1. You are developing a dissertation project on phenotypic plasticity in body size at maturity in the sexually dimorphic livebearing fish, *Limia perugia*. To get some pilot data, you enlist the help of a Bio 395 student, who you instruct to take a brood of neonates, and put each fish in its own aquarium. Half of the fish are assigned Small aquaria (2.5 gallons), and half are assigned to Large aquaria (10 gallons), though some die before reaching sexual maturity. The baby fish reach sexual maturity in 6 months, whereupon the student presents you with the following data set.

Sex$	TankSize$	LogMass
Male	L	2.204972
Male	S	2.191654
Female	L	2.386926
Female	S	2.327278
Female	S	2.979095
Male	L	1.931521
Female	S	2.61374
Female	L	2.821379
Male	S	1.272566
Male	L	1.978239
Male	S	1.759581
Female	L	3.272606
Male	S	1.229641
Female	S	2.381396
Male	S	1.940179
Female	S	2.255493

What analysis would you perform on these data and why? What are the results of your analysis? What are the effect sizes and partial effect sizes for Sex and Tank Size? What is the power of your test for each effect? What sample size would you need to achieve a power of 0.80?

Simple two way ANOVA, because the data are cross classified, and because each fish is in its own container, so the errors should be iid with no nesting (or blocking). Some of you were a little off on your $\eta^2{}_P = (SS_{Effect}/(SS_{Effect}+SS_{Error})$. The results of the analysis are…

Analysis of Variance						Power Analysis			
Source	Type III SS	df	Mean Squares	F-ratio	p-value	Eta Squared	Partial Eta Squared	Power*	N.80**
SEX$	2.465	1	2.465	19.914	0.001	0.563	0.624	.989	12
TANKSIZE$	0.427	1	0.427	3.452	0.088	0.098	0.223	.427	36
TANKSIZE$*SEX$	0.002	1	0.002	0.015	0.906	0.000	0.001		
Error	1.485	12	0.124						
Total	4.379								

* Systat's power analysis for 2-way ANOVAs assumes balanced designs, so this power analysis assumes cell Ns of 4, and the resulting power calculation is an approximation.

2. You are studying dominance behavior between large versus small female mosquitofish, which chase each other as they compete for food. Your design is to place one large female and one small female in each of 20 tanks, and conduct 10-minute focal female observations on each female in each tank. You generate the following dataset (next page, chase per minute data were log transformed).

Analyze the data two ways: 1. as a one-way ANOVA with Size as your treatment variable; and, 2. As a two-way ANOVA with Size as a fixed effect and Tank as a random effect. What effect does including Tank in your model have on your test of your fixed effect, and on the effect size of you r fixed effect? What are the differences in power between these two analyses? Why do you think this is the case?

Analyzing the data as a 1-way ANOVA is equivalent to analyzing it as a 2-sample t-test. Analyzing the data as a 2-way randomized block is equivalent to analyzing it as a paired t-test. Including tank greatly reduces the error SS, increases the partial effect size and power of the test, and reduces the p value for the test of the fixed effect. To achieve 0.80 power with an α of 0.05, one would need 11 pairs of fish if one analyzes this without Tank in the model, but only 7 pairs if one does include Tank in the model.

Analysis of Variance -- Two Sample t-test						Power Analysis			
Source	Type III SS	df	Mean Squares	F-ratio	p-value	Eta Squared	Partial Eta Squared	Power	N.80
SIZE$	1.139	1	1.139	14.480	0.001	0.276	0.276	0.962	22
Error	2.989	38	0.079						
Total	4.128								

Analysis of Variance -- Paired t-test						Power Analysis			
Source	Type III SS	df	Mean Squares	F-ratio	p-value	Eta Squared	Partial Eta Squared	Power	N.80
SIZE$	1.139	1	1.139	35.789	0.000	0.276	0.653	1.000	14
TANK$	2.384	19	0.125	3.943	0.002				
Error	0.605	19	0.032						
Total	4.128								

Tank$	Size$	Chase
A	l	0.43
A	s	0.22
B	l	0.87
B	s	0.48
C	l	0.87
C	s	-0.18
D	l	0.99
D	s	0.52
E	l	0.22
E	s	0.00
F	l	0.85
F	s	0.37
G	l	0.30
G	s	0.12
H	l	0.70
H	s	0.37
I	l	0.30
I	s	0.00
J	l	0.87
J	s	0.52
K	l	0.30
K	s	0.12
L	l	0.87
L	s	0.12
M	l	0.48
M	s	0.60
N	l	0.85
N	s	0.37
O	l	0.43
O	s	0.30
P	l	0.73
P	s	0.48
Q	l	0.12
Q	s	0.12
R	l	1.07
R	s	0.56
S	l	0.12
S	s	-0.18
T	l	0.92
T	s	0.56

3. Compare and contrast the 3 basic models for 2-way ANOVA, Model I, Model II, and Mixed Model, in terms of their assumptions, expected mean squares (and what is tested over what), and how their results are interpreted.

- Model I inferences only apply to the treatment levels of the experiment at hand, and repeats of those treatment levels
- Model II inferences apply to the population at large

Two-Way ANOVA with Replication: Model I

- $Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$

- **Source**   **Expected MSs**

  - Columns (A) $\sigma^2 + (nb/(a-1))\, \Sigma\alpha^2$
  - Rows (B) $\sigma^2 + (na/(b-1))\, \Sigma\beta^2$
  - Interaction (AxB) $\sigma^2 + (n/(a-1)(b-1))\, \Sigma(\alpha\beta)^2$
  - Error σ^2

- Main Effects and the Interaction are tested over the Error

Two-Way ANOVA with Replication: Model II

- $Y_{ijk} = \mu + A_i + B_j + (AB)_{ij} + \varepsilon_{ijk}$

- **Source**   **Expected MSs**

  - Columns (A) $\sigma^2 + n\sigma^2_{AB} + nb\sigma^2_{A}$
  - Rows (B) $\sigma^2 + n\sigma^2_{AB} + na\sigma^2_{B}$
  - Interaction (AxB) $\sigma^2 + n\sigma^2_{AB}$
  - Error σ^2

- The Interaction is tested over the Error; the Main Effects are tested over the Interaction

Two-Way ANOVA with Replication: Mixed Model (A Fixed; B Random)

- $Y_{ijk} = \mu + \alpha_i + B_j + (\alpha B)_{ij} + \varepsilon_{ijk}$

- **Source:**   **Expected MSs**
  - Columns (A) $\sigma^2 + n\sigma^2_{AB} + (nb/(a-1))\, \Sigma\alpha^2$
  - Rows (B) $\sigma^2 + na\sigma^2_{B}$
  - Interaction (AxB) $\sigma^2 + n\sigma^2_{AB}$
  - Error σ^2

- The Random Effect and Interaction are tested over the Error; the Fixed Effect is tested over the Interaction

4. What are the relative advantages and disadvantages of Nested versus Randomized Block designs? When might you use each one in your own research?

Why Design an Experiment with Nesting, Blocking or Both?

- **Nesting:** each level in the hierarchical design is tested over the level below it, and becomes the error term for the level above it.
  - Advantages: Partitioning variance among levels of a hierarchy.
    - Model II: Useful for studying sources of variability among the nested levels (e.g. spatial or temporal scales, among taxonomic levels within a phylogeny)
  - Disadvantages: Loss of power for tests of fixed effects at top level.
    - Mixed Model – Top Level is a Fixed Effect: Individual subjects generally are not the unit of replication for the level that most interests us. In the example we explored, Pools within Treatment Levels was our unit of replication, and not Subjects within Treatment Levels. Thus, we lost power for our tests of fixed effects at the top level over what we would have obtained with the same number of subjects in a completely randomized design.

Why Design an Experiment with Nesting, Blocking or Both?

- **Blocking:**
  - Randomized Block designs have benefits and costs
    - Benefits: reduces Error SS
    - Costs: reduces Error df
  - Blocking is likely to be helpful in increasing the power of tests of fixed effects if you know or have reason to suspect that variation among blocks is greater than variation within blocks
  - Repeated measures designs can be analyzed as randomized blocks, where subjects are the blocks and each subject is measured under each treatment level or combination
  - Blocking usually results in higher partial η^2 for fixed effects

5. You've got 8 growth chambers, and you want to test the effects of photoperiod (short versus long) and fertilizer (control soil versus fertilized soil) on growth in *Arabidopsis*. Each growth chamber has room for 10 plants. What would be your experimental design? Lay it out as an ANOVA table, and identify your sources of variation, degrees of freedom, expected mean squares, and F tests (what's tested over what). Assuming partial eta-squareds of 0.10, do you think you have sufficient Ns to achieve 0.80 power?

8 chambers times 10 plants is N=80, or total df=79; however, the 8 containers are the real units of replication, which gives a total df =7 from the point of view of the containers. The constraints of the chambers, chamber settings effect all subjects within chambers, versus the ability to vary soils properties within and among chambers, suggests a 2-way , split-plot ANOVA with photoperiod as the among-plot treatment, and fertilizer as the within-plot treatment. I'd assign 4 chambers to each photoperiod; and half the plants within each chamber (5 plants) to each fertilizer treatment. If a=2 photoperiods, p=4 plots per photoperiod, b=2 levels of fertilizer, and r=5 replicate subjects per fertilizer treatment, then total df=a*p*b*r-1=79. The ANOVA Table would be as follows…

Source	df	df	Expected MSs	F Test: MS/MS ; df/df
A: Photoperiod	a-1	1	σ2 + bσ2P + (rb/(a-1)) Σα2	A / Error A ; 1/6
Error A: Plot(A); Chamber(Photoperiod)	a(p-1)	6	σ2 + bσ2P	
B: Fertilizer	b-1	1	σ2+ rσ2pb + (ra/(b-1)) Σβ2	B / Error B ; 1/6
A*B: Photoperiod*Fertilizer	(a-1)(b-1)	1	σ2+ rσ2pb + (r/(a-1)(b-1)) Σ(αβ)2	A*B / Error B ; 1/6
Error B: B*Plot(A); Fertilizer*Chamber(Photoperiod)	a(b-1)(p-1)	6	σ2 + rσ2pb	
Error: Residual (due to replication within B)	apb(r-1)	64	σ2	
Total	apbr-1	79		

See Quinn & Keough (2002), Table 11.2, for an ANOVA table for this design. Regarding Power Analysis, I would approximate it either as a Two-Way ANOVA or a Generic F test in G*Power 3 or Systat. Both give similar and disappointing answers. Given that F tests for the main effects and interaction above all have 1 over 6 degrees of freedom, the sample size that matters is the number of growth chambers. If you think of it as two separate one way ANOVAs, each with two groups, the test for photoperiod is basically a two groups t test with 4 chambers in each group, and the square of the t is F with 1,6 df. The test for fertilizer is a matched pairs t test, or a 1-way randomized block with 2 fertilizers versus 8 chambers, again with 1,6 df. So chamber number really limits things. If you approximate this as a two way ANOVA for the power analysis, you'll need roughly 19 replicates per cell in the 2x2 ANOVA table, or a total of 76 growth chambers, to be able to detect a partial eta-squared of 0.1 at the .05 level with .80 power (ugh!). Hopefully, using 10 plants per chamber versus 2 gives better estimates of ANOVA Table cell means, and higher effect sizes.

Alternatively, this could have been designed as a two-way nested ANOVA, with chambers nested within cells of the ANOVA table and two chambers per ANOVA cell, i.e. per treatment combination. In this case, the number of chambers per cell is two, p=2; and, the number of plants per treatment combination within each chamber is 10, r=10. Such an ANOVA table would look like…

Source	df	df	Expected MSs	F Test: MS/MS ; df/df
A: Photoperiod	a-1	1	σ2 + rσ2C(AB) + (rb/(a-1)) Σα2	A / Chamber(A*B) ; 1/4
B: Fertilizer	b-1	1	σ2 + rσ2C(AB) + (ra/(b-1)) Σβ2	B / Chamber(A*B) ; 1/4
A*B	(a-1)(b-1)	1	σ2 + rσ2C(AB) + (r/(a-1)(b-1)) Σ(αβ)2	A*B/Chamber(A*B) ; 1/4
Chamber(A*B)	ab(p-1)	4	σ2 + rσ2C(AB)	
Residual	apb(r-1)	72	σ2	
Total	apbr-1	79		

Note that the denominator df in this design is 4 (versus 6 in the split-plot design, above), which suggests that this design would be less powerful than the split-plot (0.114 versus 0.233).

6. In the Nested ANOVA we covered in class…

Male Harassment = Male Density + Pools(Male Density) + Error

…what is the difference between testing Males over Pools(Males) versus doing an ANOVA on wading pool means? (Dataset attached)

 The two analyses give essentially the same results (given rounding error and the slight imbalance in the nested design; see below). Doing a Nested ANOVA and testing the fixed effect over the container(treatment) term is equivalent to doing a non-nested ANOVA on container means. The eta-squareds are 0.496 for the nested ANOVA versus 0.499 for the ANOVA based on wading pool means.

Analysis of Variance -- Nested					
Source	Type III SS	df	Mean Squares	F-ratio	p-value
MALES$	6.935	1	6.935	5.907	0.051
POOL$(MALES$)	7.042	6	1.174		
Error	48.859	53	0.922		
Total	62.836				

Analysis of Variance -- Wading Pool Means					
Source	Type III SS	df	Mean Squares	F-ratio	p-value
MALES$	0.911	1	0.911	5.980	0.050
Error	0.914	6	0.152		
Total	1.825				