern about heterogeneity of variance.
Use α = .05, so a P-Value of .05 or less would lead you to reject H_0. In this case, with a P-Value < .0001 the decision would be to reject H_0. [You should recognize that an F-Value of 113 is huge!]

The next step would be to conduct a post hoc test: $HSD = 3.97\sqrt{\frac{.66}{15}} = .83$

	CA	CB	NP2	NP1	PB
CA	-				
CB	4.7	-			
NP2	.5	5.3	-		
NP1	.63	5.4	.13	-	
PB	1.2	3.6	1.7	1.9	-

Thus, people in the *Complete-Before* condition recalled significantly more than people in all the other conditions. People in the *Partial-Before* condition recalled significantly more than people in all the remaining conditions (*Complete-After*, *No-Context-Hear-Once*, *No-Context-Hear-Twice*). No other comparisons were significant.

4. Generally speaking, what does a post hoc test (such as Tukey's HSD) do for you? You should know, of course, that it allows you to compare two means to determine if they are drawn from populations with different means. But you could test such a hypothesis with a t-test or an ANOVA. So, what does a post hoc test do for you *beyond* allowing you to test for the difference between two conditions? [5 pts]

A post hoc test (such as Tukey's HSD) is intended to allow you to compare two conditions, which is quite similar in purpose to a t-test. The difference is that a post hoc test protects against an inflated chance of making a Type I Error that arises because you are testing the same set of data over and over. That is, what's called the family-wise error (the chance that you might be making a Type I Error on any of the comparisons) is greater than you'd like, unless you adopt a more conservative test. For the most part, that's what a post hoc test is doing. Tests like Tukey's HSD are simply protected t-tests (protecting against making a Type I Error).

5. Hypothesis testing is essential to the research enterprise in psychology. Briefly define Type I Error, Type II Error, and power. Then, tell me as many ways as you can to increase the power of a study. [5 pts]

Type I Error arises when you choose to reject H_0 and H_0 is true. Type II Error arises when you choose to retain H_0 and H_0 is false. Power arises when you reject H_0 and H_0 is false. In order to increase power you could always increase sample size (n). You could also use a greater treatment (such as increasing the amount of drug administered, the level of stress given to participants, etc.). Finally, anything you could do to decrease variability in your participants (smaller MS_{Error}) would increase power. You could run more homogeneous groups (all male, all with IQ > 110, etc.), with the caveat that you would not be able to generalize to a wider population (which may not be that big a deal). You could also decrease random variability by providing clearer instructions, reducing distractions, etc. Note, also, that moving from an independent groups design to a repeated measures design would give you more power.

6. Briefly explain why it is essential to develop an error term (MS_{Error}) for the repeated measures design that includes only variability due to random factors? [5 pts]

In a repeated measures design, the variability among the condition means is attributable to any effects of treatment as well as random variability. However, there is no contribution of individual differences. As a result, using an error term that includes individual differences would be inappropriate. It is for that reason that we need to develop an error term (MS_{Error}) for which the only source of variability is random effects.