STATISTIKA PENELITIAN

Anova

Aryan Eka Prastya Nugraha

2017

The one-way analysis of variance (ANOVA) is an extension of the independent *t* test. It is used when the researcher is interested in whether the means from several (>2) independent groups differ. For example, if a researcher is interested in investigating whether four ethnic groups differ in their IQ scores, the one-way ANOVA can be used.

Checklist of Requirements

- In any analysis, there must be only one independent variable (e.g., ethnicity).
- There should be more than two levels for that independent variable (e.g., Australian, American, Chinese, African).
- There must be only one dependent variable.

Assumptions

- **Normality**—The dependent variable is normally distributed.
- **Homogeneity of variance**—The groups have approximately equal variance on the dependent variable.

Example

A researcher is interested in finding out whether intensity of electric shock will affect the time required to solve a set of diffcult problems. Eighteen subjects are randomly assigned to the three experimental conditions of “Low Shock,” “Medium Shock,” and “High Shock.” The total time (in minutes) required to solve all the problems is the measure recorded for each subject.

Shock Intensity					
Low		Medium		High	
s1	15	s7	30	s13	40
s2	10	s8	15	s14	35
s3	25	s9	20	s15	50
s4	15	s10	25	s16	43
s5	20	s11	23	s17	45
s6	18	s12	20	s18	40

The data set has been saved under the name **EX6.SAV**

Variables	Column(s)	Code
• SHOCK	• 1	• 1 = low, 2 = medium, 3 = high
• TIME	• 2	• Time in minutes

Testing Assumptions
Normality

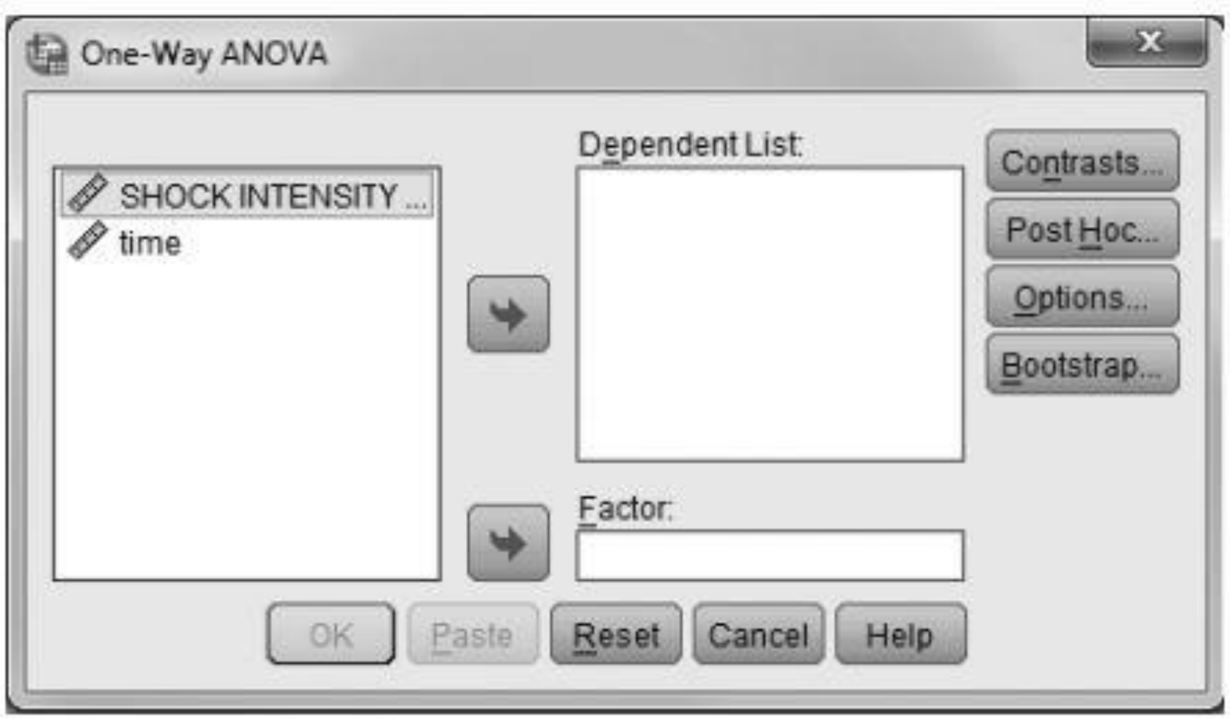
Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	*df*	Sig.	Statistic	*df*	Sig.
Time	.182	18	.118	.921	18	.136

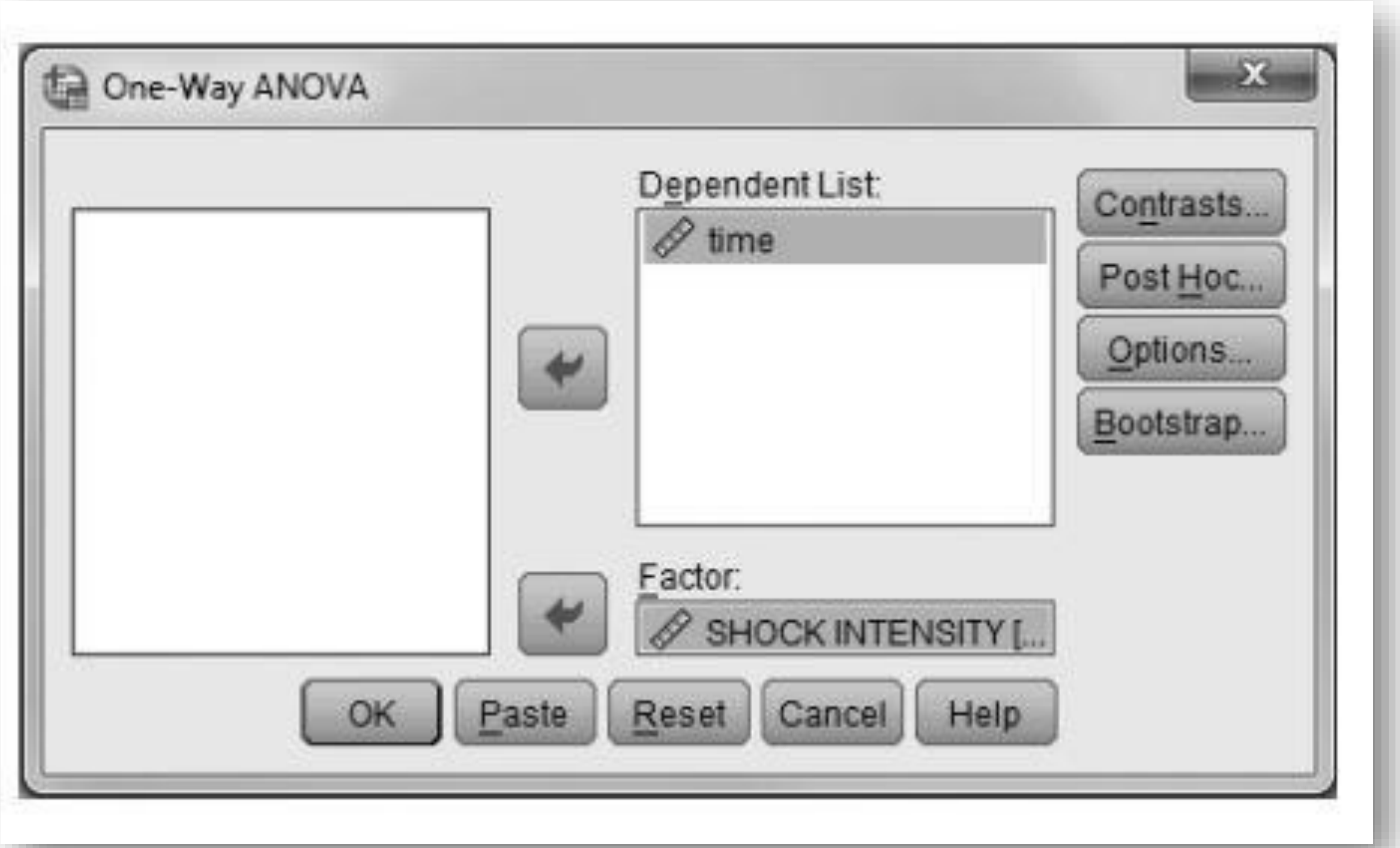
[a] Lilliefors significance correction.

Homogeneity of Variance
Windows Method: One-Way ANOVA

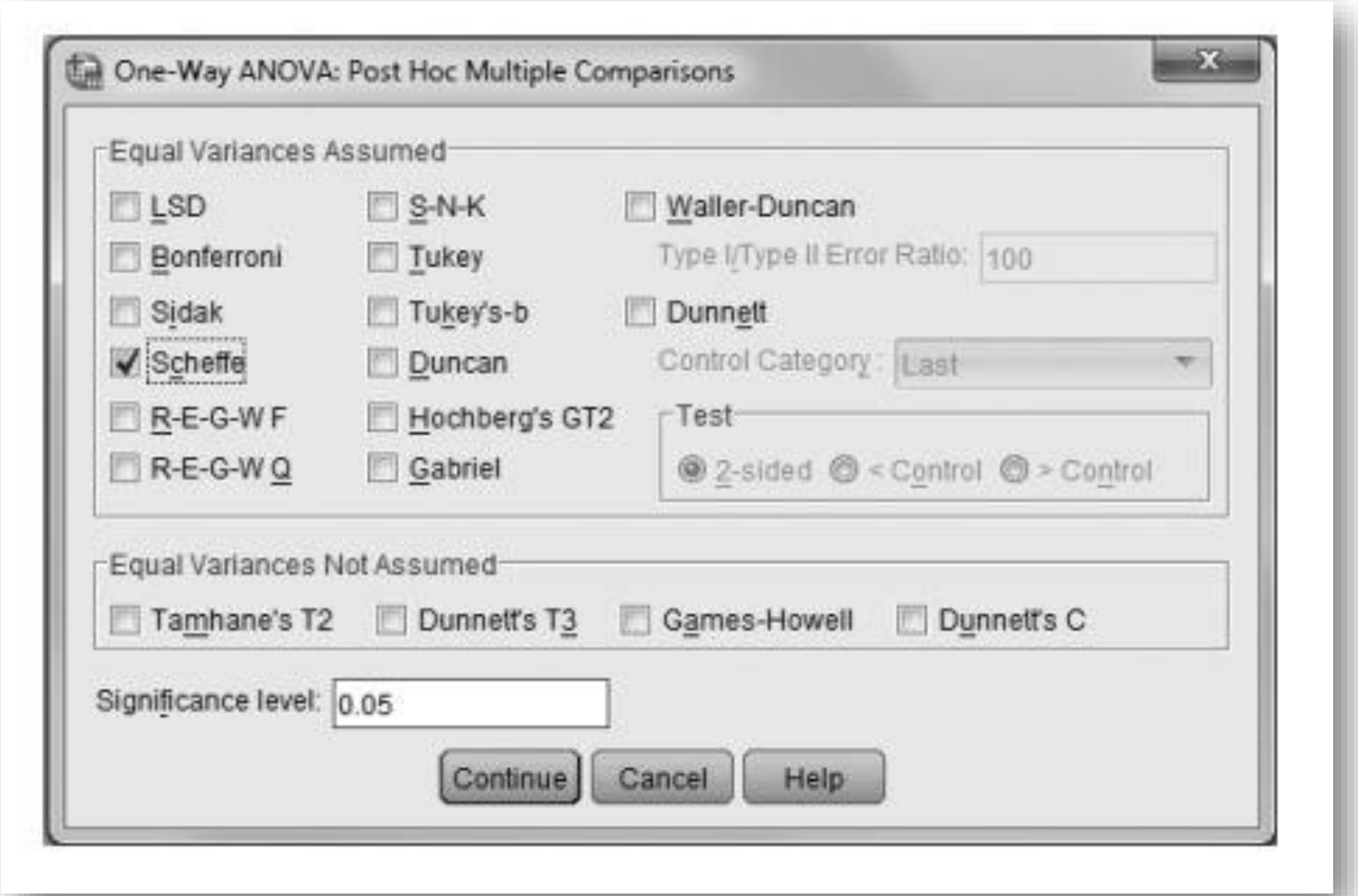
1. From the menu bar, click **Analyze**, then **Compare Means**, and then **One-Way ANOVA.** The following **One-Way ANOVA** window will open.



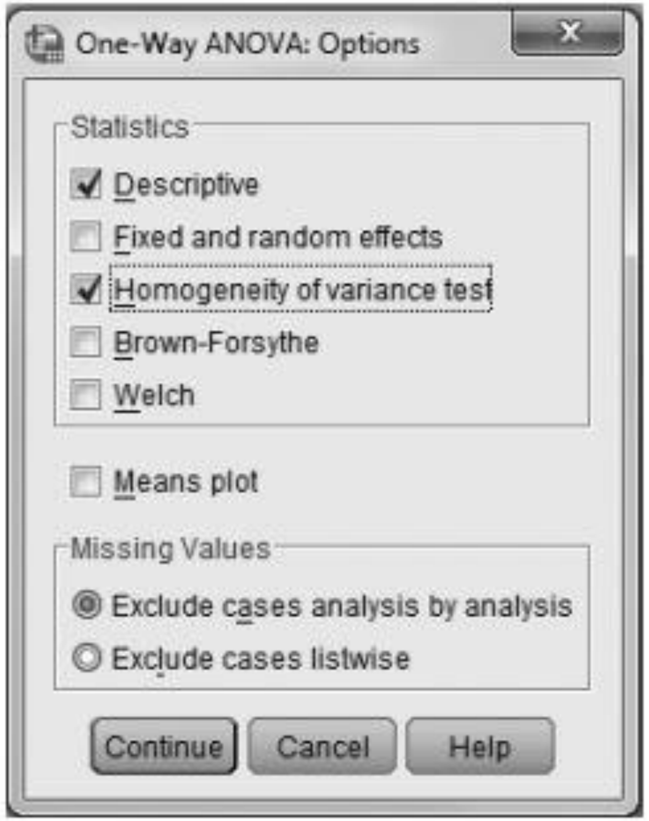
2. Transfer the dependent variable **TIME** to the **Dependent List:** field by clicking (highlight) the variable and then clicking . Transfer the independent variable **SHOCK** to the **Factor:** field by clicking (highlight) the variable and then clicking .


One-Way ANOVA
Dependent List:
time
Contrasts...
Post Hoc...
Options...
Bootstrap...
Factor:
SHOCK INTENSITY [...
OK
Paste
Reset
Cancel
Help

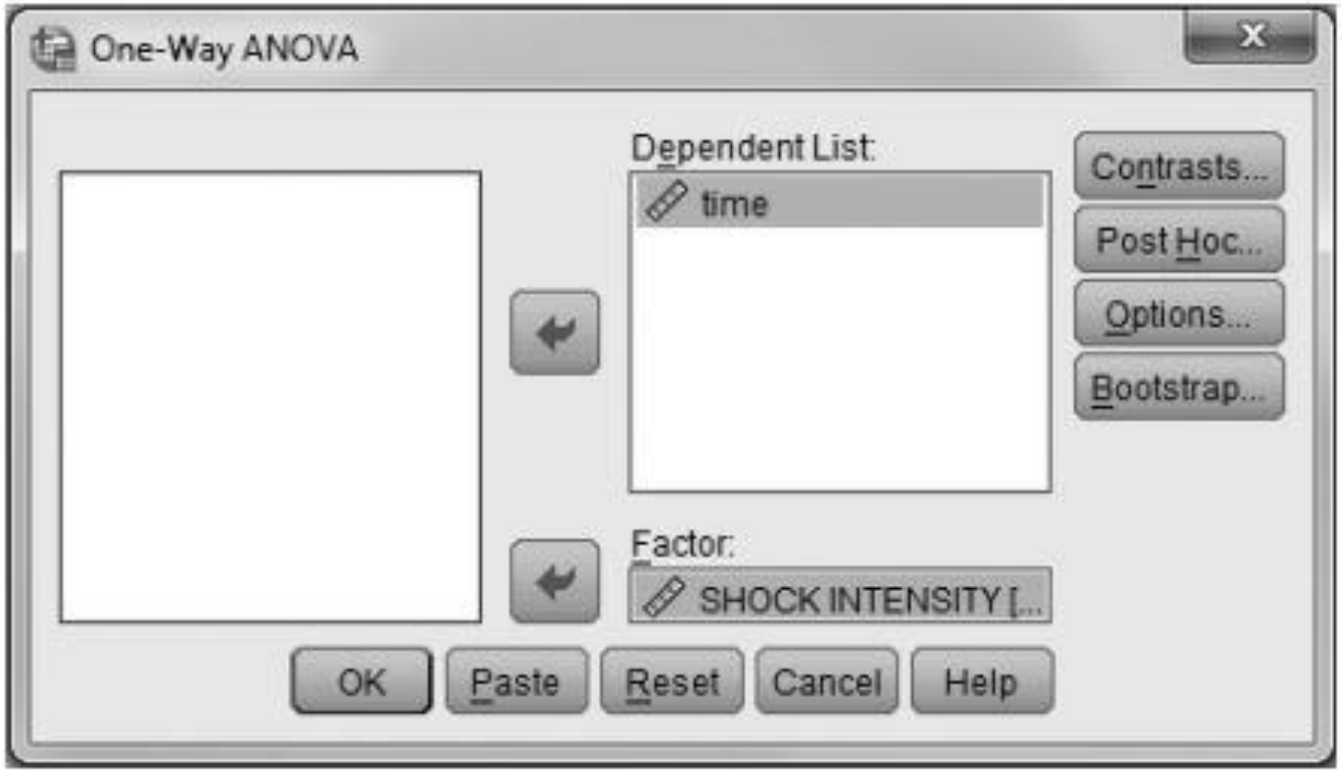
3. Since the one-way ANOVA will only perform an omnibus analysis of the *overall* differences between the three levels (low, medium, high) of the independent variable **SHOCK**, it will not analyze the differences between the *specific* shock levels. To obtain multiple comparisons between the three shock levels (low shock versus medium shock, low shock versus high shock, medium shock versus high shock), the researcher needs to perform a **post hoc** comparison test. Click Post Hoc... to achieve this. When the following **One-Way ANOVA: Post Hoc Multiple Comparisons** window opens, check the **Scheffe** field to run the Scheffé post hoc test. Next, click Continue.


One-Way ANOVA: Post Hoc Multiple Comparisons
Equal Variances Assumed
LSD
S-N-K
Waller-Duncan
Bonferroni
Tukey
Type I/Type II Error Ratio:
100
Sidak
Tukey's-b
Dunnett
Scheffe
Duncan
Control Category:
Last
R-E-G-W F
Hochberg's GT2
Test
R-E-G-W Q
Gabriel
2-sided
< Control
> Control
Equal Variances Not Assumed
Tamhane's T2
Dunnett's T3
Games-Howell
Dunnett's C
Significance level:
0.05
Continue
Cancel
Help

4. When the **One-Way ANOVA** window opens, click Options... to open the **One-Way ANOVA: Options** window. Check the **Descriptive** box and the **Homogeneity of variance test** box and then click Continue.



5. When the following **One-Way ANOVA** window opens, run the analysis by clicking OK. See Table 6.2 for the results.



SPSS Output

One-Way ANOVA Output

Descriptives

Time

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
LOW SHOCK	6	17.1667	5.11534	2.08833	11.7985	22.5349	10.00	25.00
MEDIUM SHOCK	6	22.1667	5.11534	2.08833	16.7985	27.5349	15.00	30.00
HIGH SHOCK	6	42.1667	5.11534	2.08833	36.7985	47.5349	35.00	50.00
Total	18	27.1667	12.10858	2.85402	21.1452	33.1881	10.00	50.00

Test of Homogeneity of Variances

Time

Levene Statistic	$df1$	$df2$	Sig.
.000	2	15	1.000

Anova

Time

	Sum of Squares	***df***	**Mean Square**	***F***	**Sig.**
Between Groups	2100.000	2	1050.000	40.127	.000
Within Groups	392.500	15	26.167		
Total	2492.500	17			

Post Hoc Tests

Multiple Comparisons

Dependent Variable: TIME

Scheffe

(I) SHOCK INTENSITY	(J) SHOCK INTENSITY	Mean Difference (I–J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
LOW SHOCK	MEDIUM SHOCK	–5.0000	2.95334	.269	–13.0147	3.0147
	HIGH SHOCK	–25.0000*	2.95334	.000	–33.0147	–16.9853
MEDIUM SHOCK	LOW SHOCK	5.0000	2.95334	.269	–3.0147	13.0147
	HIGH SHOCK	–20.0000*	2.95334	.000	–28.0147	–11.9853
HIGH SHOCK	LOW SHOCK	25.0000*	2.95334	.000	16.9853	33.0147
	MEDIUM SHOCK	20.0000*	2.95334	.000	11.9853	28.0147

* The mean difference significant at the .05 level.

Results and Interpretation

The assumption of **homogeneity of variance** is tested by the **Levene statistic**, which tests the hypothesis that the population variances are equal. In this example, the Levene statistic is $F = 0.000$ and the corresponding level of significance is large (i.e., $p > .05$) (see Table 6.1). Thus, the assumption of homogeneity of variance has not been violated.

The results from the analysis (Table 6.1) indicate that the intensity of the electric shock has a significant effect on the time taken to solve the problems, $F(2,15) = 40.13$, $p < .001$. The mean values for the three shock levels indicate that, as the shock level increased (from low to medium to high), so did the time taken to solve the problems (low: $M = 17.17$; medium: $M = 22.17$; high: $M = 42.17$).

Post Hoc Comparisons

While the highly significant *F*-ratio ($p < .001$) indicates that the means of the three shock levels differ significantly, it does not indicate the *location* of this difference. For example, the researcher may want to know whether the overall difference is due primarily to the difference between "Low Shock" and "High Shock" levels, or between "Low Shock" and "Medium Shock" levels, or between "Medium Shock" and "High Shock" levels. To test for differences between specific shock levels, a number of post hoc comparison techniques can be used. For this example, the more conservative **Scheffé** test was used.

In the **Multiple Comparisons** table, in the column labeled **Mean Difference (I–J)**, the mean difference values accompanied by asterisks indicate which shock levels differ significantly from each other at the 0.05 level of significance. The results indicate that the high shock level is significantly different from both the low shock and medium shock levels. The low shock level and the medium shock level do not differ significantly. These results show that the overall difference in time taken to solve complex problems between the three shock-intensity levels is due to the significantly greater amount of time taken by the subjects in the high shock condition.

Factorial Analysis of Variance

The factorial univariate ANOVA is an extension of the one-way ANOVA in that it involves the analysis of two or more independent variables. It is used in experimental designs in which every level of every factor is paired with every level of every other factor. It allows the researcher to assess the effects of each independent variable separately, as well as the joint effect or interaction of variables. Factorial designs are labeled either by the number of factors involved, or in terms of the number of levels of each factor. Thus, a factorial design with two independent variables (e.g., gender, ethnicity) and with two levels for each independent variable (male/female; Australian/Chinese) are called either a **2-way factorial** or a **2 × 2 factorial**.

Checklist of Requirements

- In any one analysis, there must be two or more independent variables (due to the complexity in interpreting higher-order interactions, most factorial designs are limited to three or four independent variables or factors).
- There can be two or more levels for each independent variable.
- There must be only one dependent variable.

Assumptions

- **Independence**—The samples are independently drawn from the source population(s).
- **Normality**—The dependent variable is normally distributed.
- **Homogeneity of variance**—The distribution of the dependent variable for one of the groups being compared has the same variance as the distribution for the other group being compared.

Example 1: Two-Way Factorial (2 × 2 Factorial)

A researcher is interested in determining the effects of two learning strategies (A and B) on the memorization of a hard versus an easy list of syllables. The factorial combination of these two independent variables (2 × 2) yields four experimental conditions: Strategy A-Easy List, Strategy A-Hard List, Strategy B-Easy List, and Strategy B-Hard List. A total of 24 subjects is randomly assigned to the four experimental conditions. The researcher recorded the total number of mistakes made by each subject.

	Strategy A	Strategy B
Easy list	s1 6	s13 20
	s2 13	s14 18
	s3 11	s15 14
	s4 8	s16 14
	s5 9	s17 12
	s6 5	s18 16
Hard list	s7 15	s19 16
	s8 17	s20 13
	s9 23	s21 15
	s10 21	s22 20
	s11 22	s23 11
	s12 20	s24 12

The data set has been saved under the name **EX7a.SAV**

Variables	Column(s)	Code
• STRATEGY	1	1 = Strategy A, 2 = Strategy B
• LIST	2	1 = Easy, 2 = Hard
• ERRORS	3–4	Number of errors made

Testing Assumptions
Normality

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	*df*	Significance	Statistic	*df*	Significance
Errors	.112	24	.200*	.975	24	.790

* This is a lower bound of the true significance.
[a] Lilliefors significance correction.

Homogeneity of Variance

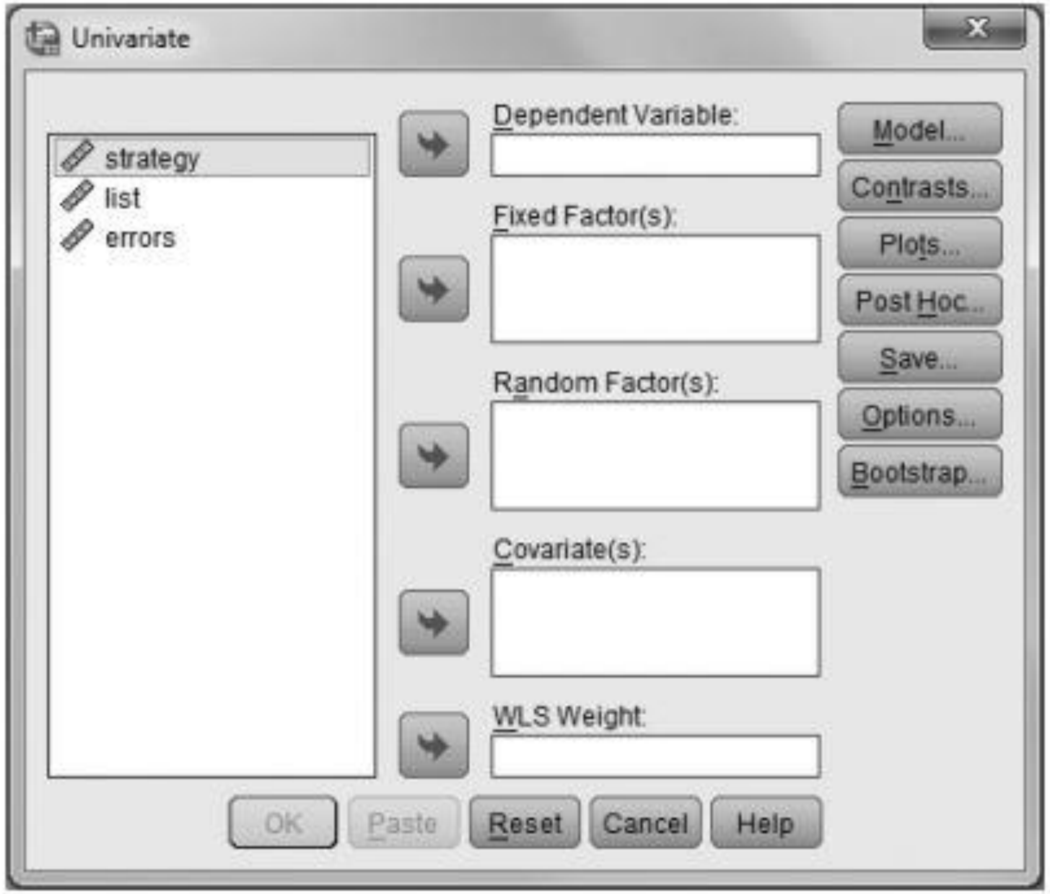
The homogeneity assumption is checked in SPSS by Levene's test

Independence

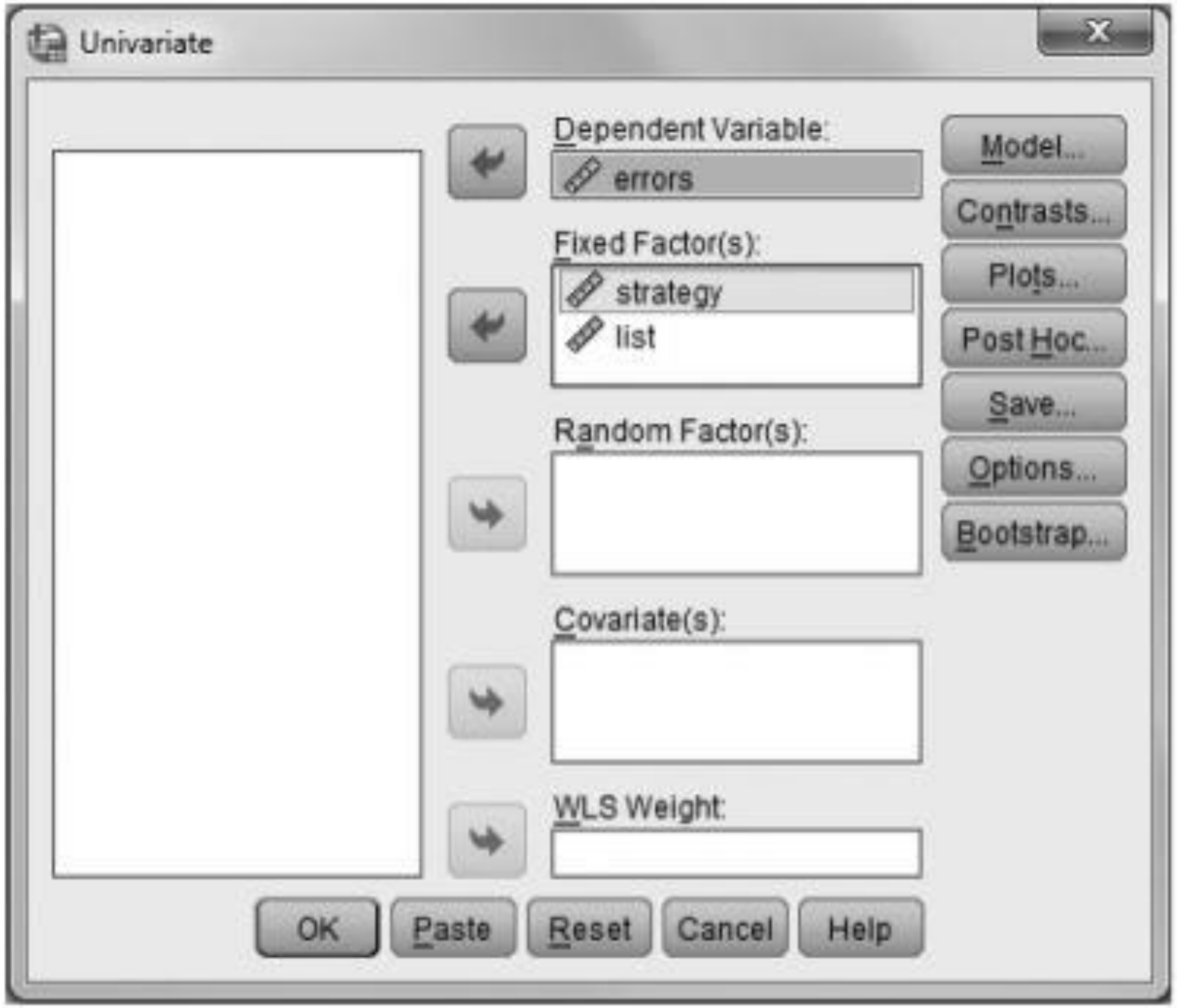
During data collection, ensure that the observations in one group are independent of the observations in the other group

Factorial ANOVA

1. From the menu bar, click **Analyze**, then **General Linear Model**, and then **Univariate**. The following **Univariate** window will open.

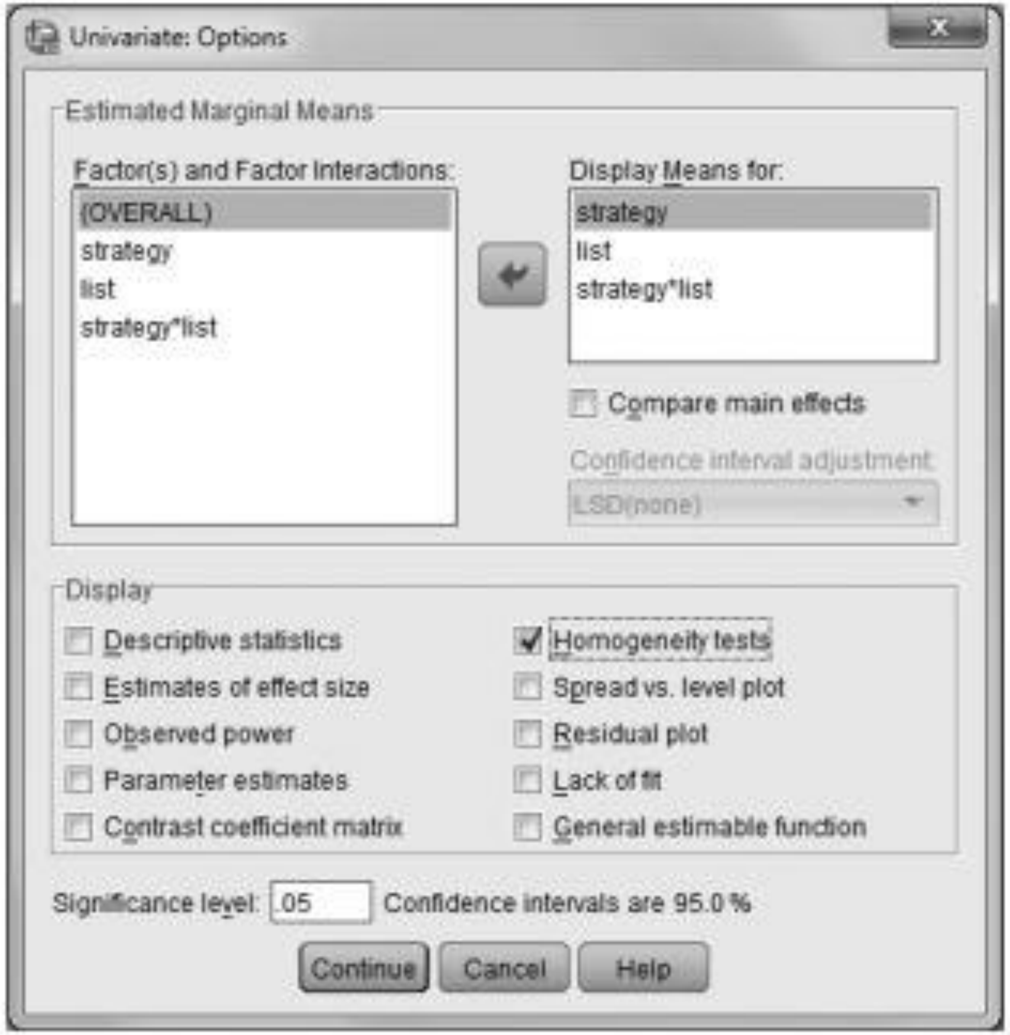


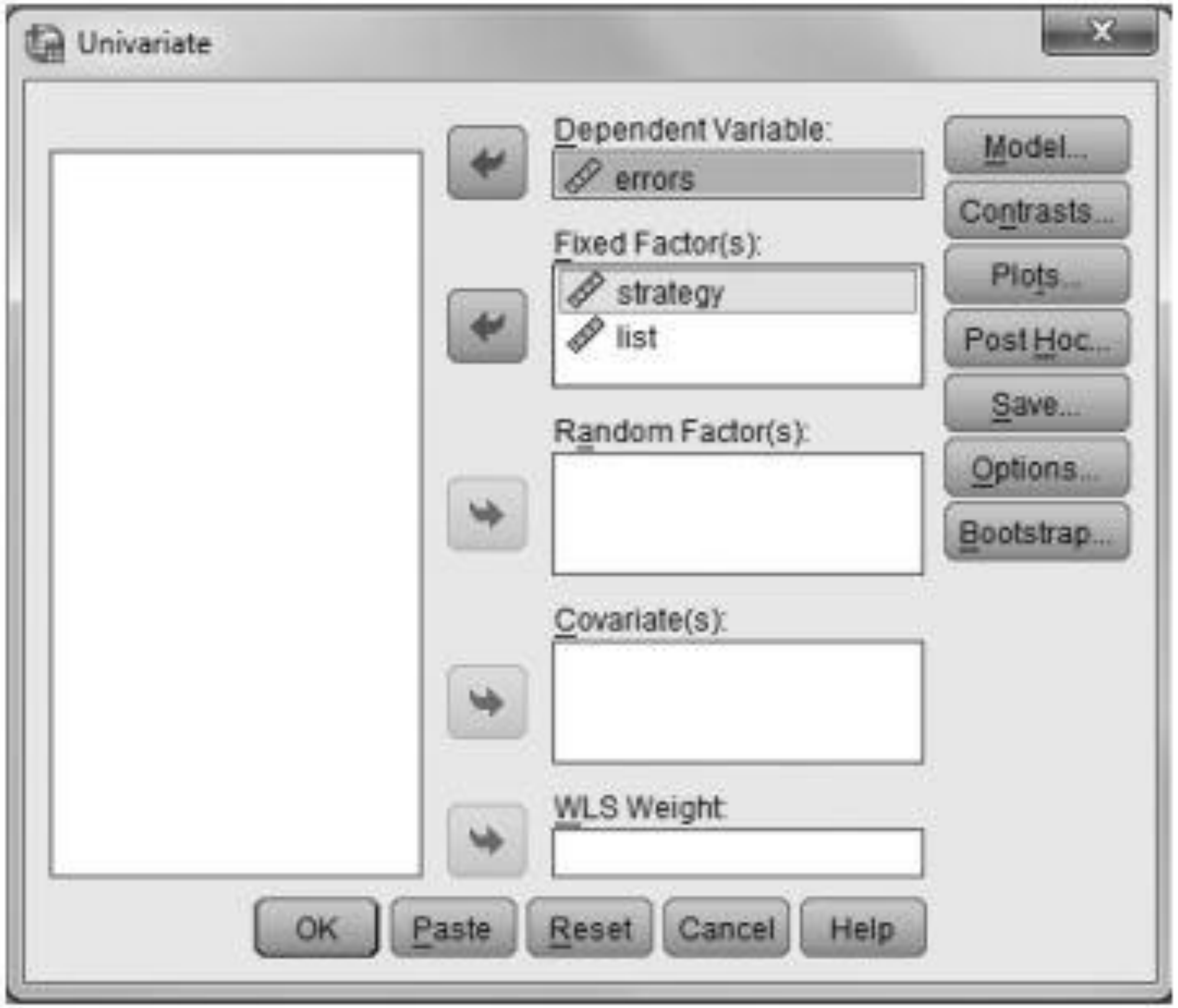
2. Transfer the **ERRORS** dependent variable of **ERRORS** to the **Dependent Variable:** field by clicking (highlight) the variable and then clicking →. Transfer the **STRATEGY** and **LIST** independent variables of **STRATEGY** and **LIST** to the **Fixed Factor(s):** field by clicking (highlight) the variables and then clicking →.
3. Click Plots... to plot a graph of the **STRATEGY*LIST** interaction. The following **Univariate: Profile Plots** window will open. Transfer the **STRATEGY** variable to the **Horizontal Axis:** field by clicking (highlight) the variable and then clicking →. Transfer the **LIST** variable to the **Separate Lines:** field by clicking (highlight) the variable and then clicking →. Next, click Add to transfer the **STRATEGY*LIST** interaction to the **Plots:** field. When this is done, click Continue.


Univariate
Dependent Variable:
errors
Fixed Factor(s):
strategy
list
Random Factor(s):
Covariate(s):
WLS Weight:
Model...
Contrasts...
Plots...
Post Hoc...
Save...
Options...
Bootstrap...
OK
Paste
Reset
Cancel
Help


Univariate: Profile Plots
Factors:
strategy
list
Horizontal Axis:
Separate Lines:
Separate Plots:
Plots:
Add
Change
Remove
strategy*list
Continue
Cancel
Help

4. Click Options... in the **Univariate** window to obtain descriptive statistics (estimated marginal means) for the full 2 x 2 **STRATEGY*LIST** interaction. When the **Univariate: Options** window opens, click (highlight) **STRATEGY**, **LIST**, and **STRATEGY*LIST** in the **Factor(s) and Factor Interactions:** field, and then click to transfer these factors and factor interaction to the **Display Means for:** field. Check the **Homogeneity tests** box. Click Continue to return to the **Univariate** window.
5. When the **Univariate** window opens, click OK to run the analysis. See Table 7.2 for the results.


Univariate: Options
Estimated Marginal Means
Factor(s) and Factor Interactions:
(OVERALL)
strategy
list
strategy*list
Display Means for:
strategy
list
strategy*list
Compare main effects
Confidence interval adjustment
LSD(none)
Display
Descriptive statistics
Estimates of effect size
Observed power
Parameter estimates
Contrast coefficient matrix
Homogeneity tests
Spread vs. level plot
Residual plot
Lack of fit
General estimable function
Significance level: .05
Confidence intervals are 95.0 %
Continue
Cancel
Help



Univariate
Dependent Variable:
errors
Fixed Factor(s):
strategy
list
Random Factor(s):
Covariate(s):
WLS Weight:
Model...
Contrasts...
Plots...
Post Hoc...
Save...
Options...
Bootstrap...
OK
Paste
Reset
Cancel
Help

SPSS Output

2 × 2 ANOVA Output

Univariate Analysis of Variance

Between-Subjects Factor

		Value Label	*N*
Strategy	1.00	STRATEGY A	12
	2.00	STRATEGY B	12
List	1.00	EASY	12
	2.00	HARD	12

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Errors

F	*df*1	*df*2	Significance
.016	3	20	.997

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

[a] Design: Intercept + strategy + list + strategy × list.

Tests of Between-Subjects Effects

Dependent Variable: Errors

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected model	372.125[a]	3	124.042	13.091	.000
Intercept	5133.375	1	5133.375	541.781	.000
Strategy	5.042	1	5.042	.532	.474
List	145.042	1	145.042	15.308	.001
Strategy * list	222.042	1	222.042	23.434	.000
Error	189.500	20	9.475		
Total	5695.000	24			
Corrected total	561.625	23			

[a] R Squared = .663 (Adjusted R Squared = .612)

Estimated Marginal Means

1. Strategy

Dependent Variable: Errors

			95% Confidence Interval	
Strategy	Mean	Std Error	Lower Bound	Upper Bound
STRATEGY A	14.167	.889	12.313	16.020
STRATEGY B	15.083	.889	13.230	16.937

2 × 2 ANOVA Output

Estimated Marginal Means

2. List

Dependent Variable: Errors

List	Mean	Std Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
EASY	12.167	.889	10.313	14.020
HARD	17.083	.889	15.230	18.937

3. Strategy * List

Dependent Variable: Errors

Strategy	List	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
STRATEGY A	EASY	8.667	1.257	6.045	11.288
	HARD	19.667	1.257	17.045	22.288
STRATEGY B	EASY	15.667	1.257	13.045	18.288
	HARD	14.500	1.257	11.879	17.121

Results and Interpretation

The assumption of **homogeneity of variance** is tested by **Levene's test of equality of error variances**, which tests the hypothesis that the population error variances are equal. In this example, the Levene statistic is $F = 0.016$ and the corresponding level of significance is large (i.e., $p > .05$) (see Table 7.2). Thus, the assumption of homogeneity of variance has not been violated.

Main Effect

The main effect of **STRATEGY** is not significant, $F(1,20) = 0.53$, $p > .05$ (see Table 7.2). From the estimated marginal means, the number of errors made by the **Strategy A group** ($M = 14.167$) is not significantly different from the number of errors produced by the **Strategy B group** ($M = 15.083$) (collapsing across the two LIST levels).

The main effect of **LIST** is significant, $F(1,20) = 15.31$, $p < .05$. From the estimated marginal means, it can be determined that the subjects produced significantly more errors in the hard list (M = 17.08) than in the easy list (M = 12.16) (collapsing across the two STRATEGY levels).

Interaction Effect

The STRATEGY*LIST interaction is signifcant, $F(1,20) = 23.43$, $p < .001$. To interpret the interaction, the task is made easier by graphing the

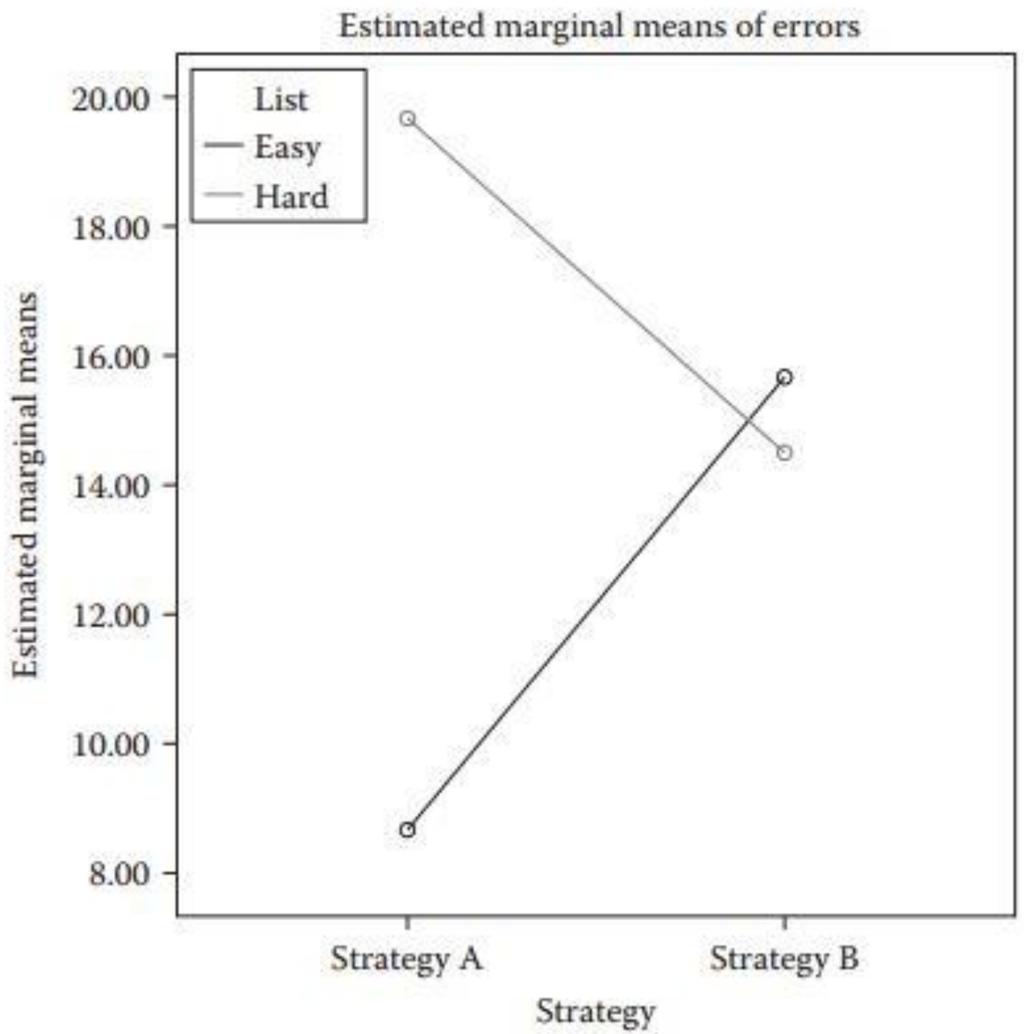


FIGURE 7.1
2 (STRATEGY) × 2 (LIST) interaction effect.

STRATEGY*LIST estimated marginal means from Table 7.2, as shown in Figure 7.1.

From Figure 7.1, it can be determined that the effect of learning strategy on the number of errors made is dependent on the difficulty of the list learned. Under strategy A, subjects made more errors on the hard list than on the easy list, but under strategy B, the effect is opposite, with subjects making more errors on the easy list than on the hard list.

Post Hoc Test for Simple Effects

The significant interaction effect indicates that the outcome of one independent variable on the dependent variable is dependent on the second independent variable, i.e., the four experimental conditions (Strategy A-Easy List, Strategy A-Hard List, Strategy B-Easy List, Strategy B-Hard List) differ significantly in affecting the number of errors made. Nonetheless, the interaction effect does not show where the differences are, i.e., between which experimental conditions. To identify specific differences, **post hoc** comparisons can be used to "tease apart" the interaction. This is equivalent to the test for simple effects, i.e., the effect of one factor (IV1) at one level of the other factor (IV2).

REFLEKSI

1. **Informasi penting hari ini**

2. **Manfaat penting dari informasi penting hari ini**

3. **Tindak lanjut yang dapat saudara lakukan**

Thank you!

Any questions?