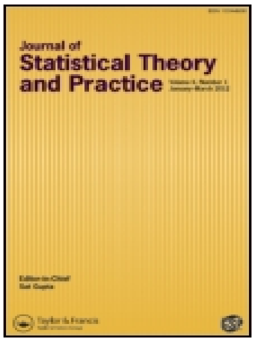

Taylor & Francis
Taylor & Francis Group

Journal of Statistical Theory and Practice

ISSN: 1559-8608 (Print) 1559-8616 (Online) Journal homepage: http://www.tandfonline.com/loi/ujsp20

Proposed Nonparametric Test for the Mixed Two-Sample Design

Rhonda C. Magel & Ran Fu

To cite this article: Rhonda C. Magel & Ran Fu (2014) Proposed Nonparametric Test for the Mixed Two-Sample Design, Journal of Statistical Theory and Practice, 8:2, 221-237, DOI: 10.1080/15598608.2014.847768

To link to this article: http://dx.doi.org/10.1080/15598608.2014.847768

Accepted author version posted online: 24 Feb 2014.
Published online: 24 Feb 2014.

Journal of Statistical Theory and Practice, 8:221–237, 2014
Copyright © Grace Scientific Publishing, LLC
ISSN: 1559-8608 print / 1559-8616 online
DOI: 10.1080/15598608.2014.847768

Proposed Nonparametric Test for the Mixed Two-Sample Design

RHONDA C. MAGEL AND RAN FU

Department of Statistics, North Dakota State University, Fargo, North Dakota, USA

A nonparametric test is proposed for a mixed design consisting of a paired sample portion and a two-independent-sample portion to test for a difference in treatment effects. The test is compared on the basis of estimated powers to a test developed by Dubnicka, Blair, and Hettmansperger (2002). Situations are found in which the proposed test has higher powers and situations are found in which the Dubnicka et al. test has higher powers.

AMS Subject Classification: 62G.

Keywords: Paired data; Independent two-sample data; Mann–Whitney test; Wilcoxon signed-rank test.

1. Introduction

In this article, we propose a test for a mixed design experiment consisting of a combination of two independent random samples and then a random sample of paired data. The proposed test is designed for testing

$$H_0 : \Delta = 0 \text{ vs. } H_1 : \Delta > 0$$

where $\Delta = \theta_1 - \theta_2$ is the difference between the two treatment means. The mixed design being considered is a combination of a paired data design and an independent two-sample design. We are also assuming that the underlying distributions being sampled from are unknown or that only rank data is available, and hence, we are considering nonparametric tests. It is, however, assumed that the two distributions being sampled have the same shape and differ only with respect to their locations.

Situations in which this design could be used may occur in many ways. To begin with, one possibility is that the researcher wishes to compare two treatments and there may be a limited number of pairs available to which the treatments may be randomly assigned. Since the number of available pairs is small, the researcher may decide to use both paired data and two-independent-sample data so that the sample size is increased. Dubnicka, Blair, and Hettmansperger (2002) give an example in which two different laser surgical methods could be used to correct an eye condition in diabetic patients. One way in which the methods will

Received 22 March 2012; accepted 6 July 2013.

Address correspondence to: Rhonda C. Magel, Department of Statistics, North Dakota State University, Fargo, ND 58108-6050, USA. Email: rhonda.magel@ndsu.edu

be compared is by using the results from a visual test given to patients 3 months after their surgery. A limited number of patients having both eyes affected by this eye condition and agreeing to be in the study are available. For these patients, one of their eyes will be randomly selected to have surgical procedure A and the other will have surgical procedure B. There are several patients available in which one eye is affected, and these patients will be randomly assigned one of the two surgical procedures.

A second way this design could occur is because of missing data. A researcher may set up a paired design experiment, and randomly assign one component of the pair to one treatment and the other component to a second treatment. It could happen that data are missing for the first treatment or the second treatment in the pair. Rather than throw out observations from pairs with missing data, one could treat the pairs with missing observations as independent samples.

A third way that this design could occur is that researchers may start out with a total of *n* subjects for comparing the two treatments. They begin by pairing the subjects based on some criteria and realize that some of the subjects can't be paired. Instead of not using some of the available subjects, the researchers may use this mixed design consisting of paired data and two-independent-sample data.

If the design consisted of only paired data, the Wilcoxon signed-rank test (Wilcoxon 1945) could be used to test for a positive difference among the treatments, and is probably the most common nonparametric test for this type of design. The Wilcoxon–Mann–Whitney test (Wilcoxon 1945; Mann and Whitney 1947) is the most commonly used nonparametric test to test for differences among two treatments when one has independent samples.

Dubnicka, Blair, and Hettmansperger (2002) proposed a test for the mixed design of paired and independent sample data, which is a combination of the Wilcoxon signed-rank test and the Wilcoxon–Mann–Whitney test. In the Dubnicka et al. (2002) test, the unstandardized versions of the Wilcoxon signed rank and the Wilcoxon–Mann–Whitney tests are added together first and then standardized. The Dubnicka test with a continuity correction factor, $\mathrm{T}^{+}(0)^{*}{}_{\mathrm{II}}$,, is given as

$$\mathrm{T}^{+}(0)^{*}{}_{\mathrm{II}} = \frac{\mathrm{T}^{+}(0) - \mathrm{E}_0\mathrm{T}^{+}(0) - \frac{1}{2}}{\sqrt{\mathrm{Var}_0\mathrm{T}^{+}(0)}}$$

where $\mathrm{T}^{+}(0) = \mathrm{S}^{+}(0) + \mathrm{U}^{+}(0)$ and

$$\mathrm{E}_0\mathrm{T}^{+}(0) = \frac{\mathrm{n}(\mathrm{n}+1)}{4} + \frac{\mathrm{n}_1\mathrm{n}_2}{2}$$

$$\mathrm{Var}_0\mathrm{T}^{+}(0) = \frac{\mathrm{n}\,(\mathrm{n}+1)\,(2\mathrm{n}+1)}{24} + \frac{\mathrm{n}_1\mathrm{n}_2(\mathrm{n}_1+\mathrm{n}_2+1)}{12} \qquad (1)$$

Here, $\mathrm{S}^{+}(0)$ is the one-sided Wilcoxon signed rank statistic and $\mathrm{U}^{+}(0)$ is the one-sided Wilcoxon–Mann–Whitney statistic. The standardized form of the Wilcoxon signed rank (Wilcoxon 1945) test statistic is given by the following:

$$\mathrm{S}^{+}(0)^{*} = \frac{\mathrm{S}^{+}(0) - \mathrm{E}_0\mathrm{S}^{+}(0)}{\sqrt{\mathrm{Var}_0\mathrm{S}^{+}(0)}} = \frac{\mathrm{S}^{+}(0) - \frac{\mathrm{n}(\mathrm{n}+1)}{4}}{\sqrt{\frac{\mathrm{n}\,(\mathrm{n}+1)\,(2\mathrm{n}+1)}{24}}} \qquad (2)$$

Here, $S^+(0)$ is the one-sided Wilcoxon signed rank statistic defined as

$$S^+(0) = \sum_{i \le j} \sum I\left(\frac{D_i + D_j}{2} > 0\right) = \sum_{i=1}^{n} R\left(|D_i|\right) I\left(D_i > 0\right) \tag{3}$$

where $I(A) = 1$ if event A occurs and $= 0$ otherwise, and $R(|D_i|)$ is the rank of $|D_i|$ among $|D_1|, \ldots, |D_n|$ Differences are first found when calculating the Wilcoxon signed rank test of the form $D_i = Y_i - X_i, i = 1, 2, \ldots, n$, where (X_i, Y_i) is the paired response for treatment i and n is the number of pairs in the sample. Dubnicka et al. (2002) have shown that the nonparametric test they developed is very efficient when compared to the more traditional test using a parametric approach and is also more robust.

The standardized form of the Wilcoxon–Mann–Whitney (Wilcoxon 1945; Mann and Whitney 1947) test statistic is given by the following:

$$U^+(0)^* = \frac{U^+(0) - E_0 U^+(0)}{\sqrt{Var_0 U^+(0)}} = \frac{U^+(0) - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} \tag{4}$$

where $U^+(0)$ is the one-sided Wilcoxon–Mann–Whitney statistic. The test statistic is given by

$$U^+(0) = \sum_{j=1}^{n_1} R_j - \frac{n_1(n_1 + 1)}{2} \tag{5}$$

where R_j is the rank of the jth value in the first sample based on the combined independent sample. The asymptotic null distributions of both $S^+(0)^*$ and $U^+(0)^*$ are standard normal (Daniel 1990).

The test statistic, $T^+(0)^*{}_{II}$, will have an asymptotic standard normal distribution when H_0 is true. H_0 is rejected when $T^+(0)^*{}_{II} > Z_\alpha$, where Z_α is the $(1 - \alpha) \times 100$ percentile of a standard normal distribution.

There are several nonparametric tests available if one wishes to compare more than two treatments in a completely randomized design. The most commonly used test is the Kruskal–Wallis nonparametric test (Kruskal and Wallis 1953), which is an extension of the Mann–Whitney test comparing two treatments. There are also tests available that test for specific types of ordering. The Jonckeere–Terpstra test is used for this type of design when one wishes to test for nondecreasing treatment effects (Jonckheere 1954; Terpstra 1952). The Mack–Wolfe test (Mack and Wolfe 1982) is designed to test for the umbrella alternative (namely, that the treatments have nondecreasing effects up to a point and then nonincreasing effects after that point). The Mack–Wolfe test has a turning point known version and a turning point unknown version.

Several nonparametric tests exist for comparing more than two treatments in a randomized complete block design. The Friedman test (Friedman 1937; 1940) is the most common of these for testing for overall differences of treatment effects. Page's test (Page 1963) tests for nondecreasing treatment effects in a randomized complete block design. The Kim and Kim test (1992) tests for the umbrella alternative for either the peak known case or the peak unknown case in this design.

Durbin (1951) introduced a nonparametric test to compare several treatments for a balanced incomplete block design. Magel, Cao, and Nnundgu (2012) developed a test for nondecreasing treatment effects in a balanced incomplete block design and also discuss and compare using the Wilcoxon signed rank test to test for nondecreasing treatment effects when there are two treatments per block.

Nonparametric tests have been developed for other types of mixed designs. Alvo (1995) developed a test to test for nondecreasing treatment effects in a randomized block design in which observations may be missing. The test is for a mixed design of complete and incomplete blocks. Magel, Terpstra, Canonizado, and Park (2010) introduced tests for mixed designs consisting of a randomized complete block portion and a completely randomized design portion. Some of their tests are for the general alternative testing differences in treatment effects. These tests are linear combinations of the Kruskal–Wallis test and the Friedman tests. Some of their tests are for the umbrella alternative, which consist of linear combinations of the Mack–Wolfe test and the Kim–Kim test. Magel, Terpstra, and Wen (2009) introduced tests for mixed designs consisting of a randomized complete block portion and a completely randomized design portion testing for nondecreasing treatment effects. These tests are linear combinations of Page's test and the Jonckheere–Terpstra test. One of the tests considered gives equal weights to the standardized versions of Page's test and the Jonckheere–Terpstra test. In other words, the tests are standardized first, added together, and then divided by $\sqrt{2}$. The second test considered adds the nonstandardized versions of Page's test and the Jonckheere–Terpsta test together and then restandardizes. The first version of the test had higher powers than the second version unless the number of blocks for the randomized complete block design is large in comparison to the sample sizes for the independent samples. This is because more weight was placed on the Jonckheere–Terpstra test unless the block size is large, and therefore, the contribution from Page's test is minimal. The first test statistic, which gives equal weight to the Jonckheere–Terspstra test and Page's test, then has higher powers. The general findings were that it was better to use the first test, which adds the standardized versions of the Jonchkeere–Terpstra and Page's test statistics together first, unless the sample size for the completely randomized design portion is one-eighth or less of the randomized complete block portion. Magel, Terpstra, Canonizado, and Park (2010) also found that it was usually better to given equal weights to both the Mack–Wolfe test and the Kim–Kim test when testing for umbrella alternatives in this mixed design.

2. Proposed Test

In this article, we are proposing a test for a mixed design consisting of paired data and independent sample data. The test we are proposing is similar to the test developed by Dubnicka et al. (2002). Instead of adding the unstandardized versions of the Wilcoxon signed rank test and the Wilcoxon–Mann–Whitney test together first and then standardizing as in the Dubnicka et al. (2002) test, the standardized versions of the Wilcoxon signed rank test and the Wilcoxon–Mann–Whitney test are added together first and then the sum is restandardized. This is the test statistic that we are proposing for the mixed design, $T^{+}(0)^{*}{}_{I}$, and it is given as

$$\mathrm{T}^{+}(0)^{*}{}_{\mathrm{I}} = \frac{\mathrm{S}^{+}(0)^{*} + \mathrm{U}^{+}(0)^{*}}{\sqrt{2}} \tag{6}$$

Under H_0, $\mathrm{T}^{+}(0)^{*}{}_{\mathrm{I}}$ will have an asymptotic standard normal distribution since the asymptotic distributions of $\mathrm{S}^{+}(0)^{*}$ and $\mathrm{U}^{+}(0)^{*}$ under H_0 are standard normal. These statistics are defined in Eqs. (2) and (4). We reject H_0 when $\mathrm{T}^{+}(0)^{*}{}_{\mathrm{I}} > \mathrm{Z}_{\alpha}$, where Z_{α} is the

$(1 - \alpha) \times 100$ percentile of a standard normal distribution. In the proposed test, the Wilcoxon signed rank test statistic and the Wilcoxon–Mann–Whitney test statistic are standardized first and then added together.

Both the proposed test for the mixed sample design and the test constructed by Dubnicka et al. (2002) use combinations of the Wilcoxon signed rank test statistic and the Wilcoxon–Mann–Whitney test statistic. They are actually both just weighted versions of the standardized Wilcoxon signed rank and the standardized Wilcoxon–Mann–Whitney statistics. This can be seen in the following.

$$\begin{aligned} T^{+}(0)^{*}{}_{I} &= \frac{S^{+}(0)^{*} + U^{+}(0)^{*}}{\sqrt{2}} \\ &= 1/\sqrt{2}\left(\left(S^{+}(0) - \mu_s\right)\right)/\sigma_s + \left(U^{+}(0) - \mu_u\right)/\sigma_u)) \\ &= 1/\sqrt{2}\left(S^{+}(0)^{*} + U^{+}(0)^{*}\right) \end{aligned} \tag{7}$$

It is noted that the proposed test, $T^{+}(0)^{*}{}_{I}$, is a linear combination of both the standardized Wilcoxon signed rank test and the standardized Wilcoxon–Mann–Whitney test with equal weight of $1/\sqrt{2}$ for both parts. The Dubnicka test without the ½ continuity correction factor is written as a weighted linear combination of the standardized Wilcoxon signed rank Test and the standardized Wilcoxon–Mann–Whitney test in Eq. (8)

$$\begin{aligned} T^{+}(0)^{*}{}_{II} &= \left(\left(S^{+}(0) + U^{+}(0)\right) - (\mu_s - \mu_u)\right)/\sqrt{}((\sigma^{2}{}_{S} + \sigma^{2}{}_{U})) \\ &= \sqrt{}\left(\sigma^{2}{}_{S}/\left(\sigma^{2}{}_{S} + \sigma^{2}{}_{U}\right)\right)^{*} S^{+}(0)^{*} + \sqrt{}\left(\sigma^{2}{}_{U}/\left(\sigma^{2}{}_{S} + \sigma^{2}{}_{U}\right)\right)^{*} U^{+}(0)^{*} \end{aligned} \tag{8}$$

where $\sigma^{2}{}_{S} = n(n+1)(2n+1)/24$ and $\sigma^{2}{}_{U} = n_1 n_2(n_1 + n_2 + 1)/12$

The Dubnicka et al. (2002) test statistic, $T^{+}(0)^{*}{}_{II}$, gives different weights to both the standardized Wilcoxon signed rank statistic and the standardized Wilcoxon–Mann–Whitney statistic. If the variance of the Wilcoxon signed rank test, $\sigma^{2}{}_{S}$, is larger, the Wilcoxon signed rank test will be given more weight. If the variance of the Wilcoxon–Mann–Whitney test is larger, the Wilcoxon–Mann $\neq$ Whitney test will be given more weight. The variance of the Wilcoxon signed rank test, $\sigma^{2}{}_{S}$, is equal to $(n(n+1)(2n+1))/24$ with n being the number of pairs. The variance of the Wilcoxon–Mann–Whitney test, $\sigma^{2}{}_{U}$, is given by $((n_1)(n_2)(n_1 + n_2 + 1))/12$ with the independent samples sizes being n_1 and n_2. When the number of pairs for the Wilcoxon signed rank test is equal to the sample size for each of the independent samples, assuming n_1 and n_2 are equal, one can see that the variance of the Wilcoxon–Mann–Whitney test is about twice the size as for the Wilcoxon signed rank test and therefore, the Wilcoxon–Mann–Whitney test would be given twice as much weight as the Wilcoxon signed rank test in the Dubnicka et al. (2002) test. When the number of pairs is two or more times higher than the sample size for each of the independent samples, the Wilcoxon signed rank test will have more weight in the Dubnicka et al. (2002) test.

The question becomes, "Which test statistic has higher powers?" In order to answer this question, a simulation study is conducted calculating estimated powers of these two tests under a variety of situations.

3. Simulation Study

A simulation study is designed to estimate and compare the powers of the proposed test and the powers of the test constructed by Dubnicka et al. (2002). The simulation study considers a variety of sample sizes for independent data and paired data; different combinations of variances for the independent data versus the paired data; three different underlying distributions; and a variety of location parameter values. Generally, a paired data experiment is used to reduce the error variation, and therefore the error variance associated with the paired data is smaller than the error variance associated with the independent-sample data. We would like to determine whether or not it makes a difference in which test statistic to use when the error variance in both the paired data and the independent-sample data is about the same and when the independent-sample data have a larger error variance. All powers are estimated based on 10,000 iterations for each combination of the sample sizes, distributions, variances, and locations parameter arrangements. Programs for simulation are coded using R 2.11.1.

The simulation study estimates powers for a variety of sample sizes. The sample sizes for both of the independent samples are always the same. Situations are considered in which the sample sizes for the independent sample data are equal to the sample size for the paired data. Situations are also considered in which the sample size for the pairs is larger than the sample sizes for the independent sample data and where the sample size for the paired data is smaller than the sample sizes for the independent data. This is done to see whether which test statistic is better depends on the sample sizes used, since the variance of the test statistic depends on the sample sizes and the weights depend on the sample sizes for each component in the Dubnicka et al. (2002) test. The following is a list of all of the sample sizes considered where $\mathrm{n_i}$ is equal to the independent sample size and $\mathrm{n_p}$ is equal to the paired sample size:

1. $\mathrm{n_i} = 10$, $\mathrm{n_p} = 5$.
2. $\mathrm{n_i} = 10$, $\mathrm{n_p} = 10$.
3. $\mathrm{n_i} = 10$, $\mathrm{n_p} = 20$.
4. $\mathrm{n_i} = 20$, $\mathrm{n_p} = 10$.
5. $\mathrm{n_i} = 30$, $\mathrm{n_p} = 10$.
6. $\mathrm{n_i} = 20$, $\mathrm{n_p} = 5$.
7. $\mathrm{n_i} = 10$, $\mathrm{n_p} = 30$.
8. $\mathrm{n_i} = 10$, $\mathrm{n_p} = 40$.
9. $\mathrm{n_i} = 5$, $\mathrm{n_p} = 20$.

The simulation study considers different combinations of variances. Situations are considered in which the variance for the completely randomized design portion is equal to the variance of the paired design portion. Situations are also considered in which the variance of the completely randomized design portion is larger than the variance of the paired design portion.

Differing underlying distributions are also considered in addition to the different variances. The underlying distributions considered include the normal distribution, the exponential distribution, and the *t*-distribution. The *t*-distribution was considered because it is a symmetric distribution, but it has larger tails than the standard normal distribution when the degrees of freedom are relatively small. We selected a *t*-distribution with 3 degrees of freedom as our baseline distribution because this will have thick tails compared with the standard normal distribution. The exponential distribution was selected because this is a nonsymmetric distribution that occurs quite often and it is light-tailed skewed.

Skewed distributions with heavier tails were not included in the study because the Wilcoxon signed rank test assumes a symmetric distribution. The stated significance values would not hold for skewed distributions that are more heavily tailed. For the normal distribution, the following combinations of variances are considered:

1. Completely randomized design variance = 1, paired design variance = 1.
2. Completely randomized design variance = 2, paired design variance = 1.
3. Completely randomized design variance = 4, paired design variance = 1.
4. Completely randomized design variance = 16, paired design variance = 1.
5. Completely randomized design variance = 49, paired design variance = 1.

We do note that it is unusual for the completely randomized design variance to be 16 or 49 times higher than the paired design variance. We included these situations in the normal case to see what would happen in the comparisons between the two test statistics if the variance difference were to get that large.

For the exponential distribution, the following combinations of variances are considered:

1. Completely randomized design variance = 1, paired design variance = 1.
2. Completely randomized design variance = 2, paired design variance = 1.
3. Completely randomized design variance = 4, paired design variance = 1.
4. Completely randomized design variance = 7, paired design variance = 1.

We do note that case 4 for the exponential distribution would be unusual. We included it here to see whether it would make a difference as to which test statistic did better.

For the *t*-distribution, the following combinations of degrees of freedom are considered:

1. Completely randomized design portion is from a *t*-distribution with 3 degrees of freedom, paired design portion is from a *t*-distribution with 3 degrees of freedom (completely randomized design variance = 3; paired design variance = 3).
2. Completely randomized design portion is from a *t*-distribution with 3 degrees of freedom, paired design portion is from a *t*-distribution with 10 degrees of freedom (completely randomized design variance = 3; paired design variance = 1.25).
3. Completely randomized design portion is from a *t*-distribution with 3 degrees of freedom, paired design portion is from a *t*-distribution with 20 degrees of freedom (completely randomized design variance = 3; paired design variance = 1.11).

We do note that the ratios between the variances are not the same as those considered in either the normal case or the exponential case, but we could not get the same ratios in considering the *t*-distribution where only the degrees of freedom were varied.

In all situations considered, significance levels of the tests are estimated by counting the number of times each test rejected the null hypothesis dividing by 10,000 when both of the means were actually equal and there is no difference between the treatments. The estimated significance levels are compared to the stated significance level, which is always 0.05. Powers are estimated for a variety of location parameter shifts.

4. Results

Selected results are given in Tables 1 through 8. When the underlying distributions are normal and exponential, three tables of estimated powers are given. For the *t*-distributions considered, two tables of estimated powers are given. In each case, one of the tables contains the estimated powers when the variance of the completely randomized design portion is equal to the variance of the paired differences portion. In the other table(s) given for each distribution, the variance of the completely randomized design portion is larger than the paired difference variance portion. Two of these second types of tables are given when the underlying distributions are normal, two are given when the underlying distributions are exponential, and one is given when the underlying distributions are t-distributions.

Table 1 gives the estimated powers when the variance of the completely randomized portion was equal to the variance of the paired design portion when the underlying distributions are normal. The levels of significance are estimated for both tests and for all sample sizes considered. These are found to all be around 0.05, which is the stated significance level that is always used. Powers are estimated for each test when the means are different for different sample size combinations of the completely randomized design portion and the paired design portion.

Table 1
Estimated powers of tests for mixed design under normal distributions (completely randomized design variance = paired design variance)

Completely randomized design variance = 1 Paired design variance = 1						
Sample size-independent pairs	$\mu_1 = 0, \mu_2 = 0$		$\mu_1 = 0.5, \mu_2 = 0$		$\mu_1 = 0.8, \mu_2 = 0$	
	Dub	New	Dub	New	Dub	New
$n_i = 10$ $n_p = 5$	0.0435	0.0509	0.3272	0.3455	0.6013	0.6333
$n_i = 10$ $n_p = 10$	0.0433	0.0477	0.4011	0.4300	0.7396	0.7693
$n_i = 10$ $n_p = 20$	0.0482	0.0485	0.5674	0.5700	0.8931	0.8958
$n_i = 20$ $n_p = 10$	0.0508	0.0522	0.5307	0.5676	0.8721	0.8961
$n_i = 30$ $n_p = 10$	0.0496	0.0508	0.6350	0.6592	0.9383	0.9463
$n_i = 20$ $n_p = 5$	0.0477	0.0494	0.4577	0.4786	0.8055	0.8136
$n_i = 10$ $n_p = 30$	0.0509	0.0484	0.6707	0.6631	0.9522	0.9472
$n_i = 10$ $n_p = 40$	0.0532	0.0523	0.7452	0.7321	0.9818	0.9766
$n_i = 5$ $n_p = 20$	0.0526	0.0505	0.4836	0.4720	0.8147	0.8006

Table 2
Estimated powers of tests for mixed design under normal distributions (completely randomized design variance = 4 × paired design variance)

Completely randomized design variance = 4
Paired design variance = 1

Sample size-independent pairs	$\mu_1 = 0, \mu_2 = 0$		$\mu_1 = 0.5, \mu_2 = 0$		$\mu_1 = 0.8, \mu_2 = 0$	
	Dub	New	Dub	New	Dub	New
$n_i = 10$ $n_i = 30$	0.0493	0.0488	0.1632	0.2173	0.2797	0.3914
$n_i = 10$ $n_p = 10$	0.0442	0.0526	0.2596	0.2890	0.4641	0.5299
$n_i = 10$ $n_p = 20$	0.0484	0.0468	0.4637	0.4068	0.8027	0.7259
$n_i = 20$ $n_p = 10$	0.0478	0.0502	0.2591	0.3499	0.4724	0.6369
$n_i = 30$ $n_p = 10$	0.0502	0.0471	0.2719	0.4006	0.5250	0.7164
$n_i = 20$ $n_p = 5$	0.0481	0.0527	0.2018	0.2659	0.3626	0.5131
$n_i = 10$ $n_p = 30$	0.0519	0.0514	0.6092	0.5078	0.9193	0.8324
$n_i = 10$ $n_p = 40$	0.0503	0.0522	0.7117	0.5915	0.9716	0.9045
$n_i = 5$ $n_p = 20$	0.0489	0.0516	0.4621	0.3702	0.7872	0.6525

When the variance of the completely randomized portion was equal to the variance of the paired design portion, the proposed test had approximately the same powers as the Dubnicka test under all the different sample size combinations considered. Recall that the sample size combinations that are considered include situations in which the sample sizes for each portion of the design are equal, combinations in which the completely randomized design portion has more observations, and combinations in which the paired design portion has more observations (see Table 1). For example, when the sample sizes for the independent data are 10 for each treatment and the sample size for the paired data is 10 pairs; when the sample sizes for the independent data are 20 for each treatment and the sample size for the paired data is 5 pairs; and when the sample size for the independent data are 5 for each treatment and the sample size for the paired data is 20 pairs, the proposed test has approximately the same powers as the Dubnicka et al. (2002) test.

When the variance of the completely randomized design portion is equal to two times the variance of the paired design portion, the significance levels of both tests are estimated for all sample size combinations considered when the underlying distributions are normal. They are all around 0.05. We compare powers of the two tests for several different sample size combinations of the independent data and the paired data. The differences of the powers between the two tests are larger in this case than they are when the variances for

Table 3
Estimated powers of tests for mixed design under normal distributions (completely randomized design variance = 49 × paired design variance)

	Completely randomized design variance = 49 Paired design variance = 1					
Sample size-independent pairs	$\mu_1 = 0, \mu_2 = 0$		$\mu_1 = 0.5, \mu_2 = 0$		$\mu_1 = 0.8, \mu_2 = 0$	
	Dub	New	Dub	New	Dub	New
$n_i = 10$ $n_p = 5$	0.0463	0.0482	0.0673	0.1203	0.0867	0.1718
$n_i = 10$ $n_p = 10$	0.0453	0.0488	0.1385	0.1779	0.2087	0.2848
$n_i = 10$ $n_p = 20$	0.0469	0.0523	0.3666	0.2743	0.6711	0.4819
$n_i = 20$ $n_p = 10$	0.0498	0.0491	0.0855	0.1714	0.1134	0.2913
$n_i = 30$ $n_p = 10$	0.0482	0.0510	0.0728	0.1739	0.0885	0.2939
$n_i = 20$ $n_p = 5$	0.0459	0.0473	0.0571	0.1168	0.0656	0.1737
$n_i = 10$ $n_p = 30$	0.0532	0.0502	0.5550	0.3575	0.8879	0.6369
$n_i = 10$ $n_p = 40$	0.0488	0.0513	0.6879	0.4431	0.9586	0.7324
$n_i = 5$ $n_p = 20$	0.0572	0.0508	0.4274	0.2771	0.7573	0.4751

both design portions are equal. When the sample size for the independent samples is greater than or equal to the sample size for the number of pairs, the proposed test has higher powers than the Dubnicka test. When the sample size for the number of pairs is two or more times greater than the sample size for the independent samples, the Dubnicka test has higher powers than the proposed test. For example, when the sample sizes for the independent data are 20 for each treatment and the sample size for the paired data is 10 pairs, the proposed test has higher estimated powers than the Dubnicka test. When the sample sizes for the independent data are 5 for each treatment and the sample size for the paired data is 20 pairs, the Dubnicka test has higher estimated powers than the proposed test in this case.

Given normal distributions, when the variance of the completely randomized design portion is equal to four times the variance of the paired design portion, the variance of the completely randomized design portion is equal to 16 times the variance of the paired design portion, and the variance of the completely randomized design portion is equal to 49 times the variance of the paired design portion; the results are consistent with the results when the variance of the completely randomized design portion was equal to two times the variance of the paired design portion. The separation between powers is more pronounced as the differences between the variances become larger. Results are given in Table 2 when the underlying distributions are normal and the variance of the completely randomized design

Table 4
Estimated powers of tests for mixed design under exponential distributions (completely randomized design variance = paired design variance)

Completely randomized design variance = 1 Paired design variance = 1						
Sample size-independent pairs	$\mu_1 = 1, \mu_2 = 1$		$\mu_1 = 1.5, \mu_2 = 1$		$\mu_1 = 1.8, \mu_2 = 1$	
	Dub	New	Dub	New	Dub	New
$n_i = 10$ $n_p = 5$	0.0444	0.0481	0.5242	0.5260	0.7987	0.8120
$n_i = 10$ $n_p = 10$	0.0463	0.0534	0.6058	0.6319	0.8879	0.8985
$n_i = 10$ $n_p = 20$	0.0503	0.0513	0.7436	0.7587	0.9588	0.9660
$n_i = 20$ $n_p = 10$	0.0515	0.0486	0.7858	0.7991	0.9722	0.9762
$n_i = 30$ $n_p = 10$	0.0512	0.0458	0.8802	0.8782	0.9928	0.9938
$n_i = 20$ $n_p = 5$	0.0474	0.0456	0.7314	0.7109	0.9430	0.9369
$n_i = 10$ $n_p = 30$	0.0541	0.0497	0.8151	0.8462	0.9835	0.9891
$n_i = 10$ $n_p = 40$	0.0505	0.0485	0.8748	0.8982	0.9958	0.9971
$n_i = 5$ $n_p = 20$	0.0467	0.0496	0.6224	0.6514	0.9006	0.9129

portion is equal to four times the variance of the paired design portion. Table 3 is also given here for illustration purposes and contains the estimated powers of the tests when the variance of the completely randomized portion is 49 times higher than the variance of the paired design portion. We recognize that this case would rarely occur, but it is given here only to illustrate how the powers of the two test statistics are related as the difference in the magnitude between the two variances changes. Examining Table 2, when $\mu_1 = 0.8$ and $\mu_2 = 0$, $n_i = 10$ and $n_p = 5$, the estimated power of the Dubnicka test is 0.2797, which is 71.46% of the estimated power of the new test, 0.3914. When $n_i = 10$ and $n_p = 10$, the estimated power of the Dubnicka test, 0.4641, is 87.58% of the estimated power of the new test, 0.5299. Examining Table 3, when the variance of the completely randomized design portion is now 49 times the variance of the paired design portion instead of only four times the variance, the estimated powers of the Dubnicka test for the same two cases, 0.0867 and 0.2087, are 50.47% and 73.28% of the estimated powers of the new test, 0.1718 and 0.2848, respectively. In Table 2, when $n_i = 5$ and $n_p = 20$, the estimated power of the new test, 0.6525, is 82.89% of the estimated power of the Dubnicka test, 0.7872. In Table 3 for the same case, the estimated power of the new test, 0.4751, is 62.74% of the estimated power of the Dubnicka test, 0.4751. The ratios of the estimated powers become more pronounced

Table 5
Estimated powers of tests for mixed design under exponential distributions (completely randomized design variance = 2 × paired design variance)

Completely randomized design variance = 2 Paired design variance = 1						
Sample size-independent pairs	$\mu_1 = 1, \mu_2 = 1$		$\mu_1 = 1.5, \mu_2 = 1$		$\mu_1 = 1.8, \mu_2 = 1$	
	Dub	New	Dub	New	Dub	New
$n_i = 10$ $n_p = 5$	0.0501	0.0527	0.2713	0.3271	0.4604	0.5773
$n_i = 10$ $n_p = 10$	0.0426	0.0501	0.3815	0.4392	0.6598	0.7188
$n_i = 10$ $n_p = 20$	0.0472	0.0514	0.6196	0.5832	0.9081	0.8692
$n_i = 20$ $n_p = 10$	0.0489	0.0500	0.4422	0.5345	0.7238	0.8372
$n_i = 30$ $n_p = 10$	0.0498	0.0512	0.5003	0.6242	0.7938	0.9014
$n_i = 20$ $n_p = 5$	0.0517	0.0525	0.3508	0.4342	0.6058	0.7210
$n_i = 10$ $n_p = 30$	0.0499	0.0490	0.7575	0.6914	0.9716	0.9388
$n_i = 10$ $n_p = 40$	0.0504	0.0492	0.8455	0.7675	0.9907	0.9721
$n_i = 5$ $n_p = 20$	0.0482	0.0481	0.5836	0.5134	0.8707	0.7966

as the variance for the completely randomized design portion increases with respect to the variance of the paired design portion.

Selected results for exponential distribution are given in Tables 4–6. As in the case for the normal distribution, situations are considered when the variance of the completely randomized portion is equal to the variance of the paired design portion. The levels of significance are estimated for both tests and for all sample sizes considered. These are found to all be around 0.05, which is the stated significance level. Powers are estimated for each test when the means are different for different sample size combinations of the completely randomized design portion and the paired design portion. When the variance of completely randomized portion is equal to the variance of the paired design portion, the proposed test has approximately the same powers as the Dubnicka test under all the different sample size combinations considered. Results are given in Table 4.

When the variance of the completely randomized design portion is equal to two times the variance of the paired design portion (Table 5), the significance levels of both tests are estimated for all sample size combinations considered when the underlying distributions are exponential. They are all around 0.05. We compare powers of the two tests for several different sample size combinations of the independent data and the paired data. The differences between the powers of the two tests is larger in this case compared with the case

Table 6

Estimated powers of tests for mixed design under exponential distributions (completely randomized design variance = 4 × paired design variance)

	Completely randomized design variance = 4 Paired design variance = 1					
Sample size-independent Pairs	$\mu_1 = 1, \mu_2 = 1$		$\mu_1 = 1.5, \mu_2 = 1$		$\mu_1 = 1.8, \mu_2 = 1$	
	Dub	New	Dub	New	Dub	New
$n_i = 10$ $n_p = 5$	0.0507	0.0518	0.1475	0.2336	0.2518	0.3924
$n_i = 10$ $n_p = 10$	0.0418	0.0521	0.2668	0.3281	0.4630	0.5532
$n_i = 10$ $n_p = 20$	0.0482	0.0504	0.5647	0.4697	0.8491	0.7469
$n_i = 20$ $n_p = 10$	0.0498	0.0516	0.2367	0.3735	0.4115	0.6455
$n_i = 30$ $n_p = 10$	0.0435	0.0512	0.2339	0.4166	0.4169	0.7042
$n_i = 20$ $n_p = 5$	0.0466	0.0456	0.1688	0.2786	0.2949	0.4803
$n_i = 10$ $n_p = 30$	0.0512	0.0482	0.7239	0.5743	0.9574	0.8620
$n_i = 10$ $n_p = 40$	0.0505	0.0541	0.8277	0.6684	0.9889	0.9245
$n_i = 5$ $n_p = 20$	0.0474	0.0509	0.5717	0.4286	0.8638	0.6910

in which the variances for both design portions are equal. Again, it is seen that when the sample size for the independent samples is equal to or greater than the sample size for the number of pairs, the proposed test has higher powers than the Dubnicka test. When the sample size for the number of pairs is at least two times the sample size for the number of independent samples, the Dubnicka test has higher powers than the proposed test. For example, when the sample size for the independent data is 30 for each treatment and the sample size for the paired data is 10 pairs, the proposed test has higher powers than the Dubnicka test. When the sample size for the independent data is 10 for each treatment and the sample size for the paired data is 30 pairs, the Dubnicka test has higher powers than the proposed test.

When the variance of the completely randomized design portion was equal to four times the variance of the paired design portion (Table 6) and when the variance of the completely randomized design portion was equal to seven times the variance of the paired design portion, the results are consistent with the results when the variance of the completely randomized design portion was equal to two times the variance of the paired design portion. As in the case when the underlying distributions were normal, the differences between the powers of the two tests became more pronounced as the differences between the variances increased. As an example, when $\mu_1 = 1.5$ and $\mu_2 = 1$, $n_i = 20$ and $n_p = 10$,

Table 7

Estimated powers of tests for mixed design under Student's *t*-distributions (completely randomized design ~ t(3); paired design ~ t(3))

	Completely randomized design ~ t(3) Paired design ~ t(3)					
Sample size-independent pairs	$\mu_1 = 0, \mu_2 = 0$		$\mu_1 = 0.5, \mu_2 = 0$		$\mu_1 = 0.8, \mu_2 = 0$	
	Dub	New	Dub	New	Dub	New
$n_i = 10$ $n_p = 5$	0.0456	0.0518	0.2448	0.2599	0.4429	0.4648
$n_i = 10$ $n_p = 10$	0.0434	0.0520	0.2920	0.3039	0.5433	0.5657
$n_i = 10$ $n_p = 20$	0.0467	0.0477	0.3873	0.3999	0.7042	0.7036
$n_i = 20$ $n_p = 10$	0.0491	0.0499	0.3985	0.4124	0.7078	0.7222
$n_i = 30$ $n_p = 10$	0.0498	0.0505	0.4894	0.4944	0.8155	0.8239
$n_i = 20$ $n_p = 5$	0.0506	0.0494	0.3332	0.3243	0.5806	0.5965
$n_i = 10$ $n_p = 30$	0.0542	0.0480	0.4577	0.4737	0.7902	0.7947
$n_i = 10$ $n_p = 40$	0.0526	0.0497	0.5253	0.5347	0.8540	0.8582
$n_i = 5$ $n_p = 20$	0.0493	0.0542	0.3509	0.3487	0.6442	0.6225

the estimated power of the Dubnicka test, 0.4422, was 82.75% of the estimated power of the new test, 0.5345, when the variance of the completely randomized design portion was two times the variance of the paired design portion (Table 5). For the same case, when the variance of the completely randomized design portion was four times the variance of the paired design portion, the estimated power of the Dubnicka test, 0.2367, was 63.37% of the estimated power of the new test, 0.3735. When $n_i = 5$ and $n_p = 20$, the estimated power of the new test, 0.5134, was 87.97% of the estimated power of the Dubnicka test when the variance of the completely randomized design portion was two times the variance of the paired design portion (Table 5). For this same case when the variance of the completely randomized design portion was four times the variance of the paired design portion (Table 6), the estimated power of the new test, 0.4286, was 74.97% of the estimated power of the Dubnicka test, 0.5717 (Table 6). The ratios of the estimated powers become more pronounced as the variance of the completely randomized design portion increases with respect to the variance of the paired design portion.

When the completely randomized portion and the paired design portion are both from a *t*-distribution with 3 degrees of freedom, the proposed test has approximately the same powers as the Dubnicka test under all the different sample size combinations considered (see Table 7).

Table 8
Estimated powers of tests for mixed design under Student's t-distributions (completely randomized design ~ t(3); paired design ~ t(10))

Completely randomized design ~ t(3)
Paired design ~ t(10)

Sample size-independent pairs	$\mu_1 = 0, n_i = 10$		$\mu_1 = 0.5, \mu_2 = 0$		$\mu_1 = 0.8, \mu_2 = 0$	
	Dub	New	Dub	New	Dub	New
$n_i = 10$ $n_p = 5$	0.0460	0.0497	0.2455	0.2807	0.4653	0.5267
$n_i = 10$ $n_p = 10$	0.0448	0.0530	0.3293	0.3558	0.6109	0.6505
$n_i = 10$ $n_p = 20$	0.0508	0.0499	0.4849	0.4729	0.8235	0.8037
$n_i = 20$ $n_p = 10$	0.0531	0.0522	0.4093	0.4681	0.7351	0.7976
$n_i = 30$ $n_p = 10$	0.0469	0.0500	0.4931	0.5467	0.8279	0.8761
$n_i = 20$ $n_p = 5$	0.0473	0.0478	0.3537	0.3788	0.6371	0.6791
$n_i = 10$ $n_p = 30$	0.0514	0.0498	0.5918	0.5624	0.9059	0.8845
$n_i = 10$ $n_p = 40$	0.0502	0.0500	0.6728	0.6486	0.9572	0.9311
$n_i = 5$ $n_p = 20$	0.0496	0.0455	0.4292	0.4048	0.7440	0.7048

When the completely randomized portion is from a t-distribution with 3 degrees of freedom and the paired design portion is from a t-distribution with 10 degrees of freedom, that is, the variance of completely randomized design portion is larger than the variance of paired design portion, the significance levels of both tests are estimated for all sample size combinations considered. These are all around 0.05 (Table 7). We compare powers of the two tests for several different sample size combinations of the independent data and the paired data. Results are similar to the results observed when the underlying distributions were normal or exponential. For example, when the sample size for the independent data is 20 for each treatment and the sample size for the paired data is 10 pairs, the proposed test has higher powers than the Dubnicka test. When the sample size for the independent data is 10 for each treatment and the sample size for the paired data is 20 pairs, the Dubnicka test had higher powers than the proposed test. As the difference between the magnitude of the variances changes, such as when the independent samples follow a t-distribution with 3 degrees of freedom and the paired sample follows a t-distribution with 20 degrees of freedom, there is more separation in powers between the two tests. We again see that the ratio between the estimated powers of the two tests becomes more pronounced as the variance of the completely randomized design portion increases with respect to the variance of the paired design portion.

5. An Example

Dubnicka et al. (2002) give an example as to when this mixed design can be used. This example was mentioned in the introduction and we continue with it here. They discuss a situation in which two surgical treatments are being compared to manage high-risk proliferative diabetic retinopathy. In some cases, diabetic patients had two eyes available for the study, and in other cases, diabetic patients only had one eye available for the study. When diabetic patients had both eyes available for the study, the red krypton treatment was randomly assigned to one of the eyes and the blue-green argon laser photocoagulation treatment was given to the other eye. In cases in which the patient only had one eye available for the study, one of the two treatments was randomly assigned. Therefore, the experiment consisted of paired data and two-independent-sample data. Visual acuity measurements were taken 3 months after treatment. Dubnicka et al. (2002) give a sample of these visual acuity measurements that were taken in Table 2 of their paper. The data given consist of 20 pairs of data and independent samples of size 10 from each of the two treatments. It is noted from the results of our simulation study that the Dubnicka test should have the higher power in this case to detect a difference between the two treatments, if there is a difference. This is because the Wilcoxon signed rank test will have more weight than the Wilcoxon–Mann–Whitney test when using the Dubnicka test with this set of sample sizes.

The value of the Wilcoxon signed rank test is found to be 135 with a mean of $20 \times 21/4$ equal to 105 and a variance of $20 \times 21 \times 41/24$ equal to 717.50. This is based on 20 pairs. The value of the Wilcoxon–Mann–Whitney test statistic is found to be 54 with a mean of $10 \times 10/2$ equal to 50 and a variance of $10 \times 10 \times 21/12$ equal to 175. This is based on two independent samples each of size 10. The standardized value of the Wilcoxon SIGNED RANK test is equal to 1.11998. The standardized value of the Wilcoxon–Mann–Whitney test statistic is 0.302372. The weight given to the standardized Wilcoxon signed rank test in the Dubnicka test statistic is the square root of $717.5/(717.5 + 175)$ equal to 0.897, and the weight given to the standardized Wilcoxon–Mann–Whitney test is the square root of $175/(717.5 + 175)$ equal to 0.443. The value of the Dubnicka test statistic (without the continuity correction factor) is 1.138. If one were testing that the mean visual acuity score is higher for the blue-green laser group, the one-sided p-value for this is given by 0.128. The value of the proposed test statistic is 1.006. The associated one-sided p-value is given by 0.157. Neither of the tests rejected the null hypothesis of no difference between the treatments, but the p-value for the Dubnicka test is lower. This would fit the findings in our simulation study.

6. Conclusions

Under normal distributions, exponential distributions, and t-distributions, if the variance of the completely randomized portion is equal to the variance of the paired portion, the proposed test has approximately the same powers as the Dubnicka test under all the different sample size combinations considered. If the variance of the completely randomized design portion is larger than the variance of the paired design portion, the conclusions are different under different sample size combinations. When the sample sizes for the independent data are equal to or larger than the sample size of the paired data, the proposed test has higher powers than the Dubnicka test. When the sample size for the paired data is two or more times larger than the sample size for the independent samples, the Dubnicka test will have higher powers than the proposed test. The powers of both tests decrease as the independent sample variance becomes larger, but the ratios of the powers of the two tests become more pronounced. This was seen in all of the distributions considered.

The reason that we are getting the conclusions that we are appears to be the weights associated with the standardized Wilcoxon signed rank test and the standardized Wilcoxon–Mann–Whitney test. The Dubnicka et al. (2001) test gives more weight to the standardized Wilcoxon–Mann–Whitney test statistic than the standardized Wilcoxon signed rank test unless the paired sample size is approximately two or more times the independent sample size. The proposed test weights both the standardized Wilcoxon signed rank test and the Wilcoxon–Mann–Whitney test equally regardless of the sample sizes. In cases in which the independent sample variance is larger than the paired variance and when the number of pairs is less than approximately twice the sample size for the independent samples, it is better to use the proposed test since the proposed test has equal weights on both the standardized Wilcoxon signed rank test and the standardized Wilcoxon–Mann–Whitney test. In this case, the Dubnicka et al. (2002) test gives more weight to the standardized Mann–Whitney–Wilcoxon test and the independent samples have a larger variance. When the number of pairs becomes approximately twice as large as the independent sample sizes, the Dubnicka test gives more weight to the Wilcoxon signed rank test with the paired population having a smaller variance, and thus, the Dubnicka test will have a higher power in this case.

References

Alvo, M., and P. Cabilio. 1995. Testing ordered alternatives in the presence of incomplete data. *J. Am. Stat. Assoc.*, 90, 1015–1024.

Daniel, W. W. 1990. *Applied nonparametric statistics*, 2nd ed. Boston, MA: PWS-Kent Publishing Company.

Dubnicka, S. R., R. C. Blair, and T. P. Hettmansperger. 2002. Rank-based procedures for mixed pairs and two-sample designs. *J. Modern Appl. Stat. Methods*, 1(1). 32–41.

Durbin, J. 1951. Incomplete blocks in ranking experiments. *Br. J. Psychol. Stat. Section*, V.4, 85–90.

Friedman, M. 1937. the use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Am. Stat. Assoc.*, 32, 675–701.

Friedman, M. 1940. A comparison of alternative tests of significance for the problem of m Rankings. *Ann. Math. Stat.*, 11, 86–92.

Jonckheere, A. R. 1954. A distribution-free k-sample test against ordered alternatives. *Biometrika*, 41, 133–145.

Kruskal, W. H., and W. A. Wallis. 1953. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.*, 58, 583–621. Addendum, 48, 907–911.

Kim, D. H., and Y. C. Kim. 1992. Distribution-free tests for umbrella alternatives in a randomized block design. *J. Nonparametric Stat.*, 1, 277–285.

Mack, G. A., and D. A. Wolfe. 1981. K-sample rank tests for umbrella alternatives. *J. Am. Stat. Assoc.*, 76, 175–181.

Magel, R., J. Terpstra, and J. Wen. 2009. Proposed tests for the nondecreasing alternative in a mixed design. *J. Stat. Manage. Systems*, 12, 963–977.

Magel, R., J. Terpstra, K. Canonizado, and J. I. Park. 2010. Nonparametric tests for mixed designs. *Commun. Stat. Simulation Comput.*, 39(6),1228–1250.

Magel, R., L. Cao, and A. Ndungu. 2012. Comparing the Durbin, Wilcoxon signed ranks test, and a proposed test in balanced incomplete block designs. *Int. J. Sci. Society*, 3, 1–16.

Mann, H. B., and D. R. Whitney. 1947. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.*, 18, 50–60.

Page, E. B. 1963. Ordered hypotheses for multiple treatments: A significance test for linear ranks. *J. Am. Stat. Assoc.*, 58, 216–230.

Terpstra, T. J. 1952. The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking. *Indagationes Math.*, 14, 327–333.

Wilcoxon, F. 1945. Individual comparisons by ranking methods. *Biometrics*, 1, 80–83.