kxk BG Factorial Designs

- expanding the 2x2 design
- F & LSD for orthogonal factorial designs
  - using the right tests for the right effects
- F & follow-up analyses for non-orthogonal designs

kxk BG Factorial Designs

We've worked extensively with the 2x2 design -- the basic factorial

Larger factorial designs are often used for the same reasons that multiple-condition 1-factor designs are used . . .

- You may need more than 2 IV conditions to properly test a RH:
  - Want multiple "experimental conditions" (qual or quant diffs)
  - Want multiple "treatment conditions " (standard vs. none, etc)
  - Want to "dissect" a multiple element treatment

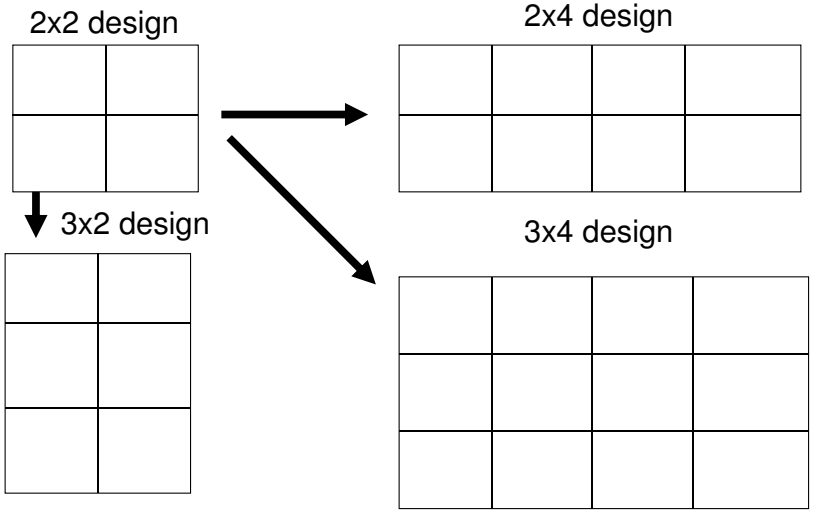
You might want to test the generalizability of an IV's effect…

- across more than 2 populations
- across more than 2 setings
- across more than 2 task-stim
- across more than 2 "eras"

Basic and Expanded Factorial Designs

The simplest factorial design is a 2x2, which can be expanded in two ways:

1) Adding conditions to one, the other, or both IVs



Adding Treatment Conditions to an IV

- Ways treatment conditions differ
  - amount of treatment
    - receiving therapy once vs. twice each week
    - getting 0, 1, 5 or 10 practice trials before testing
  - kind of treatment
    - receiving Cognitive vs. Gestalt clinical therapy
    - whether or not there is feedback on practice trials
  - combinations of treatment components
    - receiving both "talk" therapy vs. "combined drug & talk" therapy
    - receiving "10 practices without feedback" vs. "2 practices with feedback"

The "Secret" is to be sure the selection of conditions matches the research hypotheses you started with !!!

Adding Control Conditions to an IV

- "No Treatment" control
  - Asks if the Tx works "better than nothing"
- "Standard Tx" control
  - Asks if the Tx works "better than usual"
- "Best Practice" Control
  - Asks if the Tx works "better than the best known"
- "Pseudo Tx" Control
  - Asks if TX works "without a specific component"

The "Secret" is to be sure the selection of conditions matches the research hypotheses you started with !!!

Designs, RH & results are communicated by "drawing the boxes." Be sure to label each IV and specify each or its conditions.

1. Novices and Experts completed the task, either under instructions to work quickly, work accurately, to work as quickly as possible without making unnecessary errors or no instructions.

2. Folks completed a depression questionnaire either under instructions to "respond like someone with acute depression," "respond like someone with chronic depression" or "respond like someone who is trying to 'fake' being depressed". Participants were either clinical psychologists, clinical psychology grad students or volunteers from a local social club.

#1 was a 2x4 design that looks like this...

Prior Skill	Instructions: Quick	Accurate	Both	None
Novice				
Expert				

#2 was a 3 x 3 design that looks like ...

Respond like a ..	Participant: Clinician	Clin. Grad.	Soc. Club
Acute depressive			
Chronic depressive			
"Fake" depressive			

Statistical Analysis of Orthogonal kxk Designs

Only a couple of differences from the 2x2

1. Tell IVs and DV 2. Present data in table or figure
3. Determine if the interaction is significant
   - if it is, describe it in terms of one of the sets of simple effects using pairwise comparisons (PrC) to compare cell means
4. Determine whether or not the first main effect is significant
   - if so, describe it using PrC to compare 3+ marginal means
   - determine if that main effect is descriptive or misleading using the interaction PrC to compare SEs
5. Determine whether or not the second main effect is significant
   - if so, describe it using PrC to compare 3+ marginal means
   - determine if that main effect is descriptive or misleading using the interaction PrC to compare SEs

When do you need pairwise comparisons (PrC) ???

General rule: You will need PrC to compare pairs of means whenever a significant effect has k>2 conditions

#1 Whenever the interaction is significant → PrC for the cell means is needed to:

- describe the pattern of the SEs to describe the interaction pattern
- describe the pattern of the SEs to determine if corresponding ME is descriptive or misleading (necessary to do for each ME -- whether the ME is significant or not)

#2 Whenever a 3+grp ME is significant → PrC for those marginal means is needed to:

- describe the pattern of that ME

Kxk PrC follow-ups are a little different than the 2x2

- the 2x2 uses the PrC only for comparing cell means
  - describe the simple effects to explicate the interaction pattern
  - not needed for MEs , since they involve only 2 conditions
- the kxk uses the PrC for comparing cell & marginal means
- different Prcs are computed for different effects
  - if the interaction is significant, then PrC is computed to compare the cell means -- describe SEs, interaction, etc.
  - If a ME with 2 conditions is significant - no PrC needed
  - If a ME with 3 or more conditions is significant, then PrC is computed to compare the marginal means of that ME

Be sure to use the proper "n" to compute each PrC LSDmmd

- "n" = mean number of data points used to compute the means being compared (more on demo sheet)

What statistic is used for which factorial effects????

There will be 5 statistics

1. F_{TExp}
2. F_{Age}
3. Age LSD_{mmd}
4. F_{Int}
5. PrC LSDmmd

Task Experience

Age	Exp	~Exp	
5	30	30	30
10	20	30	25
15	25	30	27.5
	25	30	27.5

15 "Effects" in this study

1. Main effect of TExp
2. Main effect of Age
3. 5 vs. 10 marginals
4. 5 vs. 15 marginals
5. 10 vs. 15 marginals
6. Interaction of Age & TExp
7. 5 vs. 10 yr old for Exp
8. 5 vs. 15 yr old for Exp
9. 10 vs. 15 yr old for Exp
10. 5 vs. 10 yr old for ~Exp
11. 5 vs. 15 yr old for ~Exp
12. 10 vs. 15 yr old for ~Exp
13. Exp vs. ~Exp for 5 yr olds
14. Exp vs. ~Exp for 10 yr olds
15. Exp vs. ~Exp for 15 yr olds

Back to → 100 males and 100 females completed the task, either under instructions to work quickly, work accurately, to work as quickly as possible without making unnecessary errors or no instructions.

Prior	Instruction: Quick	Accurate	Both	None
Novice				
Expert				

For the interaction p = .03

- will we need an LSD_{mmd} to compare cell means? why or why not? Yep! sig. Int & k = 8 ! 200 / 8 = 25
- what will "n" be?

For the main effect of instruction p = .02

- will we need an LSD_{mmd} to compare marginal means? why or why not? Yep! sig. ME & k = 4 ! 200 / 4 = 50
- what will "n" be?
- will we need an LSDmmd to compare cell means? why or why not? Yep! sig. Int ! 200 / 8 = 25
- what will "n" be?

For the main effect of prior experience p = .02

- will we need an LSD_{mmd} to compare marginal means? why or why not? Nope – k = 2 !
- what will "n" be?
- will we need an LSDmmd to compare cell means? why or why not? Yep! sig. Int ! 200 / 8 = 25
- what will "n" be?

Analysis of Non-orthogonal Factorial Designs

Whether because of careful, intentional stratified sampling to match subpopulation proportions or because of sampling "exigencies" …

Many of our factor designs have unequal-n, resulting in non-orthogonality (collinearity, correlation) among our effects.

Either way, we have to be certain our variance partitioning, significance tests and interpretations properly match up!

Sum of Squares types:

Type I SS each effect controlled for previous effects
SStotal = SS(A) + SS(B|A) + SS(AB|A B) + SSerror

Type II SS each main effect controlled for other main effect
SStotal = SS(A|B) + SS(B|A) + SSerror

Type III SS each effect controlled for all other effects
Sstotal = SS(A|B AB) + SS(B|A AB) + SS(AB|A B) + Sserror

Type III SS is the most commonly used for non-orthogonal factorials!

Why are Type III SS "best" ?? Hold on, this takes a bit…

- In multiple regression the collinearity among predictors is caused by the relationship among the constructs – and is expected to replicate across samplings from the same population/setting/task-stim.
- However, for factorial designs, the collinearity among the effects is determined by relative cell sample sizes.
- So, unless there has been explicit stratified sampling, the collinearity among the effects is "happenstance" rather than reflecting the relationship between the constructs.
- Said differently, the effects "confound each other differently" depending on what the relative cell sample sizes happen to be (confounding varies with sample size!?)
- By using Type III SS – correcting each effect for the "happenstance" collinearity it has with the others – we should replicate the same corrected effects, regardless of the collinearity among the effects (that's caused by whatever the relative cell sample sizes happen to be)!

Cell means in non-orthogonal factorial designs

- cell means are calculated, analyzed and interpreted the same in orthogonal and non-orthogonal designs!
- so, the analysis and interpretation of simple effects are the same in orthogonal and non-orthogonal designs!

Marginal means – 3 kinds in a non-orthogonal design !!!

- Unweighted marginal means → computed as the average of the corresponding cell means, without regard to differential cell sample sizes - usually only used by mistake!
- Weighted marginal means → computed as average of corresponding cell means weighted by differential cell sample sizes – usually used in "Descriptives"
- Estimated marginal means → marginal means estimated from the model – used in EMMEANS significance testing

More about weighted marginal means…

Weighted marginal means are the best estimate of the mean of the population represented by the aggregate of the corresponding cell means (because of the weighting)

- However, as we've discussed, you must be very careful when producing and interpreting these marginal means, because, as an aggregate, they might not actually represent any population???
- Also ... Pay attention to this….!!!
- The weighted marginal means represent the difference between the groups *including the confounding of that group difference by the other main effect and the interaction!!!*
- Said differently… the difference between the weighted marginal means is like a simple correlation – it represents the bivariate relationship between the DV & that IV, without regard to how that relationship is confounded by other variables!

More about estimated marginal means & main effect F-tests…

Estimated marginal means are the best estimate of the mean of the population represented by that condition of the main effect, taking in to account the relationships among the DV, the IVs & their interaction!

How are they estimated?
Each factorial model has a corresponding multiple regression model (much more later…)

For a factorial design with the IVs "A" & "B" and interaction "AB"

$$DV' = b_A*A + b_B*B + b_{AB}*AB$$

For each participant, we can compute their estimated score from the model, and then compute the average estimated score for each marginal group!

The main effect F-tests are tests of these estimated marginal means! (not the weighted or unweighted marginal means) !