Repeated-Measures ANOVA

Goals

- Rationale of Repeated Measures ANOVA
  - One-way and two-way
  - Benefits
- Partitioning Variance
- Statistical Problems with Repeated-Measures Designs
  - Sphericity
  - Overcoming these problems
- Interpretation

Repeated-Measures ANOVA

	Repeated-Measures ANOVA
1. Situation/hypotheses	One or more factors dependent samples $H_0: \mu_1=\mu_2=\ldots=\mu_J$
2. Test statistic	$F = \frac{MS_{Treatment}}{MS_{Error}}$
3 .Distribution	$F_{J-1,\ (J-1)*(n-1)}$
4. Assumptions	1. Populations are normal 2. Equal population variances for each cell 3. Sphericity (if more than three conditions)

Benefits of Repeated-Measures Designs

- Sensitivity
  - Unsystematic variance is reduced
  - More sensitive to experimental effects
- Economy
  - Fewer participants are needed
  - …but participants may become fatigued or experience practice effects

An Example

- Are certain bushtucker foods (eaten by Australian Aborigines) more revolting than others?
- Four foods tasted by 8 celebrities:
  - Stick Insect
  - Kangaroo Testicle
  - Fish Eyeball
  - Witchetty Grub

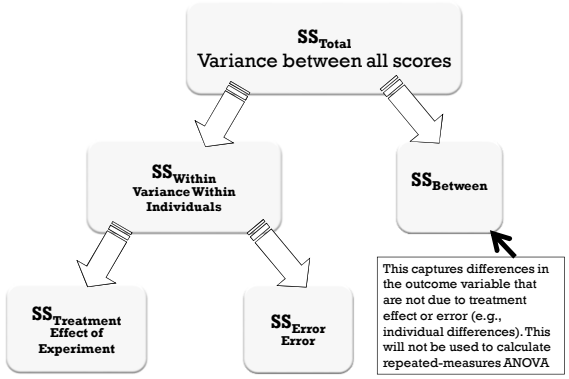



- Outcome:
  - Time to retch (seconds)

Celebrity	Stick Insect	Testicle	Fish Eye	Witchetty Grub	Mean	Variance
1	8	7	1	6	5.50	9.67
2	9	5	2	5	5.25	8.25
3	6	2	3	8	4.75	7.58
4	5	3	1	9	4.50	11.67
5	8	4	5	8	6.25	4.25
6	7	5	6	7	6.25	0.92
7	10	2	7	2	5.25	15.58
8	12	6	8	1	6.75	20.92
Mean	8.13	4.25	4.13	5.75		
Grand Mean = 5.56 Grand Variance = 8.19						

Logic of Repeated-Measures ANOVA

- The effect of the experiment is found in the within-participant variance rather than in the between-group variance
  - Within-participant variance is composed of (1) treatment effect and (2) individual differences in performance (which is "error")
- The sources of variance for the repeated-measures ANOVA are the same as those for the one-way ANOVA…but the variances reflecting the treatment effect and error are both part of the within-participant variance for repeated-measures ANOVA



Calculate SS_{Total}

8	7	1	6
9	5	2	5
6	2	3	8
5	3	1	9
8	4	5	8
7	5	6	7
10	2	7	2
12	6	8	1

Grand Mean = 5.56

Grand Variance = 8.19

This is the total amount of variability that can be explained

$$SS_{Total} = s^2_{grand}(N-1)$$
$$= 8.19(32-1)$$
$$= 253.89$$

Calculate SS_{Within}

$$SS_{Within} = s^2_{person1}(n_1 - 1) + s^2_{person2}(n_2 - 1) + s^2_{person3}(n_3 - 1) \ldots s^2_{personn}(n_n - 1)$$

This is the amount of variability that is explained by within-participant features (i.e., both treatment and error)

$$\begin{aligned} SS_{Within} &= s^2_{celebrity1}(n_1 - 1) + s^2_{celebrity2}(n_2 - 1) + s^2_{celebrity3}(n_3 - 1) + s^2_{celebrity4}(n_4 - 1) \\ &\quad + s^2_{celebrity5}(n_5 - 1) + s^2_{celebrity6}(n_6 - 1) + s^2_{celebrity7}(n_7 - 1) + s^2_{celebrity8}(n_8 - 1) \\ &= (9.67 \times 3) + (8.25 \times 3) + (7.58 \times 3) + (11.67 \times 3) \\ &\quad + (4.25 \times 3) + (0.92 \times 3) + (15.58 \times 3) + (20.92 \times 3) \\ &= 29 + 24.75 + 22.75 + 35 + 12.75 + 2.75 + 46.75 + 62.75 \\ &= 236.50 \end{aligned}$$

Calculate $SS_{Treatment}$

$$SS_{Treatment} = \sum n_i \left(\bar{x}_i - \bar{x}_{grand} \right)^2$$

$$\begin{aligned} SS_{Treatment} &= 8(8.13 - 5.56)^2 + 8(4.25 - 5.56)^2 + 8(4.13 - 5.56)^2 + 8(5.75 - 5.56)^2 \\ &= 8(2.57)^2 + 8(-1.31)^2 + 8(-1.44)^2 + 8(0.196)^2 \\ &= 83.13 \end{aligned}$$

This is the amount of variability that is explained by the foods

Calculate SS_{Error}

$$SS_{Error} = SS_{Within} - SS_{Treatment}$$

$$\begin{aligned} SS_{Error} &= SS_{Within} - SS_{Treatment} \\ &= 236.50 - 83.13 \\ &= 153.37 \end{aligned}$$

This is the amount of variability that is explained by error

Mean Squares for Repeated-Measures ANOVA

- $MS_{Treatment} = \frac{SS_{Treatment}}{df_{Treatment}} = \frac{83.13}{3} = 27.71$
- $MS_{Error} = \frac{SS_{Error}}{df_{Error}} = \frac{153.37}{21} = 7.30$
- $df_{Within} = n*(J - 1) = 8*3 = 24$
- $df_{Treatment} = J - 1 = 4 - 1 = 3$
- $df_{Error} = (J-1)*(n-1) = 3*7 = 21$

F-Ratio for Repeated-Measures ANOVA

- $F = \frac{MS_{Treatment}}{MS_{Error}} = \frac{27.71}{7.30} = 3.79$
- This is conceptually similar to the one-way ANOVA

Problems with Analyzing Repeated Measures Designs

- Same participants in all conditions
  - Scores across conditions correlate
  - Violates the assumption of independence
- Assumption of **Sphericity**
  - Sphericity refers to the condition where the variances of the differences between all possible pairs of groups (i.e., levels of the factor) are equal
  - The violation of sphericity occurs when the variances of the differences between all combinations of the groups are not equal
  - If sphericity is violated, then the variance calculations may be distorted which would result in an inflated F-ratio
  - Sphericity can be evaluated when there are three or more levels of a repeated-measures factor and, with each additional repeated-measures factor, the risk for violating sphericity increases
  - If sphericity is violated, the degrees of freedom for the repeated-measures ANOVA should be adjusted to correct for this violation
  - Sphericity is measured using Mauchly's test
    - $p < .05$ means that the sphericity assumption is violated
    - $p > .05$ means that the sphericity assumption is met

What is Sphericity?

	Testicle - Stick	Eye – Stick	Witchetty – Stick	Eye – Testicle	Witchetty – Testicle	Witchetty – Eye
1	-1	-7	-2	-6	-1	5
2	-4	-7	-4	-3	0	3
3	-4	-3	2	1	6	5
4	-2	-4	4	-2	6	8
5	-4	-3	0	1	4	3
6	-2	-1	0	1	2	1
7	-8	-3	-8	5	0	-5
8	-6	-4	-11	2	-5	-7
Variance	5.27	4.29	25.70	11.55	14.29	26.55

Corrections for Violations of Sphericity Assumption

- Three corrections:
  - Greenhouse-Geisser Correction
    - This conservative correction should be used when sphericity estimates are less than .75 or nothing is known about sphericity
  - Huynh-Feldt Correction
    - This liberal correction should be used when sphericity estimates are.75 or greater
  - Lower-Bound Correction
    - Not used very often because it is extremely conservative
- Multiply degrees of freedom by these estimates to correct for the effect of sphericity
- Another way to deal with violations of the sphericity assumption is to use a multivariate ANOVA (MANOVA) because it does not rely on a sphericity assumption

Correcting for Sphericity

Df = 3, 21

Mauchly's Test of Sphericity

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon: Greenhouse-Geisser	Epsilon: Huynh-Feldt	Epsilon: Lower-bound
Animal	.136	11.406	5	.047	.533	.666	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

3 X 0.533 = 1.59, 0.666 = 1.99, 0.333 = 1.00 **21 X** 0.533 = 11.19, 0.666 = 13.98, 0.333 = 7.00

Output

This is the amount of variability that is explained by the treatment effect

If we use the Greenhouse-Geisser correction, then we would retain the null hypothesis. However, if we had used the Huynh-Feldt, then we would have rejected the null

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Animal	Sphericity Assumed	83.125	3	27.708	2.794	.063
	Greenhouse-Geisser	83.125	1.599	52.001	2.794	.063
	Huynh-Feldt	83.125	1.997	41.619	2.794	.048
	Lower-bound	83.125	1.000	83.125	2.794	.092
Error(Animal)	Sphericity Assumed	153.375	21	7.304		
	Greenhouse-Geisser	153.375	11.190	13.707		
	Huynh-Feldt	153.375	13.981	10.970		
	Lower-bound	153.375	7.000	21.911		

This is the amount of variability that is explained by error

This is the F-ratio…notice that it is the same whether we assume the sphericity assumption is met and or violated (because the corrections only adjust degrees of freedom)

Multiple Comparison Procedures

- Compare each mean against all others (*t*-tests)
- In general terms they use a stricter criterion to accept an effect as significant
  - They control the familywise error rate
  - Bonferroni method for controlling Type I error:

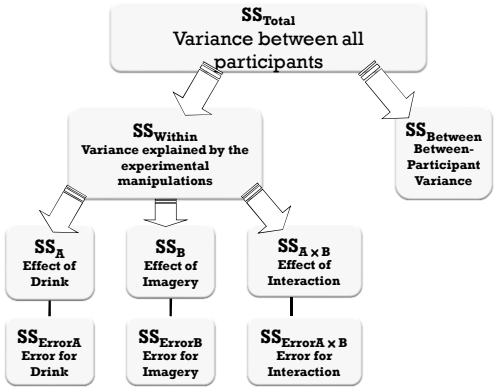
$$\text{Bonferroni } \alpha = \frac{\alpha}{\text{Number of Tests}}$$

What is Two-Way Repeated Measures ANOVA?

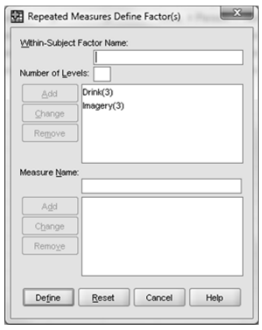
- Two factors
  - Two-way = 2 factors
  - Three-Way = 3 factors
- The same participants in *all* conditions
  - Repeated Measures = 'same participants'
  - This should be used to analyze data from a within-subjects design

An Example

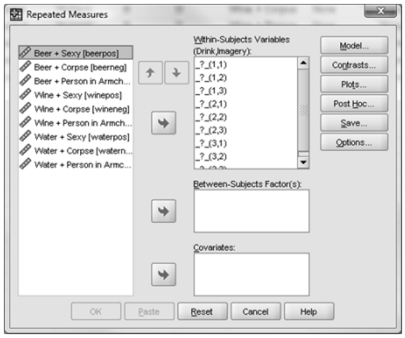
- Effects of advertising on evaluations of different drink types
  - Factor A (Drink): Beer, Wine, Water
  - Factor B (Imagery): Positive, Negative, Neutral
  - Outcome Variable: Evaluation of product from -100 (*dislike very much*) to +100 (*like very much*)



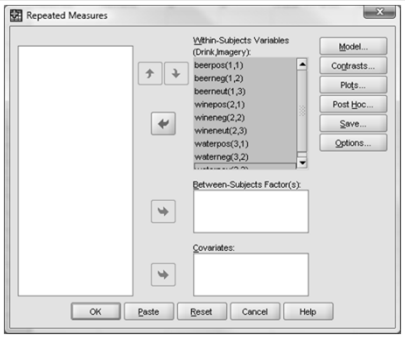
Running the Analysis: Naming factors



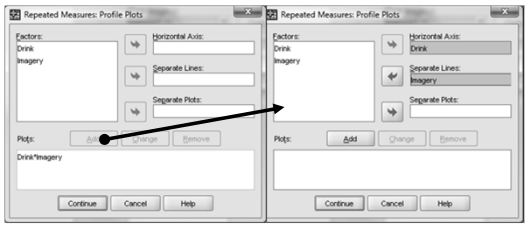
Defining Variables



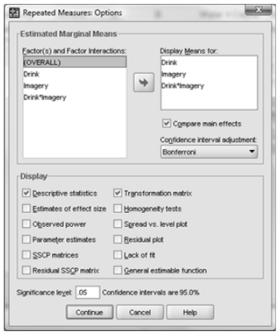
Defining Variables



Plots



Getting Means



Output: Sphericity

Mauchly's Test of Sphericity

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
DRINK	.267	23.753	2	.000	.577	.591	.500
IMAGERY	.662	7.422	2	.024	.747	.797	.500
DRINK * IMAGERY	.595	9.041	9	.436	.798	.979	.250

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the layers (by default) of the Tests of Within Subjects Effects table.

b. Design: Intercept - Within Subjects Design: DRINK+IMAGERY+DRINK*IMAGERY

The sphericity assumption is violated for the main effect of drink…so we should use the Greenhouse-Geisser correction

Output: Sphericity

Mauchly's Test of Sphericity

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
DRINK	.267	23.753	2	.000	.577	.591	.500
IMAGERY	.662	7.422	2	.024	.747	.797	.500
DRINK * IMAGERY	.595	9.041	9	.436	.798	.979	.250

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the layers (by default) of the Tests of Within Subjects Effects table.

b. Design: Intercept - Within Subjects Design: DRINK+IMAGERY+DRINK*IMAGERY

The sphericity assumption is violated for the main effect of imagery…so we should use the Greenhouse-Geisser correction

Output: Sphericity

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
DRINK	.267	23.753	2	.000	.577	.591	.500
IMAGERY	.662	7.422	2	.024	.747	.797	.500
DRINK * IMAGERY	.595	9.041	2	.436	.798	.979	.250

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the layers (by default) of the Tests of Within Subjects Effects table.

b. Design: Intercept - Within Subjects Design: DRINK+IMAGERY+DRINK * IMAGERY

The sphericity assumption is met for the interaction…so we do not need to correct the degrees of freedom

Output: Main ANOVA

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Drink	Sphericity Assumed	2092.344	2	1046.172	5.106	.011
	Greenhouse-Geisser	2092.344	1.154	1812.764	5.106	.030
	Huynh-Feldt	2092.344	1.181	1770.939	5.106	.029
	Lower-bound	2092.344	1.000	2092.344	5.106	.036
Error(Drink)	Sphericity Assumed	7785.878	38	204.892		
	Greenhouse-Geisser	7785.878	21.930	355.028		
	Huynh-Feldt	7785.878	22.448	346.836		
	Lower-bound	7785.878	19.000	409.783		
Imagery	Sphericity Assumed	21628.678	2	10814.339	122.565	.000
	Greenhouse-Geisser	21628.678	1.495	14468.490	122.565	.000
	Huynh-Feldt	21628.678	1.594	13571.496	122.565	.000
	Lower-bound	21628.678	1.000	21628.678	122.565	.000
Error(Imagery)	Sphericity Assumed	3352.878	38	88.234		
	Greenhouse-Geisser	3352.878	28.403	118.048		
	Huynh-Feldt	3352.878	30.280	110.729		
	Lower-bound	3352.878	19.000	176.467		
Drink * Imagery	Sphericity Assumed	2624.422	4	656.106	17.155	.000
	Greenhouse-Geisser	2624.422	3.194	821.778	17.155	.000
	Huynh-Feldt	2624.422	3.914	670.462	17.155	.000
	Lower-bound	2624.422	1.000	2624.422	17.155	.001
Error(Drink*Imagery)	Sphericity Assumed	2906.689	76	38.246		
	Greenhouse-Geisser	2906.689	60.678	47.903		
	Huynh-Feldt	2906.689	74.373	39.083		
	Lower-bound	2906.689	19.000	152.984		

The main effect for drink is significant

Output: Main ANOVA

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Drink	Sphericity Assumed	2092.344	2	1046.172	5.106	.011
	Greenhouse-Geisser	2092.344	1.154	1812.764	5.106	.030
	Huynh-Feldt	2092.344	1.181	1770.939	5.106	.029
	Lower-bound	2092.344	1.000	2092.344	5.106	.036
Error(Drink)	Sphericity Assumed	7785.878	38	204.892		
	Greenhouse-Geisser	7785.878	21.930	355.028		
	Huynh-Feldt	7785.878	22.448	346.836		
	Lower-bound	7785.878	19.000	409.783		
Imagery	Sphericity Assumed	21628.678	2	10814.339	122.565	.000
	Greenhouse-Geisser	21628.678	1.495	14468.490	122.565	.000
	Huynh-Feldt	21628.678	1.594	13571.496	122.565	.000
	Lower-bound	21628.678	1.000	21628.678	122.565	.000
Error(Imagery)	Sphericity Assumed	3352.878	38	88.234		
	Greenhouse-Geisser	3352.878	28.403	118.048		
	Huynh-Feldt	3352.878	30.280	110.729		
	Lower-bound	3352.878	19.000	176.467		
Drink * Imagery	Sphericity Assumed	2624.422	4	656.106	17.155	.000
	Greenhouse-Geisser	2624.422	3.194	821.778	17.155	.000
	Huynh-Feldt	2624.422	3.914	670.462	17.155	.000
	Lower-bound	2624.422	1.000	2624.422	17.155	.001
Error(Drink*Imagery)	Sphericity Assumed	2906.689	76	38.246		
	Greenhouse-Geisser	2906.689	60.678	47.903		
	Huynh-Feldt	2906.689	74.373	39.083		
	Lower-bound	2906.689	19.000	152.984		

The main effect for imagery is significant

Output: Main ANOVA

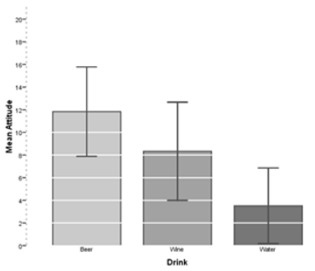
Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Drink	Sphericity Assumed	2092.344	2	1046.172	5.106	.011
	Greenhouse-Geisser	2092.344	1.154	1812.764	5.106	.030
	Huynh-Feldt	2092.344	1.181	1770.939	5.106	.029
	Lower-bound	2092.344	1.000	2092.344	5.106	.036
Error(Drink)	Sphericity Assumed	7785.878	38	204.892		
	Greenhouse-Geisser	7785.878	21.930	355.028		
	Huynh-Feldt	7785.878	22.448	346.836		
	Lower-bound	7785.878	19.000	409.783		
Imagery	Sphericity Assumed	21628.678	2	10814.339	122.565	.000
	Greenhouse-Geisser	21628.678	1.495	14468.490	122.565	.000
	Huynh-Feldt	21628.678	1.594	13571.496	122.565	.000
	Lower-bound	21628.678	1.000	21628.678	122.565	.000
Error(Imagery)	Sphericity Assumed	3352.878	38	88.234		
	Greenhouse-Geisser	3352.878	28.403	118.048		
	Huynh-Feldt	3352.878	30.280	110.729		
	Lower-bound	3352.878	19.000	176.467		
Drink * Imagery	Sphericity Assumed	2624.422	4	656.106	17.155	.000
	Greenhouse-Geisser	2624.422	3.194	821.778	17.155	.000
	Huynh-Feldt	2624.422	3.914	670.462	17.155	.000
	Lower-bound	2624.422	1.000	2624.422	17.155	.001
Error(Drink*Imagery)	Sphericity Assumed	2906.689	76	38.246		
	Greenhouse-Geisser	2906.689	60.678	47.903		
	Huynh-Feldt	2906.689	74.373	39.083		
	Lower-bound	2906.689	19.000	152.984		

The interaction is significant

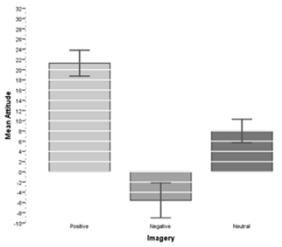
Main Effect of Drink



Needs MCPs to determine which means are different

There was a significant main effect of the type of drink for the attitude rating, $F(1.15, 21.93) = 5.11$, $p < .05$.

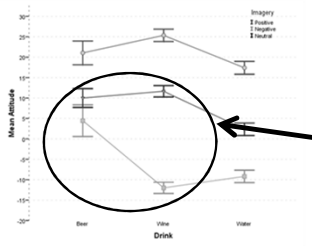
Main Effect of Imagery



Needs MCPs to determine which means are different

There was a significant main effect of the type of imagery for the attitude rating, $F(1.50, 28.40) = 122.57$, $p < .001$.

Drink by Dose Interaction



It appears that the interaction between drink and imagery is due to individuals reporting relatively positive attitudes when they drank beer while being exposed to negative imagery…but this would need to be confirmed using post-hoc tests

There was a significant interaction between the type of drink and the type of imagery for the attitude rating, $F(4, 76) = 17.16, p < .001$.