psyc3010 lecture 4

factorial between-Ps ANOVA III:
higher-order ANOVA

last lecture: factorial between-Ps ANOVA II
(following up significant effects)
next lecture: power analyses and blocking designs

Quiz 1 next Monday & Tuesday

preparation
• review Lectures 1 → 4
• organize your notes so you know where to find information
• look at quiz tips + complete the online practice questions
• complete the online Practice Quiz on Blackboard
• ?s welcome before the quiz. During the quiz, I will not be available.

taking the quiz
• opens @ 9am Sunday (27/3), closes @ 9pm Monday (28/3)
• no time restrictions + can return to active quiz
• can submit quiz only once, and must do so by closing date
• work alone
last lecture \(\rightarrow\) this lecture

- **last lecture:**
  - following up significant main effects and interactions in 2-way factorial ANOVA

- **this lecture:**
  - higher-order factorial designs
    ("complex ANOVA")

- **Next lectures:**
  - March 28\(^{th}\) guest lecture by Joanne Brown on power analyses & Blocking designs.
  - I return in Week 6.

---

topics for this week

- **introduction to higher-order designs**

- **omnibus tests in 3-way factorial ANOVA**
  - main effects, 2-way interactions, 3-way interactions

- **overview of follow-up tests in 3-way ANOVA**
  - following up main effects, 2-way interactions, and 3-way interactions

- **following up 3-way interactions - details**
  - simple interaction effects
  - simple simple effects
  - simple simple comparisons
higher-order factorial designs: an introduction

higher-order factorial designs

- more than 2 independent variables (factors)
- allow for designs with higher external validity
  → world is often more complicated than a 2 x 3...

**EXAMPLE – predicting driving performance**

- a number of variables could have an important influence:
  * age (young, old)
  * amount of alcohol drunk (0, 1, or 5 drinks)
  * gender (men, women)

  → how do these factors influence driving performance, independently and/or interactively?
notation for higher-order designs

- 2 (age) x 3 (alcohol) x 2 (gender) between-subjects design
- 12 cells [calculated as for 2-way factorial design]

<table>
<thead>
<tr>
<th></th>
<th>men</th>
<th>women</th>
</tr>
</thead>
<tbody>
<tr>
<td>no alc</td>
<td>1 drink</td>
<td>no alc</td>
</tr>
<tr>
<td>5 drinks</td>
<td></td>
<td>1 drink</td>
</tr>
<tr>
<td></td>
<td>5 drinks</td>
<td>5 drinks</td>
</tr>
<tr>
<td>old</td>
<td></td>
<td></td>
</tr>
<tr>
<td>young</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

effects in higher-order designs

- **main effects:**
  - differences between marginal means of one factor (averaging over levels of other factors)
main effects in higher-order designs

**main effect of gender**

men (averaged across alcohol and age)

women (averaged across alcohol and age)

---

effects in higher-order designs

- **main effects:**
  - differences between marginal means of one factor (averaging over levels of other factors)

- **two-way interactions:**
  - whether the effect of one factor is the same at every level of another factor (averaging over levels of a third factor)
2-way intxs in higher-order designs

*age by alcohol interaction*

<table>
<thead>
<tr>
<th>no alc</th>
<th>1 drink</th>
<th>5 drinks</th>
</tr>
</thead>
<tbody>
<tr>
<td>(averaged across men &amp; women)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

old

young

graph for 2-way interaction

1. compare line slopes – are lines are (not) parallel? (to check for interaction)

2. compare averages of red line and blue line (to check for main effect of factor represented by lines)

3. compare average for each column (to check for main effect of factor represented on X axis)
effects in higher-order designs

- **main effects:**
  - differences between marginal means of one factor (averaging over levels of other factors)

- **two-way interactions:**
  - examines whether the effect of one factor is the same at every level of another factor (averaging over levels of a third factor)

- **three-way interaction:**
  - examines whether the two-way interaction between two factors is the same at every level of the third factor
3-way intx in higher-order designs

*age x alcohol x gender* (no averaging)

**men**
- no alc
- 1 drink
- 5 drinks

**women**
- no alc
- 1 drink
- 5 drinks

old
- [Blank]
- [Blank]
- [Blank]

young
- [Blank]
- [Blank]
- [Blank]

→ **does a 2-way interaction hold at each level of the 3rd factor?**

---

**graphs for 3-way interactions**

1. plot 2-way interactions within each level of the third factor
2. check if pattern for 1st graph (simple interaction of AB at C1) is different from 2nd graph (simple interaction of AB at C2) -- If graphs are not same pattern, there is a 3-way interaction

→ difficult to interpret other 2-way interactions from graphs, let alone main effects
→ The more complex design, the more rely on statistical tests not eyeballing
partitioning variance in a 3-way ANOVA

- **Main effects**
  - Variance due to $\alpha$
  - Variance due to $\beta$
  - Variance due to $\gamma$

- **2-way interactions**
  - Variance due to $\alpha \beta$
  - Variance due to $\beta \gamma$
  - Variance due to $\alpha \gamma$

- **3-way interaction**
  - Variance due to $\alpha \beta \gamma$

- **Error/residual**
  - Variance due to $e$

---

structural models in factorial ANOVA

**2-way factorial design:**

$$X_{ijk} = \mu. + \alpha_j + \beta_k + \alpha\beta_{jk} + e_{ijk}$$

**3-way factorial design:**

$$X_{ijkl} =$$

$$\mu. + \alpha_j + \beta_k + \gamma_l + \alpha\beta_{jk} + \beta\gamma_{kl} + \alpha\gamma_{jl} + \alpha\beta\gamma_{jkl} + e_{ijkl}$$
degrees of freedom

\[ \text{df}_{\text{total}} = N - 1 \]
\[ \text{df}_j = j - 1 \]
\[ \text{df}_k = k - 1 \]
\[ \text{df}_l = l - 1 \]
\[ \text{df}_J = (j - 1)(k - 1) \]
\[ \text{df}_JL = (j - 1)(l - 1) \]
\[ \text{df}_{KL} = (k - 1)(l - 1) \]
\[ \text{df}_{JKL} = (j - 1)(k - 1)(l - 1) \]
\[ \text{df}_{\text{error}} = N - jkl \]

Regardless of the number of factors in the between-groups design, the degrees of freedom for a factor is always the number of levels minus 1.

The degrees of freedom for an interaction is the product of the degrees of freedom for the factors involved.

The degrees of freedom for error is (N - number of cells) or (n - 1) \times (number of cells).

3-way factorial designs: omnibus tests
yet another (quasi) experiment

*Reinforcement Sensitivity Theory*

- do people learn better from a particular type of reinforcement: reward versus punishment?
- impulsive personality $\rightarrow$ learn well from reward but not punishment
  anxious personality $\rightarrow$ learn well from punishment but not reward
- gender differences are also possible

---

yet another (quasi) experiment

*DV*: speed of responses in reaction time (RT) task

*3 independent factors:*

- reinforcement type
  - reward for fast responses
  - punishment for slow responses,
  - control condition (no reward or punishment)
- personality type
  - anxious personality
  - impulsive personality
- gender
  - male
  - female
yet another (quasi) experiment

7 omnibus tests baby!

– 3 main effects
  - reinforcement (reward, punishment, none)
  - personality (impulsive, anxious)
  - gender (male, female)

– 3 two-way interactions (a.k.a. first-order interactions)
  - reinforcement x personality
  - reinforcement x gender
  - personality x gender

– 1 three-way interaction (a.k.a. second-order interaction)
  - reinforcement x personality x gender

data and cell totals / means
(full layout)

<table>
<thead>
<tr>
<th>Personality</th>
<th>Males</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Reinforcement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rew</td>
<td>None</td>
<td>Pun</td>
<td>Rew</td>
<td>None</td>
<td>Pun</td>
<td>Rew</td>
<td>None</td>
<td>Pun</td>
<td>Rew</td>
<td>None</td>
<td>Pun</td>
</tr>
<tr>
<td>Impulsive</td>
<td></td>
<td>310</td>
<td>355</td>
<td>490</td>
<td>320</td>
<td>350</td>
<td>495</td>
<td>330</td>
<td>360</td>
<td>486</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxious</td>
<td></td>
<td>485</td>
<td>450</td>
<td>310</td>
<td>490</td>
<td>455</td>
<td>320</td>
<td>495</td>
<td>445</td>
<td>330</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>960</td>
<td>1065</td>
<td>1470</td>
<td>960</td>
<td>1350</td>
<td>1456</td>
<td>960</td>
<td>1350</td>
<td>1456</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>320</td>
<td>355</td>
<td>490</td>
<td>320</td>
<td>450</td>
<td>486</td>
<td>330</td>
<td>445</td>
<td>480</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Personality</th>
<th>Females</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Reinforcement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rew</td>
<td>None</td>
<td>Pun</td>
<td>Rew</td>
<td>None</td>
<td>Pun</td>
<td>Rew</td>
<td>None</td>
<td>Pun</td>
<td>Rew</td>
<td>None</td>
<td>Pun</td>
</tr>
<tr>
<td>Impulsive</td>
<td></td>
<td>310</td>
<td>450</td>
<td>490</td>
<td>320</td>
<td>455</td>
<td>486</td>
<td>330</td>
<td>445</td>
<td>480</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxious</td>
<td></td>
<td>485</td>
<td>345</td>
<td>310</td>
<td>480</td>
<td>350</td>
<td>320</td>
<td>490</td>
<td>355</td>
<td>330</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>1470</td>
<td>1350</td>
<td>960</td>
<td>1455</td>
<td>1050</td>
<td>960</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>490</td>
<td>450</td>
<td>320</td>
<td>485</td>
<td>350</td>
<td>320</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Mathematics is the language with which God has written the universe.

- Galileo Galilei, physicist and astronomer (1564-1642)
### summary table (from SPSS)

#### Tests of Between-Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reinforcement</td>
<td>51.722</td>
<td>2</td>
<td>25.861</td>
<td>.517</td>
<td>.603</td>
</tr>
<tr>
<td>Personality</td>
<td>7.111</td>
<td>1</td>
<td>7.111</td>
<td>.142</td>
<td>.709</td>
</tr>
<tr>
<td>Gender</td>
<td>53.778</td>
<td>1</td>
<td>53.778</td>
<td>1.075</td>
<td>.310</td>
</tr>
<tr>
<td>Reinforcement x Personality</td>
<td>168516.722</td>
<td>2</td>
<td>84258.361</td>
<td>1684.232</td>
<td>.000</td>
</tr>
<tr>
<td>Reinforcement x Gender</td>
<td>.056</td>
<td>2</td>
<td>.028</td>
<td>.001</td>
<td>.999</td>
</tr>
<tr>
<td>Personality x Gender</td>
<td>9538.778</td>
<td>1</td>
<td>9538.778</td>
<td>190.670</td>
<td>.000</td>
</tr>
<tr>
<td>Reinforcement x Personality x Gender</td>
<td>19015.056</td>
<td>2</td>
<td>9507.528</td>
<td>190.045</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>1200.667</td>
<td>24</td>
<td>50.028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>198383.889</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### no significant main effects

#### Tests of Between-Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reinforcement</td>
<td>51.722</td>
<td>2</td>
<td>25.861</td>
<td>.517</td>
<td>.603</td>
</tr>
<tr>
<td>Personality</td>
<td>7.111</td>
<td>1</td>
<td>7.111</td>
<td>.142</td>
<td>.709</td>
</tr>
<tr>
<td>Gender</td>
<td>53.778</td>
<td>1</td>
<td>53.778</td>
<td>1.075</td>
<td>.310</td>
</tr>
<tr>
<td>Reinforcement x Personality</td>
<td>168516.722</td>
<td>2</td>
<td>84258.361</td>
<td>1684.232</td>
<td>.000</td>
</tr>
<tr>
<td>Reinforcement x Gender</td>
<td>.056</td>
<td>2</td>
<td>.028</td>
<td>.001</td>
<td>.999</td>
</tr>
<tr>
<td>Personality x Gender</td>
<td>9538.778</td>
<td>1</td>
<td>9538.778</td>
<td>190.670</td>
<td>.000</td>
</tr>
<tr>
<td>Reinforcement x Personality x Gender</td>
<td>19015.056</td>
<td>2</td>
<td>9507.528</td>
<td>190.045</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>1200.667</td>
<td>24</td>
<td>50.028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>198383.889</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
significant 2-way interaction: personality x reinforcement

Tests of Between-Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reinforcement</td>
<td>51.722</td>
<td>2</td>
<td>25.861</td>
<td>517</td>
<td>.003</td>
</tr>
<tr>
<td>Personality</td>
<td>7.111</td>
<td>1</td>
<td>7.111</td>
<td>142</td>
<td>.709</td>
</tr>
<tr>
<td>Gender</td>
<td>53.778</td>
<td>1</td>
<td>53.778</td>
<td>1.075</td>
<td>.310</td>
</tr>
<tr>
<td>Reinforcement x Personality</td>
<td>168516.722</td>
<td>2</td>
<td>84258.361</td>
<td>1684.232</td>
<td>.000</td>
</tr>
<tr>
<td>Reinforcement x Gender</td>
<td>.056</td>
<td>2</td>
<td>.028</td>
<td>.001</td>
<td>.999</td>
</tr>
<tr>
<td>Personality x Gender</td>
<td>9538.778</td>
<td>1</td>
<td>9538.778</td>
<td>190.670</td>
<td>.000</td>
</tr>
<tr>
<td>Reinforcement x Personality x Gender</td>
<td>19015.056</td>
<td>2</td>
<td>9507.528</td>
<td>190.045</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>1200.667</td>
<td>24</td>
<td>50.028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>198383.889</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

→ effect of at least one of the two factors is different at different levels of the other factor
→ ignoring (averaging across) the third factor
### Tests of Between-Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reinforcement</td>
<td>51.722</td>
<td>2</td>
<td>25.861</td>
<td>.517</td>
<td>.603</td>
</tr>
<tr>
<td>Personality</td>
<td>7.111</td>
<td>1</td>
<td>7.111</td>
<td>.142</td>
<td>.709</td>
</tr>
<tr>
<td>Gender</td>
<td>53.778</td>
<td>1</td>
<td>53.778</td>
<td>1.075</td>
<td>.310</td>
</tr>
<tr>
<td>Reinforcement x Personality</td>
<td>168516.722</td>
<td>2</td>
<td>84258.361</td>
<td>1684.232</td>
<td>.000</td>
</tr>
<tr>
<td>Reinforcement x Gender</td>
<td>.056</td>
<td>2</td>
<td>.028</td>
<td>.001</td>
<td>.999</td>
</tr>
<tr>
<td>Personality x Gender</td>
<td>9538.778</td>
<td>1</td>
<td>9538.778</td>
<td>190.670</td>
<td>.000</td>
</tr>
<tr>
<td>Reinforcement x Personality x Gender</td>
<td>19015.056</td>
<td>2</td>
<td>9507.528</td>
<td>190.045</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>1200.667</td>
<td>24</td>
<td>50.028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>198353.889</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**significant 2-way interaction: personality x gender**

- **effect of at least one factor is different at different levels of the other factor**
- **ignoring (averaging across) the third factor**

![Graph showing mean reaction time for different levels of personality and gender](chart.png)
significant 3-way interaction: reinforcement x personality x gender

Tests of Between-Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reinforcement</td>
<td>81.722</td>
<td>2</td>
<td>25.861</td>
<td>.517</td>
<td>.603</td>
</tr>
<tr>
<td>Personality</td>
<td>7.111</td>
<td>1</td>
<td>7.111</td>
<td>.142</td>
<td>.709</td>
</tr>
<tr>
<td>Gender</td>
<td>53.778</td>
<td>1</td>
<td>53.778</td>
<td>1.075</td>
<td>.310</td>
</tr>
<tr>
<td>Reinforcement x Personality</td>
<td>168516.722</td>
<td>2</td>
<td>84258.361</td>
<td>1684.232</td>
<td>.000</td>
</tr>
<tr>
<td>Reinforcement x Gender</td>
<td>.056</td>
<td>2</td>
<td>.028</td>
<td>.001</td>
<td>.999</td>
</tr>
<tr>
<td>Personality x Gender</td>
<td>9538.778</td>
<td>1</td>
<td>9538.778</td>
<td>190.670</td>
<td>.000</td>
</tr>
<tr>
<td>Reinforcement x Personality x Gender</td>
<td>19015.056</td>
<td>2</td>
<td>9507.528</td>
<td>190.045</td>
<td>.000</td>
</tr>
<tr>
<td>Error</td>
<td>1200.667</td>
<td>24</td>
<td>50.028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>198383.889</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

so there’s this 3-way interaction…

→ means that a two-way interaction is different at different levels of the third factor
   → Simple interaction of AxB is for C1 than simple interaction for C2, etc..

→ But which simple interaction to look at? Could mean one or more of the following:

- **personality X gender** 2-way interaction is different across levels of reinforcement
- **reinforcement X personality** 2-way interaction is different across levels of gender
- **reinforcement X gender** 2-way interaction is different across levels of personality
following up significant omnibus effects in a 3-way ANOVA: overview of all tests

---

following up main effects in 3-way factorial ANOVA

- just as in a 2-way ANOVA, a significant omnibus main effect must be interpreted
  - e.g. what is the effect of Factor A, ignoring (averaging over) Factor B and Factor C?

- if the factor has > 2 levels, we don’t know exactly where the differences are

- we then use main effect comparisons (with t-tests or linear contrasts), exactly as we did in 2-way ANOVA
following up omnibus 2-way interactions in 3-way factorial ANOVA

- just as in a 2-way ANOVA, a significant omnibus 2-way interaction must be interpreted
  - e.g., is the effect of Factor A different at different levels of Factor B, and vice-versa? (ignoring Factor C)
- we then test *simple effects* (with the $F$ test), exactly as we did in 2-way ANOVA
- if you find a significant simple effect for a *factor with > 2 levels*, you follow it up with *simple comparisons* (with $t$-tests or linear contrasts), exactly as we did in 2-way ANOVA

following up a 3-way interaction in 3-way factorial ANOVA: overview of the steps

- *simple interaction effects*
  - $F$ tests
- if simple interaction effects are significant, follow up with *simple simple effects*
  - $F$ tests
- if simple simple effects are significant with > 2 levels, follow up with *simple simple comparisons*
  - $t$-tests and linear contrasts
following up a 3-way interaction, part I: simple interaction effects

so there’s this 3-way interaction...

this could mean one or more of the following:

- personality X gender 2-way interaction is different across levels of reinforcement
- reinforcement X personality 2-way interaction is different across levels of gender
- reinforcement X gender 2-way interaction is different across levels of personality

need to focus your investigation

1) go back to theory and hypotheses
2) conduct follow-up analyses to test predictions
simple interaction effects

• simple interaction effects break down the 3-way interaction into a series of 2-way interactions at each level of the third factor

→ WHY WOULD WE DO THIS?
• this gives a first close-up look at where the differences between cell means might be
• once we know this, we can follow up these simple 2-way interactions further to figure out where the differences are (simple simple effects & simple simple comparisons / contrasts)

→ just as we follow up an interaction in a 2-way design

→ in a 3-way design there are three potential follow-up steps (compared to two in a 2-way design)

the graphs depicting the 2 x 2 x 3 interaction between gender, personality, and reinforcement provide a visual representation of the simple interaction effects we would conduct: here, the simple personality x reinforcement interaction at each of the two levels of gender
**omnibus 2-way interaction in a 3-way design:** *ignores levels of the third factor*

Tests the 2-way P x R interaction with data averaged across gender (i.e., ignoring the third factor)

**simple 2-way interactions in a 3-way design:** test the 2-way interaction at each level of the third factor

Tests the 2-way P x R interaction for each gender group separately (i.e., at each level of third factor)

---

**don’t mix up your designs**

- Simple interaction effects in a 3-way design look like a series of 2-way ANOVAs at each level of the third factor (e.g., a P x R ANOVA for men, then a P x R ANOVA for women)
- But the F ratios calculated for these tests are not the same
- – Simple interaction (a) uses pooled error term; (b) is conducted after significant 3-way interaction is observed

**simple P x R 2-way interaction in a 3-way design:**
- is a follow-up test for 3-way interaction (P x R x G)
- men and women were sampled as part of the same study
- MSError taken from 3-way omnibus ANOVA table

**omnibus P x R 2-way interaction in 2-way designs:**
- separate P x R studies for men and women (e.g., at different times, in different recruiting contexts – confounds?)
- MSError taken from each 2-way omnibus ANOVA table (different tables for men and women, so different MSError)
**summary table for P x R simple interaction effects**

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>PR at G1</td>
<td>95725.00</td>
<td>2</td>
<td>47862.50</td>
<td>956.72</td>
<td>0.000</td>
</tr>
<tr>
<td>PR at G2</td>
<td>91806.78</td>
<td>2</td>
<td>45903.39</td>
<td>917.56</td>
<td>0.000</td>
</tr>
<tr>
<td>Error</td>
<td>1200.67</td>
<td>24</td>
<td>50.03</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Critical F at alpha=.05 (2,24) = 3.40

→ includes F tests for **all P x R** simple interaction effects (i.e., the P x R simple interaction effect at each level of the gender factor—men and women)

These are the SS values for the effects

Degrees of freedom for a simple interaction effect are the df for the associated interaction

df = dfPR (2 - 1)(3 - 1) = 2

SS\text{error} term (and df) is taken from the omnibus 2 x 2 x 3 ANOVA

Mean squares and F values calculated as usual

Indicates that the personality x reinforcement interaction is significant for males and females...

→ follow-up tests will identify where the differences are...
following up a 3-way interaction, part II: simple simple effects & simple simple comparisons

simple simple effects

- simple effects after an omnibus 2-way interaction examine the effect of factor A at each level of factor B

- simple simple effects are like simple effects except that they examine the effect of factor A at each level of factor B, at each level of factor C (i.e., within each combo of B & C)

- Simple simple effects differ from one-way ANOVAs because they use the $\text{MS}_{\text{error}}$ from the omnibus ANOVA table as the error term
simple simple effect # 1:
personality at each level of reinforcement
(separately for men and women)

results: for both men and women, the effect of personality
at each level of reinforcement was significant
(although opposite under neutral reinforcement!)

summary table
simple simple effects of personality,
at each level of reinforcement,
at each level of gender (males and females)

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>P at R1 at G1</td>
<td>43350.00</td>
<td>1</td>
<td>43350.00</td>
<td>866.52</td>
<td>0.000</td>
</tr>
<tr>
<td>P at R2 at G1</td>
<td>13537.50</td>
<td>1</td>
<td>13537.50</td>
<td>270.60</td>
<td>0.000</td>
</tr>
<tr>
<td>P at R3 at G1</td>
<td>4330.50</td>
<td>1</td>
<td>4330.50</td>
<td>86.56</td>
<td>0.000</td>
</tr>
<tr>
<td>P at R1 at G2</td>
<td>40837.50</td>
<td>1</td>
<td>40837.50</td>
<td>816.30</td>
<td>0.000</td>
</tr>
<tr>
<td>P at R2 at G2</td>
<td>15000.00</td>
<td>1</td>
<td>15000.00</td>
<td>299.83</td>
<td>0.000</td>
</tr>
<tr>
<td>P at R3 at G2</td>
<td>41002.7</td>
<td>1</td>
<td>41002.66</td>
<td>819.60</td>
<td>0.000</td>
</tr>
<tr>
<td>Error</td>
<td>1200.67</td>
<td>24</td>
<td>50.03</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

critical F at alpha=.05 (1,24) = 4.26

R1 = reward, R2 = neutral, R3 = punishment
G1 = men, G2 = women
simple simple effect # 2:
reinforcement at each level of personality
(separately for men and women)

results: for both men and women, the effect of reinforcement
at each level of personality was significant

summary table
simple simple effects of reinforcement,
at each level of personality,
at each level of gender (males and females)

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>R at P1 at G1</td>
<td>48350.00</td>
<td>2</td>
<td>24175.00</td>
<td>483.23</td>
<td>0.000</td>
</tr>
<tr>
<td>R at P2 at G1</td>
<td>47400.00</td>
<td>2</td>
<td>23700.00</td>
<td>473.74</td>
<td>0.000</td>
</tr>
<tr>
<td>R at P1 at G2</td>
<td>45483.56</td>
<td>2</td>
<td>22741.78</td>
<td>454.58</td>
<td>0.000</td>
</tr>
<tr>
<td>R at P2 at G2</td>
<td>46350.00</td>
<td>2</td>
<td>23175.00</td>
<td>463.24</td>
<td>0.000</td>
</tr>
<tr>
<td>Error</td>
<td>1200.67</td>
<td>24</td>
<td>50.03</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

critical F at alpha=.05 (2,24) = 3.40

P1 = impulsive, P2 = anxious
G1 = men, G2 = women
BUT...reinforcement factor has > 2
levels: how do we know which cell
means are significantly different
from each other?
**simple simple comparisons**

Like simple comparisons (or contrasts) except we compute for each level of a third factor

formulae from Lecture 3 can be used:

\[ t = \frac{L}{\sqrt{\frac{\sum a_j^2 \text{MS}_{\text{error}}}{n}}} \]

\[ L = \sum a_j \bar{X}_j \]

\[ \text{df}_{\text{error}} = N - ab \]

---

**some possible comparisons...**

R₁ and R₂ vs R₃ at P₁ for G₁ ...

R₂ vs R₃ at P₂ for G₂...
### simple simple comparisons
*for reinforcement at each level of personality (for males)*

<table>
<thead>
<tr>
<th>Males</th>
<th>Reinforcement</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rew</td>
<td>None</td>
<td>Pun</td>
</tr>
<tr>
<td>Impusivity</td>
<td>320</td>
<td>355</td>
<td>490</td>
</tr>
<tr>
<td>Contrast 1</td>
<td>1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Contrast 2</td>
<td>1</td>
<td>1</td>
<td>-2</td>
</tr>
<tr>
<td>Anxiety</td>
<td>490</td>
<td>450</td>
<td>320</td>
</tr>
<tr>
<td>Contrast 1</td>
<td>1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Contrast 2</td>
<td>1</td>
<td>1</td>
<td>-2</td>
</tr>
</tbody>
</table>

### calculations for impulsivity contrast 1

<table>
<thead>
<tr>
<th>Males</th>
<th>Reinforcement</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rew</td>
<td>None</td>
<td>Pun</td>
</tr>
<tr>
<td>Impusivity</td>
<td>320</td>
<td>355</td>
<td>490</td>
</tr>
<tr>
<td>Contrast 1</td>
<td>1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Contrast 2</td>
<td>1</td>
<td>1</td>
<td>-2</td>
</tr>
<tr>
<td>Anxiety</td>
<td>490</td>
<td>450</td>
<td>320</td>
</tr>
<tr>
<td>Contrast 1</td>
<td>1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>Contrast 2</td>
<td>1</td>
<td>1</td>
<td>-2</td>
</tr>
</tbody>
</table>

\[ L = 1(320) - 1(355) + 0(490) = -35.00 \]

\[ t'_{0.05} (24) = 2.39 \]

(with Bonferroni adjustment for 2 comparisons)
...and so on for

- impulsivity contrast 2…
- anxiety contrast 1…
- anxiety contrast 2…
- then all four contrasts
  for females…

oh, the follow-up tests you’ll do…

- omnibus tests
  - 7 (3 main effects, 3 two-way interactions, 1 three-way interaction)

- simple interaction effects
  - 2 (personality x reinforcement at each level of gender)

- simple simple effects
  - 10 (6 for personality at each level of reinforcement for men & women, 4 for reinforcement at each level of personality for men & women)

- simple simple comparisons
  - 8 (2 comparisons for each personality condition for men & women)

→ total = 27 tests! (in this study)
...but what if you don’t want to?

- conducting an exhaustive set of follow-up tests for higher-order factorial designs can inflate *familywise error rate*

  27 tests from previous slide: each has Type-1 error of .05

  → **means a familywise alpha of 27 * .05 = 1.35 (!)**

- **ultimately, there is no simple rule:** what you report depends entirely upon your research predictions

- in our case we had (implicitly) predicted the personality x reinforcement interaction, and we were going to see if this interaction was the same for males and females:
  - impulsive → reward not punishment
  - anxious → punishment but not reward
  - possible gender differences

---

**steps for following up a 3-way interaction**

1. **is 3-way significant?**
   - **YES:** calculate simple interaction effects at each level of least important factor or according to hypotheses
   - **STOP**
2. **is A x B at C1 or C2 significant?**
   - **YES:** conduct tests for simple simple comparisons
   - **STOP**
   - **NO:** does the factor have >2 levels?
     - **YES:** conduct tests for simple simple comparisons
     - **STOP**
     - **NO:** Conduct simple simple effects of key IV (A or B). Are they significant?
       - **YES:** conduct tests for simple simple comparisons
       - **STOP**
       - **NO:** **STOP**
summary

- in a 3-way ANOVA, you have main effects and 2-way interactions (like a 2-way ANOVA) plus a 3-way interaction

- the follow-up tests for omnibus main effects and two-way interactions are like in a 2-way ANOVA – main effect comparisons, simple effects, simple comparisons

- following up a 3-way interaction can be very complex and time-consuming
  - highlights the need for analyses to be driven by hypotheses

Reporting …

- some sources suggest that once you find a significant interaction you should ignore the lower order effects
  - main effects have been "qualified" by the interaction
  - 2-way interaction qualified by 3-way

- this is because the higher order interactions may require you to change the interpretation given by the lower effect alone
  - Partly depending on whether it’s a disordinal interaction

- ultimately, there is no simple rule:
  - what you report depends upon your research predictions
  - Usually, if you predict an effect, then report that effect (and any follow-up tests)
"Results indicated a significant main effect of consumption, $F(2,42) = 20.07$, $p<.001$, $\omega^2 = .34$. Linear contrasts with a Bonferroni adjustment for 2 comparisons indicated that creativity ratings were significantly lower after 2 or 4 pints than after consuming no alcohol, $t'(42) = 2.91$, $p<.05$ ($Ms = 63.75, 55.63$), and were lower after 4 pints than after 2 pints, $t'(42) = 5.63$, $p<.05$ ($Ms = 64.69, 46.56$). There was no significant main effect for distraction, indicating that creativity ratings for distracted participants’ limericks ($M = 56.46$) were not significantly different from those for controls ($M = 60.21$), $F(1,42) = 2.03$, $p = .16$, $\omega^2 = .01$. There was, however, a significant interaction between consumption and distraction, indicating that the effect of consumption was different for distracted and control participants, $F(1,42) = 11.91$, $p<.001$, $\omega^2 = .20$. The interaction is depicted in Figure 1.”

NB Interaction needs following up in results section (simple effects + simple comparisons if nec.).

Discuss: although the predicted main effect of alcohol consumption was significant, the direction of the effect was contrary to hypotheses: alcohol lowered creativity ratings. Also the predicted effect of distraction was not significant.

reporting simple effects...

• it is preferable to not report all sets of simple effects, for 2 reasons:
  a) the more simple effects we calculate, the greater our risk of making a type 1 error (see Howell, p.436)
  b) usually both sets of simple effects will communicate similar information - redundancy
    – so, in our case we would want to report either the simple effects of distraction (at each level of consumption) or the simple effects of consumption (at each level of distraction)

• ultimately, there is no simple rule: what you report depends entirely upon your research predictions.
  – Let’s say we specifically predicted that “the effect of consumption on creativity ratings will be stronger for distracted participants than for controls”. Therefore, we would want to report the simple effects for consumption (and associated simple comparisons / contrasts)
To follow up the significant two-way interaction, the simple effects of consumption were analysed at each level of distraction. There was a significant simple effect of consumption for distracted participants, $F(2,42) = 31.36, p<.001, \omega^2 = .56$, but not for controls, $F(2,42) = 0.61, p = .546, \omega^2 = .00$. The significant simple effect of consumption for distracted participants was followed up with Linear contrasts using a Bonferroni adjustment for 2 comparisons. These indicated that, for distracted participants, creativity ratings were lower after 2 or 4 pints than after consuming no alcohol, $t'(42) = 4.52, p<.001 (Ms = 66.88, 51.26)$, and also lower after 4 pints than after 2 pints, $t'(42) = 6.86, p<.001 (Ms = 66.88, 35.63)$.

Discuss: the hypothesis was confirmed that the effect of consumption on creativity will be stronger for distracted participants than for controls.

- "Writing is easy; all you do is sit staring at a blank sheet of paper until the drops of blood form on your forehead." -- Gene Fowler

- "There's nothing to writing. All you do is sit down at a typewriter and open a vein." -- Red Smith
Reporting three-way designs

- conducting an exhaustive set of follow-up tests for higher-order factorial designs can inflate familywise alpha (and is very tedious!)

- ultimately, there is no simple rule: what you report depends entirely upon your research predictions
  - in our case we had (implicitly) predicted the Personality x Reinforcement interaction, and we were going to see if this interaction was the same for males and females
    - people with an impulsive personality learn well from reward but not punishment, and people with an anxious personality learn well from punishment but not reward.
    - Possible gender differences not well understood.
  - hence our write up might have gone something like this...

Reporting

“The predicted interaction was significant, \( F(2, 24) = 1684.23, p < .001 \), but this was qualified by 3-way interaction among personality, reinforcement, and gender \( F(2, 24) = 190.67, p < .001 \). Simple interaction analyses revealed the personality x reinforcement interaction was significant for both males, \( F(2, 24) = 956.72, p < .001 \), and females, \( F(2, 24) = 917.56, p < .001 \). The simple simple effects of personality were then analysed for each level of gender and reinforcement, and Table 1 presents the relevant means. For both genders, as predicted, under punishment anxious participants were faster than impulsive participants, \( Fs > 819.58, ps < .001 \), while under reward impulsive participants were faster than anxious participants, \( Fs > 816.29, ps < .001 \). However, in the neutral reinforcement condition the gender difference emerged: impulsive males performed better than anxious males \( F(1, 24) = 270.60, p < .001 \) (Ms = 355, 450), while impulsive females performed worse than anxious females, \( F(2, 24) = 299.83, p < .001 \) (Ms = 450, 350).”

I haven’t put effect sizes. These would be required for all tests these days.
Table 1. Mean reaction time as a function of personality, reinforcement, and gender.

<table>
<thead>
<tr>
<th>Personality Type</th>
<th>Impulsive</th>
<th>Anxious</th>
</tr>
</thead>
<tbody>
<tr>
<td>Punishment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>485.33&lt;sub&gt;a&lt;/sub&gt;</td>
<td>320.00&lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
<tr>
<td>Men</td>
<td>490.00&lt;sub&gt;a&lt;/sub&gt;</td>
<td>320.00&lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
<tr>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>450.00&lt;sub&gt;a&lt;/sub&gt;</td>
<td>350.00&lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
<tr>
<td>Men</td>
<td>355.00&lt;sub&gt;a&lt;/sub&gt;</td>
<td>450.00&lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
<tr>
<td>Reward</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>320.00&lt;sub&gt;a&lt;/sub&gt;</td>
<td>485.00&lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
<tr>
<td>Men</td>
<td>320.00&lt;sub&gt;a&lt;/sub&gt;</td>
<td>490.00&lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

Note. Subscripts within the row indicate significant simple effects of personality.

Most experimental journals would also want to see standard deviations.

Education: That which discloses to the wise and disguises from the foolish their lack of understanding. -Ambrose Bierce, writer (1842-1914)
readings

*between-Ps ANOVA III: higher-order designs (this lecture)*
- Field (3rd ed): review Chapter 12
- Field (2nd ed): review Chapter 10
- Howell (all eds): Chapter 13 (pp 446-453)

*power analyses and blocking designs (next lecture)*
- Field (2nd or 3rd ed): no new readings
- Howell (all eds): Chapter 8