Using Human-Friendly Scheffé Comparisons to Explore Group Differences in One-way ANOVA

Gordon Brooks, Ph.D.
Ohio University (Athens, Ohio)

Nina Adjanin, Ph.D.
Northwest Missouri State University

Presented at
The Athens Institute for Education and Research (ATINER)
18th Annual International Conference on Statistics: Teaching, Theory & Applications
1-4 July 2024, Athens, Greece

Multiple Comparison Procedures (Post Hoc)

- Researchers commonly use MCPs following statistically significant ANOVA and main effects from Factorial ANOVA
- Most commonly these are **post hoc pairwise comparisons** (e.g., Tukey-Kramer or Games-Howell), but researchers do sometimes use **a priori contrasts** that include **non-pairwise** (complex) comparisons
- Very few use the Scheffé post hoc method because it is well-known to **lack the statistical power** of other MCPs for the pairwise post hoc comparisons that most researchers use—and most statistics programs provide only pairwise Scheffé
- The Scheffé MCP has lower power because it **adjusts for all possible comparisons**: all pairwise and non-pairwise comparisons—but researchers often don't know where to start with non-pairwise post hoc comparisons

First… a little joke…

One of the Most Embarrassing Outcomes for a Statistician…

- **Result:** The F-test for a One-Way ANOVA with five treatment groups is significant at the .05 level but **NONE of the pairwise comparisons** between the five means is statistically significant.

- **Solution:** Cry hard… then work hard… to find some **obscure, meaningless** complex (i.e., Scheffé) comparison that ***IS*** significant, such as: the average of the first three treatment means is significantly different from the average of the last two treatment means!

(from Gary Ramseyer's First Internet Gallery Of Statistics Jokes: https://about.illinoisstate.edu/gcramsey/other/)

Congruence of Scheffé with Omnibus ANOVA

- However, **only Scheffé** MCP ***guarantees congruence*** to find a statistically significant comparison when the omnibus ANOVA is statistically significant—and conversely, ***NOT*** find one when ANOVA is ***not*** significant
  - As the joke said… ANOVA can be significant, but no pairwise comparison is
  - See Kirk (2013), Maxwell, Delaney, & Kelley (2018), Keppel & Wickens (2004)

- A **maximum Scheffé contrast/comparison** can be calculated that provides the set of contrast coefficients for the means that **maximally differentiates some combination of groups** on the dependent variable
  - And there is a formula… so it is **not a lot of hard work** to calculate the ***MAX***
  - This ***maximum*** comparison has the same statistical significance as the omnibus Fisher F ANOVA and is usually a **non-pairwise, complex comparison**
  - ***Unfortunately, the hard work can be in the interpretation***

Scheffé Maximum Contrast/Comparison

Scheffé (see Keppel & Wickens, 2004; Williams, 1979)

$$c_i' = \frac{N_i(\bar{X}_i - \bar{T})}{\sqrt{SSB}}$$

Hollingsworth (1978, see also Williams, 1979)

$$c_i = \frac{\sqrt{\tilde{N}}(\bar{X}_i - \bar{T})}{\sqrt{SSB}}$$

Where:

c_i is the contrast/comparison coefficient for group *i*

N_i is the sample size in each group

$\bar{T}$ is the dependent variable grand mean for the total sample

$\bar{X}_i$ is the dependent variable mean for group *i*

SSB is the sum of squares between groups from ANOVA

$\tilde{N}$ is the harmonic mean group sample size

Maximum Contrast/Comparisons (*continued*)

For example: $N_i = 10$ for all groups $\quad \bar{T} = 49.3$

SSB = 698.4 $\quad \overline{X_i} = \{54.9, 45.9, 51.7, 44.7\}$

Therefore, the ***unscaled*** contrast coefficients, c_i, are calculated as follows:

c_1 = 10(54.9-49.3) / 26.43 = 56 / 26.43 = **2.119**

c_2 = 10(45.9-49.3) / 26.43 = -34 / 26.43 = **-1.286**

c_3 = 10(51.7-49.3) / 26.43 = 24 / 26.43 = **0.908**

c_4 = 10(44.7-49.3) / 26.43 = -46 / 26.43 = **-1.742**

Scheffé and Non-pairwise Complex Comparisons

- Unfortunately, coefficient weights from this maximum Scheffé comparison are often **uninterpretable or meaningless** from a practical or theoretical perspective (***see introductory joke***... also see Schmid, 1977).

- For example, it is **hard** to **make sense** of the maximum Scheffé coefficients from the previous slide

	Scheffé Max	Scaled Scheffé (contrasts sum to 1 and -1)
c_1 =	2.119	**0.700**
c_2 =	-1.286	**-0.425**
c_3 =	0.908	**0.300**
c_4 =	-1.742	**-0.575**

OHIO UNIVERSITY

NORTHWEST MISSOURI STATE UNIVERSITY

Robert Barcikowski & Rationale for "Human Contrasts"

```
C   This program computes one-way analysis of variance with
C   both Scheffe and Brown-Forsythe-Scheffe post hoc tests.
C   It was pieced together by Robert S. Barcikowski during
C   the last week in April, 1993 and revised April, 2000.
C*********************************************************
    DO 14 I = 1, JN
    ZMEAN(I) = (XBAR(I) - GM) / DMS
14  CONTINUE
    CALL HOLLY(LEVELS, SSB, GM, XBAR, BARCOE, RN)
    PRINT 12, (BARCOE(I), I = 1, LEVELS)
12  FORMAT(1H0,'MAXIMUM CONTRAST HAS FOLLOWING COEFFICIENTS'
   <      ,//,10F8.3,//)
    CALL HELMRT(HELM, LEVELS, ALLCON, HELNUM)
    CALL SCHEFE(BARCOE, XBAR, LEVELS, MSE, . . . , ITEST)
    CALL BFS(SDE, BARCOE, RN, LEVELS, DFB, ALPHA, APSI)
    IF (NC .EQ. 0) GO TO 23
23  STOP
    END
```

Barcikowski originally wrote the program in FORTRAN and Brooks & Adjanin converted it to R and R Shiny.

The purpose is to provide a relatively easy way (*unlike in the joke*) to find statistically significant – and ***INTERPRETABLE*** – Scheffé comparisons (and Brown-Forsythe adjustments for unequal variances, like Games-Howell)

Barcikowski "Human-Friendly" Complex Comparisons

- Barcikowski suggested a method to identify the maximum "human-friendly" comparison that approximates the maximum Scheffé comparison—**but with coefficients that are reasonably interpretable**

- Barcikowski approach tests all possible comparisons that use "reasonable" (i.e., human-friendly) ways to compare complex combinations of groups, for example:
  - Comparison of a control (or combination of treatments) group with the average of multiple treatment groups (i.e., **something versus nothing**)
  - Comparison of a low-dose treatment group with the average of higher-dose groups (i.e., **some versus more**)
  - Comparison of the average of 2 control groups with average of 3 treatment groups (***we disagree with the joke here…***)

“Human-friendly” contrasts (“***Helmert-plus***” *complex comparisons*)

Helmert:

$$\begin{cases} 1\mu_1 - \frac{1}{4}\mu_2 - \frac{1}{4}\mu_3 - \frac{1}{4}\mu_4 - \frac{1}{4}\mu_5 \\ 0\mu_1 + 1\mu_2 - \frac{1}{3}\mu_3 - \frac{1}{3}\mu_4 - \frac{1}{3}\mu_5 \\ 0\mu_1 + 0\mu_2 + 1\mu_3 - \frac{1}{2}\mu_4 - \frac{1}{2}\mu_5 \\ 0\mu_1 + 0\mu_2 + 0\mu_3 + 1\mu_4 - 1\mu_5 \end{cases}$$

Plus **(for example):**

$$\begin{cases} \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 - \frac{1}{3}\mu_3 - \frac{1}{3}\mu_4 - \frac{1}{3}\mu_5 \\ 0\mu_1 + \frac{1}{2}\mu_2 + \frac{1}{2}\mu_3 - \frac{1}{2}\mu_4 - \frac{1}{2}\mu_5 \end{cases}$$

Comparison Coefficient								
Comparison	1	2	3	4	5	6	7	8
1	0.25	0.25	0.25	0.25	-0.25	-0.25	-0.25	-0.25
2	0.33	0.33	0.33	-0.20	-0.20	-0.20	-0.20	-0.20
3	0.33	0.33	0.33	0	-0.25	-0.25	-0.25	-0.25
4	0.33	0.33	0.33	0	0	-0.33	-0.33	-0.33
5	0.50	0.50	-0.17	-0.17	-0.17	-0.17	-0.17	-0.17
6	0.50	0.50	0	-0.20	-0.20	-0.20	-0.20	-0.20
7	0.50	0.50	0	0	-0.25	-0.25	-0.25	-0.25
8	0.50	0.50	0	0	0	-0.33	-0.33	-0.33
9	0.50	0.50	0	0	0	0	-0.50	-0.50
10	1.00	-0.14	-0.14	-0.14	-0.14	-0.14	-0.14	-0.14
11	1.00	0	-0.17	-0.17	-0.17	-0.17	-0.17	-0.17
12	1.00	0	0	-0.20	-0.20	-0.20	-0.20	-0.20
13	1.00	0	0	0	-0.25	-0.25	-0.25	-0.25
14	1.00	0	0	0	0	-0.33	-0.33	-0.33
15	1.00	0	0	0	0	0	-0.50	-0.50
16	1.00	0	0	0	0	0	0	-1.00

Groups	Comparisons
3	6
4	25
5	75
6	301
7	476
8	3025

And then all permutations of these sets of coefficients… resulting in 3025 unique comparisons (in the case of 8 groups)

Barcikowski **Human-Friendly** Complex Comparisons

- Barcikowski's method identifies the **maximum comparisons** (based on contrast sum of squares explained) from among all possible **reasonably interpretable** Scheffé-like, **Human-friendly** contrasts/comparisons
  - This will include any statistically significant pairwise comparisons
  - We call them "comparisons" because they are intended for Post Hoc (even though Scheffé are typically called "contrasts")

- We have created an R Shiny web app to obtain
  - the **Scheffé, Scaled Scheffé, and Hollingsworth** maximum comparisons
  - the maximum **Barcikowski human-friendly comparison**, and all other statistically significant human-friendly comparisons
  - the relatively unknown **Brown-Forsythe adjustment** to the Scheffé MCP for **when the equal variances assumption is not met**

Purpose of Presentation

- To share results from recent research (especially, **congruence**) that supports the use of Barcikowski's Human-friendly comparisons method
- Further, we will share information about using the R Shiny App
- Finally, we share some examples of datasets from among several well-known design and analysis textbooks **that might have benefited from using Scheffé maximum comparisons** and Barcikowski human-friendly comparisons instead of focusing only on Pairwise comparisons

Congruence for Non-significance (robustness)

N	SD	Shape	Scheffé maximum comparison	Hollingsworth maximum comparison	1st most explanatory Human-friendly comparison	2nd most explanatory Human-friendly comparison
40,40,40,40	10,10,10,10	Normal	100.00%	100.00%	99.29%	98.13%
28,36,44,52	10,10,10,10	Normal	100.00%	99.70%	99.01%	97.94%
40,40,40,40	13,11,9,7	Normal	100.00%	100.00%	99.36%	98.32%
28,36,44,52	13,11,9,7	Normal	100.00%	99.64%	98.82%	97.40%
28,36,44,52	7,9,11,13	Normal	100.00%	99.76%	99.23%	98.35%
40,40,40,40	10,10,10,10	Skewed	100.00%	100.00%	99.25%	98.03%
28,36,44,52	10,10,10,10	Skewed	100.00%	99.61%	99.03%	97.95%
40,40,40,40	13,11,9,7	Skewed	100.00%	100.00%	99.34%	98.03%
28,36,44,52	13,11,9,7	Skewed	100.00%	99.69%	98.84%	97.53%
28,36,44,52	7,9,11,13	Skewed	100.00%	99.77%	99.28%	98.52%
40,40,40,40	10,10,10,10	Kurtotic	100.00%	100.00%	99.16%	98.02%
28,36,44,52	10,10,10,10	Kurtotic	100.00%	99.64%	99.01%	98.02%
40,40,40,40	13,11,9,7	Kurtotic	100.00%	100.00%	99.22%	98.01%
28,36,44,52	13,11,9,7	Kurtotic	100.00%	99.66%	98.80%	97.36%
28,36,44,52	7,9,11,13	Kurtotic	100.00%	99.74%	99.23%	98.39%

Lowest congruence

Congruence for Significance (statistical power)

Means	Scheffé maximum comparison	Hollingsworth maximum comparison	1st most explanatory Human-friendly comparison	2nd most explanatory Human-friendly comparison
50, 50, 50, 54 [1]	100.00%	100.00%	97.52%	93.31%
50, 50, 50, 58 [1]	100.00%	100.00%	98.98%	96.62%
50, 50, 54, 54 [1]	100.00%	100.00%	96.95%	91.03%
50, 50, 54, 54 [2]	100.00%	99.11%	96.17%	90.56%
50, 50, 54, 54 [3]	100.00%	100.00%	96.94%	91.18%
50, 50, 54, 58 [1]	100.00%	100.00%	98.15%	94.37%
50, 50, 58, 58 [1]	100.00%	100.00%	99.24%	97.04%
50, 54, 54, 58 [1]	100.00%	100.00%	97.53%	92.35%

Lowest congruence

4-group results were presented at American Educational Research Association (AERA) in April 2023 and 5-group results were presented at Mid-Western Educational Research Association (MWERA) in October 2023, and all results have been accepted for publication in the General Linear Model Journal (glmj.org).

2-unequal sample sizes, equal variances; 3-equal sample sizes, unequal variances

Value of Complex versus Pairwise Comparisons

- We reviewed many datasets used as examples for ANOVA in well-known **textbooks** and also many datasets provided by **R datasets** package
  - Most of these example datasets were also used to illustrate **Pairwise Multiple Comparison Procedures**

- We identified numerous examples from among these well-known datasets where the pairwise comparisons were not the most explanatory—**we will share several such examples**

- We believe there can be **value** in identifying, and making sense of, the **maximum Scheffé comparison** (which is frequently a complex comparison), or similarly, a ***Barcikowski Human-friendly comparison***

OHIO UNIVERSITY

NORTHWEST MISSOURI STATE UNIVERSITY

Example: All 25 sets of coefficients for 4 groups

	Contrast/Comparison Coefficient					Contrast/Comparison Coefficient			
Comparison	1	2	3	4	Comparison	1	2	3	4
1	1.00	-0.50	-0.50	0	13	1.00	-0.33	-0.33	-0.33
2	1.00	-0.50	0	-0.50	14	-0.33	1.00	-0.33	-0.33
3	1.00	0	-0.50	-0.50	15	-0.33	-0.33	1.00	-0.33
4	0	1.00	-0.50	-0.50	16	-0.33	-0.33	-0.33	1.00
5	-0.50	1.00	-0.50	0	17	1.00	-1.00	0	0
6	-0.50	1.00	0	-0.50	18	1.00	0	-1.00	0
7	-0.50	0	1.00	-0.50	19	1.00	0	0	-1.00
8	0	-0.50	1.00	-0.50	20	0	1.00	-1.00	0
9	-0.50	-0.50	1.00	0	21	0	1.00	0	-1.00
10	-0.50	-0.50	0	1.00	22	0	0	1.00	-1.00
11	-0.50	0	-0.50	1.00	23	0.50	0.50	-0.50	-0.50
12	0	-0.50	-0.50	1.00	24	0.50	-0.50	0.50	-0.50
					25	0.50	-0.50	-0.50	0.50

OHIO UNIVERSITY

Barcikowski's Most Explanatory Human-Friendly Comparisons (with Scheffe tests assuming equal variances)

NORTHWEST MISSOURI STATE UNIVERSITY

Stevens (2007) IRS Data (p. 97)

i Information

```
Total number of Comparisons tested:
[1] 25
```

	SSQ	Coef1	Coef2	Coef3	Coef4	diff	lwr.ci	upr.ci	pval	Cohens.d
24	**0.9298**	**0.5000**	**-0.5000**	**0.5000**	**-0.5000**	**8.0000**	**0.3066**	**15.6934**	**0.0387**	**0.9643**
10	0.8926	1.0000	-0.5000	0.0000	-0.5000	9.6000	0.1775	19.0225	0.0444	1.1571
5	0.7557	1.0000	0.0000	0.0000	-1.0000	10.2000	-0.6801	21.0801	0.0734	1.2294
16	0.7163	-0.5000	0.0000	-0.5000	1.0000	-8.6000	-18.0225	0.8225	0.0850	1.0366

Showing 1 to 4 of 4 entries

Previous 1 Next

DV	Days_to_Complete
Group	IRS_Region
1	East
2	Midwest
3	South
4	West

- **Output shows comparisons that were statistically significant at alpha = .15**
- Family 13-24
- No pairwise Scheffé SIG (but 1v4 was SIG as Tukey)
- **Scheffé MAX itself is hard to interpret**

Scheffe Tests of Maximum Comparisons assuming equal Variances

i Information

	Coef1	Coef2	Coef3	Coef4	Family	Diff	lwr.ci	upr.ci	pval	Cohens.d
ScheffeMAX	2.1190	-1.2866	0.9082	-1.7406	1,3-2,4	26.4273	2.0986	50.7559	0.0285	3.1853
ScaledMAX	0.7000	-0.4250	0.3000	-0.5750	1,3-2,4	8.7300	0.6933	16.7667	0.0285	1.0522
HollingsworthMAX	0.6701	-0.4068	0.2872	-0.5504	1,3-2,4	8.3570	0.6636	16.0504	0.0285	1.0073
BarcikowskiMAX	**0.5000**	**-0.5000**	**0.5000**	**-0.5000**	**1,3-2,4**	**8.0000**	**0.3066**	**15.6934**	**0.0387**	**0.9643**

Example contrasts from earlier in presentation

Barcikowski's Most Explanatory Human-Friendly Comparisons (with Scheffe tests assuming equal variances)

Adjanin and Brooks (2023)

i Information

```
Total number of Comparisons tested:
[1] 6
```

	SSQ	Coef1	Coef2	Coef3	diff	lwr.ci	upr.ci	pval	Cohens.d
5	**0.1166**	**-0.5000**	**1.0000**	**-0.5000**	**-0.2678**	**-0.5228**	**-0.0128**	**0.0369**	**0.4182**
2	0.1081	0.0000	1.0000	-1.0000	-0.2977	-0.5410	-0.0544	0.0116	0.4649
4	0.0690	1.0000	-1.0000	0.0000	0.2379	-0.1093	0.5851	0.2425	0.3715
6	0.0520	-0.5000	-0.5000	1.0000	0.1788	-0.0429	0.4005	0.1412	0.2792

Showing 1 to 4 of 4 entries

Previous 1 Next

Group	Pseudo-Experimental VR Group
1	Computer
2	Oculus
3	SmartPhone

Scheffe Tests of Maximum Comparisons assuming equal Variances

i Information

	Coef1	Coef2	Coef3	Family	Diff	lwr.ci	upr.ci	pval	Cohens.d
ScheffeMAX	0.4929	-6.6159	6.1230	1,3-2	1.9403	0.3621	3.5186	0.0111	3.0298
ScaledMAX	0.0745	-1.0000	0.9255	1,3-2	0.2933	0.0547	0.5318	0.0111	0.4580
HollingsworthMAX	0.2665	-0.8016	0.5351	1,3-2	0.2227	0.0305	0.4150	0.0183	0.3478
BarcikowskiMAX	**-0.5000**	**1.0000**	**-0.5000**	**2-1,3**	**-0.2678**	**-0.5228**	**-0.0128**	**0.0369**	**0.4182**

- Default example on R Shiny App website
- Oculus VR goggles vs other rather than only Oculus vs Phone
- Family comes from DescTools output

OHIO UNIVERSITY

NORTHWEST MISSOURI STATE UNIVERSITY

Additional Results

Adjanin and Brooks (2023) ***continued***

Descriptive Statistics are from `psych::describeBy()` and `psych::describe()`

Descriptive Statistics | BoxPlots

Group	n	mean	sd	median	min	max	range	skew	kurtosis	se
1	31	4.2016	0.7428	4.2500	2.7500	5.0000	2.2500	-0.5854	-0.9511	0.1334
2	62	3.9637	0.3837	4.0000	3.0000	5.0000	2.0000	-0.0546	0.4522	0.0487
3	131	4.2615	0.7077	4.5000	1.5000	5.0000	3.5000	-1.1203	1.1978	0.0618
TOTAL	**224**	**4.1708**	**0.6506**	**1.5000**	**5.0000**	**3.5000**	**0.0435**		**224.0000**	**4.1708**

Group	Pseudo-Experimental VR Group
1	Computer
2	Oculus
3	SmartPhone

Omnibus Test & Assumptions

i Information

Omnibus ANOVA | Homoscedasticity Assumption | Normality Assumption

	Test	Statistic	df1	df2	Pval
1	Fisher's F	4.5897	2.0000	221.0000	0.0111
2	**Welch's F**	**7.4848**	**2.0000**	**76.8525**	**0.0011**
3	Brown-Forsythe F	4.7669	2.0000	77.4731	0.0112
4	Kruskal-Wallis X2	17.1964	2.0000		0.0002

- Contains all information necessary for complete One-way ANOVA analysis, including descriptive statistics and tests of assumptions

One-way ANOVA results from jmv package with pairwise comparisons

Adjanin and Brooks (2023) *continued*

i Codebook for Variables

jmv::anovaOneW with Games-Howell & Assumptions | jmv::ANOVA with Helmert, Tukey, Scheffe, and Effect Sizes

Group	Pseudo-Experimental VR Group
1	Computer
2	Oculus
3	SmartPhone

ANOVA

ANOVA - DV

	Sum of Squares	df	Mean Square	F	p	η^2
Group	3.764880	2	1.8824402	4.589743	0.0111426	0.0398797
Residuals	90.641091	221	0.4101407			

CONTRASTS

Contrasts - Group

	Estimate	SE	t	p
1 - 2, 3	0.08903287	0.12516728	0.7113111	0.4776416
2 - 3	-0.29774070	0.09872192	-3.0159534	0.0028616

POST HOC TESTS

Post Hoc Comparisons - Group

Group		Group	Mean Difference	SE	df	t	p-tukey	p-scheffe	Cohen's d
1	-	2	0.23790323	0.14087412	221.0000	1.6887646	0.2117589	0.2424789	0.37147869
	-	3	-0.05983748	0.12791087	221.0000	-0.4678061	0.8864484	0.8964013	-0.09343441
2	-	3	-0.29774070	0.09872192	221.0000	-3.0159534	0.0080227	0.0115984	-0.46491310

Note. Comparisons are based on estimated marginal means

- Includes output from the jmv package in R for both Games-Howell and Tukey (and Scheffé)
- Also includes Helmert contrasts for information… but these are UNADJUSTED p values

OHIO UNIVERSITY

NORTHWEST MISSOURI STATE UNIVERSITY

Barcikowski's Most Explanatory Human-Friendly Comparisons (with Scheffe tests assuming equal variances)

Toothaker (1991) Table 3.3 (p. 72)

i Information

```
Total number of Comparisons tested:
[1] 75
```

	SSQ	Coef1	Coef2	Coef3	Coef4	Coef5	diff	lwr.ci	upr.ci	pval	Cohens.d
40	**9.1669**	**-0.5000**	**-0.5000**	**0.0000**	**0.0000**	**1.0000**	**9.9500**	**4.9738**	**14.9262**	**0.0001**	**3.7081**
60	7.6445	-0.3333	-0.3333	-0.3333	0.0000	1.0000	8.5667	3.8750	13.2583	0.0002	3.1926
7	7.5111	1.0000	0.0000	0.0000	0.0000	-1.0000	-10.4000	-16.1461	-4.6539	0.0002	3.8758
58	7.4084	-0.3333	-0.3333	0.0000	-0.3333	1.0000	8.4333	3.7417	13.1250	0.0002	3.1429
65	6.7167	-0.2500	-0.2500	-0.2500	-0.2500	1.0000	7.7750	3.2323	12.3177	0.0004	2.8976

Showing 1 to 5 of 32 entries

Previous 1 2 3 4 5 6 7 Next

Scheffe Tests of Maximum Comparisons assuming equal Variances

i Information

	Coef1	Coef2	Coef3	Coef4	Coef5	Family	Diff	lwr.ci	upr.ci	pval	Cohens.d
ScheffeMAX	-1.1354	-0.8909	0.1141	0.2227	1.6895	3,4,5-1,2	18.4076	9.3223	27.4929	0.0000	6.8601
ScaledMAX	-0.5603	-0.4397	0.0563	0.1099	0.8338	3,4,5-1,2	9.0842	4.6006	13.5678	0.0000	3.3855
HollingsworthMAX	-0.5078	-0.3984	0.0510	0.0996	0.7556	3,4,5-1,2	8.2321	4.1690	12.2952	0.0000	3.0679
BarcikowskiMAX	**-0.5000**	**-0.5000**	**0.0000**	**0.0000**	**1.0000**	**5-1,2**	**9.9500**	**4.9738**	**14.9262**	**0.0001**	**3.7081**

- Scheffé Family 345-12 not most explanatory after “humanized” (two coefficients very close to 0)
- 345-12 Family appears as 6th contrast on page 2 of results
- Barcikowski MAX does not include them (they are 0)
- Could simplify the SchefféMAX and test it specifically, but there could be multiple ways to simplify it
- The Barcikowski comparisons do this automatically

Based on Miller (1981)

OHIO UNIVERSITY

Barcikowski's Most Explanatory Human-Friendly Comparisons (with Scheffe tests assuming equal variances)

NORTHWEST MISSOURI STATE UNIVERSITY

Tamhane and Dunlop (2000) Table 12.5 (p. 479)

i Information

```
Total number of Comparisons tested:
[1] 25
```

	SSQ	Coef1	Coef2	Coef3	Coef4	diff	lwr.ci	upr.ci	pval	Cohens.d
18	**6.0495**	**-0.5000**	**-0.5000**	**0.0000**	**1.0000**	**12.5529**	**6.7574**	**18.3484**	**0.0000**	**3.0123**
23	6.0098	0.5000	0.5000	-0.5000	-0.5000	-10.2157	-14.9477	-5.4837	0.0000	2.4515
5	5.5153	1.0000	0.0000	0.0000	-1.0000	-13.8400	-20.5321	-7.1479	0.0000	3.3212
9	5.0798	1.0000	0.0000	-0.5000	-0.5000	-11.5029	-17.2984	-5.7074	0.0001	2.7604
22	4.2559	-0.3333	-0.3333	-0.3333	1.0000	9.9267	4.4626	15.3907	0.0002	2.3821

Showing 1 to 5 of 16 entries

Previous 1 2 3 4 Next

DV	Test_Scores
Group	Teaching_Method
1	1_Case
2	2_Formula
3	3_Equation
4	4_Unitary_Analysis

- Family 12-34 is 2nd most explanatory after being "Humanized"
- **One Barcikowski comparison is more explanatory than the Scheffé "family"**
- Pairwise 3rd

  from Sparks (1963)

Scheffe Tests of Maximum Comparisons assuming equal Variances

i Information

	Coef1	Coef2	Coef3	Coef4	Family	Diff	lwr.ci	upr.ci	pval	Cohens.d
ScheffeMAX	-1.5536	-0.9282	0.6731	1.8087	3,4-1,2	28.8130	16.2933	41.3327	0.0000	6.9143
ScaledMAX	-0.6260	-0.3740	0.2712	0.7288	3,4-1,2	11.6094	6.5650	16.6539	0.0000	2.7860
HollingsworthMAX	-0.5872	-0.3508	0.2544	0.6836	3,4-1,2	10.8903	6.1583	15.6223	0.0000	2.6134
BarcikowskiMAX	**-0.5000**	**-0.5000**	**0.0000**	**1.0000**	**4-1,2**	**12.5529**	**6.7574**	**18.3484**	**0.0000**	**3.0123**

OHIO UNIVERSITY

Barcikowski's Most Explanatory Human-Friendly Comparisons (with Scheffe tests assuming equal variances)

NORTHWEST MISSOURI STATE UNIVERSITY

Chick Weight (R dataset Day 21 only)

i Information

```
Total number of Comparisons tested:
[1] 25
```

	SSQ	Coef1	Coef2	Coef3	Coef4	diff	lwr.ci	upr.ci	pval	Cohens.d
7	**1.0462**	**1.0000**	**0.0000**	**-1.0000**	**0.0000**	**-92.5500**	**-167.7376**	**-17.3624**	**0.0101**	**1.4465**
9	0.9575	1.0000	0.0000	-0.5000	-0.5000	-76.6778	-140.0053	-13.3503	0.0117	1.1984
19	0.8936	-0.5000	-0.5000	1.0000	0.0000	74.0750	4.1309	144.0191	0.0340	1.1578
23	0.8275	0.5000	0.5000	-0.5000	-0.5000	-58.2028	-115.2060	-1.1995	0.0437	0.9097
1	0.7372	1.0000	-0.3333	-0.3333	-0.3333	-63.4352	-121.5461	-5.3243	0.0273	0.9915

Showing 1 to 5 of 7 entries

Previous 1 2 Next

Scheffe Tests of Maximum Comparisons assuming equal Variances

i Information

	Coef1	Coef2	Coef3	Coef4	Family	Diff	lwr.ci	upr.ci	pval	Cohens.d
ScheffeMAX	-2.7396	-0.1668	2.1586	0.7478	3,4-1,2	239.0904	52.5729	425.6079	0.0069	3.7369
ScaledMAX	-0.9426	-0.0574	0.7427	0.2573	3,4-1,2	82.2612	18.0882	146.4342	0.0069	1.2857
HollingsworthMAX	-0.7035	-0.1571	0.6650	0.1956	3,4-1,2	67.6319	14.2644	120.9993	0.0077	1.0571
BarcikowskiMAX	**1.0000**	**0.0000**	**-1.0000**	**0.0000**	**1-3**	**-92.5500**	**-167.7376**	**-17.3624**	**0.0101**	**1.4465**

4 protein diets

- Family 12-34 for SchefféMAX is 4th most explanatory after being "Humanized"
- 2 coefficients closer to 0 so BarcikowskiMAX is a different "family"
- **Sometimes PAIRWISE is the most explanatory**

OHIO UNIVERSITY

NORTHWEST MISSOURI STATE UNIVERSITY

Barcikowski's Most Explanatory Human-Friendly Comparisons (with Scheffe tests assuming equal variances)

Stevens (2007) Sesame Street Data (p. 100)

i Information

```
Total number of Comparisons tested:
[1] 25
```

	SSQ	Coef1	Coef2	Coef3	Coef4	diff	lwr.ci	upr.ci	pval	Cohens.d
9	**1.1547**	**1.0000**	**0.0000**	**-0.5000**	**-0.5000**	**-12.9782**	**-17.4947**	**-8.4617**	**.0000**	**1.3161**
23	1.0374	0.5000	0.5000	-0.5000	-0.5000	-10.0439	-13.6360	-6.4519	000	1.0185
5	0.9285	1.0000	0.0000	0.0000	-1.0000	-13.4379	-18.6064	-8.2694	0. 00	1.3627
1	0.8679	1.0000	-0.3333	-0.3333	-0.3333	-10.6083	-14.9009	-6.3157	0.00	1.0758
7	0.8058	1.0000	0.0000	-1.0000	0.0000	-12.5185	-17.6493	-7.3877	0.00	1.2695

Showing 1 to 5 of 19 entries

Previous 1 2 3 4 Next

- Another example where the "Family" changes

But…

- **An example of Violation of Homogeneity of Variances (see next slide)**

Scheffe Tests of Maximum Comparisons assuming equal Variances

i Information

	Coef1	Coef2	Coef3	Coef4	Family	Diff	lwr.ci	upr.ci	pval	Cohens.d
ScheffeMAX	-5.4001	-1.7459	3.2799	3.8661	3,4-1,2	82.7663	54.9992	110.5333	0000	8.3931
ScaledMAX	-0.7557	-0.2443	0.4590	0.5410	3,4-1,2	11.5821	7.6965	15.4678	0. 00	1.1745
HollingsworthMAX	-0.7308	-0.1918	0.4191	0.5035	3,4-1,2	10.8868	7.2308	14.5429	0.00	1.1040
BarcikowskiMAX	**1.0000**	**0.0000**	**-0.5000**	**-0.5000**	**1-3,4**	**-12.9782**	**-17.4947**	**-8.4617**	**0.000**	**1.3161**

From Educational Testing Service

OHIO UNIVERSITY

NORTHWEST MISSOURI STATE UNIVERSITY

Stevens (2007) Sesame Street Data (p. 100)

Omnibus Test & Assumptions

i Information

Omnibus ANOVA | Homoscedasticity Assumption | Normality Assumption

	Test	Statistic	df1	df2	Pval
1	Breusch-Pagan	10.7778	3.0000		0.0130
2	Levene (mean)	8.1381	3.0000	236.0000	0.0000
3	Levene (median)	7.3416	3.0000	236.0000	0.0001
4	**Levene (zero correction)**	**7.5196**			**0.0001**
5	Levene (zero removal)	7.4444			0.0001

Omnibus Test & Assumptions

i Information

Omnibus ANOVA | Homoscedasticity Assumption | Normality Assumption

	Test	Statistic	df1	df2	Pval
1	Fisher's F	23.4812	3.0000	236.0000	0.0000
2	**Welch's F**	**30.2588**	**3.0000**	**130.4942**	**0.0000**
3	Brown-Forsythe F	24.1719	3.0000	218.1689	0.0000
4	Kruskal-Wallis X2	64.6661	3.0000		0.0000

But…

- **An example of Violation of Homogeneity of Variances**
- Welch is statistically significant (as well as Brown-Forsythe omnibus test)

From Educational Testing Service

Brown-Forsythe Adjustments of Scheffe Tests for Unequal Variances

Stevens (2007) Sesame Street Data (p. 100)

i Information

Barcikowski MAX | Barcikowski 2nd | Barcikowski 3rd | Barcikowski 4th | Scaled Scheffe MAX | Hollingsworth MAX

```
COMPARISON 1 HAS COEFFICIENTS 1 0 -0.5 -0.5

BROWN-FORSYTHE SCHEFFE (UNEQUAL VARIANCES, BALANCED OR UNBALANCED GROUP SIZES))

COMPARISON IS STATISTICALLY SIGNIFICANT WHEN BROWN-FORSYTHE F (BF_F_Stat) IS LARGER THAN CONTRAST_BFCRIT
  ALPHA BF_Critical Contrast_BFcrit BF_F_Stat F_Sig Estimate
1  0.10    2.126232        6.378695  91.32938  TRUE  12.9782
2  0.05    2.674019        8.022056  91.32938  TRUE  12.9782
3  0.01    3.935294       11.805883  91.32938  TRUE  12.9782
```

Brown-Forsythe Adjustments of Scheffe Tests for Unequal Variances

i Information

Barcikowski MAX | Barcikowski 2nd | Barcikowski 3rd | Barcikowski 4th | Scaled Scheffe MAX | Hollingsworth MAX

```
Original Scheffe comparison coefficients were: -5.4 -1.746 3.28 3.866
Scaled Scheffe comparisons were original Coefficients divided by: 7.146027
SCALED SCHEFFE COMPARISON HAS COEFFICIENTS -0.756 -0.244 0.459 0.541

BROWN-FORSYTHE SCHEFFE (UNEQUAL VARIANCES, BALANCED OR UNBALANCED GROUP SIZES))

COMPARISON IS STATISTICALLY SIGNIFICANT WHEN BROWN-FORSYTHE F (BF_F_Stat) IS LARGER THAN CONTRAST_BFCRIT
  ALPHA BF_Critical Contrast_BFcrit BF_F_Stat BF_Sig Estimate
1  0.10    2.126232        6.378695  85.53974   TRUE 11.58214
2  0.05    2.674019        8.022056  85.53974   TRUE 11.58214
3  0.01    3.935294       11.805883  85.53974   TRUE 11.58214
```

- Provides Brown-Forsythe robust tests for SchefféMAX, Hollingsworth, and top 4 Barcikowski
- **Need to compare statistic to critical value (no calculation of p values available)**

From Educational Testing Service

OHIO UNIVERSITY

NORTHWEST MISSOURI STATE UNIVERSITY

Barcikowski's Most Explanatory Human-Friendly Comparisons (with Scheffe tests assuming equal variances)

chickwts (R dataset)

i Information

```
Total number of Comparisons tested:
[1] 196
```

	SSQ	Coef1	Coef2	Coef3	Coef4	Coef5	Coef6	diff	lwr.ci	upr.ci	pval	Cohens.d
193	**155.804**	**-0.3333**	**0.3333**	**0.3333**	**-0.3333**	**0.3333**	**-0.3333**	**-36.8333**	**-39.7929**	**-33.8738**	**0.0000**	**10.1916**
67	151.602	1.0000	-0.5000	-0.5000	0.0000	0.0000	0.0000	54.5000	50.0419	58.9581	0.0000	15.0799
66	150.215	-0.5000	1.0000	0.0000	-0.5000	0.0000	0.0000	-54.2500	-58.9499	-49.5501	0.0000	15.0108
174	138.806	-0.2500	0.5000	0.5000	-0.2500	-0.2500	-0.2500	-36.8750	-40.0713	-33.6787	0.0000	10.2032
19	137.809	1.0000	-1.0000	0.0000	0.0000	0.0000	0.0000	60.0000	54.6888	65.3112	0.0000	16.6018

Showing 1 to 5 of 181 entries

Previous 1 2 3 4 5 … 37 Next

Group	feed
1	casein
2	horsebean
3	linseed
4	meatmeal
5	soybean
6	sunflower

Scheffe Tests of Maximum Comparisons assuming equal Variances

i Information

	Coef1	Coef2	Coef3	Coef4	Coef5	Coef6	Family	Diff	lwr.ci	upr.ci	pval	Cohens.d
ScheffeMAX	2.0798	-1.7919	-1.3748	1.1633	-0.5346	0.4583	1,4,6-2,3,5	170.2087	157.8044	182.6130	0.0000	47.0961
ScaledMAX	0.5619	-0.4841	-0.3714	0.3143	-0.1444	0.1238	1,4,6-2,3,5	45.9857	42.6344	49.3370	0.0000	12.7241
HollingsworthMAX	0.5880	-0.5913	-0.3751	0.3620	-0.1196	0.1360	1,4,6-2,3,5	50.8744	47.1535	54.5954	0.0000	14.0768
BarcikowskiMAX	**-0.3333**	**0.3333**	**0.3333**	**-0.3333**	**0.3333**	**-0.3333**	**2,3,5-1,4,6**	**-36.8333**	**-39.7929**	**-33.8738**	**0.0000**	**10.1916**

- Four Human contrasts more explanatory than first Pairwise
- **Maybe something useful theoretically from combining 2,3,5 together vs 1,4,6**
- Or… 1 vs 23
- Or… 2 vs 14
- Or… 23 vs 1456

OHIO UNIVERSITY

NORTHWEST MISSOURI STATE UNIVERSITY

Barcikowski's Most Explanatory Human-Friendly Comparisons (with Scheffe tests assuming equal variances)

Keselman Cribbie & Holland (2004)

Pairwise multiple comparison test procedures: An update for clinical child and adolescent psychologists

i Information

```
Total number of Comparisons tested:
[1] 75
```

	SSQ	Coef1	Coef2	Coef3	Coef4	Coef5	diff	lwr.ci	upr.ci	pval	Cohens.d
15	**2.1272**	**1.0000**	**0.0000**	**0.0000**	**-0.5000**	**-0.5000**	**-1.4772**	**-2.1888**	**-0.7657**	**0.0000**	**1.7863**
66	2.0849	0.5000	0.5000	-0.3333	-0.3333	-0.3333	-1.0901	-1.6204	-0.5597	0.0000	1.3181
7	2.0036	1.0000	0.0000	0.0000	0.0000	-1.0000	-1.6555	-2.4771	-0.8339	0.0000	2.0018
40	2.0010	-0.5000	-0.5000	0.0000	0.0000	1.0000	1.4327	0.7212	2.1443	0.0000	1.7325
43	1.8900	1.0000	0.0000	-0.3333	-0.3333	-0.3333	-1.3128	-1.9837	-0.6420	0.0000	1.5875

Showing 1 to 5 of 38 entries Previous 1 2 3 4 5 … 8 Next

DV	Y
Group	Level
1	1_TwoBelow
2	2_OneBelow
3	3_Same
4	4_OneAbove
5	5_TwoAbove

Scheffe Tests of Maximum Comparisons assuming equal Variances

i Information

	Coef1	Coef2	Coef3	Coef4	Coef5	Family	Diff	lwr.ci	upr.ci	pval	Cohens.d
ScheffeMAX	-2.9633	-1.4577	0.3623	1.4269	2.6318	3,4,5-1,2	5.9176	3.3195	8.5158	0.0000	7.1556
ScaledMAX	-0.6703	-0.3297	0.0820	0.3228	0.5953	3,4,5-1,2	1.3385	0.7508	1.9262	0.0000	1.6185
HollingsworthMAX	-0.6626	-0.3259	0.0810	0.3191	0.5885	3,4,5-1,2	1.3232	0.7423	1.9042	0.0000	1.6000
BarcikowskiMAX	**1.0000**	**0.0000**	**0.0000**	**-0.5000**	**-0.5000**	**1-4,5**	**-1.4772**	**-2.1888**	**-0.7657**	**0.0000**	**1.7863**

- Family 345-12 but two coefficients close to 0
- Barcikowski does not include them
- TwoBelow vs (OneAbove & TwoAbove)
- **Below vs (Same & Above)**
- TwoBelow vs TwoAbove

OHIO UNIVERSITY

NORTHWEST MISSOURI STATE UNIVERSITY

Barcikowski's Most Explanatory Human-Friendly Comparisons (with Scheffe tests assuming equal variances)

Maxwell and Delaney (2004) Table 5.4 (p. 206)

i Information

```
Total number of Comparisons tested:
[1] 25
```

	SSQ	Coef1	Coef2	Coef3	Coef4	diff	lwr.ci	upr.ci	pval	Cohens.d
2	**1.8134**	**0.0000**	**1.0000**	**-0.5000**	**-0.5000**	**-13.3333**	**-25.6573**	**-1.0094**	**0.0308**	**1.6493**
4	1.7213	0.0000	1.0000	0.0000	-1.0000	-15.0000	-29.2304	-0.7696	0.0363	1.8554
20	1.6372	-0.3333	1.0000	-0.3333	-0.3333	-11.9444	-23.5636	-0.3253	0.0424	1.4775
13	1.4893	-0.5000	1.0000	0.0000	-0.5000	-12.0833	-24.4073	0.2406	0.0560	1.4946

Showing 1 to 4 of 4 entries

Previous 1 Next

New_Value	Original_Value
DV	SystolicBP
Group	Treatment
1	Biofeedback
2	Combination
3	Diet
4	Drug

Scheffe Tests of Maximum Comparisons assuming equal Variances

i Information

	Coef1	Coef2	Coef3	Coef4	Family	Diff	lwr.ci	upr.ci	pval	Cohens.d
ScheffeMAX	0.0458	-1.9695	0.5954	1.3283	1,3,4-2	27.2909	2.6430	51.9387	0.0264	3.3757
ScaledMAX	0.0233	-1.0000	0.3023	0.6744	1,3,4-2	13.8566	1.3420	26.3712	0.0264	1.7140
HollingsworthMAX	0.0187	-0.8041	0.2431	0.5423	1,3,4-2	11.1415	1.0790	21.2039	0.0264	1.3781
BarcikowskiMAX	**0.0000**	**1.0000**	**-0.5000**	**-0.5000**	**2-3,4**	**-13.3333**	**-25.6573**	**-1.0094**	**0.0308**	**1.6493**

Their example was a priori contrasts:

- Drug vs Bio
- Drug vs Diet
- Bio vs Diet
- **Combo vs Avg**
- Combo vs Drug was only significant Tukey

Hypothetical data

Data References and websites

- Adjanin, N., & Brooks, G. P. (2023). Witnessing the Last Tropical Glaciers: Student Use of Virtual Reality Technology to Learn about Climate Change and Protecting Endangered Environments. *TOJET: Turkish Online Journal of Educational Technology, 22*(4), 248-257. http://www.tojet.net/articles/v22i4/22424.pdf
- Bruning, J. L., & Kintz, B. L. (1997). Computational handbook of statistics (4th ed.). Addison Wesley Longman.
- ChickWeight: https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/ChickWeight.html
- Chickwts: https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/chickwts.html
- Howell, D. C. (2004). Fundamental statistics for the behavioral sciences (5th ed.). Thomson Brooks/Cole.
- InsectSprays: https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/InsectSprays.html
- Keselman, H. J., Cribbie, R. A., & Holland, B. (2004). Pairwise multiple comparison test procedures: An update for clinical child and adolescent psychologists. *Journal of Clinical Child and Adolescent Psychology, 33*(3), 623-645.
- Maxwell, S. E., & Delaney, H. D. (2004). Designing experiments and analyzing data: A model comparison perspective (2nd ed.). Lawrence Erlbaum Associates.
- Pituch, K. A., & Stevens, J. P. (2016). Applied multivariate statistics for the social sciences (6th ed.). Routledge.
- Smiles & Leniency: https://onlinestatbook.com/case_studies_rvls/smiles/index.html
- Stevens, J. P. (2007). *Intermediate statistics: A modern approach* (3rd ed.). Lawrence Erlbaum Associates.
- Toothaker, L. E. (1991). Multiple comparisons for researchers. Sage.

References

- Brooks, G. P., & Adjanin, N. (2023, April). Back to the Future: Human-friendly Scheffé Contrasts, or, the Art of Multiple Comparisons. Paper presented at the 2023 annual meeting of the American Educational Research Association, Chicago.
- Brooks, G. P., Adjanin, N., Oppong, F., & Liu, Y. (accepted for publication). Human-friendly Scheffé comparisons, or, the art of complex multiple comparisons. *General Linear Model Journal.* {glmj.org}
- Hollingsworth, H. (1978). The coefficients of the normalized maximum contrast as statistics for posttest ANOVA data interpretations. *Journal of Experimental Education, 46*(4), 4-6.
- Hollingsworth, H. (1980/1981). Maximized posttest comparisons: A clarification. *Journal of Experimental Education, 49*(2), 92-93.
- Keppel, G., & Wickens, T. D. (2004). *Design and analysis: A researcher's handbook* (4th ed.). Pearson Prentice Hall.
- Kirk, R. E. (2013). *Experimental design: Procedures for the behavioral sciences* (4th ed.). SAGE.
- Maxwell, S. E., Delaney, H. D., & Kelley, K. (2018). *Designing experiments and analyzing data: A model comparison perspective* (3rd ed.). Routledge.
- Oppong, F. A., Liu, Y., Adanin, N., & Brooks, G. P. (2023, October). "Everything Old Is New Again": Human-friendly Scheffé contrasts and "All That Jazz." Paper presented at the 2023 annual meeting of the Mid-Western Educational Research Association, Cincinnati.
- Scheffé, H. (1953). A Method for Judging all Comparisons in the Analysis of Variance. *Biometrika, 40*, 87-104.
- Schmid, J. (1977). Editor's commentary: Meaningless complex posttest comparisons. *Journal of Experimental Education, 46*(1), 4-5.
- Williams, J. D. (1979/1980). A note on maximized posttest comparisons. *Journal of Experimental Education, 48*(2), 116-118.

72x6cr-gordon-brooks.shinyapps.io/Human_Friendly_Contrasts/

https://tinyurl.com/35bkmk5u

Thank you!

SCAN ME