

Assumption (PAIRED)	Method used to test Assumption	Result of method	Decision (Met or did not meet Assumption)
Independent (x) variable can be grouped.	Observation of our dataset. Analysis of how the real-life test was conducted.	Independent variable can be grouped into new_test versus old_test treatment.	Met assumption.
Sample(s) demarcation and Independent observations within groups.	Measure of variance for the pairwise difference of scores between groups and correlation between group scores.	Sort the data and we realize that there is a strong correlation of 0.965 between new vs old but the 8.24 variance of the difference of scores indicate independence of observations. ¹ Each student should be unique.	Met assumption
Interval or ratio type continuous dependent(y) variable	Observation of our dataset. Analysis of how data was measured. Graded on a scale of a 100.	Ratio type (continuous classification) dependent variable.	Met assumption
Normality of inter-group difference or variance of scores	Compute mean, median, mode; Normality histogram with curve; Skewness and kurtosis z-score within ± 2.58 range; Shapiro-wilk (S-W) test. All tests above are performed on the attribute of scores difference (score1-score2).	S-W test with significance of $0.908 > 0.01$ (or 0.05) indicates normality of group(s) difference. Also, median==mode==mean.	Met assumption
Population normally distributed	Similar test to cell immediately above	S-W test with sig of $0.149 > 0.01$ indicates population normality. Samples show normality.	Met assumption

¹ This test helps us know that students or subjects did not mistakenly take the same test or treatment. The high correlation tempts me to think of these treatments as possibly identical but given the variance of the score differences we can postulate that the groups are different (0 variance for difference of scores could indicate same treatment). Adding our knowledge of testing, high performers will be high performers and low performers will be low performers irrespective of test type so the correlation trend is valid. Variance tells us if different treatments were applied to the students. Although, a second attempt to the old test can yield similar results as seen with the new test genre so in our case we can blindly assume the new test treatment was correctly administered.

Assumption (Independent)	Method used to test Assumption	Result of method	Decision (Met or did not meet Assumption)
Independent (x)variable can be grouped.	Observation of our dataset. Analysis of how the test was conducted.	Independent variable can be grouped into new_test versus old_test.	Met assumption
Sample(s) demarcation and Independent observations within groups.	Computation of within group variance and variance difference between groups.	variance of 87.2(old) vs 110.1(new) indicates independent observations and sample demarcation.	Met assumption
Interval or ratio type continuous dependent(y) variable	Observation of our dataset. Analysis of how data was measured.	Ratio type-continuous dependent variable.	Met assumption
Normality of our test or class samples (new test and old test samples)	Compute mean, median, mode; Normality histogram with curve; Skewness and kurtosis z- score within ± 2.58 range; Shapiro-wilk (S-W) test.	S-W test with significance of >0.01 (or 0.05) indicates normality of each class. Also, z-scores for skewness and kurtosis were within range for normality.	Met assumption
Normality of our combined samples (population slice)	Compute mean, median, mode; Normality histogram with curve; P-P/Q- Q/Box plots; Skewness and kurtosis z-score within ± 2.58 range; Shapiro-wilk (S-W) test.	S-W test with significance of $0.028 > 0.01$ indicates normality of the population slice. Also, z-scores for skewness (- 1.44) and kurtosis (- 1.04) were within range for normality. Also, P-P and Q-Q plots indicate normality.	Met assumption
Homoscedasticity	Lavene's test (variance ratio) for homoscedasticity or homogeneity of variance.	Significance of $.622 > 0.05$ hence homoscedastic or variance homogeneity.	Met assumption

Metric/Test type	Paired t-test	Independent t-test	Explanation
Type 1 error rate	Higher type I error rate.	Relatively lower rate compared to the paired t-test.	Likely to reject the null in a paired t-test due to “homogeneity” of treatment groups especially when a synchronizer is used.
Degrees of freedom	(n-1)	2(n-1)	Higher df=38 for the independent test because different groups are used hence doubling the number of participants. In the paired t-test a single group is used but passed into two treatments so while we get 40 observations, only 20 of them are unique so df=19.
Power	More powerful test of the two. Lower β so lower chance of a Type II error.	less powerful test of the two. Higher β so higher chance of a Type II error.	α and β are inversely proportional so power will be higher in the paired t-test.
Effect size	Larger effect size.	Lower effect size.	Homogeneity of the subject group from which the treatment groups are derived minimizes the unexplained error or variance hence increasing the effect size in paired t-test when compared to independent t-test. T-stats in the paired tests are likely to fall in the unlikely range so statistically significant.
Summary	A more powerful test prone to a judgement error (type I error).	A less powerful test prone to an acceptance error (type II error)	A paired t-test seems a more robust test with more power and less flaws in its design as seen by the use of a synchronizer in reducing unexplained variances or errors. A paired test is plagued by a higher possibility of a judgement error versus an acceptance error in the independent test.