

# Repeated Measures ANOVA and Mixed Model ANOVA

Comparing more than two  
measurements of the same or  
matched participants

# Data files

- Fatigue.sav
- MentalRotation.sav
- AttachAndSleep.sav
- Attitude.sav
  
- Homework:
  - wordRecall2.sav
  - Perham & Sykora 2012
- Make-up homework: Bernard et al (2012)

# One-Way Repeated Measures ANOVA

- Used when testing more than 2 experimental conditions.
- In dependent groups ANOVA, all groups are dependent: each score in one group is associated with a score in every other group. This may be because the same subjects served in every group or because subjects have been matched.

# Characteristics of Within-Subjects Designs

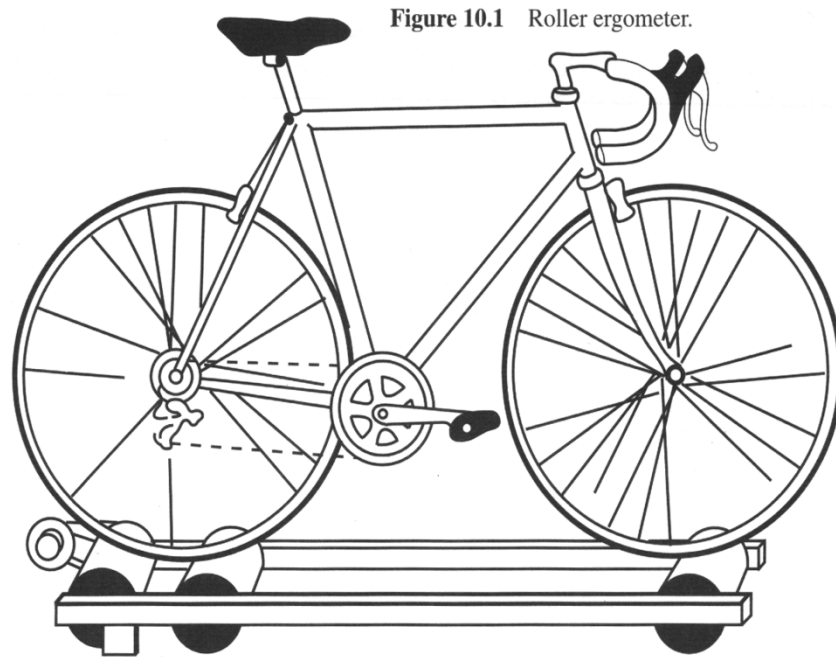
1. Each participant is exposed to all conditions of the experiment, and therefore, serves as his/her own control.
2. The critical comparison is the difference between the correlated groups on the dependent variable.
3. Susceptible to sequence effects, so the order of the conditions should be “counter-balanced”. In complete counter-balancing:
  - a. Each participant is exposed to all conditions of the experiment.
  - b. Each condition is presented an equal number of times.
  - c. Each condition is presented an equal number of times in each position.
  - d. Each condition precedes and follows each other condition an equal number of times.

## Advantages of Repeated Measures (within-subjects) over Independent Groups (between-subjects) ANOVA

- In repeated measures subjects serve as their own controls.
- Differences in means must be due to:
  - the treatment
  - variations within subjects
  - error (unexplained variation)
- Repeated measures designs are **more powerful** than independent groups designs.

## An example: Fatigue and balance (fatigue.sav)

- Example: Balance errors were measured five times, at five levels of fatigue. Fatigue is a within subjects factor with 5 levels.



Subjects rode for 15 minutes, divided into five 3-minute periods for the purpose of collecting data. Data were collected on the number of balance errors during the last minute of each 3-minute period, and resistance was increased at the end of each 3-minute period. In this design, the dependent variable is balance errors and the independent variable is increase in resistance (fatigue).

Roller Ergometer Data. Within Subjects Factor with 5 levels (3, 6, 9, 12, 15 min) – Balance errors/minute

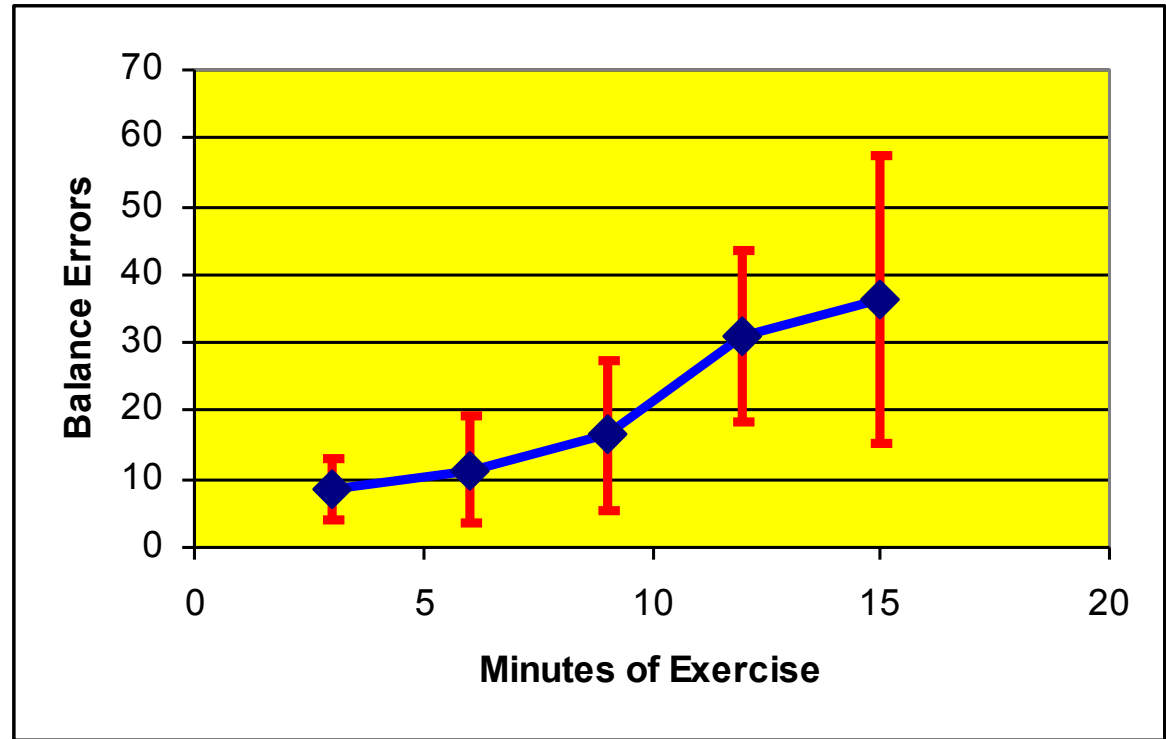
---

Subject	Minute 3	Minute 6	Minute 9	Minute 12	Minute 15
1	7	7	23	36	70
2	12	22	26	26	20
3	11	6	9	31	30
4	10	18	16	40	25
5	6	12	9	28	37
6	13	21	30	55	65
7	5	0	2	10	11
8	15	18	22	37	42
9	0	2	0	16	11
10	6	8	27	32	54

---

# Descriptives

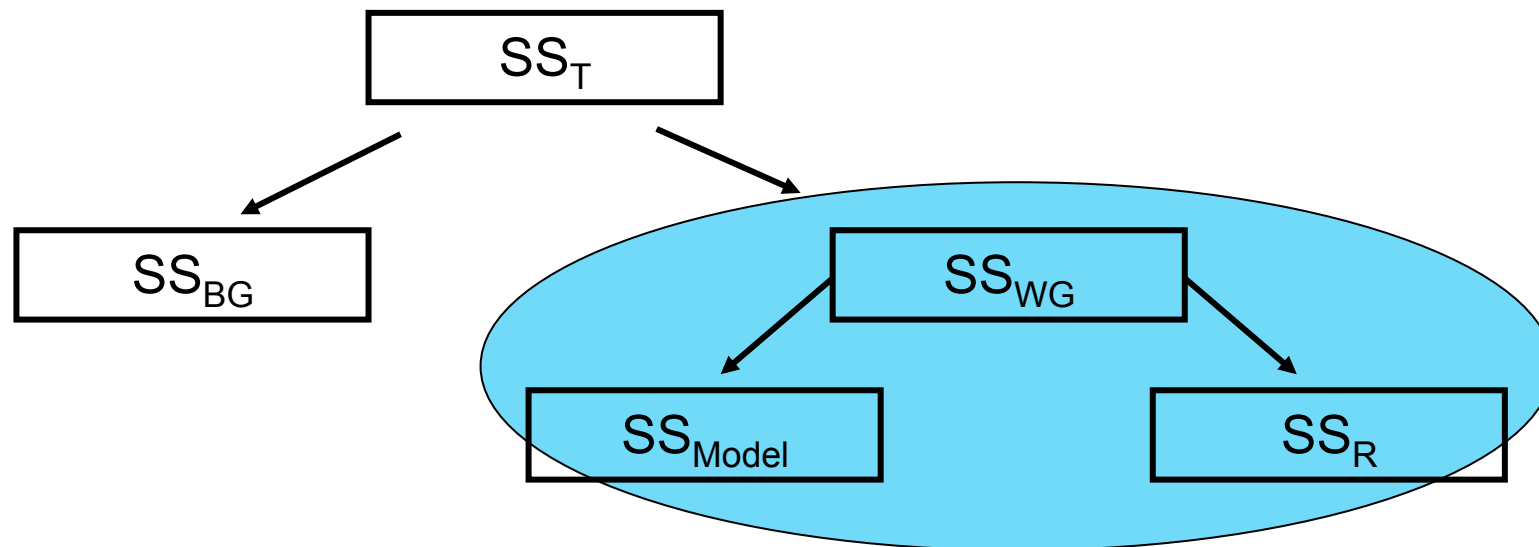
Minutes of Exercise	Balance Errors	sd
3	8.5	4.5
6	11.4	7.96
9	16.4	10.8
12	31.1	12.56
15	36.5	21.13





# Calculating One-Way Repeated Measures ANOVA

- variance is partitioned into  $SS_T$ ,  $SS_M$  and  $SS_R$
- in repeated-measures ANOVA, the model and residual sums of squares are both part of the within-group variance.



## Steps in the Analysis

- $SS_T$  = as before (squared difference between each score and the grand mean)
- $SS_{BG}$  = the squared difference between each condition mean and the grand mean multiplied by the number of subjects, summed
- $SS_{WG}$  = for each participant, difference between their individual scores and their condition mean, squared and summed
- $SS_M$  = squared difference between an individual score and the mean of the subject, summed, multiplied by number of conditions
- $SS_R = SS_{WG} - SS_M$  (the amount of within-group variation not explained by the experimental manipulation)
  
- Divide by the appropriate df:
  - (1) df for  $SS_M$  = levels of the IV minus 1 (=  $k - 1$ );
  - (2) df for  $SS_R = (k - 1) \times (n - 1)$  [ $n = \text{number of participants}$ ]
  
- $F = MS_M / MS_R$

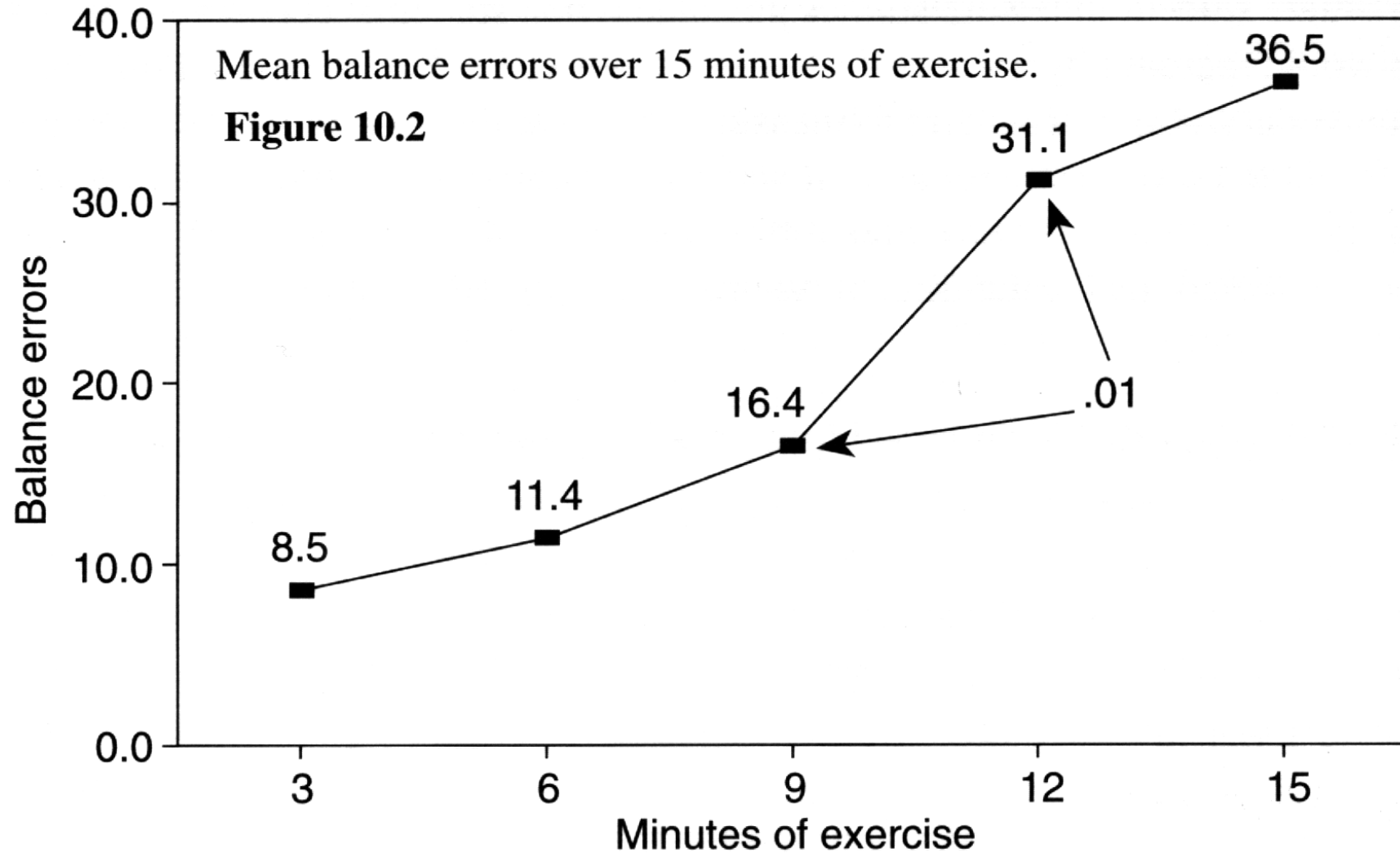
# Sphericity condition

- **Sphericity:** refers to the equality of variances of the differences between treatment levels.
  - If we were to take each pair of treatment levels and calculate the differences between each pair of scores, then it is necessary that these differences have equal variances.
- **Mauchly's test statistic**
  - If significant, the variances are significantly different from equal, and a correction must be applied to produce a valid F-ratio:  
Corrections applied to degrees of freedom to produce a valid F-ratio:
  - when G-G Sphericity Epsilon estimates  $< .75$ , use Greenhouse-Geisser estimate
  - When G-G sphericity Epsilon estimates  $> .75$ , use Huynh-Feldt estimate

# Running Repeated Measures ANOVA in SPSS

- Analyze > General Linear Model > Repeated Measures
  - Define factors
  - Options: descriptives, effect sizes and estimated marginal means for factors + pairwise comparisons for estimated marginal means
    - Estimated marginal means: population estimates calculated by adjusting the observed means for covariates
  - Contrasts (for factors with more than two levels):
    - Deviation. Compares the mean of each level (except a reference category) to the mean of all of the levels (grand mean). The levels of the factor can be in any order.
    - Simple. Compares the mean of each level to the mean of a specified level. You can choose the first or last category as the reference.
    - Difference. Compares the mean of each level (except the first) to the mean of previous levels.
    - Helmert. Compares the mean of each level of the factor (except the last) to the mean of subsequent levels.
    - Repeated. Compares the mean of each level (except the last) to the mean of the subsequent level.
    - Polynomial. Compares the linear effect, quadratic effect, cubic effect, and so on.

Polynomial, Repeated or Helmert contrast because we expect linear increase, or Bonferroni post-hoc tests



# SPSS Output

## Within-Subjects Factors

Measure: MEASURE\_1

treatmnt	Dependent Variable
1	Minute_3
2	Minute_6
3	Minute_9
4	Minute_12
5	Minute_15

## General Linear Model

### Descriptive Statistics

	Mean	Std. Deviation	N
Minute_3	8.5000	4.50309	10
Minute_6	11.4000	7.96102	10
Minute_9	16.4000	10.80329	10
Minute_12	31.1000	12.55610	10
Minute_15	36.5000	21.13055	10

### Multivariate Tests<sup>a</sup>

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power <sup>a</sup>
treatmnt	Pillai's Trace	.866	9.694 <sup>b</sup>	4.000	6.000	.009	.866	38.777	.934
	Wilks' Lambda	.134	9.694 <sup>b</sup>	4.000	6.000	.009	.866	38.777	.934
	Hotelling's Trace	6.463	9.694 <sup>b</sup>	4.000	6.000	.009	.866	38.777	.934
	Roy's Largest Root	6.463	9.694 <sup>b</sup>	4.000	6.000	.009	.866	38.777	.934

a. Computed using alpha = .05

b. Exact statistic

c.

Design: Intercept

Within Subjects Design: treatmnt

## Repeated Measure ANOVA Assumptions: Sphericity?

### Mauchly's Test of Sphericity<sup>b</sup>

Measure: MEASURE\_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon <sup>a</sup>		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
treatmnt	.024	27.594	9	.001	.371	.428	.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 average tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b.

Design: Intercept

Within Subjects Design: treatmnt

You don't want this to be significant.

Since Sphericity is violated here, we must use either the G-G or H-F adjusted ANOVAs

Mauchly's Test of Sphericity indicated that sphericity was violated [W(9) =27.59, p = .001]

This value is smaller than .75 so G-G correction is best.

## SPSS Output: Within Subjects Factors

### Tests of Within-Subjects Effects

Measure: MEASURE\_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power <sup>a</sup>
treatmnt	Sphericity Assumed	6115.880	4	1528.970	18.359	.000	.671	73.437	1.000
	Greenhouse-Geisser	6115.880	1.485	4117.754	18.359	.000	.671	27.268	.995
	Huynh-Feldt	6115.880	1.710	3575.916	18.359	.000	.671	31.400	.998
	Lower-bound	6115.880	1.000	6115.880	18.359	.002	.671	18.359	.967
Error(treatmnt)	Sphericity Assumed	2998.120	36	83.281					
	Greenhouse-Geisser	2998.120	13.367	224.289					
	Huynh-Feldt	2998.120	15.393	194.776					
	Lower-bound	2998.120	9.000	333.124					

a. Computed using alpha = .05

If Sphericity was okay then the statistics would be  $F(4,36) = 18.36$ ,  $p = .000$ , power = 1.000

But since Sphericity was violated we use the adjusted values:

$F(1.48,13.37) = 18.36$ ,  $p = .000$ , effect size or partial  $\eta^2 = .67$

(remember:  $\eta^2 = (SS_M / (SS_M + SS_R))$  .02 small, .13 medium, .26 large)



## Polynomial contrast:

### Tests of Within-Subjects Contrasts

Measure: MEASURE\_1

Source	time	Type III Sum of Squares	df	Mean Square	F	Sig.
time	Linear	5730.490	1	5730.490	25.244	.001
	Quadratic	154.350	1	154.350	2.813	.128
	Cubic	129.960	1	129.960	8.412	.018
	Order 4	101.080	1	101.080	2.823	.127
Error(time)	Linear	2043.010	9	227.001		
	Quadratic	493.864	9	54.874		
	Cubic	139.040	9	15.449		
	Order 4	322.206	9	35.801		

## SPSS Output: Between Subjects Effects

### Tests of Between-Subjects Effects

Measure: MEASURE\_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power <sup>a</sup>
Intercept	21590.420	1	21590.420	45.801	.000	.836	45.801	1.000
Error	4242.580	9	471.398					

a. Computed using alpha = .05

If we had a between subjects factor like Gender, the ANOVA results would be printed here.

# Reporting the Results

Participants' balance errors were measured after 3, 6, 9, 12 and 15 minutes of exercise on an ergometer.

The results of a One-Way Repeated Measures ANOVA show that the number of balance errors was significantly affected by fatigue,  $F(1.48, 13.36) = 18.36, p < .001$ . Since Mauchley's test of sphericity was violated, the Greenhouse-Geisser correction was used. Eta<sup>2</sup> effect size ( $\eta^2 = .67$ ) indicated that the effect of fatigue on balance errors was substantial. Bonferroni post-hoc tests comparing adjacent fatigue conditions revealed a significant difference in the number of balance errors between 9 and 12 minutes of exercise  $p = .001, \eta^2 = .78$ . No other comparisons were significant.

# Another example: visual search or mental rotation

Visual Search:

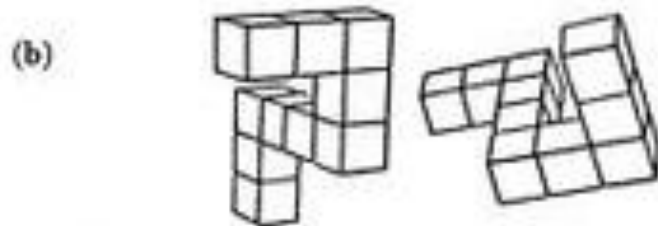
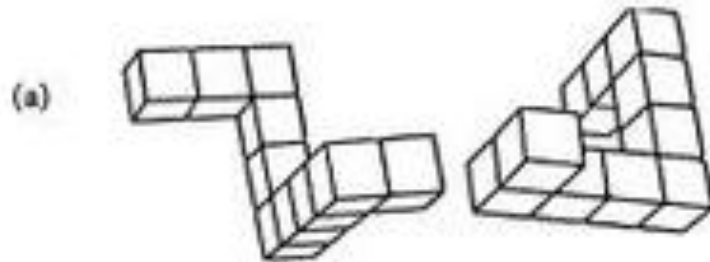
Speed of identifying a target as a function of number of distractor items: 10, 15, and 20 distractor items.

Mental Rotation:

Speed of deciding whether two 3D shapes are the same or different as a function of degree or rotation: 0, 45, 90, 135, 180, 225, 270, 315

Second variable: whether the shapes are the same or not

# Mental Rotation or Visual Search



**Mental Rotation Test**—Are these two figures the same except for their orientation?

P	P	P	P	P	P	P
P	P	P	P	P	P	P
P	P	P	P	P	P	P
P	P	P	P	P	P	P
P	P	P	P	P	P	P
P	B	P	P	P	P	P
P	P	P	P	P	P	P

	P					
		P		B		
			P			
					P	
		P				

# Sample problem: Mental Rotation or Visual search

<http://opl.apa.org/Experiments/Start.aspx?EID=30>

<http://www.gocognitive.net/demo/visual-search>

- Run the test
- Copy the results into excel or save in csv format
- Average the Reaction Times for each condition for Visual Search
- Create an SPSS data file
- Run the analysis

# Two-way repeated measures

- Two (or more) independent variables
- All are within-group variables – repeated measures
- Effects:
  - Main effect of Factor A
  - Main effect of Factor B
  - Interaction A x B
- In SPSS: define all within-group factors in General Lineral Model

# **MIXED MODEL ANOVA**



# Mixed Model ANOVA

- Two (or more) independent variables
  - Some within-subjects
  - Some between-subjects
- Effects:
  - Main effect of within-subject variable
  - Between-subject effect
  - Interaction

## **Sample Problem: Stress and partner**

The researcher conducts a study to determine whether the presence of a person's spouse while sleeping reduces the presence of sleep disturbances (reduction in deep (delta) sleep) in individuals who are stressed.

attachAndSleep.sav

# Method

**Participants.** 30 women who had recently moved to a new area to begin new jobs with their spouses. Among the women, 10 are secure, 10 are anxious, and 10 are avoidant in their attachment styles.

**Procedure.** The sleep patterns of the 30 women are monitored while they sleep alone and while they sleep with their spouses. The DV is the overall percentage of time spent in deep delta sleep.

**Design.** Two-way mixed ANOVA with one within-subjects factor and one between-groups factor. Partner-proximity (sleep with spouse vs. sleep alone) is the within-subjects factor; Attachment style is the between-subjects factor.

H1: Subjects will experience significantly greater sleep disturbances in the absence of their spouses due to the stressful nature of their present circumstances.

H2: Subjects with secure attachment styles will derive more comfort from the presence of their spouses and will experience a greater increase in deep delta sleep than subjects with insecure attachment styles.

# Data View

Attachment  
Style Key

- 1 = Secure
- 2 = Anxious
- 3 = Avoidant

	Subject	AttachStyle	AbsentDelta	PresentDelta	AVGSleep	v2
1	1	1	17.0	22.0	19.50	
2	2	1	16.0	25.0	20.50	
3	3	1	15.0	21.0	18.00	
4	4	1	18.0	23.0	20.50	
5	5	1	19.0	22.0	20.50	
6	6	1	20.0	25.0	22.50	
7	7	1	17.0	22.0	19.50	
8	8	1	15.0	21.0	18.00	
9	9	1	16.0	23.0	19.50	
10	10	1	15.0	22.0	18.50	
11	11	2	14.0	15.0	14.50	
12	12	2	15.0	17.0	16.00	
13	13	2	17.0	17.0	17.00	
14	14	2	19.0	20.0	19.50	
15	15	2	15.0	18.0	16.50	
16	16	2	14.0	15.0	14.50	
17	17	2	15.0	14.0	14.50	
18	18	2	13.0	15.0	14.00	
19	19	2	15.0	14.0	14.50	
20	20	2	15.0	16.0	15.50	
21	21	3	19.0	20.0	19.50	
22	22	3	15.0	15.0	15.00	
23	23	3	15.0	15.0	15.00	
24	24	3	17.0	18.0	17.50	
25	25	3	18.0	18.0	18.00	
26	26	3	19.0	19.0	19.00	
27	27	3	15.0	16.0	15.50	
28	28	3	15.0	17.0	16.00	
29	29	3	17.0	15.0	16.00	
30	30	3	16.0	16.0	16.00	

# Homogeneity Assessment

## Mauchly's Test of Sphericity<sup>b</sup>

Measure: MEASURE\_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon <sup>a</sup>		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Partner	1.000	.000	0	.	1.000	1.000	1.000

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 Tests of Within-Subjects Effects table.

b. Design: Intercept + AttachStyle  
Within Subjects Design: Partner

## Levene's Test of Equality of Error Variances<sup>a</sup>

	F	df1	df2	Sig.
AbsentDelta	.273	2	27	.763
PresentDelta	.622	2	27	.544

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + AttachStyle  
Within Subjects Design: Partner

# Main Analyses: Repeated Measures

Note:  
Partner "1" = Sleeping Partner Absent  
Partner "2" = Sleeping Partner Present

## 2. AttachStyle \* Partner

Measure: MEASURE\_1

Attach Style	Partner	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	16.800	.536	15.700	17.900
	2	22.600	.545	21.482	23.718
2	1	15.200	.536	14.100	16.300
	2	16.100	.545	14.982	17.218
3	1	16.600	.536	15.500	17.700
	2	16.900	.545	15.782	18.018

## Tests of Within-Subjects Effects

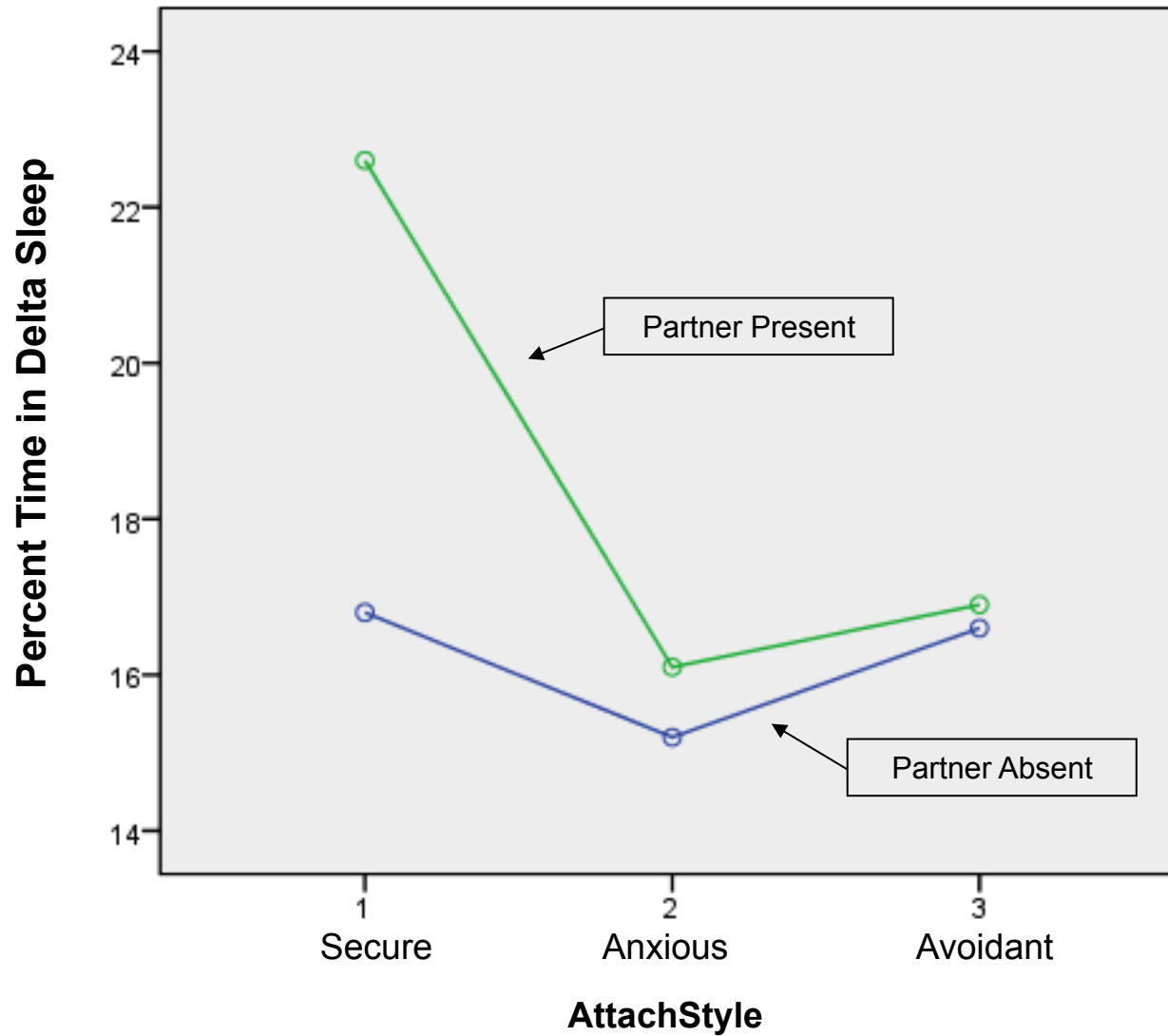
**Main effect of Partner**

Measure: MEASURE\_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Partner	Sphericity Assumed	81.667	1	81.667	90.741	.000	.771
	Greenhouse-Geisser	81.667	1.000	81.667	90.741	.000	.771
	Huynh-Feldt	81.667	1.000	81.667	90.741	.000	.771
	Lower-bound	81.667	1.000	81.667	90.741	.000	.771
Partner * AttachStyle	Sphericity Assumed	91.033	2	45.517	50.574	.000	.789
	Greenhouse-Geisser	91.033	2.000	45.517	50.574	.000	.789
	Huynh-Feldt	91.033	2.000	45.517	50.574	.000	.789
	Lower-bound	91.033	2.000	45.517	50.574	.000	.789
Error(Partner)	Sphericity Assumed	24.300	27	.900			
	Greenhouse-Geisser	24.300	27.000	.900			
	Huynh-Feldt	24.300	27.000	.900			
	Lower-bound	24.300	27.000	.900			

**Partner x Attachment Style Interaction**

Can you find the source of the interaction?



### Tests of Between-Subjects Effects

Measure:MEASURE\_1  
Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	18096.067	1	18096.067	3659.879	.000	.993
AttachStyle	175.433	2	87.717	17.740	.000	.568
Error	133.500	27	4.944			

### Estimates

Measure:MEASURE\_1

Attach Style	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Secure	19.700	.497	18.680	20.720
Anxious	15.650	.497	14.630	16.670
Avoidant	16.750	.497	15.730	17.770

### Pairwise Comparisons

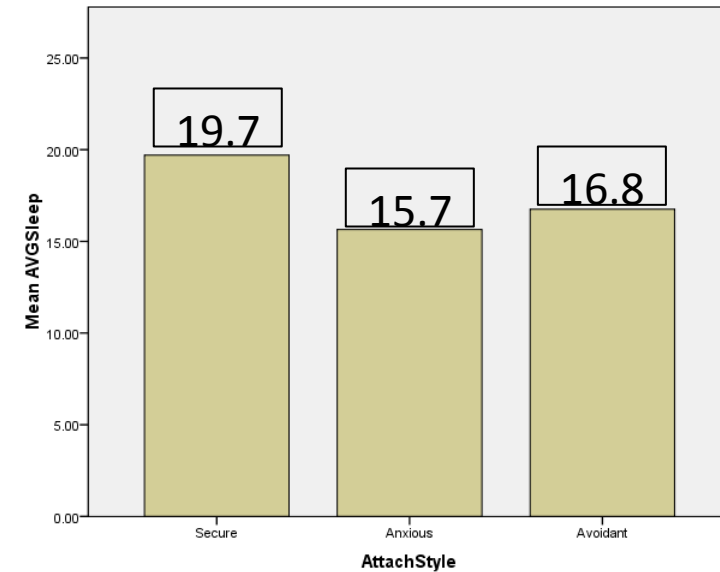
Measure:MEASURE\_1

(I) Attach Style	(J) Attach Style	Mean Difference (I-J)	Std. Error	Sig. <sup>a</sup>	95% Confidence Interval for Difference <sup>a</sup>	
					Lower Bound	Upper Bound
Secure	Anxious	4.050 <sup>*</sup>	.703	.000	2.255	5.845
	Avoidant	2.950 <sup>*</sup>	.703	.001	1.155	4.745
Anxious	Secure	-4.050 <sup>*</sup>	.703	.000	-5.845	-2.255
	Avoidant	-1.100	.703	.388	-2.895	.695
Avoidant	Secure	-2.950 <sup>*</sup>	.703	.001	-4.745	-1.155
	Anxious	1.100	.703	.388	-.695	2.895

Based on estimated marginal means

\*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni.





# Reporting

The sleep quality (percentage of time spent in delta sleep) of women with secure, anxious or avoidant attachment styles ( $N = 3 \times 10$ ) was measured when sleeping with and without their partners. If a harmonious relationship has a stress reducing effect, we expect sleep quality to improve in the presence of their partner especially for securely attached women. A  $3 \times 2$  ANOVA with Attachment Style as an independent factor and absence or Presence of Partner as a within-subjects factor was run.

The analysis revealed a main effect of Partner Presence ( $F(1, 27) = 90.74, p < .001$ ) in the predicted direction, a main effect of Attachment Style ( $F(2, 27) = 17.47, p < .001$ ) and an interaction between Partner Presence and Attachment Style ( $F(2, 27) = 50.57, p > .001$ ). As predicted, women with secure attachment styles slept better than either of the other two groups ( $p = .001$ ) and they experienced the greatest improvement in sleep quality by the presence of their partners.

# Exercise: Two-Way Repeated Measures

- attitude.sav:

The effects of advertising on various drinks. Full within-subjects design.

- Independent variable 1: type of drink (beer, wine, water)
- Independent variable 2: type of imagery associated with drink (negative, positive, neutral)
- Dependent variable: participants's rating of the drinks
- Run descriptives and a Two-way Repeated Measures ANOVA

# Homework

- Word recall 2:  
Lists of words had to be learnt: just words, words with pictures, words with pictures and sounds – number of items recalled measured
  - Run descriptives and a Repeated Measures ANOVA
  - Write up the results
- Perham & Sykora 2012: learning 8-item word lists to music
  - Independents:
    - Music (none, liked, disliked)
    - word position in list (1<sup>st</sup>, 2<sup>nd</sup> etc)

# Make-up homework: Bernard et al 2012

- Three-Way ANOVA
- Are women seen as objects?

A human face presented upside-down is more difficult to identify than an object presented upside-down. If women are seen as sexual objects, seeing a picture of a woman upright or inverted should make no difference in terms of recognition.

  - Within-subject variables
    - Gender of model in stimulus picture (men, women)
    - Orientation of picture (upright, inverted)
  - Between-subject variable:
    - Gender of participant (male, female)
  - Dependent variable: percentage of pictures correctly identified in a test where participants had to pick from a set the person whose picture they had seen