Lecture 11: Two-way (independent-measures) ANOVA

Starting from an example

		Factor *B*: Audience Condition	
		No Audience	Audience
Factor *A*: Self-Esteem	Low	Scores for a group of participants who are classified as low self-esteem and are tested with no audience.	Scores for a group of participants who are classified as low self-esteem and are tested with an audience.
	High	Scores for a group of participants who are classified as high self-esteem and are tested with no audience.	Scores for a group of participants who are classified as high self-esteem and are tested with an audience.

- Performance in a concept formation task: affected by self-esteem and the presence of audience.
- Instead of a **single factor**, we have multiple factors. So-called a **factorial design**. We have four conditions (**cells**, boxes) arranged as a matrix.

The questions behind two-way ANOVA

	No Audience	Audience	
Low Self-Esteem	$M = 7$	$M = 9$	$M = 8$
High Self-Esteem	$M = 3$	$M = 5$	$M = 4$
	$M = 5$	$M = 7$	

- **Main effect (for factor A or B)**

Does self-esteem level (A) and/or the presence of audience (B) make a difference? E.g., no matter high or low self esteem, will the presence of audience make performance better? No matter whether there is audience, will low self-esteem make performance better?

- **Interaction**

Unique combination makes a difference? E.g., does the combination of high self-esteem with absence of audience make the performance worse?

In formal language

Main effect refers to the mean difference among the levels of one factor.

$$H_1 : \mu_{A_1} \neq \mu_{A_2}$$

$$F = \frac{\text{variance for (differences between) the means for factor } A}{\text{variance (differences) expected if there is no treatment effect}}$$

$$F = \frac{\text{variance for (differences between) the row means}}{\text{variance (differences) expected if there is no treatment effect}}$$

$$H_1 : \mu_{B_1} \neq \mu_{B_2}$$

$$F = \frac{\text{variance for (differences between) the means for factor } B}{\text{variance (differences) expected if there is no treatment effect}}$$

$$F = \frac{\text{variance for (differences between) the column means}}{\text{variance (differences) expected if there is no treatment effect}}$$

Interaction

An interaction between two factors occurs whenever the mean differences between *individual treatment conditions, or cells,* are different from what would be predicted from the overall main effects of the factors.

Any "extra" mean differences that are not explained by the main effects are called an interaction, or an interaction between factors

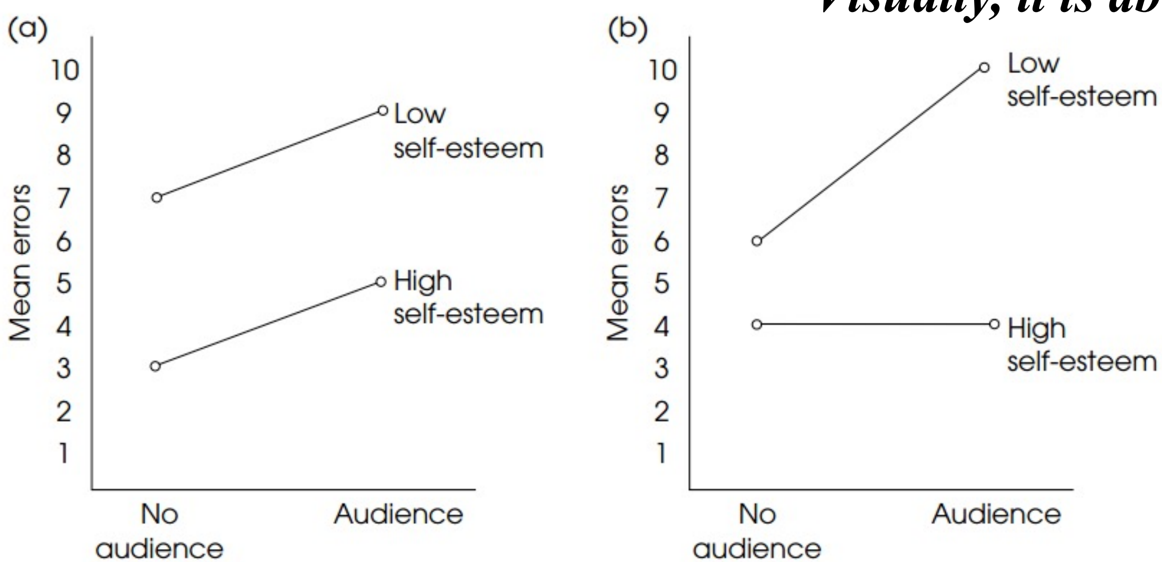
$$F = \frac{\text{variance (mean differences) not explained by the main effects}}{\text{variance (mean differences) expected if there are no treatment effects}}$$

$\mathbf{H_0}$: There is no interaction between factors A and B. All of the mean differences between treatment conditions are explained by the main effects of the two factors.
$\mathbf{H_1}$: There is an interaction between factors. The mean differences between treatment conditions are not what would be predicted from the overall main effects of the two factors.

More on interaction

When the effect of one factor depends on the different levels of a second factor, then there is an interaction between the factors.

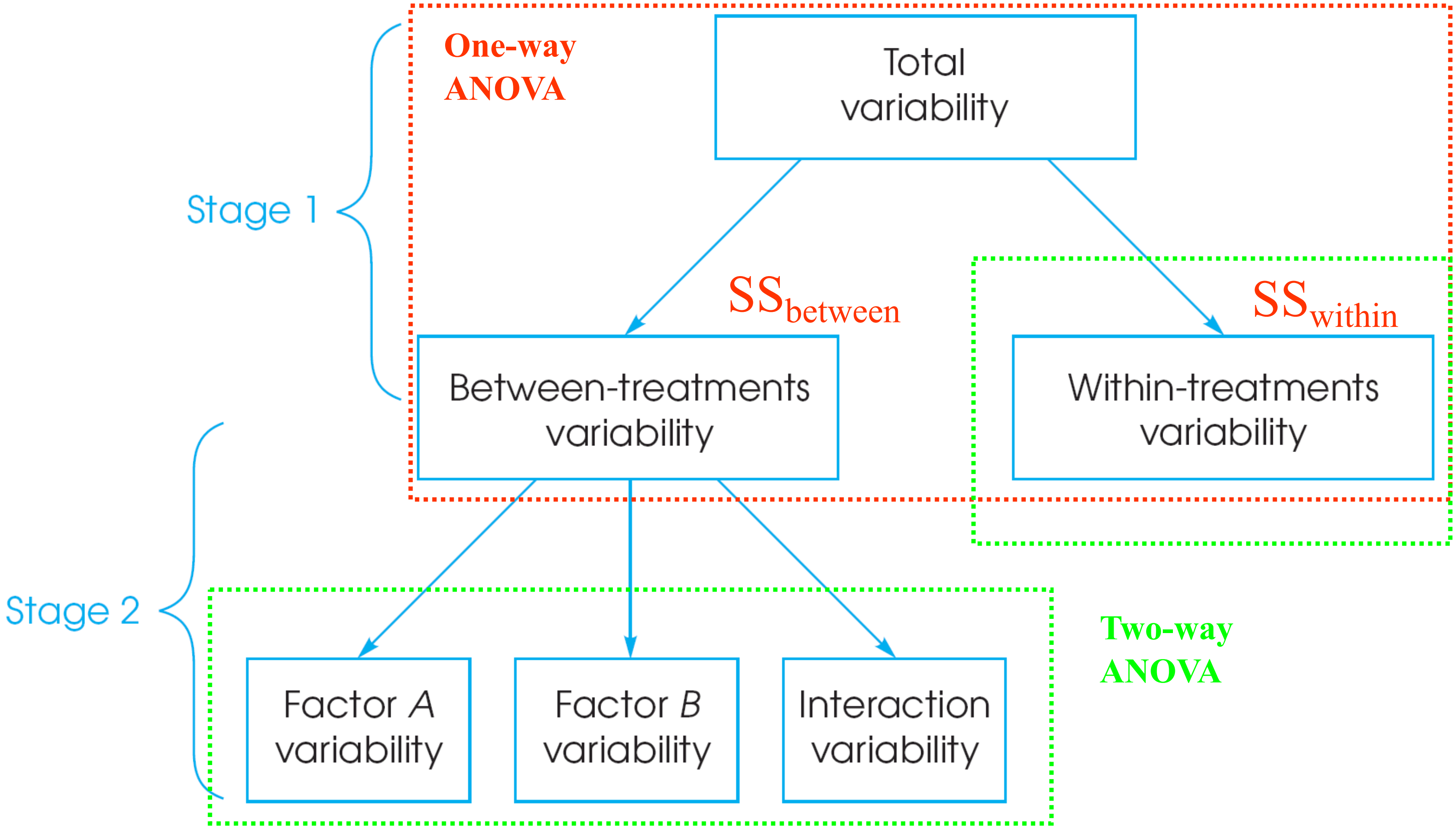
Visually, it is about parallel lines!



The effect of audience is the same with different self-esteem; the effect of self-esteem is also the same with different audience conditions。

The effect of audience depends on the self-esteem, and vice versa.

Partitioning of Variability



An example for computation

		Factor B: Text Presentation Mode			
		Paper	Computer Screen		
Factor A: Time Control	Self-regulated	11 8 9 10 7 $M = 9$ $T = 45$ $SS = 10$	4 4 8 5 4 $M = 5$ $T = 25$ $SS = 12$	$T_{row} = 70$	$N = 20$ $G = 155$ $\Sigma X^2 = 1303$
	Fixed	10 7 10 6 7 $M = 8$ $T = 40$ $SS = 14$	10 6 10 10 9 $M = 9$ $T = 45$ $SS = 12$	$T_{row} = 85$	
		$T_{col} = 85$	$T_{col} = 70$		

- Learning performance when working on computer screen or printed material, when study time is self-regulated or fixed by the researcher.
- n=5 for each condition

Computations (1)

Total variance

$$SS_{\text{total}} = \Sigma X^2 - \frac{G^2}{N}$$

$$SS_{\text{total}} = 1303 - \frac{155^2}{20}$$
$$= 1303 - 1201.25$$
$$= 101.75$$

$$df_{\text{total}} = \text{N} - 1 \ = 19$$

Within-treatment variance

$$SS_{\text{within treatments}} = \Sigma SS_{\text{each treatment}}$$

$$SS_{\text{within treatments}} = 10 + 12 + 14 + 12$$
$$= 48$$

$$df_{\text{within treatments}} = \Sigma df_{\text{each treatment}}$$

$$df_{\text{within treatments}} = 4 + 4 + 4 + 4$$
$$= 16$$

Computations (2)

Between treatment variance

$$SS_{\text{between treatments}} = \Sigma \frac{T^2}{n} - \frac{G^2}{N}$$

$$df_{\text{between treatments}} = \text{number of cells} - 1$$

$$SS_{\text{between treatments}} = \frac{45^2}{5} + \frac{25^2}{5} + \frac{40^2}{5} + \frac{45^2}{5} + \frac{155^2}{20}$$
$$= 405 + 125 + 320 + 405 - 1201.25$$
$$= 53.75$$

Factor A

$$SS_A = \Sigma \frac{T_{ROW}^2}{n_{ROW}} - \frac{G^2}{N}$$

$$SS_A = \frac{70^2}{10} + \frac{85^2}{10} - \frac{155^2}{20}$$
$$= 490 + 722.5 - 1201.25$$
$$= 11.25$$

$$df_A = \text{number of rows} - 1 \;= 1$$

Factor B

$$SS_B = \Sigma \frac{T_{COL}^2}{n_{COL}} - \frac{G^2}{N}$$

$$SS_B = \frac{85^2}{10} + \frac{70^2}{10} + \frac{155^2}{10}$$
$$= 725.5 + 490 - 1201.25$$
$$= 11.25$$

$$df_B = \text{number of columns} - 1 \;= 1$$

Computations (3)

Interaction AXB

$$SS_{A \times B} = SS_{\text{between treatments}} - SS_A - SS_B$$

$$SS_{A \times B} = 53.75 - 11.25 - 11.25$$

$$= 31.25$$

$$df_{A \times B} = df_{\text{between treatments}} - df_A - df_B$$

$$df_{A \times B} = df_A \times df_B \; = 1$$

Three separate F tests

$$MS_A = \frac{SS_A}{df_A} \quad MS_B = \frac{SS_B}{df_B} \quad MS_{A\times B} = \frac{SS_{A\times B}}{df_{A\times B}}$$

$$F_A = \frac{MS_A}{MS_{\text{within treatments}}} = \frac{11.25}{3} = 3.75$$

$$F_B = \frac{MS_B}{MS_{\text{within treatments}}} = \frac{11.25}{3} = 3.75$$

$$F_{A\times B} = \frac{MS_{A\times B}}{MS_{\text{within treatments}}} = \frac{31.25}{3} = 10.41$$

ANOVA table

Source	*SS*	*df*	*MS*	*F*
Between treatments	53.75	3		
Factor *A* (time control)	11.25	1	11.25	$F(1, 16) = 3.75$
Factor *B* (presentation)	11.25	1	11.25	$F(1, 16) = 3.75$
$A \times B$	31.25	1	31.25	$F(1, 16) = 10.42$
Within treatments	48	16	3	
Total	101.75	19		

Effect size for two-way ANOVA

$$\text{for factor } A,\ \eta^2 = \frac{SS_A}{SS_{\text{total}} - SS_B - SS_{A\times B}} = \frac{SS_A}{SS_A + SS_{\text{within treatments}}} = \frac{11.25}{11.25+48} = 0.190$$

$$\text{for factor } B,\ \eta^2 = \frac{SS_B}{SS_{\text{total}} - SS_A - SS_{A\times B}} = \frac{SS_B}{SS_B + SS_{\text{within treatments}}} = \frac{11.25}{11.25+48} = 0.190$$

$$\text{for } A\times B,\ \eta^2 = \frac{SS_{A\times B}}{SS_{\text{total}} - SS_A - SS_B} = \frac{SS_{A\times B}}{SS_{A\times B} + SS_{\text{within treatments}}} = \frac{31.25}{31.25+48} = 0.394$$

Again, these are partial eta squares. Thus, their sum will not be 100%

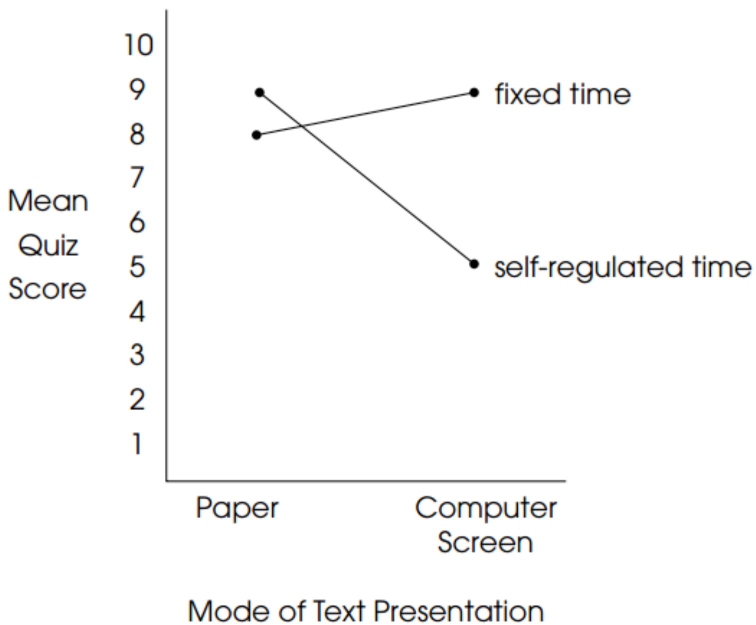
How to report

Paper	Computer Screen	
11	4	
8	4	
9	8	
10	5	$T_{row} = 70$
7	4	
$M = 9$	$M = 5$	
$T = 45$	$T = 25$	
$SS = 10$	$SS = 12$	
10	10	
7	6	
10	10	
6	10	
7	9	$T_{row} = 85$
$M = 8$	$M = 9$	
$T = 40$	$T = 45$	
$SS = 14$	$SS = 12$	
$T_{col} = 85$	$T_{col} = 70$	

$N = 20$
$G = 155$
$\Sigma X^2 = 1303$

The means and standard deviations for all treatment conditions are shown in Table 1. The two-factor analysis of variance showed no significant main effect for time control, $F(1, 16) = 3.75$, $p > .05$, $\eta^2 = 0.190$, or for presentation mode, $F(1, 16) = 3.75$, $p > .05$, $\eta^2 = 0.190$. However, the interaction between factors was significant, $F(1, 16) = 10.41$, $p < .01$, $\eta^2 = 0.394$.

Make sense of these results

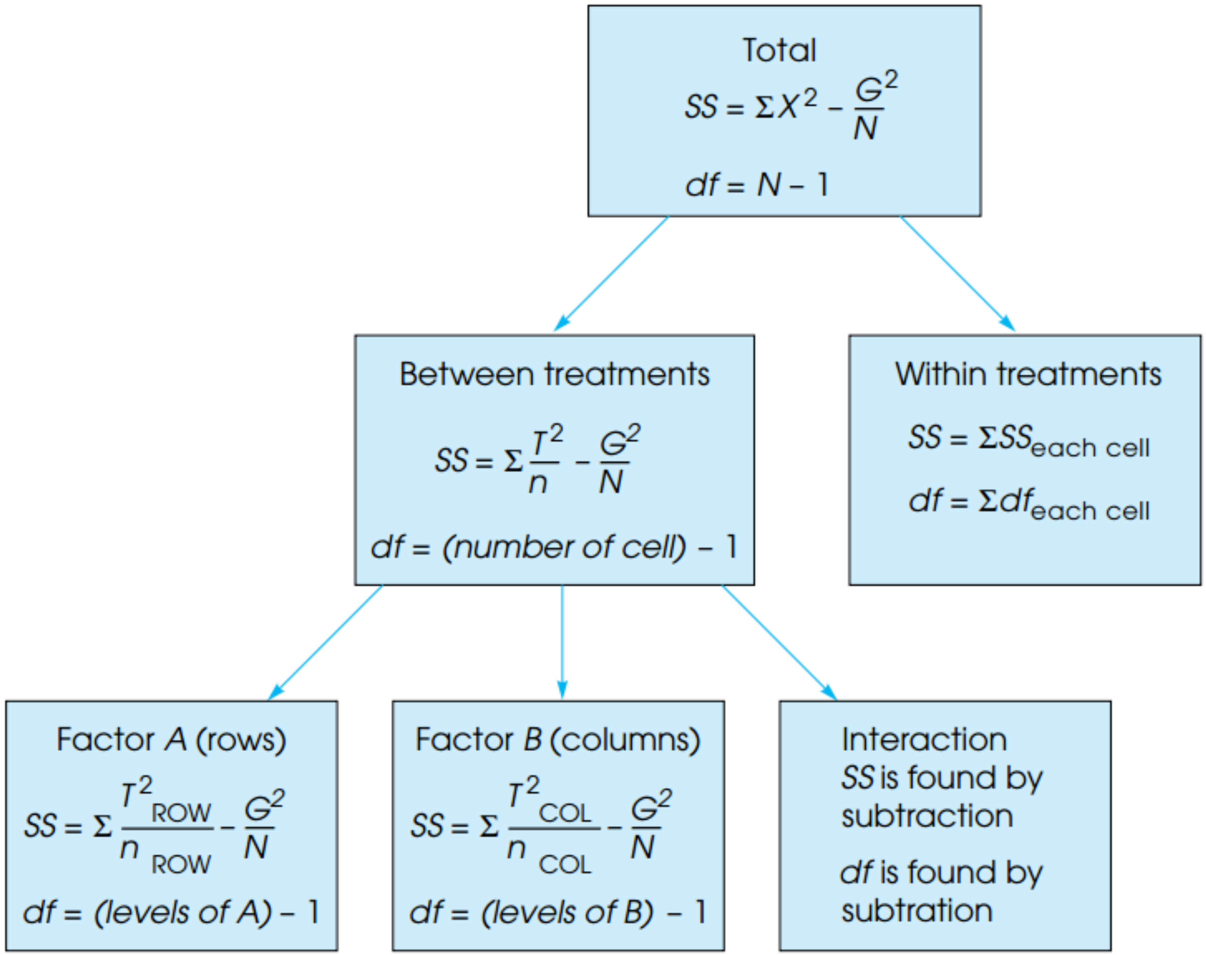


- Though the main effect of presentation mode is not significant, we will need to **be cautious to take this main effect at face value**.
- **With significant interaction**, we can only say that there is no consistent main effect.
- Here, if time is fixed, no difference between paper and computer screen. BUT, if time is self-regulated, paper produced higher quiz score than computer screen.

The assumptions of two-way ANOVA

1. The observations within each sample must be independent.
2. The populations from which the samples are selected must be normal (less of a concern if sample size is large)
3. The populations from which the samples are selected must have equal variances (homogeneity of variance).

Summary



$$MS_{\text{factor}} = \frac{SS \text{ for the factor}}{df \text{ for the factor}} \qquad MS_{\text{within}} = \frac{SS \text{ within treatments}}{df \text{ within treatments}}$$

Simple main effect

Another example

Humans have different ability to detect the scent of a chemical that is thought to have pheromonal properties. Moreover, we know gender and smoking history can affect our ability of detection. So, we designed a study with five groups of people with different smoking history (quit smoking for how long), each group further dividing into a male and a female group.

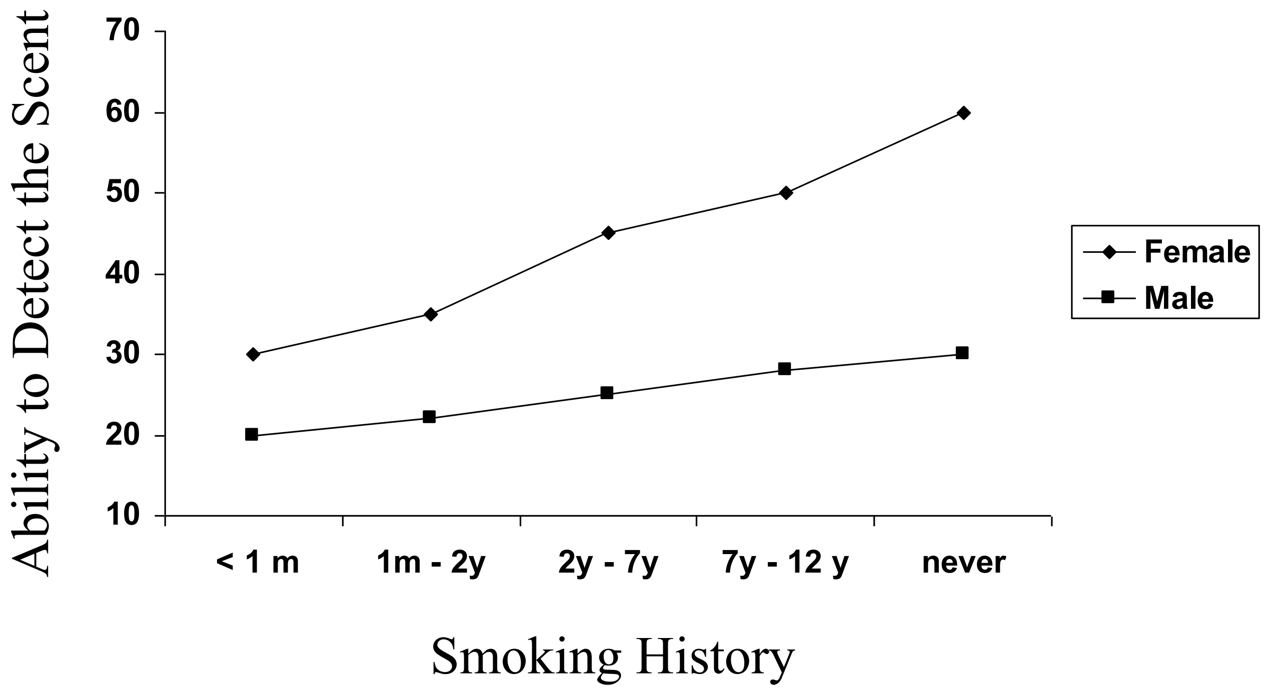
Smoking History

Gender

	never	< 1m	1 m - 2 y	2 y - 7 y	7 y - 12 y
male	*cell, or treatments*				
female					

Two-way, or two factors

ANOVA results so far



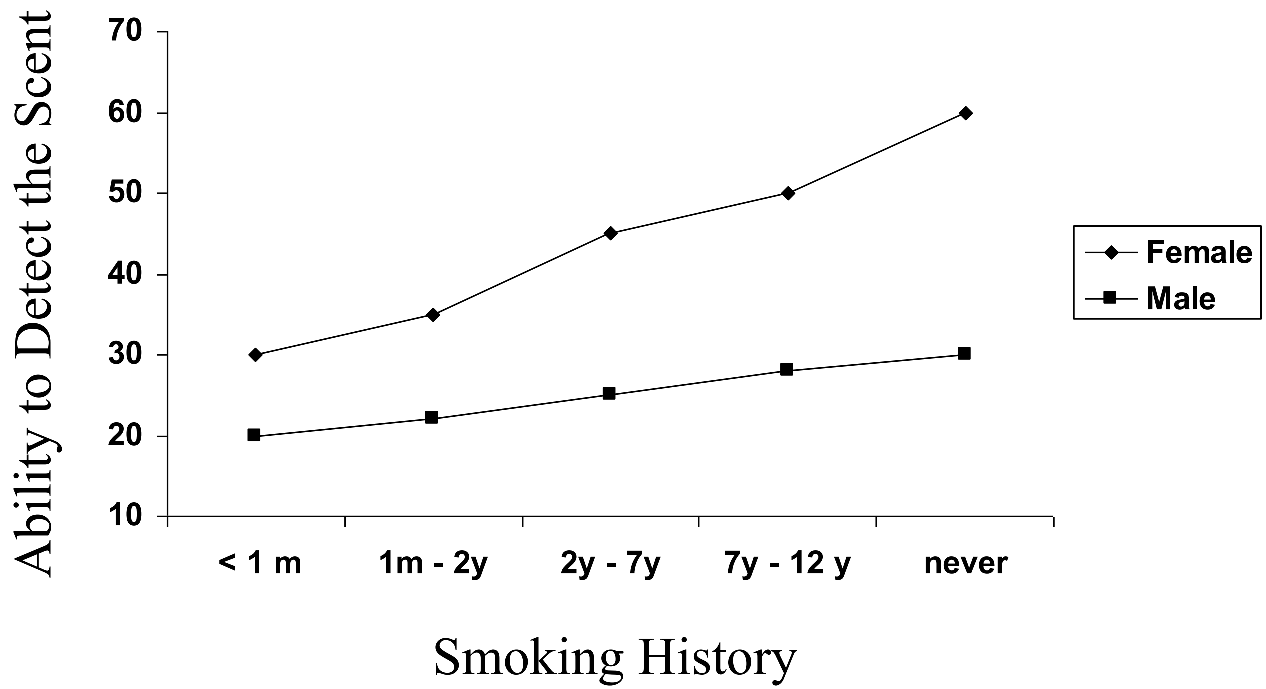
Source	*SS*	*df*	*MS*	*F*	*p*
A-gender	9025	1	9025	75.84	<.001
B-smoking history	5140	4	1285	10.80	<.001
AxB interaction	1240	4	310	2.61	.041
Error	10,710	90	119		
Total	26,115	99			

Simple main effect

Further questions:

■Fixing gender (at a specific level of gender), any difference between smoking history?

■Fixing smoking history, any difference between genders?



Simple Main Effects of <u>Gender</u>

GENDER	SMOKING HISTORY never	< 1m	1 m - 2 y	2 y - 7 y	7 y - 12 y	Sum
Male	300	200	220	250	280	1,250 (25)
Female	600	300	350	450	500	2,200 (44)
Sum	900 (45)	500 (25)	570 (28.5)	700 (35)	780 (39)	3,450

$$SS_{\text{Gender, never smoked}} \quad \frac{300^2 + 600^2}{10} - \frac{900^2}{20} = 4{,}500$$

$$SS_{\text{Gender, stopped < 1m}} \quad \frac{200^2 + 300^2}{10} - \frac{500^2}{20} = 500$$

$$SS_{\text{Gender, stopped 1m - 2y}} \quad \frac{220^2 + 350^2}{10} - \frac{570^2}{20} = 845$$

$$SS_{\text{Gender, stopped 2y - 7y}} \quad \frac{250^2 + 450^2}{10} - \frac{700^2}{20} = 2{,}000$$

$$SS_{\text{Gender stopped 7y - 12 y}} \quad \frac{280^2 + 500^2}{10} - \frac{780^2}{20} = 2{,}420$$

Simple Main Effects of <u>Gender</u>

Unchanged from normal ANOVA

- $F = MS_{effect} \,/\, MS_{within} = MS_{effect} \,/\, MSE$
- $MS_{effect} = SS \,/\, df.$ (df=1 here)
- $MS_{within} = SS_{within} / df_{within}$ (df=90 here)

	Smoking History				
SS Gender at	never	< 1m	1 m - 2 y	2 y - 7 y	7 y - 12 y
F(1, 90)	37.82	4.20	7.10	16.81	20.34
p	<.001	.043	.009	<.001	<.001

Simple Main Effects of Smoking

	SMOKING HISTORY					
GENDER	never	< 1m	1 m - 2 y	2 y - 7 y	7 y - 12 y	Sum
Male	300	200	220	250	280	1,250 (25)
Female	600	300	350	450	500	2,200 (44)
Sum	900 (45)	500 (25)	570 (28.5)	700 (35)	780 (39)	3,450

- $SS_{\text{Smoking history for men}}$

$$\frac{300^2 + 200^2 + 220^2 + 250^2 + 280^2}{10} - \frac{1{,}250^2}{50} = 680$$

- $SS_{\text{Smoking history for women}}$

$$\frac{600^2 + 300^2 + 350^2 + 450^2 + 500^2}{10} - \frac{2{,}200^2}{50} = 5{,}700$$

- Smoking history had a significant simple main effect for women, $F_{effect,woman} = F(4, 90) = (5700/4)/(10710/90) = 11.97$, $p < .001$; but not for men, $F_{effect,man} = F(4, 90) = (680/4)/(10710/90) = 1.43$, $p = .23$.

$F_{4,\ 90,0.05}$=2.46, $F_{4,\ 90,0.01}$=2.48

Multiple Comparisons Involving A Simple Main Effect

- Smoking had a significant simple main effect for women.
- But this only means that there is at least one pair-wise comparison yield significant result.
- There are 5 smoking groups. we could make 10 pairwise comparisons.
- Here we make only 4 comparisons; we compare each group of ex-smokers with those who never smoked.

Female Ex-Smokers vs. Never Smokers

It is two-sample t test again, now we use Tukey HSD test.

Compute the HSD with MS_{within} unchanged

- Using table 5: K (#group) = 5 and within-group df = 90 and α=0.05, so **q = 3.98** (when 60<df=90<120 in the table, take the conservative number 60).
- n = 10 for each cell.
- Alpha unchanged.

$$HSD = q\sqrt{MS_{within} / n} = 3.98\sqrt{119/10} = 13.73$$

Compare difference of conditions with this HSD:

Never Smoked vs Some Smoking History	$M_i - M_j$	Significant?
< 1 m	60-30=30	yes
1 m - 2 y	60-35=25	yes
2 y - 7 y	60-45=15	yes
7 y - 12 y	60-50=10	no

Female Ex-Smokers vs. Never Smokers

Two-sample t-test, but with **Bonferroni correction**.

1. The sample variance is adjusted: using the overall variance

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{MS_{within}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$\sqrt{119(1/10 + 1/10)} = 4.8785$$

Original two-sample independent t test.

MS_{within} (MS_{error}) for all treatments, instead of two

2. The alpha level is adjusted

$$\alpha' = {.05}/{4} = .0125.$$

Alpha is reduced accordingly: 4 comparisons then divided by 4.

Never Smoked vs Quit	$t(90) = \frac{M_i - M_j}{4.8785}$	p	Significant?
< 1 m	(60-30) / 4.8785=6.149	< .001	yes
1 m - 2 y	(60-35) / 4.8785=5.125	< .001	yes
2 y - 7 y	(60-45) / 4.8785=3.075	.0028	yes
7 y - 12 y	(60-50) / 4.8785=2.050	.0433	no

Presenting the Results

Participants were given a test of their ability to detect the scent of a chemical thought to have pheromonal properties in humans. Each participant had been classified into one of five groups based on his or her smoking history. A 2 x 5, Gender x Smoking History, ANOVA was employed, using a .05 criterion of statistical significance and a *MSE* of 119 for all effects tested. There were significant main effects of gender, $F(1, 90) = 75.84$, $p < .001$, and smoking history, $F(4, 90) = 10.80$, $p < .001$, as well as a significant interaction between gender and smoking history, $F(4, 90) = 2.61$, $p = .041$. As shown in Table 1, women were better detecting the scent than were men, and smoking reduced the scent detection ability, with recovery of function being greater the longer the period since the participant had last smoked.

The significant interaction was further investigated with tests of the simple main effect of smoking history. For men, the effect of smoking history fell short of statistical significance, $F(4, 90) = 1.43$, $p = .23$. For women, smoking history had a significant effect on the ability of detection, $F(4, 90) = 11.97$, $p < .001$. This significant simple main effect was followed by a set of four contrasts. Each group of female ex-smokers was compared with the group of women who had never smoked. The Bonferroni inequality was employed to cap the familywise error rate at .05 for this family of four comparisons. It was found that the women who had never smoked had a significantly better ability to detect the scent than did women who had quit smoking one month to seven years earlier, but the difference between those who never smoked and those who had stopped smoking more than seven years ago was too small to be statistically significant.