*The Winnower*

Published March 22, 2015

Updated November 16, 2015

# Spearman’s hypothesis on item-level data from Raven’s Standard Progressive Matrices: A replication and extension

Emil O. W. Kirkegaard, University of Aarhus.

**Abstract**
Item-level data from Raven’s Standard Progressive Matrices was compiled for 12 diverse groups from previously published studies. Jensen’s method (method of correlated vectors) was used on every possible pair of groups with available data (45 comparisons). The mean Jensen coefficient was about . 50. Very few were negative. Spearman’s hypothesis is confirmed for item-level data from the Standard Progressive Matrices, but interpretation is unclear.

**Key words**: Raven's Progressive Matrices, IQ, intelligence, Spearman’s hypothesis, method of correlated vectors, Jensen’s method

## 1. Introduction and method

Jensen’s method (method of correlated vectors; MCV) is a statistical method invented by Arthur Jensen (1). The purpose of it is to measure to which degree a latent variable is responsible for an observed correlation between an aggregate measure and a criteria variable. Jensen had in mind the general factor of cognitive ability (the g factor) as measured by various IQ tests and their subtests, and criteria variables such as brain size. The method, however, is applicable to any latent trait (e.g. general socioeconomic factor (2,3)). When this method is applied to group differences, particularly ethnoracial ones, it is called *Spearman’s hypothesis* (SH) because Spearman was the first to note it in his 1927 book (4).

By now, several large studies and meta-analysis of Jensen's method results for group differences have

been published (5–8). These studies generally support the hypothesis. Almost all studies use subtest loadings instead of item loadings. This is probably because psychologists are reluctant to share their data (9,10) and as a result there are few open item-level datasets available to use for this purpose, whereas subtest results are often but not always reported in papers. Furthermore, before the introduction of modern computers and the internet, it was impractical to share item-level data. There are advantages and disadvantages to using item-level data over subtest-level data. There are more items than subtests which means that the vectors will be longer and thus sampling error will be smaller. On the other hand, items are less reliable and less pure measures which introduces both error and more non-g ability variance.

The recent study by Nijenhuis et al (5) however, employed item-level data from Raven's Standard Progressive Matrices (SPM) and included a diverse set of samples (Libyan, Russian, South African, Roma from Serbia, Moroccan and Spanish). The authors did not use their collected data to its full extent, presumably because they were comparing the groups (semi-)manually. To compare all combinations with a dataset of e.g. 10 groups means that one has to do 45 comparisons.[1] However, this task can easily be overcome with programming skills, and I thus saw a research opportunity.

The authors did not provide the data in the paper despite it being easy to include it in online supplementary tables. They refused to share the data when contacted. However, the data was available from the primary studies they cited in most cases. Thus, I collected the data from their data sources. This resulted in data from 12 samples of which 10 had both difficulty and item-whole correlations data. Table 1 gives an overview of the datasets:

1 The number of possible sample comparisons is given by the formula for picking 2 out of N without order.

$$N_{pairs} = \frac{N_{samples} * (N_{samples} - 1)}{2}$$

| Short name | Race | Selection | N | Year | Ref | |
|---|---|---|---|---|---|---|
| A1 | African | Undergraduates | 173 | 2000 | Rushton and Skuy 2000 (11) | University of the Witwatersrand and the Rand Afrikaans University in Johannesburg, South Africa |
| W1 | European | Undergraduates | 136 | 2000 | Rushton and Skuy 2000 | University of the Witwatersrand and the Rand Afrikaans University in Johannesburg, South Africa |
| W2 | European | Std 7 classes | 1056 | 1992 | Owen 1992 (12) | 20 schools in the Pretoria-Witwatersrand-Vereeniging (PWV) area and 10 schools in the Cape Peninsula |
| C1 | Colored (African European) | Std 7 classes | 778 | 1992 | Owen 1992 | 20 coloured schools in the Cape Peninsula |
| I1 | Indian | Std 7 classes | 1063 | 1992 | Owen 1992 | 30 schools selected at random from the list of high schools in and around Durban |
| A2 | African | Std 7 classes | 1093 | 1992 | Owen 1992 | Three schools in the PWV area and 25 schools in KwaZulu (Natal) |
| A3 | African | First year Engineering students | 198 | 2002 | Rushton et al 2002 (13) | First-year students from the Faculties of Engineering and the Built Environment at the University of the Witwatersrand |
| I2 | Indian | First year Engineering students | 58 | 2002 | Rushton et al 2002 | First-year students from the Faculties of Engineering and the Built Environment at the University of the Witwatersrand |
| W3 | European | First year Engineering students | 86 | 2002 | Rushton et al 2002 | First-year students from the Faculties |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | | of Engineering and the Built Environment at the University of the Witwatersrand |
| R1 | Roma | Adults ages 16 to 66 | 231 | 2004.5 | Rushton et al 2007 (14) | Three communities (i.e., Drenovac, Mirijevo, and Rakovica) in the vicinity of Belgrade |
| W4 | European | Adults ages 18 to 65 | 258 | 2012 | Diaz et al 2012 (15) | Mainly from the city of Valencia |
| NA1 | North African | Adults ages 18 to 50 | 202 | 2012 | Diaz et al 2012 | Casablanca, Marrakech, Meknes and Tangiers |

*Table 1: Description of the included studies.*

## 2. Item-whole correlations and item loadings

The data in the papers did usually not contain the actual factor loadings of the items. Instead, they contained the item-whole correlations. The authors argue that one can use these because of the high correlation of unweighted means with extracted g-factors (often r=.99, e.g. (16)). Some studies did provide both loadings and item-whole correlations, yet the authors did not correlate them to see how good proxies the item-whole correlations are for the loadings. I calculated this for the 4 studies that included both metrics. Results are shown in Table 2.

| **Item-whole r / g-loading** | **W2** | **C1** | **I1** | **A2** |
|---|---|---|---|---|
| **W2** | *0.549* | 0.099 | 0.327 | 0.197 |
| **C1** | 0.695 | *0.9* | 0.843 | 0.92 |
| **I1** | 0.616 | 0.591 | *0.782* | 0.686 |
| **A2** | 0.626 | 0.882 | 0.799 | *0.981* |

*Table 2: Correlations of item-whole correlations (rows) and g-loadings of items (columns).*

Note: Within sample correlations between item-whole correlations and item factor loadings are in the diagonal, marked with italic.

As can be seen, the item-whole correlations were not in all cases great proxies for the actual loadings.

To further test this idea, I calculated the item-whole correlations and the factor loadings (first factor, minimum residuals) in the open Wicherts dataset (N=500ish, Dutch university students, see (9)) tested on Raven's Advanced Progressive Matrices. The correlation was .89. Thus, aside from the odd result in the W2 sample (N=1056), item-whole correlations were a reasonable proxy for the factor loadings.

## 3. Item difficulties across samples

If two groups are tested on the same test and this test measures the same trait in both groups, then even

if the groups have different mean trait levels, the order of difficulty of the items or subtests should be similar. Rushton et al (11,13,14) have examined this in previous studies and found it generally to be the case. Table 3 below shows all the cross-sample correlations of item difficulties.

| Sample | A1 | W1 | W2 | C1 | I1 | A2 | A3 | I2 | W3 | R1 | NA1 | W4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **A1** | 1 | 0.879 | 0.981 | 0.962 | 0.988 | 0.856 | 0.956 | 0.892 | 0.789 | 0.892 | 0.948 | 0.927 |
| **W1** | 0.879 | 1 | 0.926 | 0.792 | 0.874 | 0.652 | 0.961 | 0.973 | 0.945 | 0.695 | 0.918 | 0.946 |
| **W2** | 0.981 | 0.926 | 1 | 0.947 | 0.984 | 0.824 | 0.973 | 0.923 | 0.839 | 0.862 | 0.96 | 0.953 |
| **C1** | 0.962 | 0.792 | 0.947 | 1 | 0.976 | 0.944 | 0.89 | 0.814 | 0.695 | 0.952 | 0.918 | 0.871 |
| **I1** | 0.988 | 0.874 | 0.984 | 0.976 | 1 | 0.88 | 0.951 | 0.884 | 0.788 | 0.91 | 0.951 | 0.924 |
| **A2** | 0.856 | 0.652 | 0.824 | 0.944 | 0.88 | 1 | 0.757 | 0.682 | 0.559 | 0.968 | 0.823 | 0.761 |
| **A3** | 0.956 | 0.961 | 0.973 | 0.89 | 0.951 | 0.757 | 1 | 0.959 | 0.896 | 0.802 | 0.948 | 0.959 |
| **I2** | 0.892 | 0.973 | 0.923 | 0.814 | 0.884 | 0.682 | 0.959 | 1 | 0.924 | 0.722 | 0.913 | 0.92 |
| **W3** | 0.789 | 0.945 | 0.839 | 0.695 | 0.788 | 0.559 | 0.896 | 0.924 | 1 | 0.602 | 0.876 | 0.909 |
| **R1** | 0.892 | 0.695 | 0.862 | 0.952 | 0.91 | 0.968 | 0.802 | 0.722 | 0.602 | 1 | 0.864 | 0.804 |
| **NA1** | 0.948 | 0.918 | 0.96 | 0.918 | 0.951 | 0.823 | 0.948 | 0.913 | 0.876 | 0.864 | 1 | 0.966 |
| **W4** | 0.927 | 0.946 | 0.953 | 0.871 | 0.924 | 0.761 | 0.959 | 0.92 | 0.909 | 0.804 | 0.966 | 1 |

*Table 3: Intercorrelations between item difficulties in 12 samples.*

The unweighted mean intercorrelation is .88. This is quite remarkable given the diversity of the samples.

## 4. Item-whole correlations across samples

Given the above, one might expect similar results for the item-whole correlations. This however is not so. Results are shown in Table 4.

| Sample | A1 | W1 | W2 | C1 | I1 | A2 | A3 | I2 | W3 | R1 |
|---|---|---|---|---|---|---|---|---|---|---|
| **A1** | 1 | -0.196 | 0.588 | 0.578 | 0.729 | 0.539 | 0.267 | 0.037 | -0.302 | 0.567 |
| **W1** | -0.196 | 1 | 0.165 | -0.59 | -0.245 | -0.683 | 0.422 | 0.514 | 0.546 | -0.545 |
| **W2** | 0.588 | 0.165 | 1 | 0.442 | 0.787 | 0.292 | 0.61 | 0.248 | 0.023 | 0.393 |
| **C1** | 0.578 | -0.59 | 0.442 | 1 | 0.786 | 0.942 | 0.008 | -0.249 | -0.488 | 0.779 |
| **I1** | 0.729 | -0.245 | 0.787 | 0.786 | 1 | 0.685 | 0.424 | 0.089 | -0.325 | 0.628 |
| **A2** | 0.539 | -0.683 | 0.292 | 0.942 | 0.685 | 1 | -0.133 | -0.301 | -0.52 | 0.774 |
| **A3** | 0.267 | 0.422 | 0.61 | 0.008 | 0.424 | -0.133 | 1 | 0.262 | 0.372 | 0.02 |
| **I2** | 0.037 | 0.514 | 0.248 | -0.249 | 0.089 | -0.301 | 0.262 | 1 | 0.338 | -0.207 |
| **W3** | -0.302 | 0.546 | 0.023 | -0.488 | -0.325 | -0.52 | 0.372 | 0.338 | 1 | -0.488 |
| **R1** | 0.567 | -0.545 | 0.393 | 0.779 | 0.628 | 0.774 | 0.02 | -0.207 | -0.488 | 1 |

*Table 4: Intercorrelations between item-whole correlations in 10 samples.*

Note: The last two samples, NA1 and W4, did not have item-whole correlation data.

The reason for this state of affairs is that the factor loadings/item-whole correlations change when the group mean trait level changes. For many samples, most of the items were too easy (passing rates at or very close to 100%). When there is no variation in a variable, one cannot calculate a correlation to some other variable. This means that for a number of items for multiple samples, there was missing data for the items. Furthermore, in general, when pass rates move away from .50, the variance in the items are reduced which then reduces their correlation to the total (or their g-loading when analyzed with classical test theory approaches). Thus, observed item-wholes/g-loadings are a function of the pass rate as well as their actual g-loading.

The lack of cross-sample consistency in item-whole correlations may also explain the weak Jensen coefficients in Diaz et al (15) since they used g-loadings from another study instead of from their own samples.

## 5. Spearman’s hypothesis using one static vector of estimated factor loadings

Some of the samples had rather small sample sizes (I2, N=58, W3, N=86). Thus one might get the idea to use the item-whole correlations from one or more of the large samples for comparisons involving other groups. In fact, given the instability of item-whole correlations across sample as can be seen in Table 4, this is a bad idea. However, for sake of completeness, I calculated the results based on this method nonetheless. As the best estimate of factor loadings, I averaged the item-whole correlations data from the four largest samples (W2, C1, I1 and A2).

Using this vector of item-whole correlations, I calculated Jensen's coefficient for every possible sample comparison. Because there were 12 samples with pass rate data, this number is 66. Jensen's method was applied by subtracting the lower scoring sample’s item difficulties from the higher scoring sample’s thus producing a vector of the sample pair difference on each item. I correlated this vector with the

vector of item-whole correlations. The results are shown in Table 5.

| Sample | A1 | W1 | W2 | C1 | I1 | A2 | A3 | I2 | W3 | R1 | NA1 | W4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **A1** | | -0.149 | 0.197 | 0.422 | 0.097 | 0.83 | -0.026 | -0.116 | -0.259 | 0.797 | -0.348 | -0.316 |
| **W1** | -0.149 | | -0.307 | 0.066 | -0.139 | 0.469 | -0.291 | -0.195 | -0.401 | 0.399 | 0.307 | 0.057 |
| **W2** | 0.197 | -0.307 | | 0.561 | 0.4 | 0.865 | -0.271 | -0.28 | -0.377 | 0.83 | -0.355 | -0.456 |
| **C1** | 0.422 | 0.066 | 0.561 | | 0.534 | 0.879 | 0.227 | 0.11 | -0.058 | 0.641 | -0.233 | -0.053 |
| **I1** | 0.097 | -0.139 | 0.4 | 0.534 | | 0.884 | -0.021 | -0.104 | -0.243 | 0.826 | -0.448 | -0.287 |
| **A2** | 0.83 | 0.469 | 0.865 | 0.879 | 0.884 | | 0.655 | 0.522 | 0.318 | 0.199 | 0.421 | 0.404 |
| **A3** | -0.026 | -0.291 | -0.271 | 0.227 | -0.021 | 0.655 | | -0.173 | -0.407 | 0.613 | 0.434 | -0.425 |
| **I2** | -0.116 | -0.195 | -0.28 | 0.11 | -0.104 | 0.522 | -0.173 | | -0.372 | 0.456 | 0.331 | -0.245 |
| **W3** | -0.259 | -0.401 | -0.377 | -0.058 | -0.243 | 0.318 | -0.407 | -0.372 | | 0.233 | -0.053 | -0.112 |
| **R1** | 0.797 | 0.399 | 0.83 | 0.641 | 0.826 | 0.199 | 0.613 | 0.456 | 0.233 | | 0.357 | 0.316 |
| **NA1** | -0.348 | 0.307 | -0.355 | -0.233 | -0.448 | 0.421 | 0.434 | 0.331 | -0.053 | 0.357 | | -0.026 |
| **W4** | -0.316 | 0.057 | -0.456 | -0.053 | -0.287 | 0.404 | -0.425 | -0.245 | -0.112 | 0.316 | -0.026 | |

*Table 5: Jensen's coefficients of group differences across 12 samples using 1 static item-whole correlations.*

As one can see, the results are all over the place. The mean coefficient is .12.

## 6. Spearman’s hypothesis using a variable vector of estimated factor loadings

Since item-whole correlations varied from sample to sample, another idea is to use the samples’ item-whole correlations. I used the unweighted mean of the item-whole correlations for each item (te Nijenhuis et al (5) used a weighted mean). In some cases, only one sample has item-whole correlations for some items (because the other sample had no variance on the item, i.e. 100% or 0% got it right). In these cases, one can choose to use the value from the remaining sample, or one can ignore the item and use Jensen's method on the remaining items. Not knowing which was the best method, I calculated results using both methods, they are shown in Table 6 and 7.

| Sample | A1 | W1 | W2 | C1 | I1 | A2 | A3 | I2 | W3 | R1 |
|---|---|---|---|---|---|---|---|---|---|---|
| **A1** | | 0.787 | 0.388 | 0.292 | 0.046 | 0.7 | 0.484 | 0.409 | 0.368 | 0.709 |
| **W1** | 0.787 | | 0.799 | 0.499 | 0.756 | 0.515 | 0.786 | 0.405 | 0.604 | 0.545 |
| **W2** | 0.388 | 0.799 | | 0.675 | 0.631 | 0.852 | 0.429 | 0.467 | 0.504 | 0.791 |
| **C1** | 0.292 | 0.499 | 0.675 | | 0.517 | 0.88 | 0.324 | 0.3 | -0.095 | 0.673 |
| **I1** | 0.046 | 0.756 | 0.631 | 0.517 | | 0.842 | 0.473 | 0.397 | 0.31 | 0.789 |
| **A2** | 0.7 | 0.515 | 0.852 | 0.88 | 0.842 | | 0.571 | 0.428 | -0.034 | 0.217 |
| **A3** | 0.484 | 0.786 | 0.429 | 0.324 | 0.473 | 0.571 | | 0.379 | 0.661 | 0.64 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **I2** | 0.409 | 0.405 | 0.467 | 0.3 | 0.397 | 0.428 | 0.379 | | 0.603 | 0.44 |
| **W3** | 0.368 | 0.604 | 0.504 | -0.095 | 0.31 | -0.034 | 0.661 | 0.603 | | 0.201 |
| **R1** | 0.709 | 0.545 | 0.791 | 0.673 | 0.789 | 0.217 | 0.64 | 0.44 | 0.201 | |

*Table 6: Jensen's coefficients of group differences across 10 samples using variable item-whole correlations, method 1.*

| **Sample** | **A1** | **W1** | **W2** | **C1** | **I1** | **A2** | **A3** | **I2** | **W3** | **R1** |
|---|---|---|---|---|---|---|---|---|---|---|
| **A1** | | 0.421 | 0.397 | 0.326 | 0.056 | 0.717 | 0.483 | 0.3 | 0.147 | 0.739 |
| **W1** | 0.421 | | 0.72 | 0.14 | 0.523 | 0.178 | 0.703 | 0.443 | 0.65 | 0.354 |
| **W2** | 0.397 | 0.72 | | 0.675 | 0.631 | 0.852 | 0.443 | 0.533 | 0.502 | 0.791 |
| **C1** | 0.326 | 0.14 | 0.675 | | 0.517 | 0.88 | 0.385 | 0.191 | -0.082 | 0.673 |
| **I1** | 0.056 | 0.523 | 0.631 | 0.517 | | 0.842 | 0.507 | 0.403 | 0.279 | 0.789 |
| **A2** | 0.717 | 0.178 | 0.852 | 0.88 | 0.842 | | 0.618 | 0.303 | 0.018 | 0.217 |
| **A3** | 0.483 | 0.703 | 0.443 | 0.385 | 0.507 | 0.618 | | 0.418 | 0.554 | 0.667 |
| **I2** | 0.3 | 0.443 | 0.533 | 0.191 | 0.403 | 0.303 | 0.418 | | 0.579 | 0.348 |
| **W3** | 0.147 | 0.65 | 0.502 | -0.082 | 0.279 | 0.018 | 0.554 | 0.579 | | 0.078 |
| **R1** | 0.739 | 0.354 | 0.791 | 0.673 | 0.789 | 0.217 | 0.667 | 0.348 | 0.078 | |

*Table 7: Jensen's coefficients of group differences across 10 samples using variable item-whole correlations, method 2.*

Nearly all results are positive using either method. The results are slightly stronger when ignoring items where both samples do not have item-whole correlation data. A better way to visualize the results is to use a histogram with an empirical density curve, as shown in Figure 1 and 2.

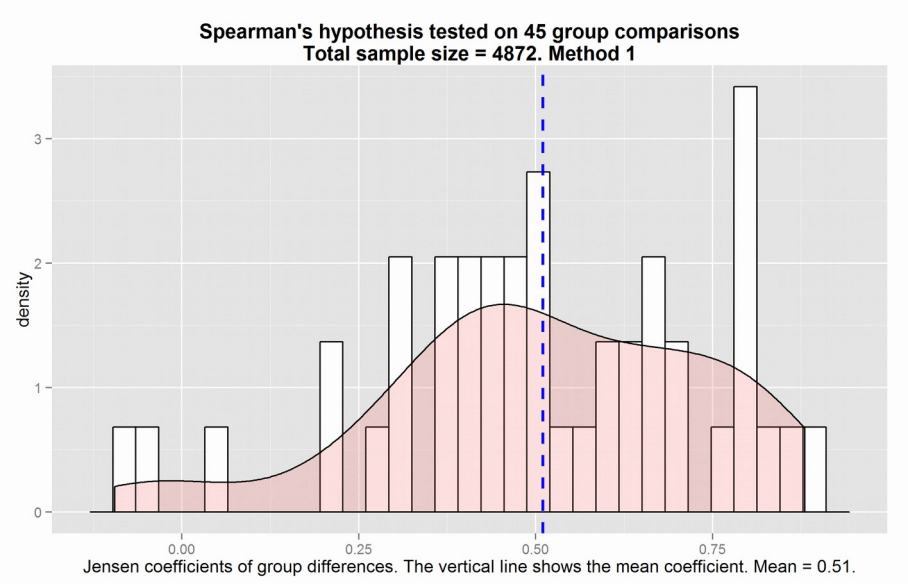



*Figure 1: Histogram of Jensen's method coefficients using method 1.*

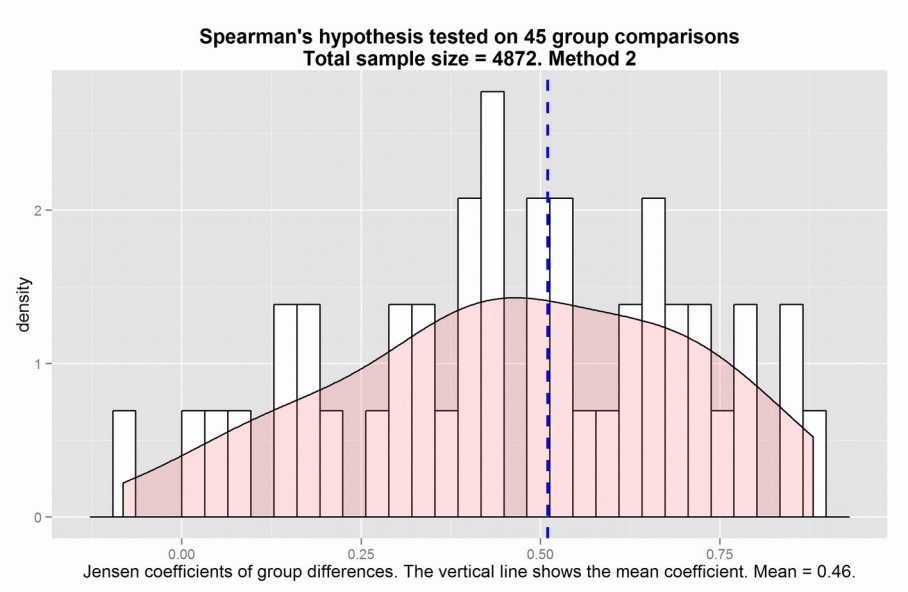



*Figure 2: Histogram of Jensen's method coefficients using method 2.*

The mean result for method 1/2 was .51/.46. Almost all coefficients were positive, only 2/1 was negative for method 1/2.

For the remainder of the paper, only results using method 1 are used.

## 7. Mean Jensen's coefficient value by sample and moderator analysis

It is interesting to examine the mean Jensen coefficient by sample. They are shown in Table 8.

| Sample | mean | sd | median |
|---|---|---|---|
| **A1** | 0.465 | 0.235 | 0.409 |
| **W1** | 0.633 | 0.151 | 0.604 |
| **W2** | 0.615 | 0.175 | 0.631 |
| **C1** | 0.452 | 0.286 | 0.499 |
| **I1** | 0.529 | 0.257 | 0.517 |
| **A2** | 0.552 | 0.312 | 0.571 |
| **A3** | 0.527 | 0.149 | 0.484 |
| **I2** | 0.425 | 0.081 | 0.409 |
| **W3** | 0.347 | 0.278 | 0.368 |
| **R1** | 0.556 | 0.226 | 0.64 |

*Table 8: Mean Jensen's coefficient by sample.*

There is no obvious racial pattern. Instead, one might expect the relatively lower result of some samples to be due to sampling error. Jensen's coefficient is extra sensitive to sampling error. If so, the mean coefficient should be higher for the larger samples. To see if this was the case, I calculated the rank-order correlation between sample size and sample mean coefficient, r=.45. Rank-order was used

because the effect of sample size on sampling error is non-linear. Figure 3 shows the scatter plot of this.




*Figure 3: Sample size as a moderating variable at the sample mean-level.*

One can also examine sample size as a moderating variable as the comparison-level. This increases the number of datapoints to 45. I used the harmonic mean[2] of the 2 samples as the sample size metric. Figure 4 shows a scatter plot of this.

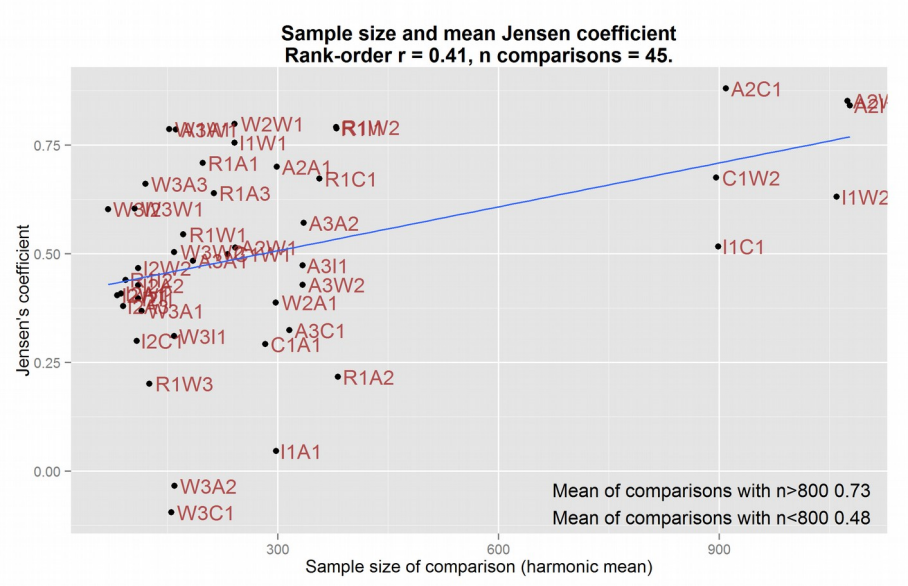



*Figure 4: Sample size as a moderator variable at the comparison-level.*

We can see in the plot that the results from the 6 largest comparisons (harmonic mean sample size>800) have a mean of .73. For the smaller studies (harmonic mean sample size<800), the results have a mean of .48. The results from the smaller studies vary more, as expected with their higher sampling error, and they are on average weaker, as expected from the increased sampling error in the item-whole

2 The harmonic mean is given by the number of numbers divided by the sum of reciprocal values of each number.

$$mean_{harmonic} = \frac{n}{\sum_{i=1}^{n} x_i^{-1}}$$

correlation and pass rate difference vectors.

I also examined the group difference size as a moderator variable. I computed this as the difference between the mean item difficulty by the groups. However, it had a near-zero relationship to the coefficients (rank-order r=.03).

## 8. Discussion and conclusion

Spearman's hypothesis has been decisively confirmed using classical test theory item-level data from Raven's Standard Progressive Matrices. The analysis presented here can easily be extended to cover more datasets, as well as item-level data from other IQ tests. Researchers should compile such data into open datasets so they can be used for future studies.

It is interesting to note the consistency of results within and across samples that differ in race. Race differences in general intelligence as measured by the SPM appear to be just like those within races.

Still, the interpretation of the results is contested (5,17,18).

### 8.1. Limitations

- Some authors refused to share data when contacted. They usually required co-authorship to be willing to share their data. For this reason, the meta-analysis was unnecessarily limited.
- Item-level g-loadings were not available for most studies, so instead item-whole correlations were used as a proxy. These seem to be fairly useful as proxies, but this has not been extensively tested.

### Supplementary material and acknowledgments

R source code and datasets can be found at the Open Science Framework repository https://osf.io/ef6vb/

### References


1. Jensen AR. The g factor: the science of mental ability. Westport, Conn.: Praeger; 1998.

2. Kirkegaard EOW. The international general socioeconomic factor: Factor analyzing international rankings. Open Differ Psychol [Internet]. 2014 Sep 8 [cited 2014 Oct 13]; Available from: http://openpsych.net/ODP/2014/09/the-international-general-socioeconomic-factor-factor-analyzing-international-rankings/

3. Kirkegaard EOW. Crime, income, educational attainment and employment among immigrant groups in Norway and Finland. Open Differ Psychol [Internet]. 2014 Oct 9 [cited 2014 Oct 13]; Available from: http://openpsych.net/ODP/2014/10/crime-income-educational-attainment-and-employment-among-immigrant-groups-in-norway-and-finland/

4. Spearman C. The abilities of man. 1927 [cited 2015 Nov 15]; Available from: http://doi.apa.org/psycinfo/1927-01860-000

5. te Nijenhuis J, Al-Shahomee AA, van den Hoek M, Grigoriev A, Repko J. Spearman's hypothesis

tested comparing Libyan adults with various other groups of adults on the items of the Standard Progressive Matrices. Intelligence. 2015 May;50:114–7.

6. te Nijenhuis J, David H, Metzen D, Armstrong EL. Spearman's hypothesis tested on European Jews vs non-Jewish Whites and vs Oriental Jews: Two meta-analyses. Intelligence. 2014 May;44:15–8.

7. te Nijenhuis J, van den Hoek M, Armstrong EL. Spearman's hypothesis and Amerindians: A meta-analysis. Intelligence. 2015 May;50:87–92.

8. Jensen AR. The nature of the black–white difference on various psychometric tests: Spearman's hypothesis. Behav Brain Sci. 1985 Jul;8(02):193.

9. Wicherts JM, Bakker M. Publish (your data) or (let the data) perish! Why not publish your data too? Intelligence. 2012 Mar;40(2):73–6.

10. Wicherts JM, Borsboom D, Kats J, Molenaar D. The poor availability of psychological research data for reanalysis. Am Psychol. 2006;61(7):726–8.

11. Rushton JP, Skuy M. Performance on Raven's Matrices by African and White University Students in South Africa. Intelligence. 2000;28(4):251–65.

12. Owen K. The suitability of Raven's Standard Progressive Matrices for various groups in South Africa. Personal Individ Differ. 1992;13(2):149–59.

13. Rushton JP, Skuy M, Fridjhon P. Jensen Effects among African, Indian, and White engineering students in South Africa on Raven's Standard Progressive Matrices. Intelligence. 2002 Sep;30(5):409–23.

14. Rushton JP, Čvorović J, Bons TA. General mental ability in South Asians: Data from three Roma (Gypsy) communities in Serbia. Intelligence. 2007 Jan;35(1):1–12.

15. Diaz A, Sellami K, Infanzón E, Lanzón T, Lynn R. A comparative study of general intelligence in Spanish and Moroccan samples. Span J Psychol. 2012 Jul;15(2):526–32.

16. Kirkegaard EOW, Nordbjerg O. Validating a Danish translation of the International Cognitive Ability Resource sample test and Cognitive Reflection Test in a student sample. Open Differ Psychol [Internet]. 2015 Jul 31 [cited 2015 Aug 6]; Available from: http://openpsych.net/ODP/2015/07/validating-a-danish-translation-of-the-international-cognitive-ability-resource-sample-test-and-cognitive-reflection-test-in-a-student-sample/

17. Raven J. Testing the SpearmanJensen Hypothesis Using the Items of the RPM [Internet]. 2010 [cited 2015 Nov 16]. Available from: http://www.eyeonsociety.co.uk/resources/testingSJHyp.pdf

18. Repko J. Spearman's hypothesis tested with Raven's Progressive Matrices: A psychometric meta-analysis. 2011 [cited 2015 Nov 16]; Available from: http://dare.uva.nl/cgi/arno/show.cgi?fid=347520