Orthogonal Comparisons A Teaching Example

Keith McNeil
New Mexico State University

When the omnibus one-way Analysis of Variance (ANOVA) is found to be significant, the research question that "at least two populations have different means" can be accepted, but is found to be lacking. (What most textbooks fail to mention is that this means that the one-way ANOVA question is a fruitless question.) Most textbooks turn to post-hoc analyzes as a way to determine "where the significance is." But that journey is often muddled by: a) discussion of a myriad of post-hoc procedures, b) insufficient parallel examples, c) downplay of the value of planned comparisons, and d) failure to tie orthogonal comparisons to the two-way ANOVA. This paper will attempt to alleviate the above issues, with various examples of four groups.

Suppose that a researcher is interested in comparing four different treatments, and is encouraged to "first conduct the one-way ANOVA." The research hypothesis being tested here is, "For the population, at least two of these treatments are differentially effective." Given that the omnibus F is significant, the researcher can conclude, "For the population, at least two of these treatments are differentially effective." Note that which treatments are different cannot be specified. Nor can the more effective treatment be specified. The omnibus one-way F can be called a non-specific, non-directional research hypothesis, yielding little (or no) information.

Post-Hoc Comparisons

The myriad of post-hoc comparisons have been developed to attempt to rectify the non-specificity problem. These procedures protect the Type I error, some with orthogonal comparisons. It is this family of orthogonal comparisons on which the remainder of the paper will focus.

Orthogonal Comparisons

A comparison is said to be orthogonal if the set of contrast coefficients sum to zero, and if the sum of cross products with all other orthogonal comparisons also sums to zero. The set of contrast coefficients for RH1 in Exhibit 1 meets both criteria, as the set of coefficients sums to 0 $(1 + 0 + 0 + -1 = 0)$, and the sum of the cross products of set 1 with set 2 also sums to 0 $[(1 \times 0) + (0 \times 1) + (0 \times -1) + (-1 \times 0) = 0]$. Each orthogonal comparison is a t-test question, either comparing one group to another (as in RH1 and RH2), or some combination of groups to some other combination of groups (as in RH3). With four groups, there is three degrees of freedom associated with the Between groups sum of squares. The three orthogonal contrasts identify three ways this sum of squares can be partitioned. It should be noted here that there are many (infinite?) ways that the sum of squares can be partitioned--some more meaningful for how the four groups were determined.

An example of when research hypothesis 1 (RH1), RH2, and RH3 might be of interest is when a researcher is studying two classes of each of two teachers, one in the AM and one in the PM. Let's assume that M1 is Teacher A, AM; M4 is Teacher A, PM. RH1 could be: "There is a difference in the effectiveness of Teacher A in the PM from that in the AM." Further assume that M1 is Teacher B, AM and M3 is Teacher B, PM. RH2 could be: "There is a difference in the effectiveness of Teacher B in the PM from that in the AM." While RH1 and RH2 both compare teacher effectiveness of AM and PM, the comparisons are on different teachers, so what is found with RH1 (Teacher A) will not have a bearing on what is found with RH2 (Teacher B). In this case, the data to determine the answer to RH1 is different from that determining the answer to RH2. (The data doesn't have to be different in order for orthogonality to hold, as evidenced by RH3, but it certainly clarifies the issue). RH3 compares the effectiveness of Teacher A (averaged over AM and PM) with the effectiveness of Teacher B (averaged over AM and PM). Logically, the outcome of RH1 (the relative effectiveness of Teacher A at AM and PM), and the outcome of RH2 (the relative effectiveness of Teacher B at AM and PM) does not impinge on the overall effectiveness of Teacher A as compared to Teacher B.

Exhibit 1 One Possible Set Of Contrast Coefficients With Four Groups: Non-Directional Hypotheses

	M1	M2	M3	M4
RH1 Non-directional: M1 not equal M4 SH: M1 = M4 OR 1*M1-1*M4 = 0	1	0	0	-1
RH2 Non-directional: M2 not equal M3 SH: M2 = M3 OR 1*M2 -1*M3 = 0	0	1	-1	0
RH3 Non-directional: M1+M4 not equal M2+M3 SH: M1+M4 = M2+M3 OR 1*M1 + 1*M4 -1*M2 -1*M3 = 0	1	-1	-1	1

Directional, Planned Orthogonal Comparisons

The above research hypotheses were non-directional, which is to say that differences were expected, but not directionally specified. For RH1, if the orthogonal contrast if found to be significant, then the conclusion is simply a restatement of the research hypothesis, "There is a difference in the effectiveness of Teacher A in the PM from that in the AM." While we know now that "groups M1 and M4 are different," we do not know how they are different. A directional conclusion can be made if the direction was posited in the research hypothesis before the data were looked at (preferably before the data were collected). Orthogonal contrasts specified before data collection are referred to as planned comparisons, and may be directional. Directional conclusions cannot be made from any post-hoc comparisons, only from planned comparisons. Exhibit 2 contains the same set of orthogonal comparisons as in Exhibit 1, but here as planned comparisons with expectations: (RH1') Teacher A being more effective in the AM than the PM, (RH2'), Teacher B being more effective in the AM than the PM, and (RH3') Teacher A being more effective than Teacher B (averaging over AM and PM classes).

Notice that the statistical hypothesis (SH) is the same in Exhibit 1 and Exhibit 2, and the orthogonal coefficients are the same. Again, what is different is the expected direction, and the permissible conclusion.

RH4, RH5, and RH6 in Exhibit 3 are another set of three orthogonal contrasts. While RH5 and RH2 are exactly the same, RH4 and RH6 are different from RH1 and RH3. The coefficients within RH4, RH5, and RH6 all add up to zero, and the sum of the cross products add up to zero, thus RH4, RH5, and RH6 constitute a different set of three orthogonal contrasts. Which set a researcher should use depends on the design of the study and the questions one has of the groups. Indeed, there are many other sets of orthogonal contrasts. As in all research, the questions should guide the analysis. With post-hoc comparisons, the researcher is limited to one less question than there are groups.

An example of when RH4, RH5, and RH6 might be of interest is when a researcher is testing the effectiveness of three different New treatments (M1, M2, and M3) and one Comparison treatment (M4). Since there are four groups, three orthogonal questions can be asked, and if the questions are asked before inspection of the data, Directional Research Hypotheses can be tested. Indeed, if a New treatment is being researched,we should expect it to be better than the Existing treatment. RH4 determines if the avererage

Exhibit 2 One Possible Set Of Contrast Coefficients With Four Groups: Directional Hypotheses

	M1	M2	M3	M4
RH1' Directional: M1 > M4 SH: M1 = M4 OR 1*M1-1*M4 = 0	1	0	0	-1
RH2' Directional: M2 > M3 SH: M2 = M3 OR 1*M2 -1*M3 = 0	0	1	-1	0
RH3' Directional: M1+M4 > M2+M3 SH: M1+M4 = M2+M3 OR 1*M1 + 1*M4 -1*M2 -1*M3 = 0	1	-1	-1	1

Exhibit 3 Another Possible Set Of Contrast Coefficients With Four Groups

M1 = New Treatment #1 M2 = New Treatment #2 M3 = New Treatment #3 M4 = Existing Treatment

		M1	M2	M3	M4
RH4	Non-directional: (M1+M2+M3)/3 not equal M4 Directional: (M1+M2+M3)/3 > M4				
SH	(M1+M2+M3)/3 = M4 OR 1*M1 + 1*M2 + 1*M3 - 3*M4 = 0	1	1	1	-3
RH5	Non-directional: M2 not equal M3 Directional: M2 > M3				
SH:	M2 = M3 OR 1*M2 - 1*M3 = 0	0	1	-1	0
RH6	Non-directional: M1 not equal (M2+M3)/2 Directional: M1 > (M2+M3)/2				
SH:	M1 = (M2+M3)/2 OR 2*M1 -1*M2 -1*M3 = 0	2	-1	-1	0

of the three New treatments is better than the one Comparison treatment. RH5 tests if New treatment 2 is better than New treatment 3. Finally, RH6 tests if New treatment 1 is better than the average of the other New treatments. As should now be clear, the design of the research, and the desired conclusion(s) determine the choice of the hypotheses, and whether the hypotheses are directional or non-directional. No one choice is always correct; the choice will depend on the research questions!

Pictorial Representation of Orthogonal Comparisons

Notice that RH5 and RH2 are the same, in terms of contrast coefficients. Since the two Exhibits were discussed with different samples, the research hypotheses may have seemed different. But in both cases, M2 was contrasted with M3. The sum of squares due to the four groups, though, was partitioned in different ways, as depicted in the Venn diagram in Figure 1. Figure 1a illustrates the one-way partitioning of sum of squares, into Within groups and Between groups. Note that the Between groups is between the four groups. Figure 1b illustrates the contrasts in Exhibit 1. About one-half of the Between groups sum of squares is due to RH2, and about one-fourth is due to RH1 and one-fourth to RH3. If the contrasts in Exhibit 3 were applied to the same data as in Exhibit 1, then Figure 1c might result. Note that since RH5 and RH2 are the same contrast, the sum of squares attributable to those contrasts is the same. But since RH4 and RH6 are different from RH1 and RH3, the sum of squares partitioned to these hypotheses will likely be different. RH6 is shown to account for none of the sum of squares in Figure 1c, while RH4 accounts for one-half of the Between groups sum of squares.

Source Tables

Another way to comprehend the different comparisons depicted in the Exhibits and in Figure 1 is through the source tables in Tables 1 through 3. Table 1 contains the one-way results, with the Total sum of squares being partitioned into just Between and Within. The four groups account for 40% of the Total sum of squares.Table 2 contains the partitioning depicted in Exhibit 1. Notice that the Total and Within sum of squares is the same as in Table 1, but the sum of squares due to Between groups has been further partitioned into the three comparisons. The RH2 comparison accounts for half of the sum of squares due to groups (20/40--hence half the overlapped area in Figure 1b). Since all of the F values in Table 2 fall beyond the critical value, all of these comparisons would be significant. Table 3 reflects the contrasts in Exhibit 3 and Figure 1c. Again notice that the sum of squares for RH2 in Table 2 and RH5 in Table 3 is the same. RH4 and RH6 are different from RH1 and RH3, and therefore the sum of squares is different. RH6 accounts for none of the sum of squares and is therefore not significant.

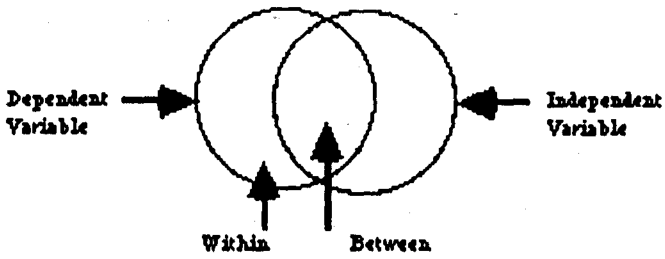
Example of Non-orthogonal Hypotheses

The reader may wonder why each of the New treatments in Exhibit 3 were not compared to the Existing treatment. These may be interesting research hypotheses, but they are not orthogonal. Exhibit 4 contains the hypotheses and orthogonal coefficients. While the coefficients do sum to zero within each of the hypotheses, the sum of the cross products is not zero. Think of it this way--if we start out by assuming all four treatments are equal, but find one inferior to another, isn't it likely that that one will be inferior to one of the others as well? In this case, the results from one hypothesis have a bearing on the results from another. Once we know the answer to one hypothesis, we have an inkling as to the answer to the other hypothesis. Additionally these hypotheses as a set are of little value, because they do not lead to a conclusive

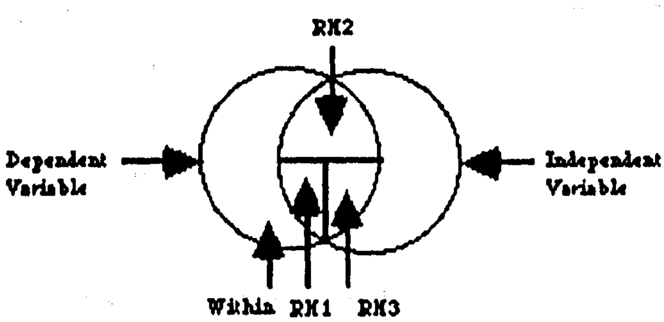
Figure 1 Hypothetical Sample Means and Venn Diagrams

	AM	PM
TEACHER A	7.5 (M1)	12.5 (M4)
TEACHER B	10.0 (M2)	20.0 (M3)

1A One Way Analysis



1B Exhibit 1 or 2 Analysis



1C Exhibit 3 Analysis

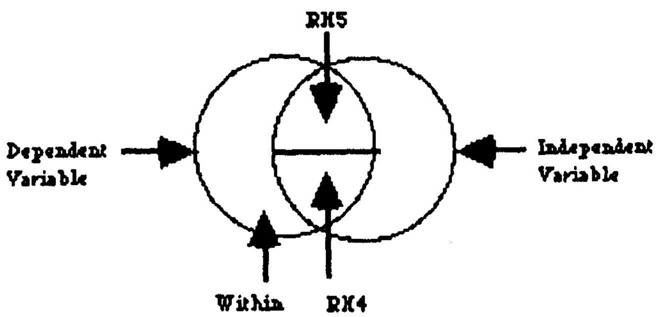


Table 1 One-way Source Table

SOURCE	SS	df	MS	F	F_{cv}
BETWEEN	40	3	13.33	13.33	2.76
WITHIN	60	60	1.00		
TOTAL	100	63			

Note. All F_{cv} are at alpha = .05.

Table 2 Exhibit 1 Source Table

SOURCE	SS	df	MS	F	F_{cv}
RH1	10	1	10	10.0	4.0
RH2	20	1	20	20.0	4.0
RH3	10	1	10	10.0	4.0
WITHIN	60	60	1		
TOTAL	100	63			

Note. All F_{cv} are at alpha = .05.

Table 3 Exhibit 3 Source Table

SOURCE	SS	df	MS	F	F_{cv}
RH4	20	1	20	20.0	4.0
RH5	20	1	20	20.0	4.0
RH6	0	1	0	0.0	4.0
WITHIN	60	60	1		
TOTAL	100	63			

Note. All F_{cv} are at alpha = .05.

answer. Suppose that all of the New treatments were better than the Existing treatment. Which New treatment would you recommend? The set of orthogonal hypotheses in Exhibit 3, on the other hand, lead to such a definite recommendation.

Trend Analysis

When the treatments are ordered on some underlying continuum, one may want to investigate the trends in the data as in Exhibit 5. That is, does the criterion increase linearly with an increase in the underlying continuum (as in RH10), or is there a minimum performance as in RH11? (By reversing all the weights in RH11, one could investigate maximum performance.) Finally, with four groups there may be a quadratic trend as in RH12. Note that the coefficients for RH10, RH11, and RH12 all add to zero, and that the cross products all add to zero. Therefore, RH10, RH11, and RH12 constitute another set of orthogonal contrasts for four groups.

Two Factors

Now suppose that the four groups differ not on just one underlying factor as in the above examples, but on two underlying factors. Exhibit 6 posits the following example of two groups getting the New treatment and two groups getting the Comparison treatment. Thus the first underlying factor is treatment: New vs. Comparison.
One of the New treatment groups is in the AM and one is in the PM. One of the Comparison treatment groups is in the AM and one is in the PM. Thus, the second factor is time of treatment: AM vs. PM.

What would be the research hypotheses of interest with this design? One probably would want to compare the New treatments to the Comparison treatments, and

Exhibit 4 Another Possible Set of Contrast Coefficients With Four Groups: Non-Orthogonal

M1 = New Treatment #1   M2 = New Treatment #2   M3 = New Treatment #3   M4 = Existing Treatment

		M1	M2	M3	M4
RH4	Non-directional: M1 not equal M4 Directional: M1 > M4				
SH	M1 = M4 OR 1*M1 + 0*M2 + 0*M3 - 1*M4 = 0	1	0	0	-1
RH5	Non-directional: M2 not equal M4 Directional: M2 > M4				
SH	M2 = M4 OR 1*M2 - 1*M4 = 0	0	1	0	-1
RH6	Non-directional: M3 not equal M4 Directional: M3 > M4				
SH	M3 = M4 OR 1*M3 -1*M4 = 0	0	0	1	-1

possibly the AM treatments to the PM treatments. These two hypotheses will be developed first, and then we will turn our attention to the third orthogonal comparison.

The Non-directional Research Hypothesis for treatment would be: "The two treatments, averaged across the two different time periods, are not equally effective," resulting in the orthogonal coefficients for RH13 in Exhibit 6. One could have stated this Research Hypothesis with a directional expectation, resulting in the same set of orthogonal coefficients. The Non-directional Research Hypothesis for time of treatment would be RH14: "The two time periods, averaged across the two different treatments, are not equally effective." Again, one could have stated this hypothesis with a directional expectation. Notice that the coefficients for RH14 are orthogonal to those for RH13. RH13 and RH14 are referred to as "main effects" hypotheses within the Analysis of Variance framework. Unless stated directionally a priori, they are always tested in a non-directional fashion.

Given the above two orthogonal contrasts, the third orthogonal contrast would have to be that specified in RH15. The non-directional research hypothesis associated with these coefficients is: "The difference between AM New treatment and PM New treatment is different from the difference between AM Comparison treatment and PM Comparison treatment." Again, one could have stated this hypothesis with a directional expectation. (For example, "The difference between AM New treatment and PM New treatment is different from the difference between AM Comparison treatment and PM comparison treatment." Again, one could have

Exhibit 5 One Possible Set of Contrast Coefficients With Four Groups: Trend Analysis

		M1	M2	M3	M4
RH10 linear trend	Non-directional: -3M1 -1M2 +1M3 +3M4 not equal 0 Directional: -3M1 -1M2 +1M3 +3M4 > 0				
SH	-3M1 -1M2 +1M3 +3M4 = 0 OR -3*M1 -1*M2 +1*M3 +3*M4 = 0	-3	-1	1	3
RH11 quadratic trend	Non-directional: M1 -M2 -M3 + M4 not equal 0 Directional: M1 -M2 -M3 + M4 > 0				
SH	M1 -M2 -M3 + M4 = 0 OR 1*M1 -1*M2 -1*M3 +1*M4 = 0	1	-1	-1	1
RH12 cubic trend	Non-directional: -M1 +3M2 -3M3 +M4 not equal 0 Directional: -M1 +3M2 -3M3 +M4 > 0				
SH	-M1 +3M2 -3M3 +M4 = 0 OR -1*M1 +3*M2 -3*M3 +1*M4 = 0	-1	3	-3	1

Exhibit 6 One Possible Set of Contrast Coefficients: Two-Way Analysis Of Variance

M1 = New treatment, AM　　M2 = New treatment, PM
M3 = Comparison treatment, AM　　M4 = Comparison treatment, PM

		M1	M2	M3	M4
RH13	Non-directional: The two treatments, averaged across the two different time periods, are not equally effective (M1+M2)/2 not equal (M3+M4)/2 Directional: The New treatment, averaged across the two different time periods, is more effective than the Comparison treatment (M1+M2)/2 > (M3+M4)/2				
SH	(M1+M2)/2 = (M3+M4)/2 OR (M1+M2) = (M3+M4) OR (M1+M2) - (M3+M4) = 0 OR 1*M1 +1*M2 -1*M3 -1*M4 = 0	1	1	-1	-1
RH14	Non-directional: The two time periods, averaged across the two treatments, are not equally effective (M1+M3)/2 not equal (M2+M4)/2 Directional: The AM period, averaged across the two different treatments, is more effective than the PM period (M1+M3)/2 > (M2+M4)/2				
SH	(M1+M3)/2 = (M2+M4)/2 OR (M1+M3) = (M2+M4) OR (M1+M3) - (M2+M4) = 0 OR 1*M1 -1*M2 +1*M3 -1*M4 = 0	1	-1	1	-1
RH15	Non-directional: The difference in effectiveness of the AM New treatment and the PM New treatment is different from the difference between the AM Comparison treatment and the PM Comparison treatment (M1 - M2) not equal (M3 - M4) Directional: The difference in effectiveness of the AM New treatment and the PM New treatment is greater than the difference between the AM Comparison treatment and the PM Comparison treatment (M1 - M2) > (M3 - M4)				
SH	The difference in effectiveness of the AM New treatment and the PM New treatment is the same as the difference between the AM Comparison treatment and the PM Comparison treatment (M1 - M2) = (M3 - M4) OR (M1 - M2) - (M3 - M4) = 0 OR 1*M1 -1*M2 -1*M3 +1*M4 = 0	1	-1	-1	1

stated this hypothesis with a directional expectation. (For example, "The difference between AM New treatment and PM New treatment is greater than the difference between AM Comparison treatment and PM Comparison treatment.") RH15 is referred to in the ANOVA literature as the "interaction" hypothesis.

An alternative way of stating this hypothesis is by looking at the differences within time, rather than within treatment: "The difference between AM New Treatment and AM Comparison Treatment is greater than the difference between PM New Treatment and PM Comparison Treatment. Both statements yield the same orthogonal coefficients, since they are the same question.

Summary

Discussing various sets of orthogonal comparisons for four groups should help illustrate the fact that there are many possible contrasts. The "appropriate contrast" depends on the design of the study and the research hypotheses of the researcher. While four groups were chosen for all the examples, the same conclusions can be developed for other numbers of groups. Four groups, though, does make the link to two-way ANOVA easy.

Few statistical texts make the link between orthogonal comparisons and the two-way ANOVA. Few also encourage directional hypothesis testing when there is one degree of freedom, as in the planned orthogonal comparisons. The reader is reminded that although all these orthogonal comparisons (as well as many others) can be made on these four groups, only some of the comparisons make sense for any one design. For instance, trend analysis is appropriate to neither the teacher-time design in Exhibits 1-4, nor the two-way ANOVA design in Exhibit 6.