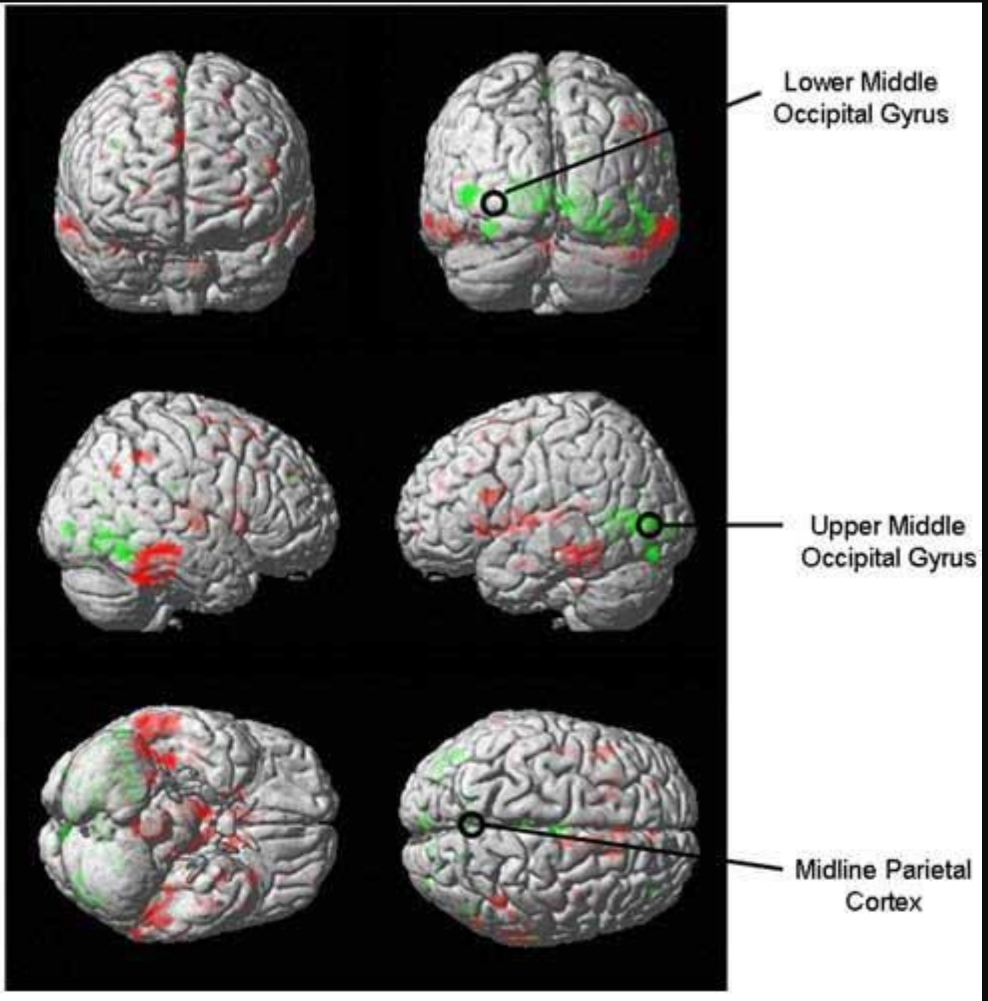

Brain Region Affected by Sleep Deprivation Identified

One-factor ANOVA example, plus effect source and size

Reminders

- Quiz 8 due Monday before midnight.
- Make sure to read chapter 16.

Inferential Statistics

<u>Continuing Analysis of Variance (ANOVA)</u>

- ✓ One factor *F* test
- ✓ Sources of variability
- ○ Example
- • Estimating general effect size with squared curvilinear correlation (η^2)
- • Localizing effects with Tukey's HSD test
- • Estimating specific effect size with Cohen's *d*
- • Introducing repeated measures ANOVA

Last time: One-factor ANOVA: Mean squares

$$F = \frac{variance\ between\ groups}{variance\ within\ groups}$$

$$variance = MS = \frac{SS}{df}$$

To calculate an *F* ratio, we will need to find two types of mean squares:

- Mean square for between groups: $\mathbf{MS_{between}}$
- Mean square for within groups: $\mathbf{MS_{within}}$

- Each of these values will have their own sum of squares (SS) and degrees of freedom (df) associated with them .

Last time: One-factor ANOVA: Calculating F

Total variability

Variability **between** groups (subjects in *different* groups receive different experimental treatments)

Variability **within** groups (subjects within the *same* group receive the same experimental treatment)

$$MS_{between} = \frac{SS_{between}}{df_{between}}$$

$$MS_{within} = \frac{SS_{within}}{df_{within}}$$

$$F = \frac{MS_{between}}{MS_{within}}$$

Last time: One-factor ANOVA: Example

Let's suppose we're interested in the effect of sleep deprivation on working memory. 9 participants are randomly assigned to 3 different sleep deprivation conditions: 0, 24 and 48 hours. We measure working memory loss (on a 10-point scale, with 10 being the greatest loss) for all participants.

- Is there an effect of sleep deprivation at the 0.05 level?

0 hours (control)	24 hours	48 hours
1	3	8
1	4	9
1	2	7
$\sum$ 3	9	24

All scores are **independent**: no repeated measures.

Last time: Hypothesis Testing: Sleep deprivation

1. State your research problem. Identify the population, distribution, and assumptions.
    - Does sleep deprivation affect short-term memory?
    - There are three independent samples, one independent variable – use one-factor ANOVA.
    - All ANOVAs are **non-directional**. They look for a <u>difference</u> among the samples.
    - We will assume a normal distribution.
    - No population standard deviation, so estimate mean squares.
    - Variability should be equal in all groups.

Last time: Hypothesis Testing: Sleep deprivation

2. State statistical hypotheses (both null and alternative) about the population.

- H_0: $\mu_1 = \mu_2 = \mu_3$
- H_1: The null is false.

Last time: One-factor ANOVA: Degrees of freedom

$$variability = MS = \frac{SS}{df}$$

Degrees of freedom are the number of deviation scores free to vary in your SS term. You lose 1 degree of freedom when comparing a score to a mean within your sample.

All ANOVAs are non-directional (positive difference).

$$MS_{between} = \frac{SS_{between}}{df_{between}}$$

$df_{between} = k - 1$ (where k = number of groups)

$$MS_{within} = \frac{SS_{within}}{df_{within}}$$

$df_{within} = N - k$ (where N = total number of scores)

SS_{total}

$df_{total} = N - 1$ (where N = total number of scores)

$$df_{total} = df_{between} + df_{within}$$

Hypothesis Testing: Sleep deprivation

$df_{between} = k - 1 = 2$

$df_{within} = N - k = 6$

TABLE B.3 THE *F* DISTRIBUTIONS

WITHIN-GROUPS *df*	SIGNIFICANCE (*p*) LEVEL	BETWEEN-GROUPS DEGREES OF FREEDOM 1	2	3	4	5	6
1	.01	4,052	5,000	5,404	5,625	5,764	5,859
	.05	162	200	216	225	230	234
	.10	39.9	49.5	53.6	55.8	57.2	58.2
2	.01	98.50	99.00	99.17	99.25	99.30	99.33
	.05	18.51	19.00	19.17	19.25	19.30	19.33
	.10	8.53	9.00	9.16	9.24	9.29	9.33
3	.01	34.12	30.82	29.46	28.71	28.24	27.91
	.05	10.13	9.55	9.28	9.12	9.01	8.94
	.10	5.54	5.46	5.39	5.34	5.31	5.28
4	.01	21.20	18.00	16.70	15.98	15.52	15.21
	.05	7.71	6.95	6.59	6.39	6.26	6.16
	.10	4.55	4.33	4.19	4.11	4.05	4.01
5	.01	16.26	13.27	12.06	11.39	10.97	10.67
	.05	6.61	5.79	5.41	5.19	5.05	4.95
	.10	4.06	3.78	3.62	3.52	3.45	3.41
6	.01	13.75	10.93	9.78	9.15	8.75	8.47
	.05	5.99	5.14	4.76	4.53	4.39	4.28
	.10	3.78	3.46	3.29	3.18	3.11	3.06
7	.01	12.25	9.55	8.45	7.85	7.46	7.19
	.05	5.59	4.74	4.35	4.12	3.97	3.87
	.10	3.59	3.26	3.08	2.96	2.88	2.83
8	.01	11.26	8.65	7.59	7.01	6.63	6.37
	.05	5.32	4.46	4.07	3.84	3.69	3.58
	.10	3.46	3.11	2.92	2.81	2.73	2.67
9	.01	10.56	8.02	6.99	6.42	6.06	5.80
	.05	5.12	4.26	3.86	3.63	3.48	3.37
	.10	3.36	3.01	2.81	2.69	2.61	2.55
10	.01	10.05	7.56	6.55	6.00	5.64	5.39
	.05	4.97	4.10	3.71	3.48	3.33	3.22
	.10	3.29	2.93	2.73	2.61	2.52	2.46
11	.01	9.65	7.21	6.22	5.67	5.32	5.07
	.05	4.85	3.98	3.59	3.36	3.20	3.10
	.10	3.23	2.86	2.66	2.54	2.45	2.39

Hypothesis Testing: Sleep deprivation

3. Characterize the sampling distribution. Form a decision rule, stating the critical values that will decide when we reject the null hypothesis.

- Degrees of freedom = (2,6)
- $\alpha = 0.05$
- Find F_{critical} from the F table.
- Reject H_0 at .05 level of significance if $F \geq$ **5.14**

Hypothesis Testing: Sleep deprivation

4. Collect a random sample and calculate your test statistic.

	0 hours (control)	24 hours	48 hours
	1	3	8
	1	4	9
	1	2	7
$\sum$	3	9	24

One-factor ANOVA: Calculating F

Steps for using a one-factor ANOVA

1. Confirm assumptions of test
   - ✓ Normal population (or large enough sample for Central Limit Theorem).
   - ✓ Samples are independent (each participant receives one treatment only
   - ✓ The variances in the population are approximately equal.

2. Find the $df_{between}$, df_{within}, df_{total} = 2, 6 and 8 (respectively).

One-factor ANOVA: Calculating *F*

<u>Steps for using a one-factor ANOVA</u>

3. Calculate the $SS_{between}$, SS_{within} , (and SS_{total})

$$SS_{between} = \sum \frac{T^2}{n} - \frac{G^2}{N}$$

$$SS_{within} = \sum X^2 - \sum \frac{T^2}{n}$$

	0 hours (control)	24 hours	48 hours
	1	3	8
	1	4	9
	1	2	7
$\sum$	3	9	24

One-factor ANOVA: Calculating *F*

Progress check

What does T represent?

$$\sum \frac{T^2}{n} - \frac{G^2}{N}$$

A. The total number of groups.
B. The total number of scores.
C. The sum of treatment scores in a group.
D. The sum of treatment scores in all groups.
E. I Totally forget.

One-factor ANOVA: Calculating *F*

Progress check

What does T represent?

$$\sum \frac{T^2}{n} - \frac{G^2}{N}$$

A. The total number of groups.
B. The total number of scores.
C. The sum of treatment scores in a group.
D. The sum of treatment scores in all groups.
E. I Totally forget.

C. Each group will have its own sum of treatment scores.

One-factor ANOVA: Calculating *F*

Steps for using a one-factor ANOVA

3. Calculate the $SS_{between}$, SS_{within} , (and SS_{total})

$$SS_{between} = \sum \frac{T^2}{n} - \frac{G^2}{N}$$

$$SS_{within} = \sum X^2 - \sum \frac{T^2}{n}$$

Where T = the treatment group score **total,**
n = number of observations per condition,
G = **Grand** total of scores, and
N = grand (combined) number of observations.

0 hours (control)	**24 hours**	**48 hours**
1	3	8
1	4	9
1	2	7
$\sum$ 3	9	24

One-factor ANOVA: Calculating *F*

Steps for using a one-factor ANOVA

3. Calculate the $SS_{between}$, SS_{within} , (and SS_{total})

$$SS_{between} = \sum \frac{T^2}{n} - \frac{G^2}{N} = \left[\frac{3^2}{3} + \frac{9^2}{3} + \frac{24^2}{3}\right] - \frac{36^2}{9} = [3 + 27 + 192] - 144 = \mathbf{78}$$

	0 hours (control)	24 hours	48 hours
	1	3	8
	1	4	9
	1	2	7
$\sum$	3	9	24

One-factor ANOVA: Calculating *F*

Steps for using a one-factor ANOVA

3. Calculate the $SS_{between}$, SS_{within}, (and SS_{total})

$$SS_{within} = \sum X^2 - \sum \frac{T^2}{n} = [1^2 + 1^2 + 1^2 + 3^2 + 4^2 + 2^2 + 8^2 + 9^2 + 7^2] - [3 + 27 + 192] = 226 - 222 = \mathbf{4}$$

	0 hours (control)	24 hours	48 hours
	1	3	8
	1	4	9
	1	2	7
$\sum$	3	9	24

One-factor ANOVA: Calculating *F*

Steps for using a one-factor ANOVA

4. Estimate the mean squares: $MS_{between}$ and MS_{within}.

$$MS_{between} = \frac{SS_{between}}{df_{between}} = \frac{78}{2} = \mathbf{39}$$

$$MS_{within} = \frac{SS_{within}}{df_{within}} = \frac{4}{6} = \mathbf{.667}$$

One-factor ANOVA: Calculating *F*

<u>Steps for using a one-factor ANOVA</u>

5. Calculate the *F* value.

$$F = \frac{MS_{between}}{MS_{within}} = \frac{39}{.667} = \mathbf{58.47}$$

Hypothesis Testing: Sleep deprivation

5. Compare the test statistic with the predictions made in your hypotheses. Decide whether to reject or fail to reject the null hypothesis.

- Reject H_0 at .05 level of significance if $F \geq$ **5.14**
- $F = 58.47$
- Decision: Reject H_0 at the 0.05 level of significance because the F value is greater than 5.14.

Source	***SS***	***df***	***MS***	***F***
Between	78	2	39	58.47**
Within	4	6	.667	
Total	82	8		

** Significant at the .01 level

Hypothesis Testing: Sleep deprivation

Interpretation

We rejected the null hypothesis [H_0: $\mu_1 = \mu_2 = \mu_3$] as mean working memory scores for participants sleep deprived for 0 ($M = 1$), 24 ($M = 3$), and 48 ($M = 8$) hours differ significantly [$F(2,6) = 58.47$; **MSE = .667**; $p < 0.05$]. <u>Sleep deprivation appears to affect working memory</u>.

MSE is the mean square error (MS_{within}). This just represents the error that should be present, without a treatment effect.

Note: We don't know <u>how large</u> of an effect we have found, only that it is statistically significant.

We also can't be sure of <u>which mean(s)</u> is responsible for the difference. The ANOVA only tells us a difference exists.

Hypothesis Testing: Sleep deprivation

Progress check

You read a research paper that reports: “$F(3,44) = 14.21$; MSE = 1.22; $p < 0.05$” in the results section. Assuming independent samples, how many participants were in the study?

A. 44
B. 45
C. 46
D. 47
E. 48

Hypothesis Testing: Sleep deprivation

Progress check

You read a research paper that reports: "$F(3,44) = 14.21$; $MSE = 1.22$; $p < 0.05$" in the results section. Assuming independent samples, how many participants were in the study?

A. 44

B. 45

C. 46

D. 47

E. 48

E. (df between, df within).
(k-1, N-k). 3+1+44=48

Hypothesis Testing: Sleep deprivation

Estimating general effect size

We can ask how much of the variance in the dependent variable can be explained by the presence of the independent variable.

The *squared curvilinear correlation coefficient* (η^2), called "eta-squared", expresses the proportion of variance explained by the independent variable. (With a small effect, less of the variance should be explained by the independent variable.)

$$\eta^2 = \frac{SS_{between}}{SS_{total}}$$

Hypothesis Testing: Sleep deprivation

$$\eta^2 = \frac{SS_{between}}{SS_{total}}$$

Progress check

In theory, what is the highest possible value for η^2?

A. 1
B. 100
C. Whatever $SS_{between}$ is.
D. Whatever SS_{total} is.
E. There should be no upper limit.

Hypothesis Testing: Sleep deprivation

$$\eta^2 = \frac{SS_{between}}{SS_{total}}$$

Progress check

In theory, what is the highest possible value for η^2?

A. 1
B. 100
C. Whatever $SS_{between}$ is.
D. Whatever SS_{total} is.
E. There should be no upper limit.

A. A value of 1 would mean that all of the variability in the scores is explained by the independent variable.

Hypothesis Testing: Sleep deprivation

$$\eta^2 = \frac{SS_{between}}{SS_{total}}$$

The estimated effect of sleep deprivation on short-term memory.

Source	*SS*	*df*	*MS*	*F*
Between	78	2	39	58.47**
Within	4	6	.667	
Total	82	8		

$$\eta^2 = \frac{78}{82} = .95$$

.01 Small
.09 Medium
.25 Large

This is a (very) large effect.

Hypothesis Testing

Understanding our result

The *F* ratio tells us that an effect likely exists. η^2 tells us how large the *overall* effect is. However, we still don't know *where* the effect comes from. That is, we don't know *which* of the groups provides the between group variance that makes the test significant.

- H_0: $\mu_1 = \mu_2 = \mu_3$
- H_1: The null is false.

How do we *localize* the effect and measure its *specific* size?

One-factor ANOVA: Finding the effect

Isolating the real effect

- Why not just use *t* tests?

- You wouldn't typically use multiple *t* tests, as this would inflate the probability of a type I error. This is a problem of multiple comparisons.

  - For each *t* test, the false alarm rate is 5%.
  - For <u>multiple</u> *t* tests, the false alarm rate is:

$$1 - (1 - \alpha)^c$$

Where α is the chosen level of significance and *c* is the number of independent *t* tests.

One-factor ANOVA: Finding the effect

However, you can do a **Bonferroni Correction** that adopts a much more conservative α and use that in your *t* tests. [It is so conservative that it increases your type II error rate.]

Original alpha (family-wise, related to all original estimates)

$$\alpha_{pc} = \frac{\alpha_{fw}}{c}$$

New alpha for paired comparisons

Number of comparisons you make

One-factor ANOVA: Finding the effect

A more typical (better) approach is to use Tukey's HSD test

- The Tukey HSD (Honestly Significant Difference) test compares all possible *absolute differences* in group <u>means</u> against a critical value that maintains the same type I error probability (i.e. same alpha level).

- The critical value (HSD) is

$$HSD = q\sqrt{\frac{MS_{within}}{n}}$$

Studentized range statistic, from Table G.

One-factor ANOVA: Finding the effect

$$HSD = q\sqrt{\frac{MS_{within}}{n}}$$

Tukey's HSD test

- Where is our effect significant?

	0 hours (control)	24 hours	48 hours
	1	3	8
	1	4	9
	1	2	7
$\bar{X}$	1	3	8

q is found in Table G using the number of groups (k = 3) and the df_{within} (N-k = 9 – 3 = 6).

One-factor ANOVA: Finding the effect

Critical Values of Studentized Range Distribution(q) for Familywise ALPHA = .05.

Denominator	Number of Groups (a.k.a. Treatments)							
DF	3	4	5	6	7	8	9	10
1	26.976	32.819	37.081	40.407	43.118	45.397	47.356	49.070
2	8.331	9.798	10.881	11.734	12.434	13.027	13.538	13.987
3	5.910	6.825	7.502	8.037	8.478	8.852	9.177	9.462
4	5.040	5.757	6.287	6.706	7.053	7.347	7.602	7.826
5	4.602	5.218	5.673	6.033	6.330	6.582	6.801	6.995
6	4.339	4.896	5.305	5.629	5.895	6.122	6.319	6.493
7	4.165	4.681	5.060	5.359	5.606	5.815	5.997	6.158
8	4.041	4.529	4.886	5.167	5.399	5.596	5.767	5.918
9	3.948	4.415	4.755	5.024	5.244	5.432	5.595	5.738
10	3.877	4.327	4.654	4.912	5.124	5.304	5.460	5.598

q is found in Table G using the number of groups (k = 3) and the df_{within} (N-k = 9 – 3 = 6).

In this case, *q* (at .05 level of significance) is **4.34**.

One-factor ANOVA: Finding the effect

$$HSD = q\sqrt{\frac{MS_{within}}{n}}$$

Tukey's HSD test

- Where is our effect significant?

	0 hours (control)	24 hours	48 hours
	1	3	8
	1	4	9
	1	2	7
$\bar{X}$	1	3	8

q = 4.34
n = 3
MS_{within} = .667

$$HSD = 4.34\sqrt{\frac{.667}{3}} = \mathbf{2.05}$$

One-factor ANOVA: Finding the effect

Tukey's HSD test

$$HSD = q\sqrt{\frac{MS_{within}}{n}}$$

- Where is our effect significant?

	0 hours (control)	24 hours	48 hours
	1	3	8
	1	4	9
	1	2	7
$\bar{X}$	1	3	8

$|\bar{X}_1 - \bar{X}_2|$ = 2

$|\bar{X}_2 - \bar{X}_3|$ = 5*

$|\bar{X}_1 - \bar{X}_3|$ = 7*

$$HSD = 4.34\sqrt{\frac{.667}{3}} = \mathbf{2.05}$$

One-factor ANOVA: Standardized effect size

Tukey's HSD test told us that we have significant differences between groups 2 and 3 ($|3 - 8| = 5$) and groups 1 and 3 ($|1 - 8| = 7$).

We can adapt Cohen's *d* to give us a standardized (unit-free) **effect size**.

$$d = \frac{mean\ difference}{standard\ deviation} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{MS_{within}}}$$

Groups 1 and 3 $\quad d = \frac{8-1}{\sqrt{.667}} = \frac{7}{.82} = 8.54$

Groups 2 and 3 $\quad d = \frac{8-3}{\sqrt{.667}} = \frac{5}{.82} = 6.10$

These would count as (very) large effects (see page 262).

Earlier: Hypothesis Testing: Sleep deprivation

Interpretation

We rejected the null hypothesis [H_0: $\mu_1 = \mu_2 = \mu_3$] as mean working memory scores for participants sleep deprived for 0 ($M = 1$), 24 ($M = 3$), and 48 ($M = 8$) hours differ significantly [$F(2,6) = 58.47$; **MSE = .667**; $p < 0.05$]. **Sleep deprivation appears to affect working memory**.

MSE is the mean square error (MS_{within}). This just represents the error that should be present, without a treatment effect.

Note: We don't know how large of an effect we have found, only that it is statistically significant.

$$\eta^2 = \frac{SS_{between}}{SS_{total}}$$

$$HSD = q\sqrt{\frac{MS_{within}}{n}}$$

We also can't be sure of which mean(s) is responsible for the difference. The ANOVA only tells us a difference exists.

Hypothesis Testing: Sleep deprivation

Intuition check

If we had used the *same participants* in all three conditions of our experiment, but without changing the values, how would this affect our experiment?

A. It would increase our variability.
B. It would decrease our variability.
C. It would have no effect.

Hypothesis Testing: Sleep deprivation

<u>Intuition check</u>

If we had used the *same participants* in all three conditions of our experiment, but without changing the values, how would this affect our experiment?

A. It would increase our variability.
B. It would decrease our variability.
C. It would have no effect.

B. It should reduce the variability due to individual differences.

Repeated measures ANOVA

As with our repeated measures *t* test, using repeated measures in an ANOVA should result in much less individual variability.

- This is also called a "within subjects ANOVA" because the we can examine the effect of treatment within the same subjects.
- The variability in scores *between groups* is assumed to be due to the **error + treatment** (if present).
- By (null) hypothesis, the *between group* variability should be small (no treatment effect) and only due to *error* variability.
- The amount of variability due to subjects should be captured by MS_{within}. We can quantify this and remove it from the MS_{within}.
- Remember what our independent *F* ratio consists of:

$$F = \frac{MS_{between}}{MS_{within}}$$

Remove variability due to subjects from here.

Repeated measures ANOVA

$$F = \frac{MS_{between}}{MS_{within}}$$

Mean square **within** groups:

$$MS_{within} = \frac{SS_{within}}{df_{within}}$$

All error (variability)

$$SS_{within} = \cancel{SS_{subject}} + SS_{error}$$

$$SS_{within} - SS_{subject} = SS_{error}$$

Where $SS_{subject}$ = individual variability

SS_{error} = all other error

Repeated measures ANOVA

If we calculate the $SS_{subject}$ then we can subtract this from the SS_{within} and all that remains is the SS_{error} (all residual variability).

$$SS_{within} = SS_{subject} + SS_{error}$$

$$SS_{error} = SS_{within} - SS_{subject}$$

$$MS_{error} = \frac{SS_{error}}{df_{error}}$$

$$F = \frac{MS_{between}}{MS_{error}}$$

Repeated measures ANOVA: *df*

$$MS_{within} = \frac{SS_{within}}{df_{within}}$$

$$SS_{within} = SS_{subject} + SS_{error}$$

Since we have some new *SS* terms, we will need to determine their degrees of freedom:

df_{total} = N – 1 (N = number of scores)
$df_{between}$ = k – 1 (k = number of treatment conditions)
df_{within} = N – k

$df_{subject}$ = n – 1 (n = number of subjects)
df_{error} = df_{within} - $df_{subject}$

Repeated Measures One-factor ANOVA

Calculating the $SS_{subject}$

	0 hours (control)	24 hours	48 hours	$T_{Subject}$	$\bar{X}_{Subject}$
	1	3	8	12	4
	1	4	9	14	4.67
	1	2	7	10	3.33
$\bar{X}$	1	3	8	36	4

Definition formula

$$SS_{subject} = k \sum(\bar{X}_{subject} - \bar{X}_{grand})^2$$

Where k = number of groups,
$\bar{X}_{subject}$ is the mean for each subject, and
$\bar{X}_{grand}$ is the mean of **all** scores recorded.

Computation formula

$$\sum \frac{T^2_{Subject}}{k} - \frac{G^2}{N}$$

Where $T_{subject}$ = the subject score **total,**
k = number of groups,
G = **Grand** total of scores, and
N = grand (combined) sample size for all conditions.

Repeated measures ANOVA

The test assumes that using the same people *actually* reduces variability. What other design issues need to be checked to ensure that the individual differences really will be smaller when using the same people?

- **Carryover effects**: Something about being in one treatment condition may affect your performance in the other conditions. Examples are the *practice* and *fatigue* effects. [Typically a problem in same-day experiments.]

- If the experiment runs over a longer period (i.e. weeks or months), you need to worry about **history** and **maturation** effects.
    - Did something external happen in the world between sessions that may affect the participants responses?
    - Did the participant physiologically change over time in a significant way?

Next time

- More repeated measures ANOVAs.