**E1-a**

1. Means of pre-test scores of the three methods are very similar to each other.
2. The standard deviation of method A is the smallest among the three methods, while the standard deviation of method C is the largest.
3. The numbers of participants are the same that were assigned to each method.



**E1-b**

1. Means of post-test scores of the three methods are very different. Method A has the smallest mean post-test scores, while method C has the largest.
2. Post-test scores under method B have the smallest standard deviation, while that under method A have the biggest standard deviation.
3. Again, the numbers of participants are the same that were assigned to each method.
4. Mean post-test scores under method A and B decreased compared with pre-test condition, while mean post-test score under method C increased.
5. Standard deviations in all three groups decreased a lot compared with pre-test condition.



**E2-a**

It looks to me that there is a linear relationship between X and Y.



**E2-b**

I obtained the Pearson correlation coefficient between pre and post-test scores. The correlation between pre and post-test scores is .63, and the p-value is smaller than .05, which indicates that pre and post-test scores have a significantly positive correlation.

This notable correlation between the DV and the covariate indicates that pre-test score could be a reasonable covariate.



**E2-c**

According to the graph below, it looks like the slopes of regression lines for method B and method C are pretty parallel. Even though the regression line for method A is not parallel to that of method B, they do not cross within the range of the X axis. The regression line for method A does not cross that of method C, either. It is quite possible that the interaction effects between pre-test and method are not statistically significant. Therefore, it looks to me that the parallel slopes assumption is reasonable.



**E3-a**

According to the tables below, the one-way ANOVA shows that there’s no statistically significant difference between the means of pre-test scores across different levels of methods (p-value = .9980), and that is to say the pre-test scores are independent of methods. Therefore, this assumption of ANCOVA models is met.



**E3-b**

1. According to E2-a, there’s a linear relationship between X and Y.
2. According to E2-b, there’s a significant correlation between X and Y, indicating that X could be a reasonable covariate.
3. According to E2-c, the assumption of homogeneity of regression slopes is met.
4. According to E3-a, the assumption of independence of the covariate and treatment effect is met.
5. Therefore, the ANCOVA model seems appropriate for the data.

**E4-a**

According to the tables below, training type has a statistically significant main effect on post-test scores (p-value<.0001).



**E4-b**

According to tables and the graph below, method A and method B do not have significant pairwise mean difference. Method A and method C as well method B and method C have significant pairwise mean difference.



**E5-a**

According to the tables below, the model with X (pre-test scores), methods, and their interaction effect as predictor, account for a significant proportion of Y (post-test scores), because the p-value is smaller than .0001. Also, the value of R-square given in the table below means approximately 90% of the variance in Y (post-test scores) is accounted for by the model.



**E5-b**

According to the Type I SS test, the main effect of pre-test scores is statistically significant given that no other predictors are in the model (p-value <.0001).

**E5-c**

According to the Type I SS test, the main effect of method is statistically significant given that pre-test scores is included in the model (p-value<.0001).



**E6-a**

The estimated within-group variance of E4-(a) is 545.7, while that of E5-(a) is 176.5, which is much smaller than the former. Therefore, the pre-test score is effective in reducing the error variance.

**E4-a:**



**E5-a**



**E6-b**

The reduction in the error variance does affect the Tukey pairwise comparisons. Before the reduction, group means of method A and method B are not significantly different. However, after the reduction of error variance, all three groups have statistically significant different group means.

 