

Post Hoc Test Example

The results from the one-way ANOVA do not indicate which of the three groups differ from one another, so, in many cases, it is of interest to follow the analysis with a post hoc test or a planned comparison among particular means. If several comparisons between pairs of means are made, it is a good idea to use a test, such as the Tukey, that controls for alpha inflation.

For the present example, there are not too many possible comparisons to make (just 3), so alpha inflation will not be a tremendous issue. Also, in the case of this artificial example, two of the means are exactly equal and should not require a statistical test. Nevertheless, to illustrate the use of the Tukey post hoc approach. I've also included the Games-Howell test in SPSS, because very small sample sizes may be particularly sensitive to the homogeneity of variance assumption.

SPSS

To obtain the post hoc tests with syntax, the POST HOC subcommand is added to the one-way procedure.

```
oneway recall by groups
  /posthoc=tukey gh.
```

Post Hoc Tests

Multiple Comparisons

Dependent Variable: recall number of new terms recalled

	(I) groups learning strategy experimental groups	(J) groups learning strategy experimental groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	1.00 reading only	2.00 retrieval practice	-3.00000 *	.85635	.011	-5.2846	-.7154
		3.00 concept mapping	.00000	.85635	1.000	-2.2846	2.2846
	2.00 retrieval practice	1.00 reading only	3.00000 *	.85635	.011	.7154	5.2846
		3.00 concept mapping	3.00000 *	.85635	.011	.7154	5.2846
	3.00 concept mapping	1.00 reading only	.00000	.85635	1.000	-2.2846	2.2846
		2.00 retrieval practice	-3.00000 *	.85635	.011	-5.2846	-.7154
Games-Howell	1.00 reading only	2.00 retrieval practice	-3.00000	1.00000	.056	-6.0861	.0861
		3.00 concept mapping	.00000	.94868	1.000	-3.0903	3.0903
	2.00 retrieval practice	1.00 reading only	3.00000	1.00000	.056	-.0861	6.0861
		3.00 concept mapping	3.00000 *	.54772	.002	1.3978	4.6022
	3.00 concept mapping	1.00 reading only	.00000	.94868	1.000	-3.0903	3.0903
		2.00 retrieval practice	-3.00000 *	.54772	.002	-4.6022	-1.3978

*. The mean difference is significant at the .050 level.

Homogeneous Subsets

recall number of new terms recalled

groups learning strategy experimental groups	N	Subset for alpha = .050	
		1	2
Tukey HSD ^a			
1.00 reading only	5	6.0000	
3.00 concept mapping	5	6.0000	
2.00 retrieval practice	5		9.0000
Sig.		1.000	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 5.000.

R

```
#using the full ANOVA command in less R gives the Tukey follow-up tests  
> ANOVA(recall ~ groups)
```

TUKEY MULTIPLE COMPARISONS OF MEANS

Family-wise Confidence Level: 0.95

```
-----  
      diff   lwr   upr p adj  
2-1  3.00  0.72  5.28 0.01  
3-1  0.00 -2.28  2.28 1.00  
3-2 -3.00 -5.28 -0.72 0.01
```

```
> #get the Games-Howell test
```

```
> library(rstatix)
```

```
> games_howell_test(d, recall ~ groups, conf.level = 0.95, detailed = FALSE)
```

```
.y.      group1 group2 estimate conf.low conf.high p.adj p.adj.signif  
* <chr>  <chr>  <chr>      <dbl>  <dbl>  <dbl>  <dbl>  <chr>  
1 recall 1      2          3  -0.0861  6.09 0.056 ns  
2 recall 1      3          0  -3.09   3.09 1      ns  
3 recall 2      3          -3  -4.60  -1.40 0.002 **
```

Write-up

The Tukey post hoc tests indicated that student recall in the retrieval practice differed significantly from recall in the control, $p = .01$, and concept mapping, $p = .01$, conditions. The Games-Howell test, however, suggested that the difference between control and retrieval practice conditions was only marginally significant, $p = .056$.