Universiteit Leiden

Psychologie
Faculteit der Sociale Wetenschappen

Violation of Homogeneity of Variances: A Comparison Between Welch T Test and the Permutation Test

Ruchella Kock

Bachelor thesis Psychology
Institute of Psychology
Faculty of Social and Behavioural Sciences – Leiden University
Date: 4 – 22 – 2019
Bachelorproject number: 34
Student number: 1815458
First examiner: Julian Karch

Contents

Introduction

In psychological research parametric tests are used more often than non-parametric tests (Edgington, 1974; Goodwin & Goodwin, 1985; Skidmore & Thompson, 2010). The t-test is one of these parametric tests that is widely used in research. It is used to statistically test the differences between means. There are many different types of t-test s such as the Student t test, Welch t-test and Yuen t-test (Student, 1908; Welch, 1947; Yuen, 1974). The tests differ in terms of the assumptions they make. The three central assumptions of the t-test are independence, homogeneity of variances and normality. The permutation test is a non-parametric tests also used to compare means. It was first discussed by Fisher (1937). In the permutation test all possible permutations are calculated for a sample, this forms a test distribution. The null hypothesis can be tested under this distribution (Howell, 2009).

All statistical tests may lead to wrong conclusions, we distinguish between two types of errors. The type I error and type II error. Type I error is when H_0 is rejected when it should not have been. Type II error is when H_0 is not rejected when it should have been. We want both of these errors to be as low as possible. As the type II error decreases the power of a test increases. The power is the probability that H_0 is rejected when H_1 is true (Howell, 2009).

In this thesis we will compare the Welch t-test with the permutation test in terms of type I error and power. More specifically we will focus on the comparison between the two tests when the assumption of homogeneity is violated. Welch t-test was chosen because it provides more reliable type I error rates when the assumption of homogeneity of variance is not met. Compared to Student's t-test, Welch's t-test loses some statistical power. However, this loss is very small. (Delacre, Lakens, & Leys, 2017).

In the following two sections the assumptions of the two tests are discussed in depth. This is followed by a review of existing literature comparing the two tests. Then the research questions and hypothesis are discussed. Followed by a description of the simulation study and the results of the study. Finally we will end with a discussion of the findings and conclude with which test performs better when the assumption of homogeneity is violated.

Assumptions of the t test

The t-test has a number of assumptions. First of all, it assumes independent errors. This means that the residuals should not be able to be predicted above chance. This assumption would be violated if within a group, one participant can influence another one. This assumption cannot be tested. It must be controlled for based on the design of the study.

Secondly, it assumes that the the sampling distribution is normal. Many other parametric tests have this assumption. The reason why the normal distribution was chosen, can be explained by the central limit theorem (CLT). The CLT says that when starting with random and independent samples, the distribution of sample means will be approximately normally distributed if the sample size is large enough. As a simple example, say we flip a coin N number of times. Then we repeat the procedure many times. If N is large enough then our sampling distribution will be normal. The question remains as to what is large enough. In general it depends on what the sample looks like. However, it is believed that if $N \geq 30$ the sampling distribution of the mean is normal. Thus, if the sample size is greater than or equal to 30 then this assumption is fulfilled. The distribution of the data can be visually tested with a Q-Q/P-P plot. However, more often statistical tests are used namely, the Kolmogorov–Smirnov-test, the Shapiro-Wilk's W test or Skewness and Kurtosis can be checked. (Howell, 2009).

Another assumption is that there are no outliers. As the means of the groups are compared an outlier can greatly skew the mean which can lead to incorrect conclusions. There are many different ways to detect and deal with outliers. Some tests that can be used are Grubbs' test and Dixon's test. For a more detailed analysis on outliers and testing read Quesenberry and David (1961) or David and Paulson (1965).

Finally, there is the assumption of homogeneity of variances. Variance (σ^2) referes to the way the scores are distributed around the mean. Homogeneity of variances means that the variances across groups are considered equal. This assumption is important because if the scores in the one group were spread differently, compared to the second group before any treatment was given, then these groups are no longer comparable (Salkind, 2010). The null hypothesis when testing this assumption is $H_0 : \sigma_1^2 = \sigma_2^2$ or $H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1$. It is most commonly tested using the levene's test (Schultz, 1985). However, other tests include Hartley's F-max, Cochran's and Barlett's test (Conover, Johnson, & Johnson, 1981).

There are many studies that show that the t-test is robust against violations of the assumptions (e.g., Sawilowsky & Blair, 1992; Bradley, 1978). In general it is believed that the t-test is robust against non-normality if the samples size is greater or equal to 30. The t-test is believed to be robust against violation of the assumption of homogeneity if the group sizes are approximately equal. However, when the assumption of homogeneity is violated, the Welch t-test can be used. It is robust against this violation (Howell, 2009; Delacre et al., 2017).

Assumptions of the permutation test

There are two kinds of probability models namely the randomization model and the population model. In the randomization model, the subjects are randomly assigned to a condition. In the population model subjects are randomly sampled from a population (Ernst, 2004). The name permutation test is often used to refer to both the randomization model and population model because in many cases they can be equivalent to each other. The two tests are also referred to as randomization test and permutation test (Nichols & Holmes, 2002).

The randomization and permutation tests assume exchangeability. This assumption has different implications for the tests. One implication is the stable unit treatment value assumption (SUTVA). In Rubin (1980) he explained the idea of SUTVA. In an experiment, subjects/units i can be exposed to treatment j. Therefore, Y_{i_j} is the observed effect of unit i in treatment j. In this experiment each unit is only part of one treatment group at a time.

Thus, Y_{i_1} and Y_{i_2} cannot be observed at the same time, we have to make inferences about the value that was not observed. The effect of treatment 1 on unit i should be independent of the effect on other units in any treatment group, otherwise SUTVA will be violated.

In the randomization model exchangeability should be a given because participants are randomly assigned to the groups and should therefore be thought of as interchangeable. However, this cannot always be assumed for the population model.

For the population model the random assignment was not possible therefore exchangeability cannot be directly assumed. Thus, there is the assumption that the distributions of the two groups have approximately the same shape (Nichols & Holmes, 2002).

These permutation tests are non-parametric tests. However, there are parametric permutation tests such as the permutation t-test s which then takes the assumptions of the t-test (Toothaker, Wisconsin Univ., & for Cognitive Learning., 1972; Mendeş & Akkartal, 2010).

To conclude, there is a subtle difference between the two tests in terms of who the population is. In this thesis the randomization model is used. Thus, the assumption of exchangeability is met. The randomization model is chosen because the population model is often not used in psychological studies. Convenient sampling is used instead (Fife, 2013). In the randomization model convenient sampling can be used as long as the participants are randomly distributed between groups. In the population model convenient sampling cannot be used because the researchers are not sampling from the population.

Literature review

In psychology the t-test is often used more than permutation tests (e.g., Goodwin & Goodwin, 1985). However, often times we do not know the extent to which the assumptions are met in the population (Hunter & May, 1993). Compared to the t test, the permutation test has less assumptions. It assumes exchangeability. However, the randomization test automatically meets this assumption. Thus, the question of how the two tests compare in terms of type I error and power remains. In this section we review some literature comparing the two tests.

Toothaker et al. (1972) wrote a dissertation on comparing the permutation t-test with Student's t-test and the Mann Whitney U test. He performed a simulation study using sample sizes ranging from 2 to 5. The study concluded that the permutation t-test does not outperform Student's t-test and Mann Whitney U test and the latter two should be preferred when comparing means.

Ludbrook and Dudley (1998) compared permutation tests with t-test and F-test in Biomedical Research. They found that researchers in this field often choose an F-test or t-test instead of a permutation test even if the assumptions are not met. They conclude that exact permutation or randomization tests should be preferred in biomedical research.

Hughes (2010) conducted a simulation study, she compared the two sample t-test with the two sample exact permutation test. She used 6 non-normal distributions, tested at 3 different significance levels and the sample sizes ranged from 2 to 6. She concluded that the permutation test should be preferred, especially if power is very important for a study.

Most relevant to this thesis is the simulation study performed by Mendeş and Akkartal (2010). They compared the ANOVA F-test and Welch t-test with the permutation F-test and the permutation Welch t test. They used 3 different distributions, 5 different group sizes ranging from 5 to 15 and 3 different group variances namely, equal variances ($\sigma_1^2 = 1, \sigma_2^2 = 1, \sigma_3^2 = 1$), a small deviation ($\sigma_1^2 = 1, \sigma_2^2 = 1, \sigma_3^2 = 4$) and a larger deviation ($\sigma_1^2 = 1, \sigma_2^2 = 1, \sigma_3^2 = 9$). By comparing these groups they observed the effects of non-normality and heterogeneity. They concluded that when the assumption of homogeneity and normality is violated, the permutation F-test should be used. When the assumption of normality is violated but equal variances are assumed then the permutation Welch t-test should be used.

In all these papers very small sample sizes were used, the largest group size being 15. This is not representative of current psychological studies. We can see in the study from

Kühberger, Fritz, and Scherndl (2014) that only 14.9% of studies had a sample size of 15 or smaller. Moreover, only one study compared the two tests when the homogeneity assumption is violated. Mendeş and Akkartal (2010) also looked at the effect of different group sizes. However, the largest deviation between groups was 10. Larger deviations when comparing the two tests have not been studied.

This thesis aims to observe the effects on both tests when the assumption of homogeneity is violated. None of the tests assume homogeneity. Nevertheless, it is interesting to see if both tests perform equally well even when there are heterogenous variances. The assumption of homogeneity is not widely explored. Most literature observe the effect of non normality (e.g., Hughes, 2010; Weber & Sawilowsky, 2009). Moreover, the t-test is robust against violation of homogeneity of variances when the group sizes are approximately equal (Howell, 2009). Therefore, there should be more research on the effect of different group sizes. We look at small deviations as well as large deviations between group sizes. The goal of this thesis is to provide relevant results that can be used in psychological research. To achieve this goal sample sizes that are often used in psychology will be chosen and the randomization test which is more common in psychological research will be used.

Research questions

To be able to compare the Welch t-test with the permutation test when the variances are not equal, we look at different circumstances that affect the tests. The tests are compared in terms of type I and typeII error. In this section, the questions and sub-questions and hypothesis are discussed.

The research question of this thesis is: How does the permutation test compare to the Welch t-test under violation of the assumption of homogeneity of variances? To answer this question the following sub-questions will be explored.

- How does the permutation test compare to Welch t-test under no violation of the assumption of homogeneity of variances?

- What is the effect of sample size on the performance of the permutation test and the Welch *t*-test under violation of the assumption of homogeneity of variances?
- What is the effect of unequal group sizes on the performance of the permutation test and the Welch *t*-test under violation of the assumption of homogeneity of variances?

We hypothesize that due to the robustness of the Welch *t*-test against violation of homogeneity (Delacre et al., 2017), the Welch *t*-test (in further sections refered to as *t*-test) outperforms the permutation test.

Methods

To compare the *t*-test and permutation test, a simulation study is conducted using the programming language R (R Core Team, 2018). The type I error and power of the Welch *t*-test and permutation test are computed and compared against each other.

Sample sizes

Sample sizes that are relevant to psychology were chosen with the data provided by Kühberger et al. (2014). They investigated whether effect size is independent from sample size in psychological research. To do this they randomly sampled 1000 articles from 22 different psychological disciplines. All these articles were published in 2007. They excluded many articles because they did not meet their criteria. The final data set contained 531 articles. From these articles 529 sample sizes were reported and analysed in this study to choose the sample sizes.

From these articles we selected 3 different sample sizes. First of all, a small sample size used in psychology, namely $N = 10$. Despite this being a small sample size, it was reported in 13 articles. Less than 10% of the articles have a sample size that is smaller than 10 (8.9%). Secondly, a medium sample size that is commonly used was chosen. This is $N = 60$. It was reported 11 times and almost half of the sample sizes are smaller than or equal to 60 (47.6%). The third sample size is an extremely large sample size, namely $N = 1000$. Only 10% of the sample sizes that were reported were larger than 1000. These three sample sizes and deviations of them will be used in 3 separate simulations.

Each simulation was performed with one constant sample size (*samp1*), which is either 10, 60 or 1000, and deviations of these sample sizes. To determine the effect of unequal group sizes 7 different group sizes were chosen. The first group sizes are equal this is called this is condition 1. The other groups (*samp2*) have a downwards or upwards deviation from *samp1*. The size of *samp2* have a small deviation of 25% (condition 2a and 2b), a medium deviation of 50% (condition 3a and 3b) and a large deviation of 75% (condition 4a and 4b). This creates 24 different conditions each with slightly different group sizes (see Table 1).

Simulation

This section describes the simulation. Data was simulated with a normal distribution. The mean was 0 and the standard deviation was 1. When testing for type I error both means were kept equal. If the p-value of either test is smaller than $\alpha = 0.05$, then there is a type I error. The power is calculated with 1 - type II error. Type II error was simulated with 0 as the mean of one group and 1 as the mean of the other group. This is a mean that deviates by one standard deviation. If the p-value of either test is larger than $\alpha = 0.05$, then there is a type II error.

To simulate the violation of the assumption of homogeneity of variances, the standard deviation (σ) was altered. This is because variance is the squared standard deviation (σ^2). When there is no violation the variances of both groups are equal ($\sigma_1^2 = 1 : \sigma_2^2 = 1$). However, when the assumption is violated, the two variances are not equal. Six different deviations were chosen to simulate this. A downwards and upwards deviation of 25%, 50%, 75% and an upwards deviation of 300% (see Table 2). Each condition was performed with all seven deviations. The simulation was essentially performed 7 * 8 times for all 3 sample sizes.

We also varied the effect size (ES). The ES is the standardized mean difference between two groups (Coe, 2002). If there is a strong effect, the ES will be large. This means that the probability that the statistical test is significant is also large. Therefore, different effect sizes have different implications. In this thesis Cohen's three benchmark effect sizes were chosen, namely a small ES of 0.2, a medium ES of 0.5 and a large ES of 0.8 (Cohen, 2013).

Each simulation was repeated 10000 times. The number of type I and type II errors that occurred in 10000 were recorded for each condition. The data was simulated using `rnorm()`. The Welch *t*-test was performed using the `t.test()` formula in R with the argument `var.equal` set to False. The permutation test was performed using the library *perm* (Fay & Shaw, 2010). The Monte Carlo sampling technique was used during the permutation test. Ideally all permutations are performed in a permutation test. However, with larger sample sizes the number of permutations become very large. Therefore, the Monte Carlo sampling technique should be used. This technique randomly chooses test statistics from the permutation distribution. From this random sample the *p*-value for the permutation test can be calculated (Ernst, 2004; Hastings, 1970). We can conclude that one test outperforms another when the type I and type II error for one test is smaller than or equal to 0.05 and the type I and type II error for the other test is larger than 0.05. If both tests have a type I and type II error smaller or equal to 0.05, then we look at how big the difference is between the two tests. The test with the smaller errors outperforms the other. The code for the simulation is included on `https://github.com/rushkock/sim_study_thesis/tree/master/src/simulation`.

Results

The results consisted of 224 design elements (7 group ratios x 4 effect sizes x 8 standard deviations). This was performed 3 times with each sample size (10, 60 and 1000). The end results consisted of 672 different design rows. The results can be found in Table 3, Table 4 and Table 5 in the appendix or on https://github.com/rushkock/sim_study_thesis/tree/master/src/features/csv. The data analysis was performed using python (Team, 2015). The code for the data analysis can be found on: https://github.com/rushkock/sim_study_thesis/.

Pattern analysis

First, the most desirable conditions where there is no violation of homogeneity and the group sizes are equal were analyzed. These are the baseline conditions. The results are in Table 6 in the appendix. For a small sample size (sample 1 (*samp1*) = 10 and sample 2 (*samp2*) = 10), and medium (0.5) and large (0.8) ES, the *t*-test significantly outperforms the permutation test ($p < .001$). For all other conditions there were no statistically significant difference. This suggests that at baseline the tests perform equally well. However, when the sample size is small the typeII error of the *t*-test may be better than the permutation test. Next we analyzed for each sample size (*samp1* = 10, 60 and 1000) how many times there is a statistically significant difference between the *t*-test and the permutation test and under which circumstances this happens.

***samp1* = 10.** The results for the simulation where *samp1* = 10 are in Table 3. The *t*-test performs better than the permutation test 97 out of 224 times. The permutation test performs better 94 times out of 224. In 33 conditions there were no differences between the tests. When the group ratios are equal (*samp1* = 10 and *samp2* = 10) and when there is no violation of homogeneity (standard deviation 1 (SD1) = 1.0 and standard deviation 2 (SD2) = 1.0), the number of statistically significant differences between the tests is small (see Figure 1 and Figure 2). No additional patterns were found (see Figure 3 and Figure 4).

***samp1* = 60.** The results for the simulation where *samp1* = 60 are in Table 4. The *t*-test performs better 85 times and the permutation test performs better 76 times. There are more conditions that found no statistically significant differences between the two tests (63 conditions). Similar to the pattern in sample size 10, when the group ratios or standard deviations are equal the number of statistically significant differences are small (Table 4). This effect is stronger in sample size 60. The permutation test never outperforms the *t*-test in these conditions. Moreover, an interesting pattern is seen in effect size 0.8. The permutation test does not outperform the t-test when *samp1* is 60 and *samp2* is smaller than 60. The standard deviations that accompany these group ratios are larger than 1 (SD1 = 1.25, 1.50, 1.75 or 3.0 and SD2 = 1.0).

***samp1* = 1000.** The results for the simulation where *samp1* = 1000 are in Table 5. There are 155 conditions with no differences between the two tests. In 34 conditions the permutation tests is better and in 35 conditions the *t*-test performs better. When the effect size is medium (0.5) or large (0.8) there are no statistically significant differences between the tests. Similar to sample size 10 and 60, there are no differences when the group sizes are equal or when homogeneity is not violated. With effect size 0.2 there are only 6 statistically significant differences when sample size of group 2 is larger than sample size of group 1 (*samp1* = 1000 and *samp2* = 1250, 1500 or 1750). Whereas there are 21 statistically significant differences when the sample size of group 1 is larger than that of group 2 (*samp1* = 1000 and *samp2* = 250, 500, 750). Again the permutation test never outperforms the *t*-test when *samp2* is larger than *samp1*. Similar with sample size 60 the standard deviations that accompany these groups are larger than 1 (SD1 = 1.25, 1.50, 1.75 or 3.0 and SD2 = 1.0, The same figures as in *samp1* = 10 can be found for *samp1* = 60 and 1000 at:
`https://github.com/rushkock/sim_study_thesis/tree/master/reports/figures`).

To conclude the pattern analysis, as the sample size gets larger the tests become less significantly different. This is especially the case for larger effect sizes. Overall, when the sample size of group 2 is smaller than the sample size of group 1, there are less statistically significant differences found between the two tests. However, when there is a statistically significant difference, the *t*-test performs better than the permutation test. In the next section we group by standard deviations and look at each group separately.

No violation of homogeneity

In this section we explore the subquestion: How does the permutation test compare to Welch t-test under no violation of the assumption of homogeneity of variances? Homogeneity is not violated in the conditions where the standard deviations are equal. These conditions contain equal standard deviations but different effect sizes and sample sizes. The sample size of group 1 is always held equal, we refer to this group size as *samp1*. The size of group 2 deviates from group 1 with 25%, 50% or 75% this is *samp2* (see Table 1). This is repeated 3 times for *samp1* = 10, 60 and 1000.

***samp1* = 10.** The only statistically significant differences were found when *samp2* is smaller than *samp1* (*samp1* = 10, *samp2* = 3, 5 or 8) or when *samp1* and *samp2* are equal. No pattern was found, among the significant differences the permutation test outperformed the t-test in 2 cases whereas the t-test outperformed the permutation test in 4 cases (see Figure 5). All values for typeI error were around 0.05. For typeII error the largest statistically significant difference between the two tests was 0.02 in favor of the permutation test. Overall both tests performed equally well (Table 3).

***samp1* = 60.** The only statistically significant differences were found in ES 0.5 and 0.8. From the 3 conditions where there was a significant difference the *samp2* was always smaller than *samp1* (*samp1* = 60 and *samp2* = 45, 30, 15). The permutation test outperformed the t-test in all 3 conditions (see Figure 6).

***samp1* = 1000.** There were no statistically significant differences between the two tests. (see Figure 7).

To conclude, as indicated by the pattern analysis, when the sample size increases there are less statistically significant differences between the two tests. When there is no violation of homogeneity the only differences were found when the *samp2* is smaller than *samp1* or when they are equal. For sample size 10 the t-test seems to do slightly better. For sample size 60 the permutation test does slightly better. However, there are not many large differences which indicates the tests perform equally well when homogeneity is not violated.

Violation of homogeneity

In this section we look at the conditions with a violation of homogeneity. SD2 is always 1.0, SD1 deviates from SD2 with 25% (SD1 = 0.75 or 1.25), 50% (SD1 = 0.50 or 1.50), 75% (SD1 = 0.25 or 1.25) or 300% (SD1 = 3.0). If the tests are affected by a violation of homogeneity, then the most differences should be found in the condition with a violation of 300%. Here we compare *samp1* and *samp2* for each deviation of SD2. However, we summarise the main finding for all SD2 conditions where homogeneity is violated. All the figures similar to those presented when there was no violation of homogeneity can be found at: `https://github.com/rushkock/sim_study_thesis/tree/master/reports/figures/Hetero`

***samp1* = 10.** In most conditions there was a statistically significant difference between the tests, especially when the deviation of SD2 gets larger. For typeI error (ES = 0.0) and a deviation of SD1 that is smaller than SD2 (SD1 = 0,25, 0.50, 0.75 and SD2 = 1.0), the typeI error of the permutation test greatly exceeds $\alpha = 0.05(\pm 0.01)$ when samp2 is smaller than samp1 (samp1 = 10 and samp2 = 8, 5 or 3). However, when samp1 is smaller than samp2 (samp1 = 10 and samp2 = 13, 15 or 18), the typeI error of the permutation test is a lot smaller than $\alpha = 0.05(\pm 0.01)$. The *t*-test in contrast always performs around $\alpha = 0.05(\pm 0.01)$. If SD1 is larger than SD2 (SD1 = 1.25, 1.50, 1.75 or 3.0 and SD2 = 1.0), the opposite occures (see Table 7).

For the typeII error (effect size 0.2, 0.5, 0.8), most conditions were significantly different. When SD1 is smaller than SD2, the permutation test performs better in the conditions where *samp2* is smaller than *samp1* (*samp1* = 10 and *samp2* = 8, 5, 3). The *t*-test performs better when *samp2* is larger than *samp1* (*samp1* = 10 and *samp2* = 13, 15, 18). When SD2 is larger the opposite happens (see Table 7).

***samp1* = 60.** The same pattern as in *samp1* = 10 was found for typeI and typeII error in *samp1* = 60. The only exception is that in a lot of the conditions with effect size 0.8 the only statistically significant differences were found when *samp1* is larger than *samp2* (*samp1* = 60 and *samp2* = 45, 30, 15). If SD1 is larger than SD2 (SD1 = 1.25, 1.50, 1.75 or 3.0 and SD2 = 1.0) then the t-test performs better in all these conditions. If SD2 is larger than SD1 (SD1 = 0.75, 0.50 or 0.25 and SD2 = 1.0) then the permutation test performs better in all these conditions (see Table 8).

***samp1* = 1000.** The only statistically significant differences were found for effect size 0.0 and 0.2. For typeI error (effect size 0.0) the same pattern as in the previous sample sizes were found (see Table 9).

In effect size 0.2 (typeII error) the only significant differences were found when *samp1* is larger than *samp2* (*samp1* = 1000 and *samp2* = 750, 500 or 250). When SD1 is larger than SD2 (SD1 = 1.0 and SD2 = 1.25, 1.50 or 1.75) the permutation test always outperform the *t*-test. When SD2 is larger than SD1 (SD1 = 1.0 and SD2 = 1.25, 1.50, 1.75 or 3.0) the *t*-test always outperform the permutation test (see Table 9).

In conclusion, as the sample sizes get larger less statistically significant differences are found between the tests, this is especially the case for effect size 0.8. When the assumption of homogeneity is violated the permutation test can have a really low or really high typeI error depending on the condition. Whereas the *t*-test almost always performs around $\alpha = 0.05(\pm 0.01)$ For both typeI and typeII, the difference between the tests, depend on which sample size is group 1 and which sample size is group 2. This is also the case for the standard deviations. (For more clarification see the figures on `https://github.com/rushkock/sim_study_thesis/tree/master/reports/figures`).

Discussion

First of all, we explore the subquestion "How does the permutation test compare to Welch *t*-test under no violation of the assumption of homogeneity of variances?". We compared the tests in the most desirable conditions. This is where homogeneity is not violated and the group sizes are equal. The differences between these tests are small even when there was a significant difference (see Table 6). We can conclude that in the desirable conditions both tests perform equally well but the *t*-test may have an advantage in small sample sizes and the permutation test for medium sample sizes. Next, the conditions where homogeneity is not violated is analyzed for each group size, not only the equal groups. Overall, there were a few statistically significant differences between the tests. As the sample sizes gets larger there were less differences. For sample size 1000 there were no statistically significant differences at all. For sample size 10 and its deviations, the t-test has more significant differences than the permutation test. For sample size 60 and its deviations, the permutation test has more significant differences than the *t*-test. However, there were not many differences. We may conclude that when homogeneity is not violated the tests perform equally well for both typeI and typeII error.

Second, we explore the subquestion "What is the effect of sample size on the performance of the permutation test and the Welch *t*-test under violation of the assumption of homogeneity of variances?". The simulation was performed 3 times each time with one constant group size namely, sample size 10, 60 or 1000. The other group sizes were deviations from these sample sizes. To explore this subquestion we compare each simulation with each other. In sample size 10 most conditions were statistically significant when homogeneity was violated. There were 191 statistically significant differences. In sample size 60 and 1000 there were less statistically significant differences. This is expected because as the sample sizes get larger the difference between the means get larger, it is harder for the tests to commit a typeI or typeII error. Each test then perform equally well when the sample size and effect size is large. Aside from less significant differences across the sample sizes we see the same pattern for each sample size. This pattern is discussed in more details in the next subquestion.

We investigate the subquestion: "What is the effect of unequal group sizes on the performance of the permutation test and the Welch *t*-test under violation of the assumption of homogeneity of variances?". There was one condition with equal group sizes. For the other conditions the group sizes deviated upwards or downwards with 25%, 50% or 75%. This created the 3 different group ratios and 6 conditions where there are unequal group sizes. The size of the deviation did not matter. However, a clear and consistent pattern was found.

The permutation test had typeI errors that largely deviated from $\alpha = 0.05(\pm 0.01)$. They were either very high or very low. In contrast the t-test always performed around $\alpha = 0.05(\pm 0.01)$. This suggests that violation of homogeneity affects the permutation test but not the Welch t-test. This can be explained by the Fisher-Behren problem, where it is difficult to find differences in groups where the variance is unknown (Fenstad & Skovlund, 1992). The Welch t-test is protected against the Fisher-Behren problem but the permutation test is not. Therefore, homogeneity is required by the permutation test, this is well known but often not reported (Murphy, 1976; Boik, 1987; Zimmerman, 1996). The consequence of this is that the permutation test performs unreliably this can also explain the patterns found in typeII errors.

For typeII error which test performed better depended on which sample size *samp1* and *samp2* had and whether the SD1 was larger or smaller than SD2 (SD2 = 1.0 and SD1 deviates with 25%, 50%, 75% or 300%). To investigate this effect further we performed the simulation again with the sample size of the groups reversed. We repeated it 1000 times. Sample size of group 2 (*samp2*) was held constant, it was always 60. We choose the medium sample because this simulation was performed only 1000. The means of the medium sample size are large enough for the tests to find differences even when the simulation did not run for 10000 times. However, they are not too large that there would be no differences in effect size 0.5 and 0.8. The sample size of group 1 (*samp1*) was deviated with an upwards or downwards deviation of 25%, 50% or 75% (This was the same as the deviations of *samp2* in the original simulation, see Table 1).

The results of this simulation show the exact same pattern but in the opposite direction (Table 10). For typeI error the permutation tests shows deviations a lot larger or smaller than $\alpha = 0.05(\pm 0.01)$. For typeI error when SD1 is larger than SD2 (SD1 = 1.25, 1.50, 1.75 or 3.0 and SD2 = 1.0), the permutation test performed better in the conditions where *samp1* is larger than *samp2* (*samp1* = 75, 90 or 105 and *samp2* = 60). The t-test performs better when *samp2* is larger than *samp1* (*samp1* = 45, 30 or 15 and *samp2* = 60). When SD2 is larger the opposite happens.

In the previous simulations, *samp1* was held constant. When SD1 is larger than SD2, the permutation test performed better when *samp1* is larger than *samp2* (*samp1* = 60 and *samp2* = 45, 30 or 15, see Table 7). This indicates that which test performs better depends on which sample size is larger. This may also be related to the unreliable performance of the permutation test when the assumption of homogeneity is violated.

Limitations

The biggest limitation in this study is that violation of homogeneity was not considered an assumption of the permutation test when designing the study. Furthermore, in this study many conditions were used and this is a limitation because some conditions become redundant. An example of this is using both upwards and downwards deviation of group sizes, whereas for both these deviations the group ratio stays the same. It also makes data analysis more complicated.

Conclusion

As the sample sizes get larger the tests become less significantly different. When there is no violation of homogeneity both tests perform equally well. For both typeI and typeII error which test performs better depends on which size group 1 and group 2 has. When homogeneity is violated the typeI error of the Welch t-test is around $= 0.05(\pm 0.01)$. However, there are large deviations from $= 0.05(\pm 0.01)$ for the permutation test. This can be explained with the Fisher-Behren problem where the permutation tests acts unreliably when variances are unknown. Thus, when there is a violation of homogeneity the Welch t-test should be chosen.

References

Boik, R. J. (1987). The fisher-pitman permutation test: A non-robust alternative to the normal theory f test when variances are heterogeneous. *British Journal of Mathematical and Statistical Psychology*, *40*(1), 26–42.

Bradley, J. V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology*, *31*(2), 144-152. doi: 10.1111/j.2044-8317.1978.tb00581.x

Coe, R. (2002). It's the effect size, stupid: What effect size is and why it is important.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Conover, W. J., Johnson, M. E., & Johnson, M. M. (1981). A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics*, *23*(4), 351-361. doi: 10.1080/00401706.1981.10487680

David, H., & Paulson, A. (1965). The performance of several tests for outliers. *Biometrika*, *52(3/4)*(3/4), 429–436. doi: 10.2307/2333695

Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use welch's t-test instead of student's t-test. *International Review of Social Psychology*, *30*(1), 92–101. doi: http://doi.org/10.5334/irsp.82

Edgington, E. S. (1974). A new tabulation of statistical procedures used in apa journals. *American Psychologist*, *29*(1), 25-26. doi: http://dx.doi.org/10.1037/h0035846

Ernst, M. D. (2004, 11). Permutation methods: A basis for exact inference. *Statist. Sci.*, *19*(4), 676–685. doi: 10.1214/088342304000000396

Fay, M. P., & Shaw, P. A. (2010). Exact and asymptotic weighted logrank tests for interval censored data: The interval R package. *Journal of Statistical Software*, *36*(2), 1–34. Retrieved from http://www.jstatsoft.org/v36/i02/

Fenstad, G. U., & Skovlund, E. (1992). The behrens-fisher problem; a comparison of test properties when observations are not necessarily normal. *Statistical Research Report*.

Fife, D. A. (2013). *The achilles heel of psychology: How convenience sampling affects parameter estimates*. The University of Oklahoma.

Fisher, R. A. (1937). *The design of experiments*. Edinburgh, London: Oliver And Boyd.

Goodwin, L. D., & Goodwin, W. L. (1985). An analysis of statistical techniques used in the journal of educational psychology, 1979-1983. *Educational Psychologist*, *20*(1), 13-21. doi: 10.1207/s15326985ep2001_3

Hastings, W. K. (1970, 04). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*(1), 97-109. doi: 10.1093/biomet/57.1.97

Howell, D. C. (2009). *Statistical methods for psychology*. Cengage Learning.

Hughes, M. B. (2010). *Comparison of size/level and power of two-sample t-test to that of* *two-sample exact permutation test under various population shapes and small sample sizes*. Stephen F. Austin State University.

Hunter, M. A., & May, R. B. (1993). Some myths concerning parametric and nonparametric tests. *Canadian Psychology/Psychologie canadienne*, *34*(4), 384–389. doi: http://dx.doi.org/10.1037/h0078860

Kühberger, A., Fritz, A., & Scherndl, T. (2014, 09). Publication bias in psychology: A diagnosis based on the correlation between effect size and sample size. *PLOS ONE*, *9*(9), 1-8. doi: 10.1371/journal.pone.0105825

Ludbrook, J., & Dudley, H. (1998). Why permutation tests are superior to t and f tests in biomedical research. *The American Statistician*, *52*(2), 127–132. doi: 10.1080/00031305.1998.10480551

Mendeş, M., & Akkartal, E. (2010). Comparison of anova f and welch tests with their respective permutation versions in terms of type i error rates and test power. *Kafkas Univ Vet Fak Derg*, *16*(5), 711–716. doi: 10.9775/kvfd.2009.1507

Murphy, B. (1976). Comparison of some two sample means tests by simulation. *Communication in Statistics-Simulation and Computation*, *5*(1), 23–32. doi: 10.1080/03610917608812004

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human brain mapping*, *15*(1), 1–25. doi: https://doi.org/10.1002/hbm.1058

Quesenberry, C. P., & David, H. A. (1961). Some tests for outliers. *Biometrika*, *48(3/4)*(3/4), 379–390. doi: 10.2307/2332759

R Core Team. (2018). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Rubin, D. B. (1980). Randomization analysis of experimental data: The fisher randomization test comment. *Journal of the American Statistical Association*, *75*(371), 591–593. doi: 10.2307/2287653

Salkind, N. J. (2010). *Encyclopedia of research design* (Vol. 1). Sage.

Sawilowsky, S. S., & Blair, R. C. (1992). A more realistic look at the robustness and type ii error properties of the t test to departures from population normality. *Psychological bulletin*, *111*(2), 352-360. doi: http://dx.doi.org/10.1037/0033-2909.111.2.352

Schultz, B. B. (1985, 12). Levene's Test for Relative Variation. *Systematic Biology*, *34*(4), 449-456. doi: 10.1093/sysbio/34.4.449

Skidmore, S. T., & Thompson, B. (2010). Statistical techniques used in published articles: A historical review of reviews. *Educational and Psychological Measurement*, *70*(5), 777-795. doi: 10.1177/0013164410379320

Student. (1908). The probable error of a mean. *Biometrika*, *6*(1), 1–25. doi: doi:10.2307/2331554

Team, P. C. (2015). Python: A dynamic, open source programming language (3.7 ed.) [Computer software manual]. Retrieved from https://www.python.org/

Toothaker, L. E., Wisconsin Univ., M. R., & for Cognitive Learning., D. C. (1972). *An empirical investigation of the permutation t-test as compared to student's t-test and the mann-whitney u-test.* Washington, D.C.: ERIC Clearinghouse.

Weber, M., & Sawilowsky, S. (2009). Comparative power of the independent t, permutation t, and wilcoxontests. *Journal of Modern Applied Statistical Methods*, *8*(1), 3. doi: 10.22237/jmasm/1241136120

Welch, B. L. (1947). The generalization of 'student's' problem when several different population variances are involved. *Biometrika*, *34(1/2)*(1/2), 28–35. doi: doi:10.2307/2332510

Yuen, K. (1974, 04). The two-sample trimmed t for unequal population variances. *Biometrika*, *61*(1), 165-170. doi: 10.1093/biomet/61.1.165

Zimmerman, D. W. (1996). A note on homogeneity of variance of scores and ranks. *The Journal of Experimental Education*, *64*(4), 351-362. doi: 10.1080/00220973.1996.10806603

Table 1

Methods

Sample Size	Group Ratios
Small N = 10	
Condition 1	$N_1 = 10 : N_2 = 10$
Condition 2a	$N_1 = 10 : N_2 = 8$
Condition 2b	$N_1 = 10 : N_2 = 13$
Condition 3a	$N_1 = 10 : N_2 = 5$
Condition 3b	$N_1 = 10 : N_2 = 15$
Condition 4a	$N_1 = 10 : N_2 = 3$
Condition 4b	$N_1 = 10 : N_2 = 18$
Medium N = 60	
Condition 1	$N_1 = 60 : N_2 = 60$
Condition 2a	$N_1 = 60 : N_2 = 45$
Condition 2b	$N_1 = 60 : N_2 = 75$
Condition 3a	$N_1 = 60 : N_2 = 30$
Condition 3b	$N_1 = 60 : N_2 = 90$
Condition 4a	$N_1 = 60 : N_2 = 15$
Condition 4b	$N_1 = 60 : N_2 = 105$
Large N = 1000	
Condition 1	$N_1 = 1000 : N_2 = 1000$
Condition 2a	$N_1 = 1000 : N_2 = 750$
Condition 2b	$N_1 = 1000 : N_2 = 1250$
Condition 3a	$N_1 = 1000 : N_2 = 500$
Condition 3b	$N_1 = 1000 : N_2 = 1500$
Condition 4a	$N_1 = 1000 : N_2 = 250$
Condition 4b	$N_1 = 1000 : N_2 = 1750$

Table 2

Standard Deviations

Standard Deviation
$\sigma_1 = 1 : \sigma_2 = 1$
$\sigma_1 = 1 : \sigma_2 = 0.75$
$\sigma_1 = 1 : \sigma_2 = 1.25$
$\sigma_1 = 1 : \sigma_2 = 0.50$
$\sigma_1 = 1 : \sigma_2 = 1.50$
$\sigma_1 = 1 : \sigma_2 = 0.25$
$\sigma_1 = 1 : \sigma_2 = 1.75$
$\sigma_1 = 1 : \sigma_2 = 3$

Table 3

Simulation results for sample size 10 and its deviations

samp1	*samp2*	es	sd1	sd2	perm	*t*-test	pval	typeI	sig	dif
10	10	.0	1.00	1	.049	.050	1.000	True	False	-0.001
10	8	.0	1.00	1	.044	.046	1.000	True	False	-0.002
10	13	.0	1.00	1	.050	.052	1.000	True	False	-0.003
10	5	.0	1.00	1	.052	.052	1.000	True	False	-0.000
10	15	.0	1.00	1	.048	.051	1.000	True	False	-0.003
10	3	.0	1.00	1	.048	.066	0.000	True	True	-0.018
10	18	.0	1.00	1	.050	.051	1.000	True	False	-0.001
10	10	.2	1.00	1	.935	.933	0.493	False	False	0.002
10	8	.2	1.00	1	.936	.933	1.000	False	False	0.003
10	13	.2	1.00	1	.931	.928	1.000	False	False	0.003
10	5	.2	1.00	1	.932	.934	1.000	False	False	-0.002
10	15	.2	1.00	1	.930	.926	1.000	False	False	0.004
10	3	.2	1.00	1	.943	.933	0.008	False	True	0.010
10	18	.2	1.00	1	.924	.922	1.000	False	False	0.003
10	10	.5	1.00	1	.826	.820	0.000	False	True	0.006
10	8	.5	1.00	1	.834	.831	1.000	False	False	0.004
10	13	.5	1.00	1	.800	.795	0.365	False	False	0.005
10	5	.5	1.00	1	.863	.869	1.000	False	False	-0.006
10	15	.5	1.00	1	.790	.786	1.000	False	False	0.003
10	3	.5	1.00	1	.892	.888	1.000	False	False	0.005
10	18	.5	1.00	1	.770	.771	1.000	False	False	-0.002
10	10	.8	1.00	1	.619	.612	0.000	False	True	0.007
10	8	.8	1.00	1	.667	.658	0.000	False	True	0.009

10	13	.8	1.00	1	.574	.570	1.000	False	False	0.004
10	5	.8	1.00	1	.730	.748	0.000	False	True	-0.017
10	15	.8	1.00	1	.543	.542	1.000	False	False	0.001
10	3	.8	1.00	1	.807	.818	0.340	False	False	-0.011
10	18	.8	1.00	1	.513	.518	1.000	False	False	-0.005
10	10	.0	0.75	1	.049	.051	0.002	True	True	-0.002
10	8	.0	0.75	1	.055	.050	0.000	True	True	0.006
10	13	.0	0.75	1	.042	.052	0.000	True	True	-0.010
10	5	.0	0.75	1	.068	.053	0.000	True	True	0.015
10	15	.0	0.75	1	.037	.047	0.000	True	True	-0.010
10	3	.0	0.75	1	.076	.066	0.012	True	True	0.010
10	18	.0	0.75	1	.033	.048	0.000	True	True	-0.015
10	10	.2	0.75	1	.934	.932	0.117	False	False	0.002
10	8	.2	0.75	1	.924	.932	0.000	False	True	-0.007
10	13	.2	0.75	1	.940	.926	0.000	False	True	0.013
10	5	.2	0.75	1	.913	.933	0.000	False	True	-0.020
10	15	.2	0.75	1	.941	.926	0.000	False	True	0.015
10	3	.2	0.75	1	.908	.924	0.000	False	True	-0.016
10	18	.2	0.75	1	.942	.920	0.000	False	True	0.022
10	10	.5	0.75	1	.816	.814	1.000	False	False	0.001
10	8	.5	0.75	1	.832	.845	0.000	False	True	-0.013
10	13	.5	0.75	1	.816	.792	0.000	False	True	0.024
10	5	.5	0.75	1	.836	.876	0.000	False	True	-0.040
10	15	.5	0.75	1	.811	.781	0.000	False	True	0.031
10	3	.5	0.75	1	.856	.891	0.000	False	True	-0.035
10	18	.5	0.75	1	.802	.750	0.000	False	True	0.052
10	10	.8	0.75	1	.620	.615	0.001	False	True	0.004

10	8	.8	0.75	1	.640	.661	0.000	False	True	-0.022
10	13	.8	0.75	1	.582	.549	0.000	False	True	0.033
10	5	.8	0.75	1	.693	.771	0.000	False	True	-0.078
10	15	.8	0.75	1	.565	.513	0.000	False	True	0.052
10	3	.8	0.75	1	.763	.837	0.000	False	True	-0.073
10	18	.8	0.75	1	.546	.476	0.000	False	True	0.069
10	10	.0	1.25	1	.045	.047	0.153	True	False	-0.002
10	8	.0	1.25	1	.040	.048	0.000	True	True	-0.008
10	13	.0	1.25	1	.058	.053	0.001	True	True	0.005
10	5	.0	1.25	1	.039	.050	0.000	True	True	-0.012
10	15	.0	1.25	1	.055	.047	0.000	True	True	0.007
10	3	.0	1.25	1	.030	.056	0.000	True	True	-0.027
10	18	.0	1.25	1	.061	.050	0.000	True	True	0.011
10	10	.2	1.25	1	.932	.930	0.046	False	True	0.002
10	8	.2	1.25	1	.941	.933	0.000	False	True	0.008
10	13	.2	1.25	1	.923	.929	0.000	False	True	-0.006
10	5	.2	1.25	1	.947	.936	0.000	False	True	0.011
10	15	.2	1.25	1	.919	.932	0.000	False	True	-0.013
10	3	.2	1.25	1	.961	.930	0.000	False	True	0.031
10	18	.2	1.25	1	.910	.926	0.000	False	True	-0.016
10	10	.5	1.25	1	.821	.817	0.000	False	True	0.005
10	8	.5	1.25	1	.852	.834	0.000	False	True	0.017
10	13	.5	1.25	1	.786	.798	0.000	False	True	-0.012
10	5	.5	1.25	1	.884	.869	0.000	False	True	0.015
10	15	.5	1.25	1	.775	.799	0.000	False	True	-0.024
10	3	.5	1.25	1	.924	.886	0.000	False	True	0.038
10	18	.5	1.25	1	.753	.790	0.000	False	True	-0.037

10	10	.8	1.25	1	.622	.616	0.000	False	True	0.006
10	8	.8	1.25	1	.663	.640	0.000	False	True	0.023
10	13	.8	1.25	1	.556	.581	0.000	False	True	-0.025
10	5	.8	1.25	1	.752	.726	0.000	False	True	0.025
10	15	.8	1.25	1	.516	.554	0.000	False	True	-0.038
10	3	.8	1.25	1	.857	.814	0.000	False	True	0.043
10	18	.8	1.25	1	.479	.534	0.000	False	True	-0.056
10	10	.0	0.50	1	.049	.047	0.074	True	False	0.002
10	8	.0	0.50	1	.072	.054	0.000	True	True	0.017
10	13	.0	0.50	1	.033	.048	0.000	True	True	-0.014
10	5	.0	0.50	1	.106	.057	0.000	True	True	0.049
10	15	.0	0.50	1	.029	.050	0.000	True	True	-0.022
10	3	.0	0.50	1	.136	.068	0.000	True	True	0.068
10	18	.0	0.50	1	.020	.048	0.000	True	True	-0.028
10	10	.2	0.50	1	.924	.926	0.027	False	True	-0.002
10	8	.2	0.50	1	.915	.935	0.000	False	True	-0.020
10	13	.2	0.50	1	.948	.928	0.000	False	True	0.020
10	5	.2	0.50	1	.877	.933	0.000	False	True	-0.057
10	15	.2	0.50	1	.950	.923	0.000	False	True	0.028
10	3	.2	0.50	1	.849	.924	0.000	False	True	-0.076
10	18	.2	0.50	1	.961	.916	0.000	False	True	0.046
10	10	.5	0.50	1	.819	.823	0.005	False	True	-0.004
10	8	.5	0.50	1	.807	.846	0.000	False	True	-0.039
10	13	.5	0.50	1	.824	.782	0.000	False	True	0.042
10	5	.5	0.50	1	.796	.892	0.000	False	True	-0.096
10	15	.5	0.50	1	.826	.753	0.000	False	True	0.074
10	3	.5	0.50	1	.798	.902	0.000	False	True	-0.104

10	18	.5	0.50	1	.848	.740	0.000	False	True	0.108
10	10	.8	0.50	1	.612	.620	0.000	False	True	-0.008
10	8	.8	0.50	1	.616	.680	0.000	False	True	-0.064
10	13	.8	0.50	1	.599	.530	0.000	False	True	0.069
10	5	.8	0.50	1	.661	.801	0.000	False	True	-0.140
10	15	.8	0.50	1	.591	.482	0.000	False	True	0.108
10	3	.8	0.50	1	.698	.860	0.000	False	True	-0.162
10	18	.8	0.50	1	.582	.432	0.000	False	True	0.150
10	10	.0	1.50	1	.051	.052	1.000	True	False	-0.001
10	8	.0	1.50	1	.036	.044	0.000	True	True	-0.008
10	13	.0	1.50	1	.063	.052	0.000	True	True	0.011
10	5	.0	1.50	1	.031	.048	0.000	True	True	-0.017
10	15	.0	1.50	1	.066	.049	0.000	True	True	0.017
10	3	.0	1.50	1	.026	.058	0.000	True	True	-0.031
10	18	.0	1.50	1	.078	.053	0.000	True	True	0.024
10	10	.2	1.50	1	.931	.931	1.000	False	False	-0.000
10	8	.2	1.50	1	.942	.930	0.000	False	True	0.013
10	13	.2	1.50	1	.918	.931	0.000	False	True	-0.013
10	5	.2	1.50	1	.957	.941	0.000	False	True	0.016
10	15	.2	1.50	1	.903	.924	0.000	False	True	-0.022
10	3	.2	1.50	1	.972	.936	0.000	False	True	0.036
10	18	.2	1.50	1	.897	.930	0.000	False	True	-0.033
10	10	.5	1.50	1	.813	.812	1.000	False	False	0.001
10	8	.5	1.50	1	.856	.832	0.000	False	True	0.024
10	13	.5	1.50	1	.776	.808	0.000	False	True	-0.032
10	5	.5	1.50	1	.892	.857	0.000	False	True	0.035
10	15	.5	1.50	1	.752	.799	0.000	False	True	-0.047

10	3	.5	1.50	1	.937	.887	0.000	False	True	0.050
10	18	.5	1.50	1	.725	.799	0.000	False	True	-0.073
10	10	.8	1.50	1	.619	.616	0.139	False	False	0.003
10	8	.8	1.50	1	.680	.640	0.000	False	True	0.039
10	13	.8	1.50	1	.550	.596	0.000	False	True	-0.045
10	5	.8	1.50	1	.765	.706	0.000	False	True	0.059
10	15	.8	1.50	1	.497	.567	0.000	False	True	-0.070
10	3	.8	1.50	1	.862	.786	0.000	False	True	0.076
10	18	.8	1.50	1	.464	.565	0.000	False	True	-0.101
10	10	.0	0.25	1	.059	.052	0.000	True	True	0.007
10	8	.0	0.25	1	.086	.052	0.000	True	True	0.034
10	13	.0	0.25	1	.029	.047	0.000	True	True	-0.018
10	5	.0	0.25	1	.155	.057	0.000	True	True	0.098
10	15	.0	0.25	1	.022	.048	0.000	True	True	-0.025
10	3	.0	0.25	1	.222	.065	0.000	True	True	0.158
10	18	.0	0.25	1	.015	.048	0.000	True	True	-0.033
10	10	.2	0.25	1	.916	.927	0.000	False	True	-0.011
10	8	.2	0.25	1	.894	.934	0.000	False	True	-0.040
10	13	.2	0.25	1	.944	.921	0.000	False	True	0.023
10	5	.2	0.25	1	.824	.939	0.000	False	True	-0.115
10	15	.2	0.25	1	.959	.920	0.000	False	True	0.040
10	3	.2	0.25	1	.771	.933	0.000	False	True	-0.162
10	18	.2	0.25	1	.973	.919	0.000	False	True	0.054
10	10	.5	0.25	1	.804	.820	0.000	False	True	-0.015
10	8	.5	0.25	1	.784	.855	0.000	False	True	-0.072
10	13	.5	0.25	1	.837	.787	0.000	False	True	0.050
10	5	.5	0.25	1	.728	.895	0.000	False	True	-0.167

10	15	.5	0.25	1	.847	.751	0.000	False	True	0.096
10	3	.5	0.25	1	.710	.911	0.000	False	True	-0.201
10	18	.5	0.25	1	.865	.715	0.000	False	True	0.150
10	10	.8	0.25	1	.608	.635	0.000	False	True	-0.027
10	8	.8	0.25	1	.603	.707	0.000	False	True	-0.104
10	13	.8	0.25	1	.612	.533	0.000	False	True	0.079
10	5	.8	0.25	1	.586	.817	0.000	False	True	-0.230
10	15	.8	0.25	1	.620	.477	0.000	False	True	0.143
10	3	.8	0.25	1	.628	.881	0.000	False	True	-0.253
10	18	.8	0.25	1	.614	.392	0.000	False	True	0.222
10	10	.0	1.75	1	.051	.051	1.000	True	False	0.000
10	8	.0	1.75	1	.038	.048	0.000	True	True	-0.009
10	13	.0	1.75	1	.060	.046	0.000	True	True	0.014
10	5	.0	1.75	1	.028	.047	0.000	True	True	-0.019
10	15	.0	1.75	1	.074	.048	0.000	True	True	0.025
10	3	.0	1.75	1	.018	.048	0.000	True	True	-0.030
10	18	.0	1.75	1	.087	.050	0.000	True	True	0.037
10	10	.2	1.75	1	.932	.932	1.000	False	False	-0.000
10	8	.2	1.75	1	.950	.936	0.000	False	True	0.013
10	13	.2	1.75	1	.910	.929	0.000	False	True	-0.019
10	5	.2	1.75	1	.962	.938	0.000	False	True	0.024
10	15	.2	1.75	1	.896	.929	0.000	False	True	-0.033
10	3	.2	1.75	1	.977	.936	0.000	False	True	0.041
10	18	.2	1.75	1	.872	.921	0.000	False	True	-0.049
10	10	.5	1.75	1	.820	.820	1.000	False	False	-0.001
10	8	.5	1.75	1	.849	.820	0.000	False	True	0.029
10	13	.5	1.75	1	.759	.802	0.000	False	True	-0.043

10	5	.5	1.75	1	.903	.856	0.000	False	True	0.047
10	15	.5	1.75	1	.744	.808	0.000	False	True	-0.064
10	3	.5	1.75	1	.946	.882	0.000	False	True	0.065
10	18	.5	1.75	1	.718	.810	0.000	False	True	-0.092
10	10	.8	1.75	1	.610	.614	0.001	False	True	-0.005
10	8	.8	1.75	1	.679	.632	0.000	False	True	0.047
10	13	.8	1.75	1	.530	.594	0.000	False	True	-0.064
10	5	.8	1.75	1	.776	.693	0.000	False	True	0.083
10	15	.8	1.75	1	.504	.600	0.000	False	True	-0.096
10	3	.8	1.75	1	.874	.768	0.000	False	True	0.106
10	18	.8	1.75	1	.445	.582	0.000	False	True	-0.136
10	10	.0	3.00	1	.059	.054	0.000	True	True	0.005
10	8	.0	3.00	1	.038	.050	0.000	True	True	-0.013
10	13	.0	3.00	1	.082	.052	0.000	True	True	0.031
10	5	.0	3.00	1	.023	.050	0.000	True	True	-0.027
10	15	.0	3.00	1	.106	.054	0.000	True	True	0.052
10	3	.0	3.00	1	.009	.049	0.000	True	True	-0.040
10	18	.0	3.00	1	.126	.049	0.000	True	True	0.078
10	10	.2	3.00	1	.930	.935	0.000	False	True	-0.005
10	8	.2	3.00	1	.945	.927	0.000	False	True	0.017
10	13	.2	3.00	1	.890	.933	0.000	False	True	-0.043
10	5	.2	3.00	1	.970	.933	0.000	False	True	0.037
10	15	.2	3.00	1	.868	.928	0.000	False	True	-0.060
10	3	.2	3.00	1	.987	.934	0.000	False	True	0.053
10	18	.2	3.00	1	.834	.930	0.000	False	True	-0.096
10	10	.5	3.00	1	.815	.826	0.000	False	True	-0.012
10	8	.5	3.00	1	.862	.829	0.000	False	True	0.033

10	13	.5	3.00	1	.744	.821	0.000	False	True	-0.077
10	5	.5	3.00	1	.911	.830	0.000	False	True	0.081
10	15	.5	3.00	1	.704	.814	0.000	False	True	-0.110
10	3	.5	3.00	1	.963	.861	0.000	False	True	0.101
10	18	.5	3.00	1	.663	.815	0.000	False	True	-0.152
10	10	.8	3.00	1	.610	.631	0.000	False	True	-0.021
10	8	.8	3.00	1	.690	.632	0.000	False	True	0.058
10	13	.8	3.00	1	.514	.619	0.000	False	True	-0.105
10	5	.8	3.00	1	.798	.658	0.000	False	True	0.139
10	15	.8	3.00	1	.471	.620	0.000	False	True	-0.149
10	3	.8	3.00	1	.908	.705	0.000	False	True	0.203
10	18	.8	3.00	1	.407	.606	0.000	False	True	-0.199

- "*samp1*" and "*samp2*" are the sizes of the two groups.
- "es" is the effect size. An effect size of 0.0 represents a typeI error this is also visible in the column "typeI". Effect size 0.2, 0.5 and 0.8 represent typeII errors.
- "sd1" and "sd2" are the standard deviations of the two groups.
- The column perm contains the number of errors for the permutation test.
- The "*t*-test" column contains the number of errors for the *t*-test.
- The column "pval" gives the *p*-value from the mcnemar test comparing the permutation test with the *t*-test.
- The column "sig" is True if there is a significant difference between the two tests and False if there is no significant difference.
- The column "dif" is the difference between errors for the *t*-test minus the errors of the permutation test. Thus, a negative value indicates that the permutation test outperforms the *t*-test.

Table 4

Simulation results for sample size 60 and its deviations

samp1	samp2	es	sd1	sd2	perm	t_test	pval	typeI	sig	dif
60	60	.0	1.00	1	.049	.050	1.000	True	False	-0.000
60	45	.0	1.00	1	.051	.050	1.000	True	False	0.000
60	75	.0	1.00	1	.050	.050	1.000	True	False	-0.000
60	30	.0	1.00	1	.049	.050	1.000	True	False	-0.001
60	90	.0	1.00	1	.050	.050	1.000	True	False	-0.001
60	15	.0	1.00	1	.049	.051	1.000	True	False	-0.001
60	105	.0	1.00	1	.050	.051	1.000	True	False	-0.000
60	60	.2	1.00	1	.808	.806	0.335	False	False	0.001
60	45	.2	1.00	1	.826	.825	1.000	False	False	0.001
60	75	.2	1.00	1	.796	.797	1.000	False	False	-0.001
60	30	.2	1.00	1	.852	.851	1.000	False	False	0.000
60	90	.2	1.00	1	.775	.773	1.000	False	False	0.002
60	15	.2	1.00	1	.896	.899	1.000	False	False	-0.004
60	105	.2	1.00	1	.764	.765	1.000	False	False	-0.001
60	60	.5	1.00	1	.225	.224	0.335	False	False	0.001
60	45	.5	1.00	1	.292	.291	1.000	False	False	0.002
60	75	.5	1.00	1	.186	.184	1.000	False	False	0.001
60	30	.5	1.00	1	.396	.404	0.009	False	True	-0.008
60	90	.5	1.00	1	.153	.152	1.000	False	False	0.001
60	15	.5	1.00	1	.600	.619	0.000	False	True	-0.018
60	105	.5	1.00	1	.135	.134	1.000	False	False	0.001
60	60	.8	1.00	1	.009	.009	1.000	False	False	0.000
60	45	.8	1.00	1	.019	.020	1.000	False	False	-0.001

60	75	.8	1.00	1	.004	.004	1.000	False	False	0.000
60	30	.8	1.00	1	.055	.056	1.000	False	False	-0.001
60	90	.8	1.00	1	.003	.003	1.000	False	False	-0.000
60	15	.8	1.00	1	.217	.240	0.000	False	True	-0.023
60	105	.8	1.00	1	.002	.002	1.000	False	False	-0.000
60	60	.0	0.75	1	.051	.051	1.000	True	False	-0.000
60	45	.0	0.75	1	.058	.047	0.000	True	True	0.011
60	75	.0	0.75	1	.042	.047	0.000	True	True	-0.005
60	30	.0	0.75	1	.076	.051	0.000	True	True	0.025
60	90	.0	0.75	1	.040	.053	0.000	True	True	-0.012
60	15	.0	0.75	1	.090	.042	0.000	True	True	0.048
60	105	.0	0.75	1	.036	.051	0.000	True	True	-0.015
60	60	.2	0.75	1	.804	.804	1.000	False	False	0.000
60	45	.2	0.75	1	.810	.832	0.000	False	True	-0.022
60	75	.2	0.75	1	.807	.786	0.000	False	True	0.021
60	30	.2	0.75	1	.826	.870	0.000	False	True	-0.044
60	90	.2	0.75	1	.801	.765	0.000	False	True	0.036
60	15	.2	0.75	1	.835	.908	0.000	False	True	-0.073
60	105	.2	0.75	1	.804	.759	0.000	False	True	0.044
60	60	.5	0.75	1	.227	.227	1.000	False	False	0.001
60	45	.5	0.75	1	.280	.309	0.000	False	True	-0.029
60	75	.5	0.75	1	.183	.166	0.000	False	True	0.017
60	30	.5	0.75	1	.366	.438	0.000	False	True	-0.073
60	90	.5	0.75	1	.168	.137	0.000	False	True	0.030
60	15	.5	0.75	1	.526	.664	0.000	False	True	-0.138
60	105	.5	0.75	1	.138	.108	0.000	False	True	0.030
60	60	.8	0.75	1	.009	.009	1.000	False	False	0.000

60	45	.8	0.75	1	.020	.025	0.000	False	True	-0.004
60	75	.8	0.75	1	.004	.004	0.992	False	False	0.001
60	30	.8	0.75	1	.058	.086	0.000	False	True	-0.028
60	90	.8	0.75	1	.002	.001	1.000	False	False	0.001
60	15	.8	0.75	1	.192	.319	0.000	False	True	-0.127
60	105	.8	0.75	1	.001	.001	1.000	False	False	0.000
60	60	.0	1.25	1	.051	.052	1.000	True	False	-0.000
60	45	.0	1.25	1	.043	.051	0.000	True	True	-0.008
60	75	.0	1.25	1	.050	.046	0.000	True	True	0.005
60	30	.0	1.25	1	.032	.047	0.000	True	True	-0.015
60	90	.0	1.25	1	.058	.049	0.000	True	True	0.010
60	15	.0	1.25	1	.024	.051	0.000	True	True	-0.027
60	105	.0	1.25	1	.068	.051	0.000	True	True	0.017
60	60	.2	1.25	1	.805	.805	1.000	False	False	0.001
60	45	.2	1.25	1	.839	.822	0.000	False	True	0.017
60	75	.2	1.25	1	.776	.788	0.000	False	True	-0.012
60	30	.2	1.25	1	.884	.850	0.000	False	True	0.034
60	90	.2	1.25	1	.762	.788	0.000	False	True	-0.026
60	15	.2	1.25	1	.935	.890	0.000	False	True	0.045
60	105	.2	1.25	1	.740	.774	0.000	False	True	-0.034
60	60	.5	1.25	1	.230	.229	1.000	False	False	0.001
60	45	.5	1.25	1	.294	.271	0.000	False	True	0.023
60	75	.5	1.25	1	.174	.186	0.000	False	True	-0.012
60	30	.5	1.25	1	.425	.370	0.000	False	True	0.055
60	90	.5	1.25	1	.142	.164	0.000	False	True	-0.022
60	15	.5	1.25	1	.659	.567	0.000	False	True	0.092
60	105	.5	1.25	1	.127	.156	0.000	False	True	-0.028

60	60	.8	1.25	1	.008	.008	1.000	False	False	0.000
60	45	.8	1.25	1	.021	.018	0.000	False	True	0.003
60	75	.8	1.25	1	.005	.006	0.992	False	False	-0.001
60	30	.8	1.25	1	.060	.045	0.000	False	True	0.015
60	90	.8	1.25	1	.003	.004	1.000	False	False	-0.001
60	15	.8	1.25	1	.248	.184	0.000	False	True	0.064
60	105	.8	1.25	1	.002	.002	1.000	False	False	-0.001
60	60	.0	0.50	1	.054	.054	1.000	True	False	0.000
60	45	.0	0.50	1	.073	.049	0.000	True	True	0.024
60	75	.0	0.50	1	.038	.056	0.000	True	True	-0.017
60	30	.0	0.50	1	.112	.053	0.000	True	True	0.058
60	90	.0	0.50	1	.028	.051	0.000	True	True	-0.023
60	15	.0	0.50	1	.181	.049	0.000	True	True	0.132
60	105	.0	0.50	1	.022	.050	0.000	True	True	-0.028
60	60	.2	0.50	1	.804	.804	1.000	False	False	-0.001
60	45	.2	0.50	1	.793	.838	0.000	False	True	-0.044
60	75	.2	0.50	1	.814	.778	0.000	False	True	0.036
60	30	.2	0.50	1	.777	.873	0.000	False	True	-0.096
60	90	.2	0.50	1	.827	.755	0.000	False	True	0.072
60	15	.2	0.50	1	.746	.913	0.000	False	True	-0.167
60	105	.2	0.50	1	.830	.722	0.000	False	True	0.108
60	60	.5	0.50	1	.226	.226	1.000	False	False	-0.001
60	45	.5	0.50	1	.266	.329	0.000	False	True	-0.063
60	75	.5	0.50	1	.196	.163	0.000	False	True	0.033
60	30	.5	0.50	1	.337	.495	0.000	False	True	-0.158
60	90	.5	0.50	1	.169	.112	0.000	False	True	0.058
60	15	.5	0.50	1	.444	.714	0.000	False	True	-0.270

60	105	.5	0.50	1	.150	.082	0.000	False	True	0.068
60	60	.8	0.50	1	.009	.009	1.000	False	False	0.000
60	45	.8	0.50	1	.018	.030	0.000	False	True	-0.012
60	75	.8	0.50	1	.004	.002	0.008	False	True	0.002
60	30	.8	0.50	1	.052	.111	0.000	False	True	-0.059
60	90	.8	0.50	1	.002	.001	0.196	False	False	0.001
60	15	.8	0.50	1	.159	.396	0.000	False	True	-0.237
60	105	.8	0.50	1	.001	.001	1.000	False	False	0.000
60	60	.0	1.50	1	.047	.047	1.000	True	False	-0.000
60	45	.0	1.50	1	.037	.049	0.000	True	True	-0.012
60	75	.0	1.50	1	.057	.048	0.000	True	True	0.009
60	30	.0	1.50	1	.026	.050	0.000	True	True	-0.025
60	90	.0	1.50	1	.069	.049	0.000	True	True	0.019
60	15	.0	1.50	1	.012	.051	0.000	True	True	-0.039
60	105	.0	1.50	1	.077	.049	0.000	True	True	0.029
60	60	.2	1.50	1	.807	.807	1.000	False	False	0.000
60	45	.2	1.50	1	.847	.820	0.000	False	True	0.027
60	75	.2	1.50	1	.771	.798	0.000	False	True	-0.027
60	30	.2	1.50	1	.901	.845	0.000	False	True	0.057
60	90	.2	1.50	1	.739	.784	0.000	False	True	-0.044
60	15	.2	1.50	1	.956	.880	0.000	False	True	0.075
60	105	.2	1.50	1	.722	.788	0.000	False	True	-0.066
60	60	.5	1.50	1	.222	.222	1.000	False	False	-0.000
60	45	.5	1.50	1	.308	.270	0.000	False	True	0.037
60	75	.5	1.50	1	.175	.197	0.000	False	True	-0.021
60	30	.5	1.50	1	.446	.342	0.000	False	True	0.104
60	90	.5	1.50	1	.150	.187	0.000	False	True	-0.037

60	15	.5	1.50	1	.701	.506	0.000	False	True	0.195
60	105	.5	1.50	1	.124	.174	0.000	False	True	-0.049
60	60	.8	1.50	1	.008	.008	1.000	False	False	0.000
60	45	.8	1.50	1	.021	.015	0.000	False	True	0.005
60	75	.8	1.50	1	.005	.006	0.196	False	False	-0.001
60	30	.8	1.50	1	.062	.034	0.000	False	True	0.029
60	90	.8	1.50	1	.003	.004	1.000	False	False	-0.001
60	15	.8	1.50	1	.266	.139	0.000	False	True	0.127
60	105	.8	1.50	1	.001	.004	0.000	False	True	-0.003
60	60	.0	0.25	1	.053	.052	0.575	True	False	0.001
60	45	.0	0.25	1	.084	.048	0.000	True	True	0.036
60	75	.0	0.25	1	.030	.050	0.000	True	True	-0.019
60	30	.0	0.25	1	.156	.052	0.000	True	True	0.104
60	90	.0	0.25	1	.020	.048	0.000	True	True	-0.028
60	15	.0	0.25	1	.294	.051	0.000	True	True	0.242
60	105	.0	0.25	1	.013	.050	0.000	True	True	-0.037
60	60	.2	0.25	1	.804	.808	0.000	False	True	-0.004
60	45	.2	0.25	1	.778	.845	0.000	False	True	-0.067
60	75	.2	0.25	1	.831	.777	0.000	False	True	0.054
60	30	.2	0.25	1	.727	.875	0.000	False	True	-0.148
60	90	.2	0.25	1	.849	.739	0.000	False	True	0.110
60	15	.2	0.25	1	.654	.918	0.000	False	True	-0.264
60	105	.2	0.25	1	.863	.707	0.000	False	True	0.156
60	60	.5	0.25	1	.225	.229	0.000	False	True	-0.003
60	45	.5	0.25	1	.252	.340	0.000	False	True	-0.088
60	75	.5	0.25	1	.201	.151	0.000	False	True	0.050
60	30	.5	0.25	1	.311	.521	0.000	False	True	-0.209

60	90	.5	0.25	1	.179	.096	0.000	False	True	0.083
60	15	.5	0.25	1	.375	.740	0.000	False	True	-0.365
60	105	.5	0.25	1	.155	.060	0.000	False	True	0.094
60	60	.8	0.25	1	.008	.008	1.000	False	False	-0.000
60	45	.8	0.25	1	.021	.038	0.000	False	True	-0.017
60	75	.8	0.25	1	.003	.002	0.067	False	False	0.001
60	30	.8	0.25	1	.045	.140	0.000	False	True	-0.095
60	90	.8	0.25	1	.002	.000	0.196	False	False	0.001
60	15	.8	0.25	1	.125	.449	0.000	False	True	-0.325
60	105	.8	0.25	1	.001	.000	1.000	False	False	0.001
60	60	.0	1.75	1	.050	.050	1.000	True	False	0.000
60	45	.0	1.75	1	.033	.048	0.000	True	True	-0.016
60	75	.0	1.75	1	.065	.050	0.000	True	True	0.015
60	30	.0	1.75	1	.018	.049	0.000	True	True	-0.031
60	90	.0	1.75	1	.075	.049	0.000	True	True	0.026
60	15	.0	1.75	1	.009	.052	0.000	True	True	-0.043
60	105	.0	1.75	1	.083	.046	0.000	True	True	0.037
60	60	.2	1.75	1	.806	.807	1.000	False	False	-0.001
60	45	.2	1.75	1	.859	.824	0.000	False	True	0.035
60	75	.2	1.75	1	.759	.793	0.000	False	True	-0.034
60	30	.2	1.75	1	.910	.834	0.000	False	True	0.076
60	90	.2	1.75	1	.732	.793	0.000	False	True	-0.060
60	15	.2	1.75	1	.971	.875	0.000	False	True	0.096
60	105	.2	1.75	1	.706	.790	0.000	False	True	-0.085
60	60	.5	1.75	1	.228	.228	1.000	False	False	-0.001
60	45	.5	1.75	1	.315	.264	0.000	False	True	0.051
60	75	.5	1.75	1	.174	.203	0.000	False	True	-0.029

60	30	.5	1.75	1	.456	.313	0.000	False	True	0.143
60	90	.5	1.75	1	.143	.193	0.000	False	True	-0.050
60	15	.5	1.75	1	.734	.474	0.000	False	True	0.260
60	105	.5	1.75	1	.120	.184	0.000	False	True	-0.064
60	60	.8	1.75	1	.009	.009	1.000	False	False	0.000
60	45	.8	1.75	1	.020	.012	0.000	False	True	0.008
60	75	.8	1.75	1	.006	.007	0.115	False	False	-0.001
60	30	.8	1.75	1	.062	.028	0.000	False	True	0.034
60	90	.8	1.75	1	.003	.005	0.000	False	True	-0.003
60	15	.8	1.75	1	.281	.101	0.000	False	True	0.180
60	105	.8	1.75	1	.002	.005	0.000	False	True	-0.003
60	60	.0	3.00	1	.052	.051	0.575	True	False	0.001
60	45	.0	3.00	1	.030	.054	0.000	True	True	-0.024
60	75	.0	3.00	1	.076	.050	0.000	True	True	0.025
60	30	.0	3.00	1	.012	.052	0.000	True	True	-0.040
60	90	.0	3.00	1	.101	.052	0.000	True	True	0.049
60	15	.0	3.00	1	.001	.050	0.000	True	True	-0.049
60	105	.0	3.00	1	.122	.048	0.000	True	True	0.075
60	60	.2	3.00	1	.815	.818	0.000	False	True	-0.003
60	45	.2	3.00	1	.870	.812	0.000	False	True	0.058
60	75	.2	3.00	1	.754	.808	0.000	False	True	-0.054
60	30	.2	3.00	1	.937	.824	0.000	False	True	0.113
60	90	.2	3.00	1	.706	.804	0.000	False	True	-0.098
60	15	.2	3.00	1	.990	.840	0.000	False	True	0.151
60	105	.2	3.00	1	.677	.810	0.000	False	True	-0.133
60	60	.5	3.00	1	.224	.227	0.000	False	True	-0.003
60	45	.5	3.00	1	.321	.239	0.000	False	True	0.081

60	75	.5	3.00	1	.168	.217	0.000	False	True	-0.049
60	30	.5	3.00	1	.501	.272	0.000	False	True	0.229
60	90	.5	3.00	1	.134	.212	0.000	False	True	-0.078
60	15	.5	3.00	1	.826	.338	0.000	False	True	0.488
60	105	.5	3.00	1	.111	.202	0.000	False	True	-0.091
60	60	.8	3.00	1	.008	.008	1.000	False	False	-0.000
60	45	.8	3.00	1	.021	.012	0.000	False	True	0.009
60	75	.8	3.00	1	.005	.009	0.000	False	True	-0.004
60	30	.8	3.00	1	.065	.016	0.000	False	True	0.049
60	90	.8	3.00	1	.003	.009	0.000	False	True	-0.006
60	15	.8	3.00	1	.330	.031	0.000	False	True	0.298
60	105	.8	3.00	1	.002	.006	0.000	False	True	-0.004

See Table 3 for explanation on column names

Table 5

Simulation results for sample size 1000 and its deviations

samp1	samp2	es	sd1	sd2	perm	t_test	pval	typeI	sig	dif
1000	1000	.0	1.00	1	.049	.049	1.000	True	False	0.000
1000	750	.0	1.00	1	.048	.048	1.000	True	False	0.000
1000	1250	.0	1.00	1	.048	.048	1.000	True	False	0.000
1000	500	.0	1.00	1	.046	.046	1.000	True	False	0.000
1000	1500	.0	1.00	1	.047	.048	1.000	True	False	-0.000
1000	250	.0	1.00	1	.052	.052	1.000	True	False	-0.000
1000	1750	.0	1.00	1	.048	.048	1.000	True	False	0.000
1000	1000	.2	1.00	1	.006	.006	1.000	False	False	0.000
1000	750	.2	1.00	1	.013	.013	1.000	False	False	0.000
1000	1250	.2	1.00	1	.002	.002	1.000	False	False	0.000
1000	500	.2	1.00	1	.047	.047	1.000	False	False	0.000
1000	1500	.2	1.00	1	.002	.002	1.000	False	False	0.000
1000	250	.2	1.00	1	.193	.194	1.000	False	False	-0.000
1000	1750	.2	1.00	1	.001	.001	1.000	False	False	-0.000
1000	1000	.5	1.00	1	.000	.000	1.000	False	False	0.000
1000	750	.5	1.00	1	.000	.000	1.000	False	False	0.000
1000	1250	.5	1.00	1	.000	.000	1.000	False	False	0.000
1000	500	.5	1.00	1	.000	.000	1.000	False	False	0.000
1000	1500	.5	1.00	1	.000	.000	1.000	False	False	0.000
1000	250	.5	1.00	1	.000	.000	1.000	False	False	0.000
1000	1750	.5	1.00	1	.000	.000	1.000	False	False	0.000
1000	1000	.8	1.00	1	.000	.000	1.000	False	False	0.000
1000	750	.8	1.00	1	.000	.000	1.000	False	False	0.000

1000	1250	.8	1.00	1	.000	.000	1.000	False	False	0.000
1000	500	.8	1.00	1	.000	.000	1.000	False	False	0.000
1000	1500	.8	1.00	1	.000	.000	1.000	False	False	0.000
1000	250	.8	1.00	1	.000	.000	1.000	False	False	0.000
1000	1750	.8	1.00	1	.000	.000	1.000	False	False	0.000
1000	1000	.0	0.75	1	.053	.053	1.000	True	False	0.000
1000	750	.0	0.75	1	.064	.052	0.000	True	True	0.011
1000	1250	.0	0.75	1	.046	.053	0.000	True	True	-0.007
1000	500	.0	0.75	1	.073	.048	0.000	True	True	0.025
1000	1500	.0	0.75	1	.037	.050	0.000	True	True	-0.012
1000	250	.0	0.75	1	.099	.050	0.000	True	True	0.049
1000	1750	.0	0.75	1	.034	.050	0.000	True	True	-0.016
1000	1000	.2	0.75	1	.006	.006	1.000	False	False	0.000
1000	750	.2	0.75	1	.016	.019	0.000	False	True	-0.003
1000	1250	.2	0.75	1	.002	.002	1.000	False	False	0.000
1000	500	.2	0.75	1	.045	.065	0.000	False	True	-0.020
1000	1500	.2	0.75	1	.002	.001	1.000	False	False	0.001
1000	250	.2	0.75	1	.170	.261	0.000	False	True	-0.091
1000	1750	.2	0.75	1	.001	.001	1.000	False	False	0.000
1000	1000	.5	0.75	1	.000	.000	1.000	False	False	0.000
1000	750	.5	0.75	1	.000	.000	1.000	False	False	0.000
1000	1250	.5	0.75	1	.000	.000	1.000	False	False	0.000
1000	500	.5	0.75	1	.000	.000	1.000	False	False	0.000
1000	1500	.5	0.75	1	.000	.000	1.000	False	False	0.000
1000	250	.5	0.75	1	.000	.000	1.000	False	False	0.000
1000	1750	.5	0.75	1	.000	.000	1.000	False	False	0.000
1000	1000	.8	0.75	1	.000	.000	1.000	False	False	0.000

1000	750	.8	0.75	1	.000	.000	1.000	False	False	0.000
1000	1250	.8	0.75	1	.000	.000	1.000	False	False	0.000
1000	500	.8	0.75	1	.000	.000	1.000	False	False	0.000
1000	1500	.8	0.75	1	.000	.000	1.000	False	False	0.000
1000	250	.8	0.75	1	.000	.000	1.000	False	False	0.000
1000	1750	.8	0.75	1	.000	.000	1.000	False	False	0.000
1000	1000	.0	1.25	1	.050	.050	1.000	True	False	0.000
1000	750	.0	1.25	1	.043	.049	0.000	True	True	-0.006
1000	1250	.0	1.25	1	.056	.051	0.000	True	True	0.005
1000	500	.0	1.25	1	.035	.048	0.000	True	True	-0.013
1000	1500	.0	1.25	1	.062	.050	0.000	True	True	0.012
1000	250	.0	1.25	1	.026	.051	0.000	True	True	-0.025
1000	1750	.0	1.25	1	.066	.053	0.000	True	True	0.014
1000	1000	.2	1.25	1	.005	.005	1.000	False	False	0.000
1000	750	.2	1.25	1	.015	.013	0.001	False	True	0.002
1000	1250	.2	1.25	1	.003	.004	1.000	False	False	-0.001
1000	500	.2	1.25	1	.048	.034	0.000	False	True	0.014
1000	1500	.2	1.25	1	.002	.002	1.000	False	False	-0.000
1000	250	.2	1.25	1	.213	.142	0.000	False	True	0.072
1000	1750	.2	1.25	1	.001	.001	1.000	False	False	-0.000
1000	1000	.5	1.25	1	.000	.000	1.000	False	False	0.000
1000	750	.5	1.25	1	.000	.000	1.000	False	False	0.000
1000	1250	.5	1.25	1	.000	.000	1.000	False	False	0.000
1000	500	.5	1.25	1	.000	.000	1.000	False	False	0.000
1000	1500	.5	1.25	1	.000	.000	1.000	False	False	0.000
1000	250	.5	1.25	1	.000	.000	1.000	False	False	0.000
1000	1750	.5	1.25	1	.000	.000	1.000	False	False	0.000

1000	1000	.8	1.25	1	.000	.000	1.000	False	False	0.000
1000	750	.8	1.25	1	.000	.000	1.000	False	False	0.000
1000	1250	.8	1.25	1	.000	.000	1.000	False	False	0.000
1000	500	.8	1.25	1	.000	.000	1.000	False	False	0.000
1000	1500	.8	1.25	1	.000	.000	1.000	False	False	0.000
1000	250	.8	1.25	1	.000	.000	1.000	False	False	0.000
1000	1750	.8	1.25	1	.000	.000	1.000	False	False	0.000
1000	1000	.0	0.50	1	.051	.051	1.000	True	False	0.000
1000	750	.0	0.50	1	.065	.046	0.000	True	True	0.019
1000	1250	.0	0.50	1	.038	.053	0.000	True	True	-0.015
1000	500	.0	0.50	1	.109	.050	0.000	True	True	0.059
1000	1500	.0	0.50	1	.031	.054	0.000	True	True	-0.022
1000	250	.0	0.50	1	.174	.050	0.000	True	True	0.124
1000	1750	.0	0.50	1	.020	.048	0.000	True	True	-0.028
1000	1000	.2	0.50	1	.006	.006	1.000	False	False	0.000
1000	750	.2	0.50	1	.016	.023	0.000	False	True	-0.008
1000	1250	.2	0.50	1	.002	.002	1.000	False	False	0.001
1000	500	.2	0.50	1	.040	.086	0.000	False	True	-0.045
1000	1500	.2	0.50	1	.002	.001	0.992	False	False	0.001
1000	250	.2	0.50	1	.142	.320	0.000	False	True	-0.179
1000	1750	.2	0.50	1	.000	.000	1.000	False	False	0.000
1000	1000	.5	0.50	1	.000	.000	1.000	False	False	0.000
1000	750	.5	0.50	1	.000	.000	1.000	False	False	0.000
1000	1250	.5	0.50	1	.000	.000	1.000	False	False	0.000
1000	500	.5	0.50	1	.000	.000	1.000	False	False	0.000
1000	1500	.5	0.50	1	.000	.000	1.000	False	False	0.000
1000	250	.5	0.50	1	.000	.000	1.000	False	False	0.000

1000	1750	.5	0.50	1	.000	.000	1.000	False	False	0.000
1000	1000	.8	0.50	1	.000	.000	1.000	False	False	0.000
1000	750	.8	0.50	1	.000	.000	1.000	False	False	0.000
1000	1250	.8	0.50	1	.000	.000	1.000	False	False	0.000
1000	500	.8	0.50	1	.000	.000	1.000	False	False	0.000
1000	1500	.8	0.50	1	.000	.000	1.000	False	False	0.000
1000	250	.8	0.50	1	.000	.000	1.000	False	False	0.000
1000	1750	.8	0.50	1	.000	.000	1.000	False	False	0.000
1000	1000	.0	1.50	1	.051	.051	1.000	True	False	0.000
1000	750	.0	1.50	1	.035	.048	0.000	True	True	-0.013
1000	1250	.0	1.50	1	.062	.050	0.000	True	True	0.012
1000	500	.0	1.50	1	.026	.050	0.000	True	True	-0.024
1000	1500	.0	1.50	1	.072	.051	0.000	True	True	0.020
1000	250	.0	1.50	1	.012	.054	0.000	True	True	-0.042
1000	1750	.0	1.50	1	.076	.052	0.000	True	True	0.024
1000	1000	.2	1.50	1	.008	.008	1.000	False	False	0.000
1000	750	.2	1.50	1	.015	.011	0.000	False	True	0.003
1000	1250	.2	1.50	1	.004	.005	0.575	False	False	-0.001
1000	500	.2	1.50	1	.048	.026	0.000	False	True	0.022
1000	1500	.2	1.50	1	.002	.003	0.115	False	False	-0.001
1000	250	.2	1.50	1	.234	.103	0.000	False	True	0.132
1000	1750	.2	1.50	1	.001	.002	0.992	False	False	-0.001
1000	1000	.5	1.50	1	.000	.000	1.000	False	False	0.000
1000	750	.5	1.50	1	.000	.000	1.000	False	False	0.000
1000	1250	.5	1.50	1	.000	.000	1.000	False	False	0.000
1000	500	.5	1.50	1	.000	.000	1.000	False	False	0.000
1000	1500	.5	1.50	1	.000	.000	1.000	False	False	0.000

1000	250	.5	1.50	1	.000	.000	1.000	False	False	0.000
1000	1750	.5	1.50	1	.000	.000	1.000	False	False	0.000
1000	1000	.8	1.50	1	.000	.000	1.000	False	False	0.000
1000	750	.8	1.50	1	.000	.000	1.000	False	False	0.000
1000	1250	.8	1.50	1	.000	.000	1.000	False	False	0.000
1000	500	.8	1.50	1	.000	.000	1.000	False	False	0.000
1000	1500	.8	1.50	1	.000	.000	1.000	False	False	0.000
1000	250	.8	1.50	1	.000	.000	1.000	False	False	0.000
1000	1750	.8	1.50	1	.000	.000	1.000	False	False	0.000
1000	1000	.0	0.25	1	.050	.050	1.000	True	False	0.000
1000	750	.0	0.25	1	.088	.052	0.000	True	True	0.036
1000	1250	.0	0.25	1	.031	.050	0.000	True	True	-0.019
1000	500	.0	0.25	1	.140	.047	0.000	True	True	0.093
1000	1500	.0	0.25	1	.020	.051	0.000	True	True	-0.031
1000	250	.0	0.25	1	.282	.053	0.000	True	True	0.229
1000	1750	.0	0.25	1	.012	.051	0.000	True	True	-0.039
1000	1000	.2	0.25	1	.006	.006	1.000	False	False	0.000
1000	750	.2	0.25	1	.015	.027	0.000	False	True	-0.012
1000	1250	.2	0.25	1	.003	.001	0.002	False	True	0.002
1000	500	.2	0.25	1	.040	.105	0.000	False	True	-0.066
1000	1500	.2	0.25	1	.001	.000	1.000	False	False	0.001
1000	250	.2	0.25	1	.116	.374	0.000	False	True	-0.258
1000	1750	.2	0.25	1	.001	.000	1.000	False	False	0.000
1000	1000	.5	0.25	1	.000	.000	1.000	False	False	0.000
1000	750	.5	0.25	1	.000	.000	1.000	False	False	0.000
1000	1250	.5	0.25	1	.000	.000	1.000	False	False	0.000
1000	500	.5	0.25	1	.000	.000	1.000	False	False	0.000

1000	1500	.5	0.25	1	.000	.000	1.000	False	False	0.000
1000	250	.5	0.25	1	.000	.000	1.000	False	False	-0.000
1000	1750	.5	0.25	1	.000	.000	1.000	False	False	0.000
1000	1000	.8	0.25	1	.000	.000	1.000	False	False	0.000
1000	750	.8	0.25	1	.000	.000	1.000	False	False	0.000
1000	1250	.8	0.25	1	.000	.000	1.000	False	False	0.000
1000	500	.8	0.25	1	.000	.000	1.000	False	False	0.000
1000	1500	.8	0.25	1	.000	.000	1.000	False	False	0.000
1000	250	.8	0.25	1	.000	.000	1.000	False	False	0.000
1000	1750	.8	0.25	1	.000	.000	1.000	False	False	0.000
1000	1000	.0	1.75	1	.051	.051	1.000	True	False	0.000
1000	750	.0	1.75	1	.034	.048	0.000	True	True	-0.014
1000	1250	.0	1.75	1	.061	.048	0.000	True	True	0.013
1000	500	.0	1.75	1	.020	.052	0.000	True	True	-0.032
1000	1500	.0	1.75	1	.072	.044	0.000	True	True	0.028
1000	250	.0	1.75	1	.006	.050	0.000	True	True	-0.044
1000	1750	.0	1.75	1	.090	.050	0.000	True	True	0.041
1000	1000	.2	1.75	1	.006	.006	1.000	False	False	0.000
1000	750	.2	1.75	1	.015	.010	0.000	False	True	0.004
1000	1250	.2	1.75	1	.004	.005	0.040	False	True	-0.002
1000	500	.2	1.75	1	.043	.018	0.000	False	True	0.026
1000	1500	.2	1.75	1	.002	.004	0.115	False	False	-0.001
1000	250	.2	1.75	1	.249	.076	0.000	False	True	0.173
1000	1750	.2	1.75	1	.001	.004	0.000	False	True	-0.002
1000	1000	.5	1.75	1	.000	.000	1.000	False	False	0.000
1000	750	.5	1.75	1	.000	.000	1.000	False	False	0.000
1000	1250	.5	1.75	1	.000	.000	1.000	False	False	0.000

1000	500	.5	1.75	1	.000	.000	1.000	False	False	0.000
1000	1500	.5	1.75	1	.000	.000	1.000	False	False	0.000
1000	250	.5	1.75	1	.000	.000	1.000	False	False	0.000
1000	1750	.5	1.75	1	.000	.000	1.000	False	False	0.000
1000	1000	.8	1.75	1	.000	.000	1.000	False	False	0.000
1000	750	.8	1.75	1	.000	.000	1.000	False	False	0.000
1000	1250	.8	1.75	1	.000	.000	1.000	False	False	0.000
1000	500	.8	1.75	1	.000	.000	1.000	False	False	0.000
1000	1500	.8	1.75	1	.000	.000	1.000	False	False	0.000
1000	250	.8	1.75	1	.000	.000	1.000	False	False	0.000
1000	1750	.8	1.75	1	.000	.000	1.000	False	False	0.000
1000	1000	.0	3.00	1	.054	.054	1.000	True	False	0.000
1000	750	.0	3.00	1	.026	.051	0.000	True	True	-0.025
1000	1250	.0	3.00	1	.071	.049	0.000	True	True	0.022
1000	500	.0	3.00	1	.010	.051	0.000	True	True	-0.041
1000	1500	.0	3.00	1	.094	.051	0.000	True	True	0.043
1000	250	.0	3.00	1	.000	.051	0.000	True	True	-0.051
1000	1750	.0	3.00	1	.112	.047	0.000	True	True	0.065
1000	1000	.2	3.00	1	.008	.008	1.000	False	False	0.000
1000	750	.2	3.00	1	.014	.007	0.000	False	True	0.007
1000	1250	.2	3.00	1	.004	.006	0.003	False	True	-0.002
1000	500	.2	3.00	1	.047	.012	0.000	False	True	0.035
1000	1500	.2	3.00	1	.002	.006	0.000	False	True	-0.004
1000	250	.2	3.00	1	.271	.023	0.000	False	True	0.248
1000	1750	.2	3.00	1	.001	.004	0.000	False	True	-0.003
1000	1000	.5	3.00	1	.000	.000	1.000	False	False	0.000
1000	750	.5	3.00	1	.000	.000	1.000	False	False	0.000

1000	1250	.5	3.00	1	.000	.000	1.000	False	False	0.000
1000	500	.5	3.00	1	.000	.000	1.000	False	False	0.000
1000	1500	.5	3.00	1	.000	.000	1.000	False	False	0.000
1000	250	.5	3.00	1	.000	.000	1.000	False	False	0.000
1000	1750	.5	3.00	1	.000	.000	1.000	False	False	0.000
1000	1000	.8	3.00	1	.000	.000	1.000	False	False	0.000
1000	750	.8	3.00	1	.000	.000	1.000	False	False	0.000
1000	1250	.8	3.00	1	.000	.000	1.000	False	False	0.000
1000	500	.8	3.00	1	.000	.000	1.000	False	False	0.000
1000	1500	.8	3.00	1	.000	.000	1.000	False	False	0.000
1000	250	.8	3.00	1	.000	.000	1.000	False	False	0.000
1000	1750	.8	3.00	1	.000	.000	1.000	False	False	0.000

See Table 3 for explanation on column names

Table 6

Desirable conditions

samp1	*samp2*	es	sd1	sd2	perm	*t*-test	pval	typeI	sig	dif
10	10	0.0	1.0	1	0.0491	0.0505	1.00	True	False	-0.0014
10	10	0.2	1.0	1	0.9346	0.9330	0.49	False	False	0.0016
10	10	0.5	1.0	1	0.8261	0.8201	0.00	False	True	0.0060
10	10	0.8	1.0	1	0.6190	0.6122	0.00	False	True	0.0068
60	60	0.0	1.0	1	0.0494	0.0498	1.00	True	False	-0.0004
60	60	0.2	1.0	1	0.8075	0.8063	0.34	False	False	0.0012
60	60	0.5	1.0	1	0.2254	0.2242	0.34	False	False	0.0012
60	60	0.8	1.0	1	0.0091	0.0090	1.00	False	False	0.0001
1000	1000	0.0	1.0	1	0.0488	0.0488	1.00	True	False	-0.0000
1000	1000	0.2	1.0	1	0.0057	0.0057	1.00	False	False	-0.0000
1000	1000	0.5	1.0	1	0.0000	0.0000	1.00	False	False	-0.0000
1000	1000	0.8	1.0	1	0.0000	0.0000	1.00	False	False	-0.0000

See Table 3 for explanation on column names

Table 7

samp1 = 10

	ES = 0.0	ES = 0.2	ES = 0.5	ES = 0.8
SD1 = 1.0				
samp2 perm	3			5
samp2 *t*-test		3	10	10 - 8
SD1 = 0.75				
samp2 perm	10 - 13 - 15 - 18	8 - 5 - 3	8 - 5 -3	8 - 5 - 3
samp2 *t*-test	8 - 5 - 3	13 - 15 - 18	13 - 15 - 18	10 - 13 - 15 - 18
SD1 = 1.25				
samp2 perm	8 - 5 - 3	13 - 15 - 18	13 - 15 - 18	13 - 15 - 18
samp2 *t*-test	13 - 15 - 18	10 - 8 - 5 - 3	10 - 8 - 5 - 3	10 - 8 - 5 - 3
SD1 = 0.50				
samp2 perm	13 - 15 - 18	10 - 8 - 5 - 3	10 - 8 - 5 -3	10 - 8 - 5 - 3
samp2 *t*-test	8 - 5 - 3	13 - 15 - 18	13 - 15 - 18	13 - 15 - 18
SD1 = 1.50				
samp2 perm	8 - 5 - 3	13 - 15 - 18	13 - 15 - 18	13 - 15 - 18
samp2 *t*-test	13 - 15 - 18	8 - 5 - 3	8 - 5 - 3	8 - 5 - 3
SD1 = 0.25				
samp2 perm	13 - 15 - 18	10 - 8 - 5 - 3	10 - 8 - 5 -3	10 - 8 - 5 - 3
samp2 *t*-test	10 - 8 - 5 - 3	13 - 15 - 18	13 - 15 - 18	13 - 15 - 18
SD1 = 1.75				
samp2 perm	8 - 5 - 3	13 - 15 - 18	13 - 15 - 18	10 - 13 - 15 - 18
samp2 *t*-test	13 - 15 - 18	8 - 5 - 3	8 - 5 - 3	8 - 5 - 3
SD1 = 3.0				
samp2 perm	8 - 5 - 3	10 -13 - 15 - 18	10 -13 - 15 - 18	10 -13 - 15 - 18

samp2 t-test	10 -13 - 15 - 18	8 - 5 - 3	8 - 5 - 3	8 - 5 - 3

Each cell contains the sample sizes of group 2 (*samp2*) that either the permutation test performed statistically significant (*samp2* perm) or the t-test performed statistically significant (*samp2* t-test)

The sample size of group 1 (*samp1*) is given in the caption

Each column is seperated by effect size (ES)

Each row is seperated by standard deviation (SD1)

If a sample size for group 2 (*samp2*) is missing it indicates that there were no statistically significant difference between the permutation test and the t-test for this condition

Table 8

samp1 = 60

	ES = 0.0	ES = 0.2	ES = 0.5	ES = 0.8
SD1 = 1.0				
samp2 perm			30 - 15	15
samp2 *t*-test				
SD1 = 0.75				
samp2 perm	75 - 90 - 105	45 - 30 - 15	45 - 30 - 15	45 - 30 - 15
samp2 *t*-test	45 - 30 - 15	75 - 90 - 105	75 - 90 - 105	
SD1 = 1.25				
samp2 perm	45 - 30 - 15	75 - 90 - 105	75 - 90 - 105	
samp2 *t*-test	75 - 90 - 105	45 - 30 - 15	45 - 30 - 15	45 - 30 -15
SD1 = 0.50				
samp2 perm	75 - 90 - 105	45 - 30 - 15	45 - 30 - 15	45 - 30 - 15
samp2 *t*-test	45 - 30 - 15	75 - 90 - 105	75 - 90 - 105	75
SD1 = 1.50				
samp2 perm	45 - 30 - 15	75 - 90 - 105	75 - 90 - 105	
samp2 *t*-test	75 - 90 - 105	45 - 30 - 15	45 - 30 - 15	45 - 30 - 15
SD1 = 0.25				
samp2 perm	75 - 90 - 105	60 - 45 - 30 - 15	60 - 45 - 30 - 15	45 - 30 - 15
samp2 *t*-test	45 - 30 - 15	75 - 90 - 105	75 - 90 - 105	
SD1 = 1.75				
samp2 perm	45 - 30 - 15	75 - 90 - 105	75 - 90 - 105	90 - 105
samp2 *t*-test	75 - 90 - 105	45 - 30 - 15	45 - 30 - 15	45 - 30 - 15
SD1 = 3.0				
samp2 perm	45 - 30 - 15	60 - 75 - 90 - 105	60 - 75 - 90 - 105	75 - 90 - 105

samp2 t-test	75 - 90 - 105	45 - 30 - 15	45 - 30 - 15	45 - 30 -15

see Table 7 for explanation of cells and columns.

Table 9

samp1 = 1000

	ES = 0.0	ES = 0.2	ES = 0.5	ES = 0.8
SD1 = 1.0				
samp2 perm				
samp2 *t*-test				
SD1 = 0.75				
samp2 perm	1250 - 1500 - 1750	750 - 500 - 250		
samp2 *t*-test	750 - 500 - 250			
SD1 = 1.25				
samp2 perm	750 - 500 - 250			
samp2 *t*-test	1250 - 1500 - 1750	750 - 500 - 1000		
SD1 = 0.50				
samp2 perm	1250 - 1500 - 1750	750 - 500 - 250		
samp2 *t*-test	750 - 500 - 250			
SD1 = 1.50				
samp2 perm	750 - 500 - 250			
samp2 *t*-test	1250 - 1500 - 1750	750 - 500 - 250		
SD1 = 0.25				
samp2 perm	1250 - 1500 - 1750	750 - 500 - 250		
samp2 *t*-test	750 - 500 - 250	1250		
SD1 = 1.75				
samp2 perm	750 - 500 - 250			
samp2 *t*-test	1250 - 1500 - 1750			
SD1 = 3.0				
samp2 perm	750 - 500 - 250	1250 - 1500 - 1750		

samp2 t-test	1250 - 1500 - 1750	750 - 500 - 250

see Table 7 for explanation of cells and columns.

Table 10

samp1 = 60 Switched condition

	ES = 0.0	ES = 0.2	ES = 0.5	ES = 0.8
SD1 = 1.0				
samp2 perm				15
samp2 *t*-test				
SD1 = 0.75				
samp2 perm		75 - 90 - 105	90 - 105	
samp2 *t*-test	90	30 - 15	45 - 30 - 15	30 - 15
SD1 = 1.25				
samp2 perm		30 - 15	45 - 30 - 15	30 - 15
samp2 *t*-test	15	90 - 105	105	
SD1 = 0.50				
samp2 perm	30 - 15	75 - 90 - 105	75 - 90 - 105	
samp2 *t*-test	75 - 90 - 105	45 - 30 - 15	45 - 30 - 15	30 - 15
SD1 = 1.50				
samp2 perm	105	45 - 30 - 15	45 - 30 - 15	30 - 15
samp2 *t*-test	30 - 15	75 - 90 - 105	75 - 90 - 105	
SD1 = 0.25				
samp2 perm	30 - 15	75 - 90 - 105	75 - 90 - 105	
samp2 *t*-test	75 - 90 - 105	30 - 15	45 - 30 - 15	30 - 15
SD1 = 1.75				
samp2 perm	90 - 105	45 - 30 - 15	45 - 30 - 15	30 - 15
samp2 *t*-test	45 - 30 - 15	75 - 90 - 105	75 - 90 - 105	
SD1 = 3.0				
samp2 perm	75 - 90 - 105	45 - 30 - 15	45 - 30 - 15	45 - 30 - 15

samp2	t-test	45 - 30 - 15	75 - 90 - 105	75 - 90 - 105

see Table 7 for explanation of cells and columns.

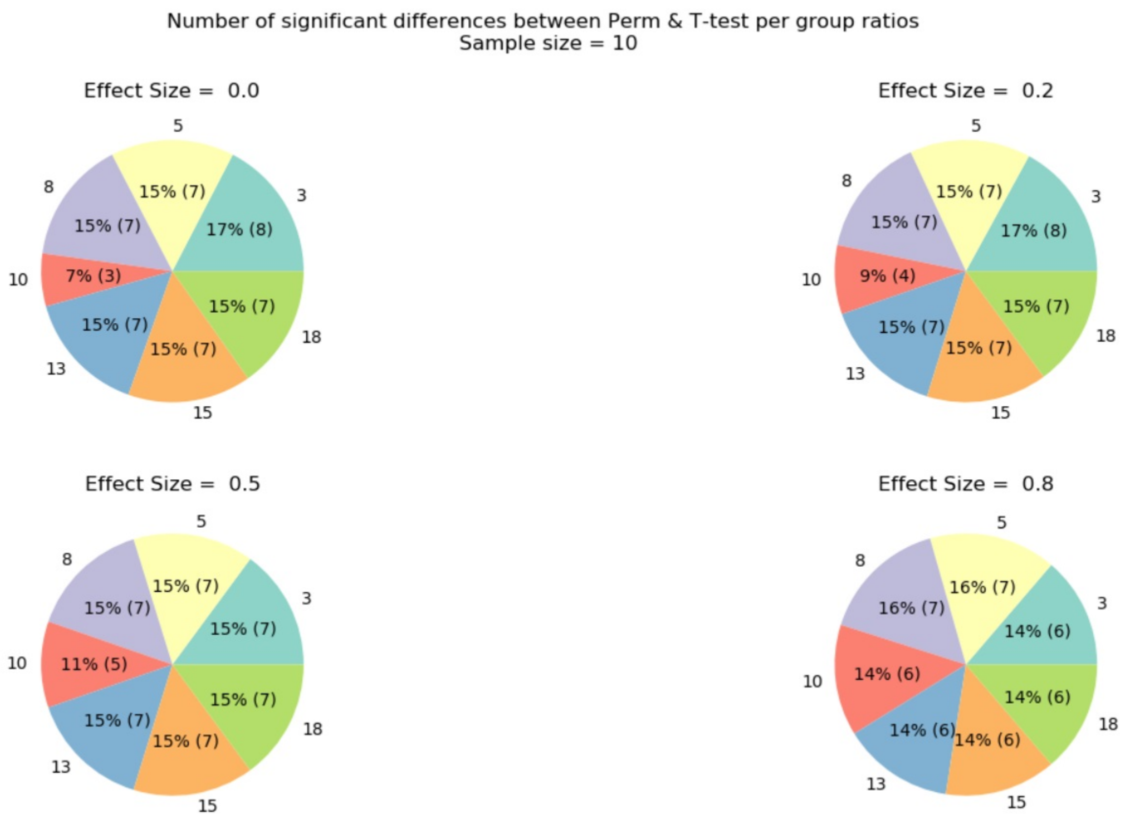

Figure 1. Number of significant differences between permutation test and t-test for each group ratio.

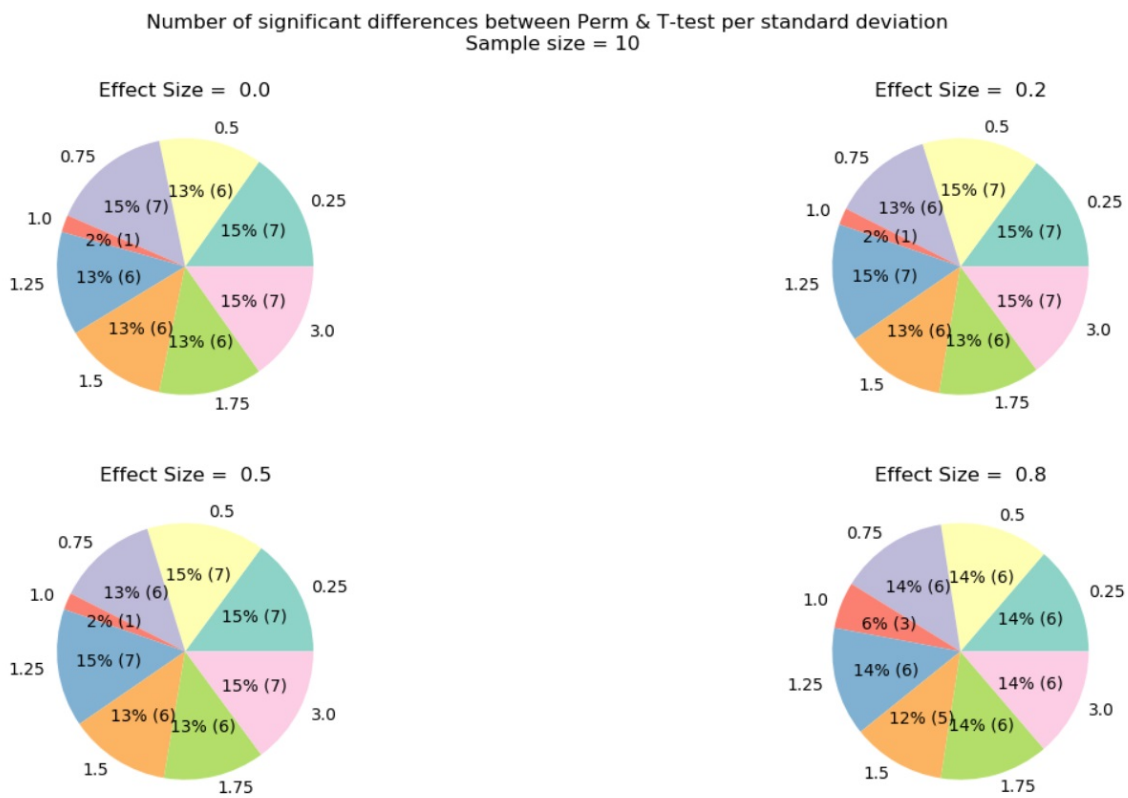

Figure 2. Number of significant differences between permutation test and t-test for each standard deviation.

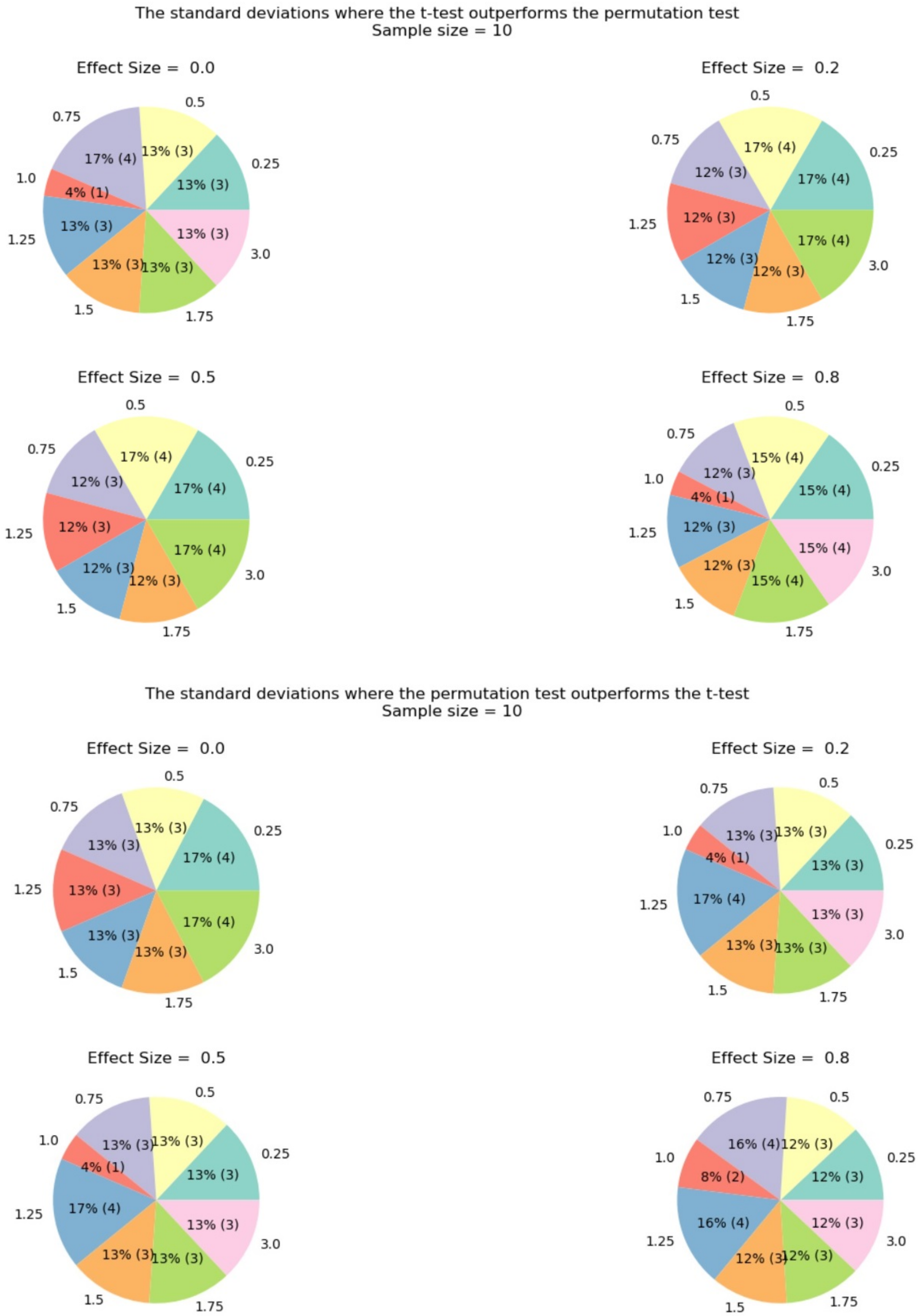

Figure 3. ! insert caption

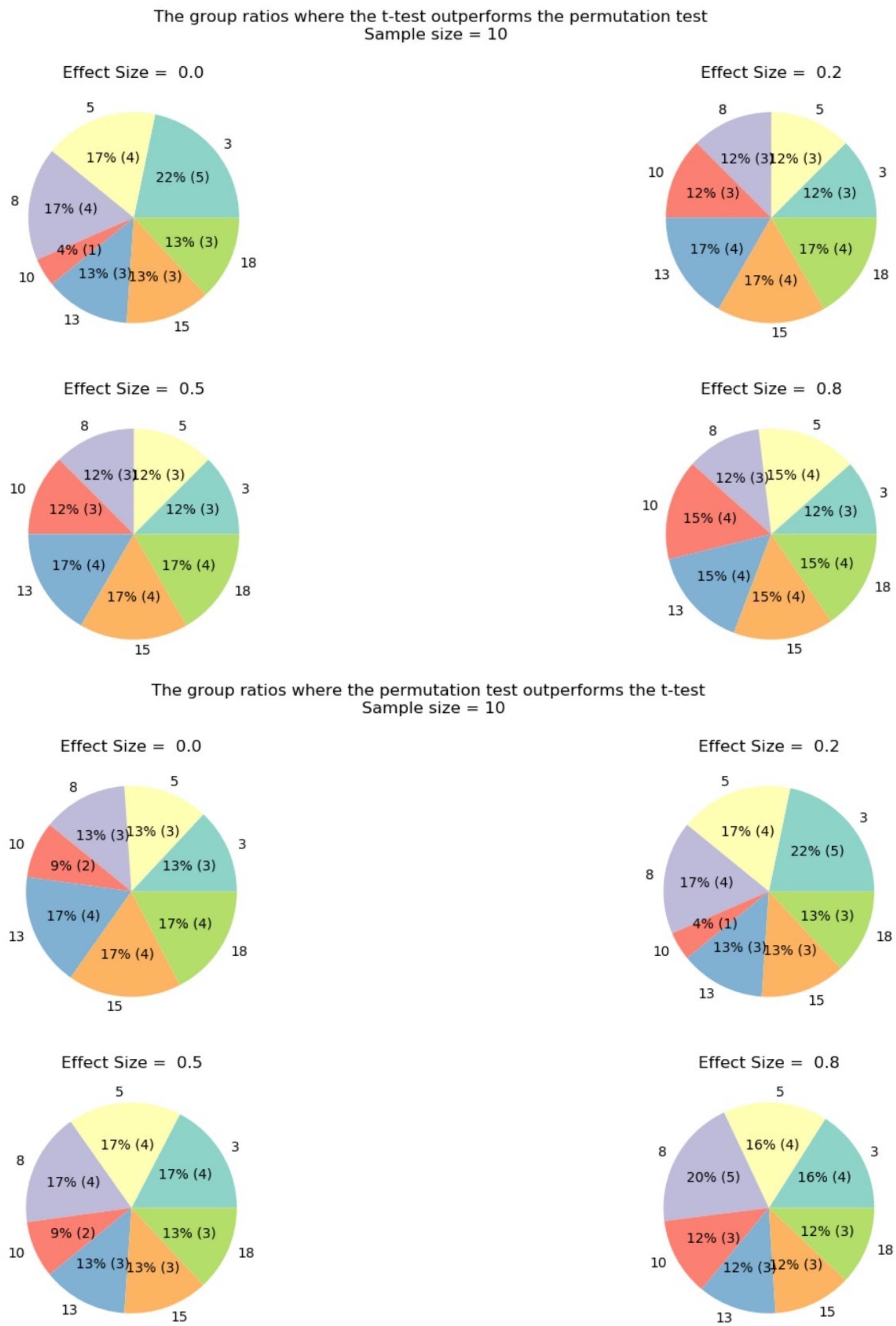

Figure 4. Insert caption

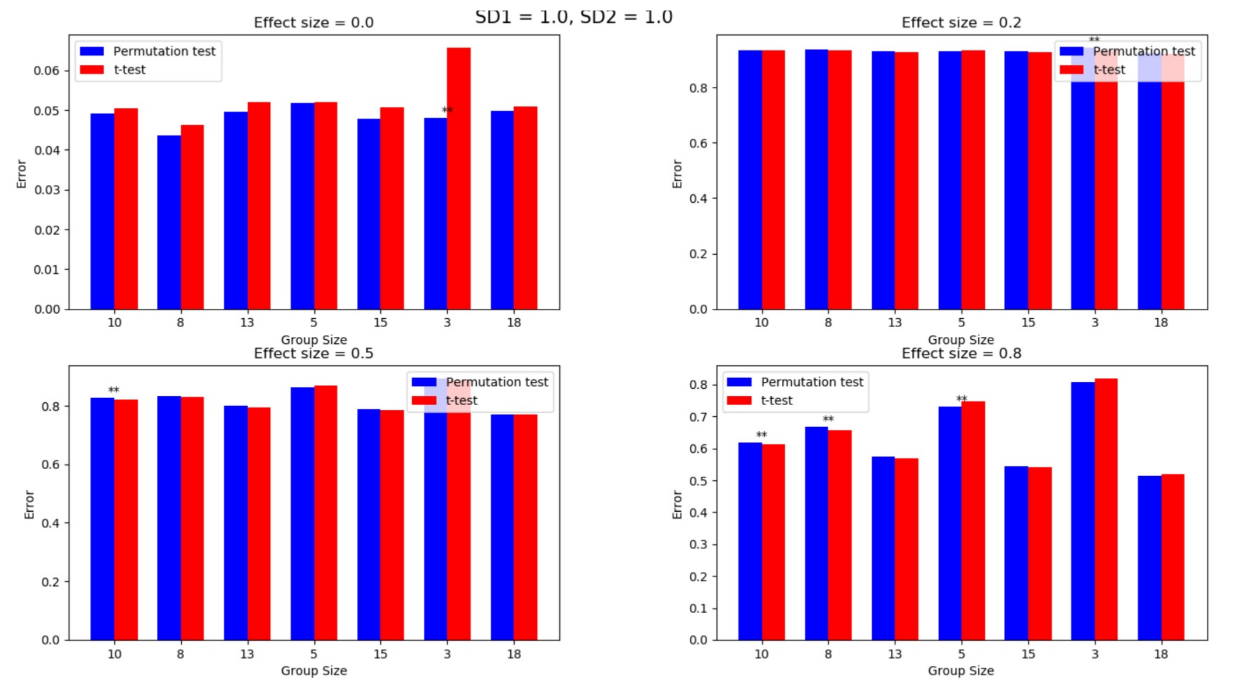

Figure 5. Significant differences between the two tests for sample size of group 1 = 10 and its deviations visible on x-axis

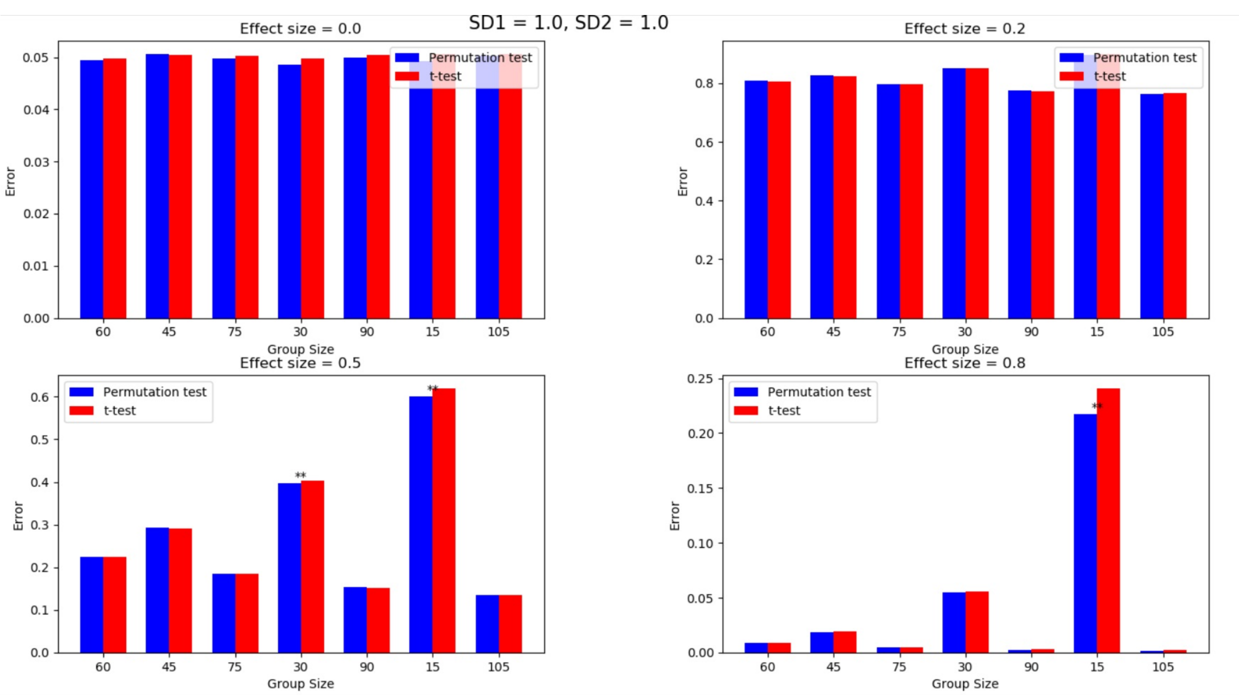

Figure 6. Significant differences between the two tests for sample size of group 1 = 60 and its deviations visible on x-axis

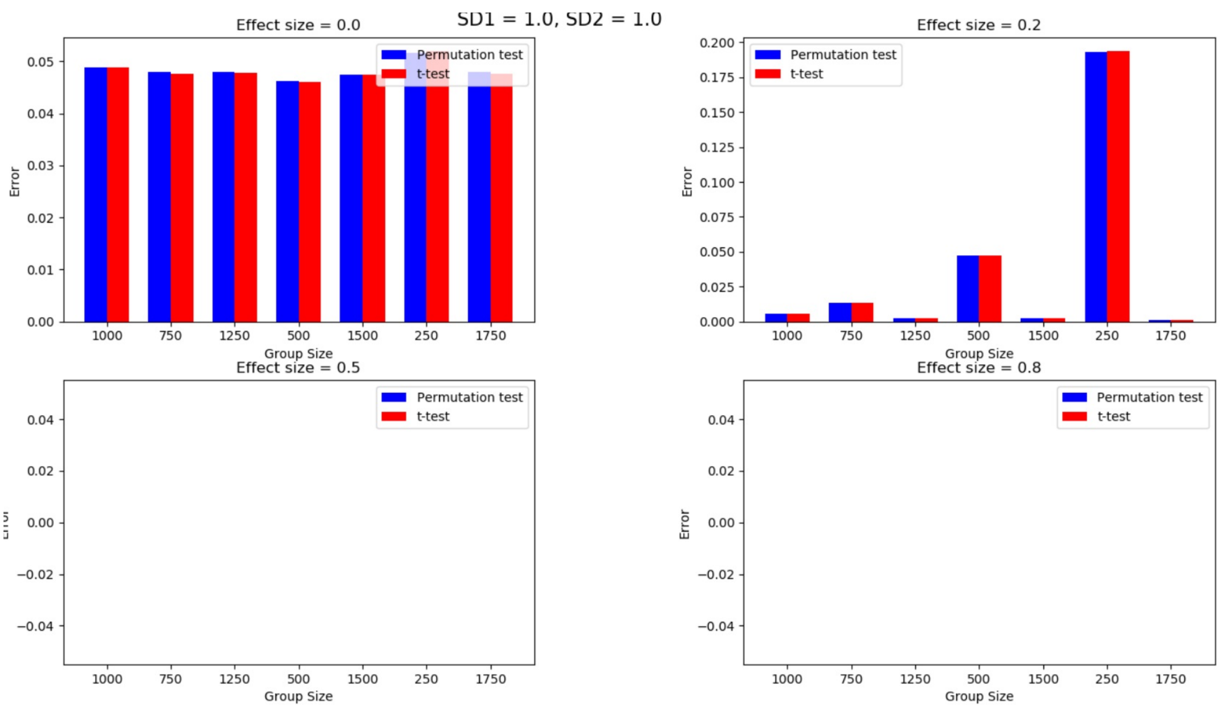

Figure 7. Significant differences between the two tests for sample size of group 1 = 1000 and its deviations visible on x-axis