Research article.
Published in *Open Differential Psychology* 8th of September, 2014
Submitted 20th of July, 2014

The international general socioeconomic factor: Factor analyzing international rankings

Emil O. W. Kirkegaard[1]

Abstract
Many studies have examined the correlations between national IQs and various country-level indexes of well-being. The analyses have been unsystematic and not gathered in one single analysis or dataset. In this paper I gather a large sample of country-level indexes and show that there is a strong general socioeconomic factor (S factor) which is highly correlated (.86-.87) with national cognitive ability using either Lynn and Vanhanen's dataset or Altinok's. Furthermore, the method of correlated vectors shows that the correlations between variable loadings on the S factor and cognitive measurements are .99 in both datasets using both cognitive measurements, indicating that it is the S factor that drives the relationship with national cognitive measurements, not the remaining variance.

Keywords: National IQ, social progress index, democracy ranking, intelligence, g-factor, group differences, general socioeconomic factor, method of correlated vectors, psychoinformatics

1 Introduction

Recently, research has been done on general socioeconomic factors. Gregory Clark argues that there is a general socioeconomic factor which underlies performance at the individual level.[1] Moving one step up, earlier I found evidence that among 71 Danish immigrant groups ranked on 4 different measures of socioeconomic variables (crime, use of social benefits, income, education attainment), there is a large (40% of variance explained) general socioeconomic factor.[2] Given that groups can generally be considered collections of individuals, this leads to the expectation that there may be a general socioeconomic factor at the country level as well. The general mental ability factor at the individual level has been termed "g" (often italicized "*g*"[3]), while the national-level group equivalent has been termed "G" ("big g factor").[4, 5] Keeping in line with this terminology, one might refer to the general socioeconomic factor at the individual level as "s factor" and the group level version "S factor" (or "big s").

There are by now many different national measures of country well-being. Perhaps the most common is Human Development Index published by the United Nations, but there are plenty of others e.g. Social Progress Index, Democracy Ranking, Quality of Life Index, Where-to-be-born Index, Democracy Index. Previous studies have correlated some of these with national IQs, but not in a systematic manner (see review in [6]). Most of these national indexes have subcomponents that can be analyzed as well to see whether there is common variance, or whether the presence of these variables on the index is merely a function of their desirability as judged by the authors. The index scores are usually not outputs from a factor analysis, but usually some weighted average of the subcomponents based on how important the authors thought they were.

The goals of this study were thus: 1) gather a large collection of country well-being indexes

[1]Corresponding author: emil@emilkirkegaard.dk

and their subcomponents, 2) perform factor analyses[2] and other tests to check whether there is an S factor in the data and how reliably it can be measured, and 3) examine how national cognitive measurements relate with this.

2 Building the dataset

After I began the task of merging all the datasets into one, the main reason why others had not done it before became apparent: it is a lot of work to combine datasets manually. There are two reasons. Firstly, the different datasets do not contain exactly the same countries, so one has to rearrange them so the two lists match up. Secondly, often the same country is spelled different ways giving problems with alphabetic sorting of the countries. For instance, South Korea is called "South Korea", "Korea, Republic of", "S. Korea", "Korea Rep.", "Korea South" and variations in other languages as well.

To avoid this intensive manual labor, I wrote a Python script[3] that can combine two datasets into one. I then used this script to combine datasets for, among others, GDP per capita (from the International Monetary Fund, 2014)[7], national IQs from Lynn and Vanhanen (2012)[6], Human Development Index, Social Progress Index, Democracy Ranking, Quality-of-Life Index, Where-to-be-born index, Democracy Index and over 250 other variables concerning countries. The entire dataset as well as the Python script is available at http://emilkirkegaard.dk/megadataset.

After inspecting the dataset, I decided to limit my analysis to the subcomponents of the Social Progress Index and Democracy Ranking. This is because both of these are very comprehensive, each having more than 40 subcomponents that cover a wide spectrum of living. Furthermore, these big composites usually included many of the more narrow national rankings, so using both would result in double-sampling some measurements. Furthermore, using two precompiled datasets allows me to avoid the objection that I put the datasets together in such a way as to bias the results towards the hypotheses I am testing.

2.1 Social progress Index (SPI) and Democracy Ranking (DR)

The SPI is very comprehensive and is a function of 54 basic subcomponents.[8] The structure is complicated and best shown visually, see Figure 1. A 56 page description of the SPI is found in the methodological report.[9]

The DR also has a large structure. It has 6 overall dimensions: political system, economy, environment, gender equality, health and knowledge. These have a total of 42 subcomponents.[10]

The DR sample of subcomponents overlaps heavily with that of SPI. As with the SPI, there is a 50 page methodological report that explains in further detail how the index works.[11] The main differences between SPI and DR is that DR includes economic variables while the SPI only includes social and political. The SPI authors explicitly state they want to "move beyond GDP" (p. 8). However, the correlation between GDP per capita (IMF, 2013) and SPI is .860. The reader can judge for themselves how well they succeeded.

2.2 Should one combine the SPI and DR?

One could combine the subcomponents from SPI and DR into one dataset by removing the duplicate variables. I did not do this for three reasons. Firstly, it is useful to have two separate datasets to test hypotheses on. If an hypothesis by chance happens to be confirmed in the first dataset, it is unlikely that it will also be confirmed in the second. Secondly, the indexes are already extremely comprehensive. Thirdly, it would reduce the sample size and result in a lower case-variable ratio. The case-variable ratio is already small: 132/54=2.44 in the SPI and 115/42=2.74 in the DR. Small

[2]For the purposes of this article, principal components analysis is considered a factor analytic method, even though some researchers argue that PCA should not be considered as such.

[3]Python is a free and versatile yet easy to use general purpose programming language. Learn more at https://www.python.org/. I used the Anaconda package. https://store.continuum.io/cshop/anaconda/

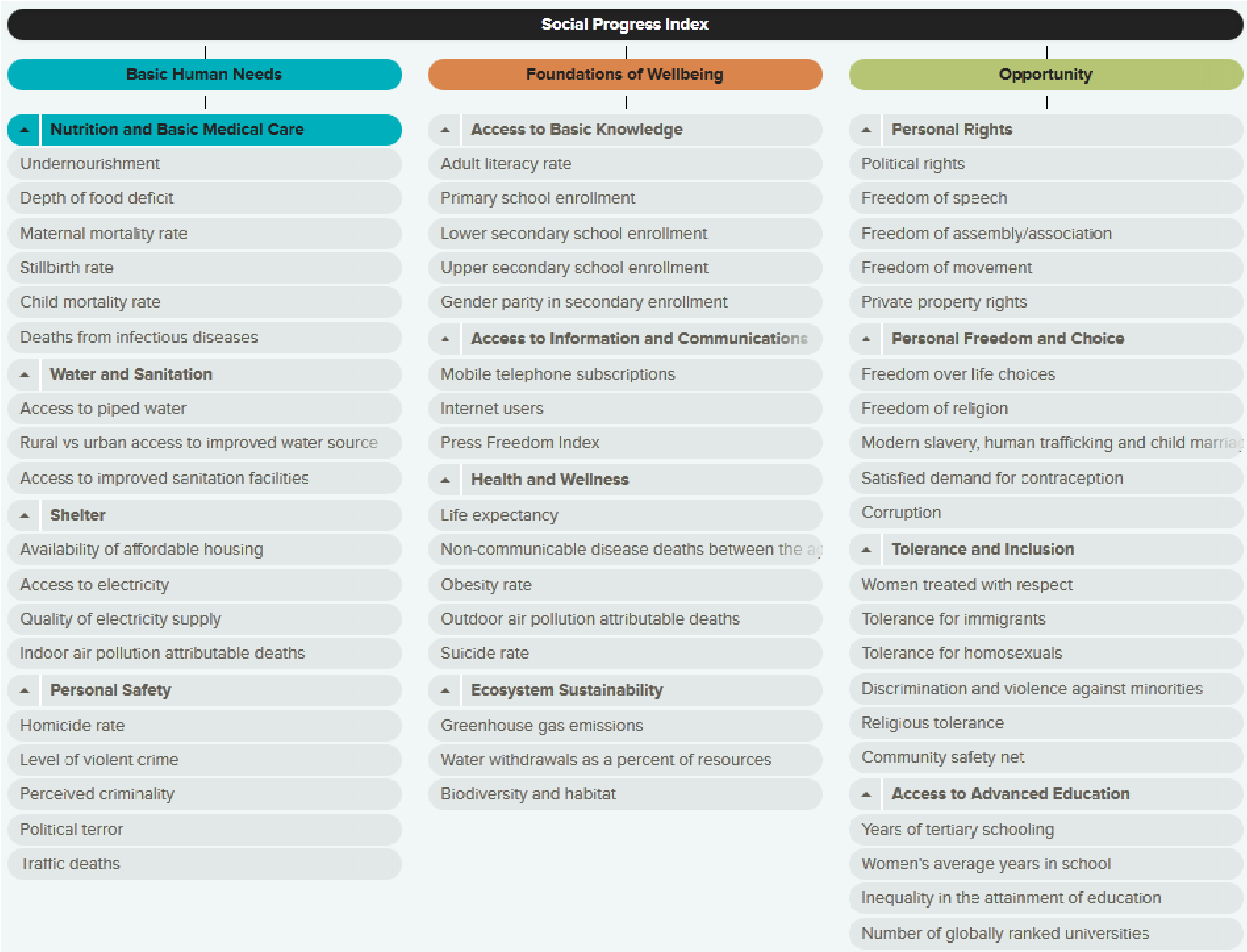

Figure 1: The structure of the Social Progress Index. Source: http://socialprogressimperative.org/

case-variable ratios can make factor analytic results unstable. Nathan Zhao has compiled an excellent overview of recommendations and practices concerning the case-variable ratio.[12] He mentions that in 26% of a sample of studies using principal components analysis (PCA) in PsychINFO, the case-variable ratio was between 2 and 5 as it is with my two datasets. If the datasets were combined, the ratio would be much lower.

3 Initial analyses

I used R to do the analyses.[4]

Since I was primarily interested in the first factor, all analyses extracted only the first factor unless otherwise noted.

The SPI dataset had missing data for some cases, between 1 and 5 missing values. I decided to examine the effect of using the reduced sample with complete data vs. imputing the means in the cases with missing values. The subsample with complete data has N=78, while the full sample has N=132. In other words, using only the reduced sample implies a sample size decrease of 41%.

I performed PCA on both the reduced and the means-imputed datasets to examine the effect of the procedure. The Pearson correlation of factor loadings is 0.996, indicating that the procedure does not alter the structure of the data much. Similarly, the congruence coefficient is 1.0. I performed KMO tests (a measure of sampling adequacy) on both samples which showed that reducing the

[4]R is a free, powerful, and yet easy to use programming language designed for data mining and statistics. See http://www.r-project.org/

sample reduces the KMO (0.899 to 0.809). In comparison, KMO in the DR dataset is 0.884. All values are considered 'meritorious'.[13, p. 225] All further analysis uses the means-imputed dataset.

Bartlett's test (tests whether the data is suitable for factor analysis) are extremely significant in all three datasets ($p<0.00001$).

3.1 First factors from different methods

Sometimes factor methods give different results. For this reason, I compared the first factors extracted using PCA as well as minimum residuals (MinRes/MR), weighted least squares (WLS), generalized least squares (GLS), principal axis factoring (PAF), minimized chi square (MinChi/MC) and maximum likelihood estimation (ML). Tables 1, 2 and 3, show factor score intercorrelations, factor loadings intercorrelation and factor congruence coefficients, respectively.

Method	**PC1**	**MR1**	**WLS1**	**GLS1**	**PA1**	**ML1**	**MC1**
PC1		0.964	0.999	0.999	0.997	0.995	0.964
MR1	0.983		0.968	0.968	0.968	0.978	1
WLS1	0.999	0.985		1	0.998	0.996	0.968
GLS1	0.999	0.985	1		0.998	0.996	0.968
PA1	0.995	0.982	0.997	0.997		0.997	0.968
ML1	0.986	0.995	0.987	0.987	0.986		0.978
MC1	0.983	1	0.985	0.985	0.982	0.995	

Table 1: Factor score intercorrelations. SPI below the diagonal, DR above.

Method	**PC1**	**MR1**	**WLS1**	**GLS1**	**PA1**	**ML1**	**MC1**
PC1		0.993	1	1	1	0.999	0.993
MR1	0.997		0.993	0.993	0.994	0.996	1
WLS1	1	0.997		1	1	0.999	0.993
GLS1	1	0.997	1		1	0.999	0.993
PA1	1	0.997	1	1		0.999	0.994
ML1	0.996	1	0.997	0.997	0.997		0.996
MC1	0.997	1	0.997	0.997	0.997	1	

Table 2: Factor loadings intercorrelations. SPI below the diagonal, DR above.

Method	**PC1**	**MR1**	**WLS1**	**GLS1**	**PA1**	**ML1**	**MC1**
PC1		0.997	1	1	1	1	0.997
MR1	0.997		0.997	0.997	0.997	0.998	1
WLS1	1	0.997		1	1	1	0.997
GLS1	1	0.997	1		1	1	0.997
PA1	1	0.997	1	1		1	0.997
ML1	0.997	1	0.997	0.997	0.997		0.998
MC1	0.997	1	0.997	0.997	0.997	1	

Table 3: Factor congruence coefficients. SPI below the diagonal, DR above.

As can be seen, these are extremely similar across method and dataset.

3.2 Size of the first factor x factor method

Jensen and Weng[14] note that PCA tends to overestimate the variance accounted for by the first factor. Table 4 shows the size of the first factor by factor method. There is not much evidence that PCA overestimates the size.

Dataset/method	PCA	MR	WLS	GLS	PA	ML	MC
SPI	0.42	0.41	0.42	0.42	0.41	0.4	0.41
DR	0.47	0.44	0.47	0.47	0.46	0.46	0.44

Table 4: Variance accounted for by the first factor by factor method.

3.3 Mean loadings x factor method

Major[15] report that PCA tends to overestimate factor loadings. Table 5 shows the mean absolute (because some were negative) loading using each method. There is not much evidence that PCA overestimates loadings.

Dataset/method	PCA	MR	WLS	GLS	PA	ML	MC
SPI	0.603	0.582	0.600	0.600	0.595	0.580	0.582
DR	0.653	0.624	0.650	0.650	0.644	0.641	0.624

Table 5: Mean absolute loading by factor method.

Generally, factor method makes little difference for the first factor in the full analyses.

4 Subset x whole comparisons: How many variables are necessary for reliable measurement?

Since I find that, regardless of method and dataset used, the first factor is a large factor accounting for about 40-47% of the variance, it would be interesting to know how many variables one needs to measure it well. To find out, I sampled subsets of variables at random from the datasets, extracted the first factor, and then correlated the scores from it with the scores of the first factor using all the variables. I repeated the sampling 1000 times to reduce sampling error to almost zero. I used the 7 different methods mentioned earlier.

Results are shown in Figures 2 and 3. Each step shows the change in correlation from adding another random variable. They show that the number of variables necessary to reliably estimate the first factor from all variables well is small. For PCA, picking one variable at random gives 0.603 which is the same as the average (absolute) loading on the first factor in the full analysis. Using the first factor of 5 variables yields an average correlation of 0.877.

Since PCA is usually singled out for criticism, I compared PCA with the average of all the other methods. As one can see, PCA is a bit higher for the small samples. However, even the largest difference is quite small. The difference in the SPI dataset with 3 variables is 0.042.

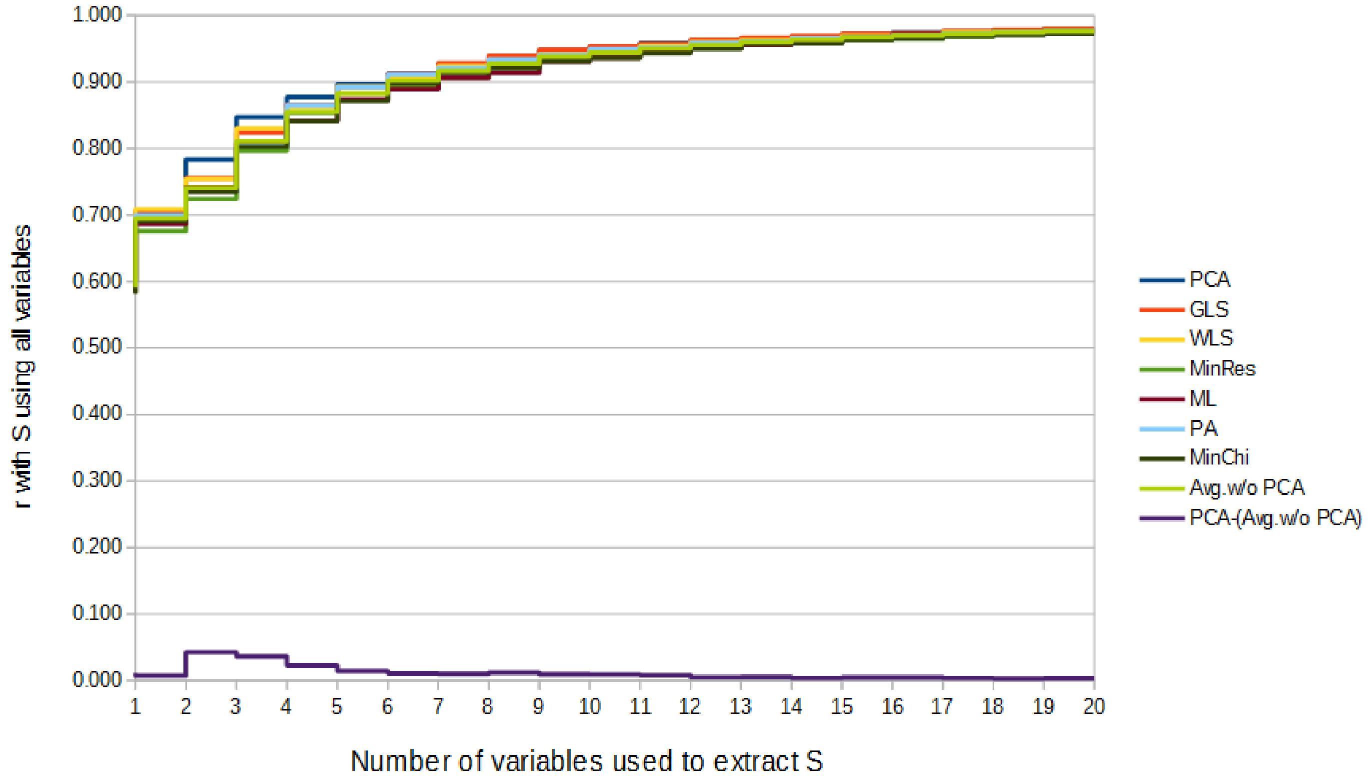

Figure 2: Subset factor correlation with S factor in SPI dataset. The plot shows the average correlation of the first factor within a sample of X variables with the first factor of all 54 variables.

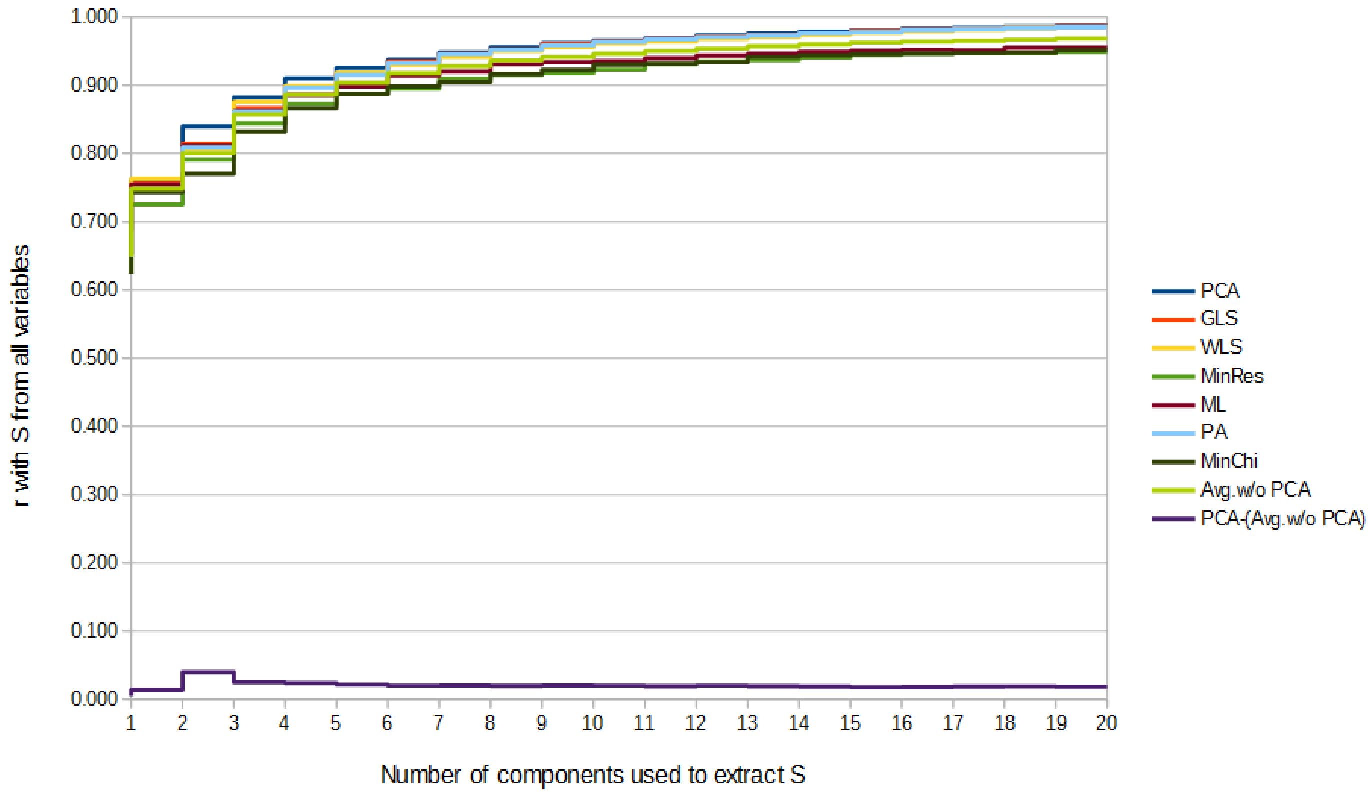

Figure 3: Subset factor correlation with S factor in DR dataset. The plot shows the average correlation of the first factor within a sample of X variables with the first factor of all 54 variables.

5 Subset x subset comparisons

How well does the first factor from a random subset correlate with that of another random subset, with no overlap in variables? I tested this in both datasets with subset sizes 5 and 10 with PCA. Results are shown in Table 6.

N subsets / Dataset	SPI	DR
N=5	0.758	0.804
N=10	0.874	0.902

Table 6: First factors from different and random subsets of the same dataset.

Results are quite high. They show that as long as one picks a reasonable number of them, e.g. 10., it is not so important which subset of national indexes one chooses since they measure to a large degree a common factor.

6 Factor loadings with different methods x number of variables

I found before that PCA did not yield higher loadings than other methods. However, that conclusion was based on analyzing the entire datasets with a large number of variables. There is evidence that PCA yields higher loadings when the number of variables is small.[16, 17] To see if this true for my international data, I extracted the mean absolute loading from 1000 analyses of each sample size 2 to 20 with each of 7 different methods. Results are shown in Figures 4 and 5.

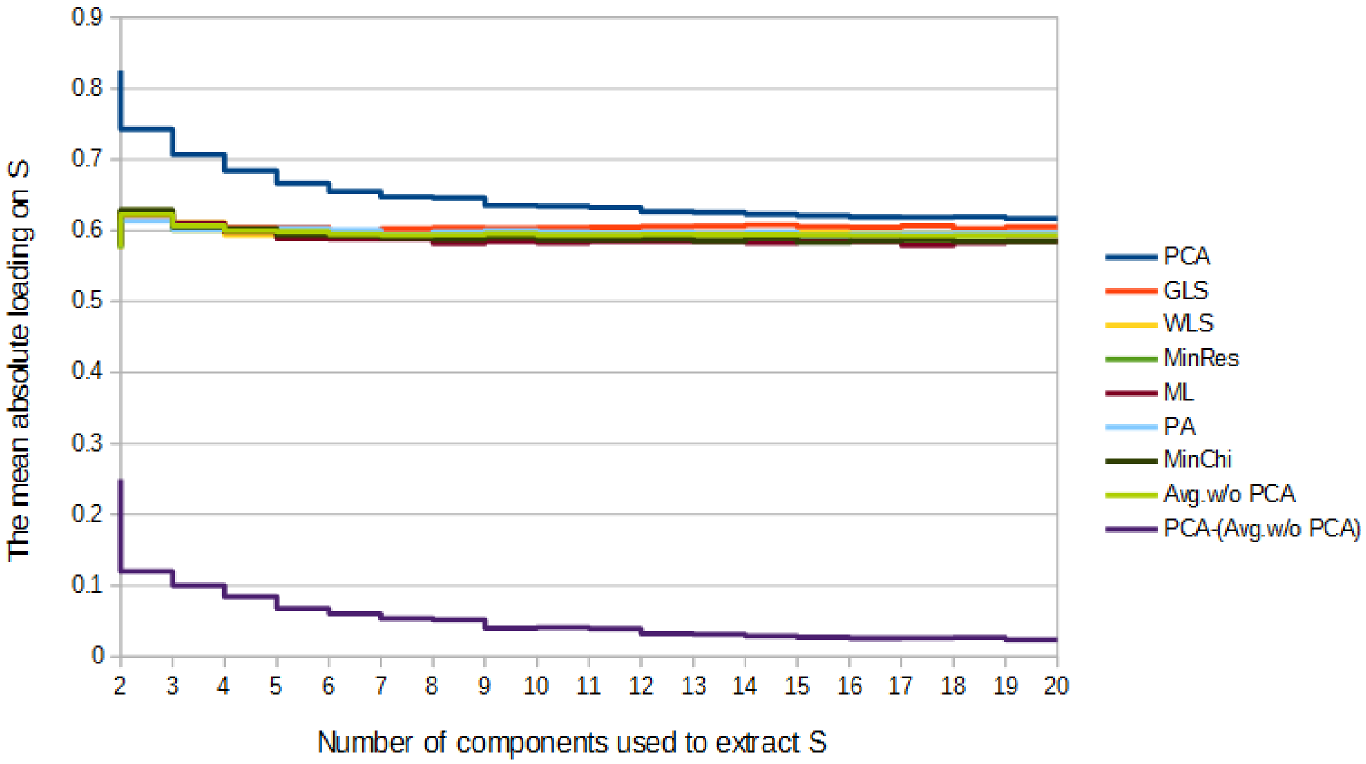

Figure 4: Mean absolute loadings on the first factor with different methods and different sample sizes of variables. Results for SPI dataset.

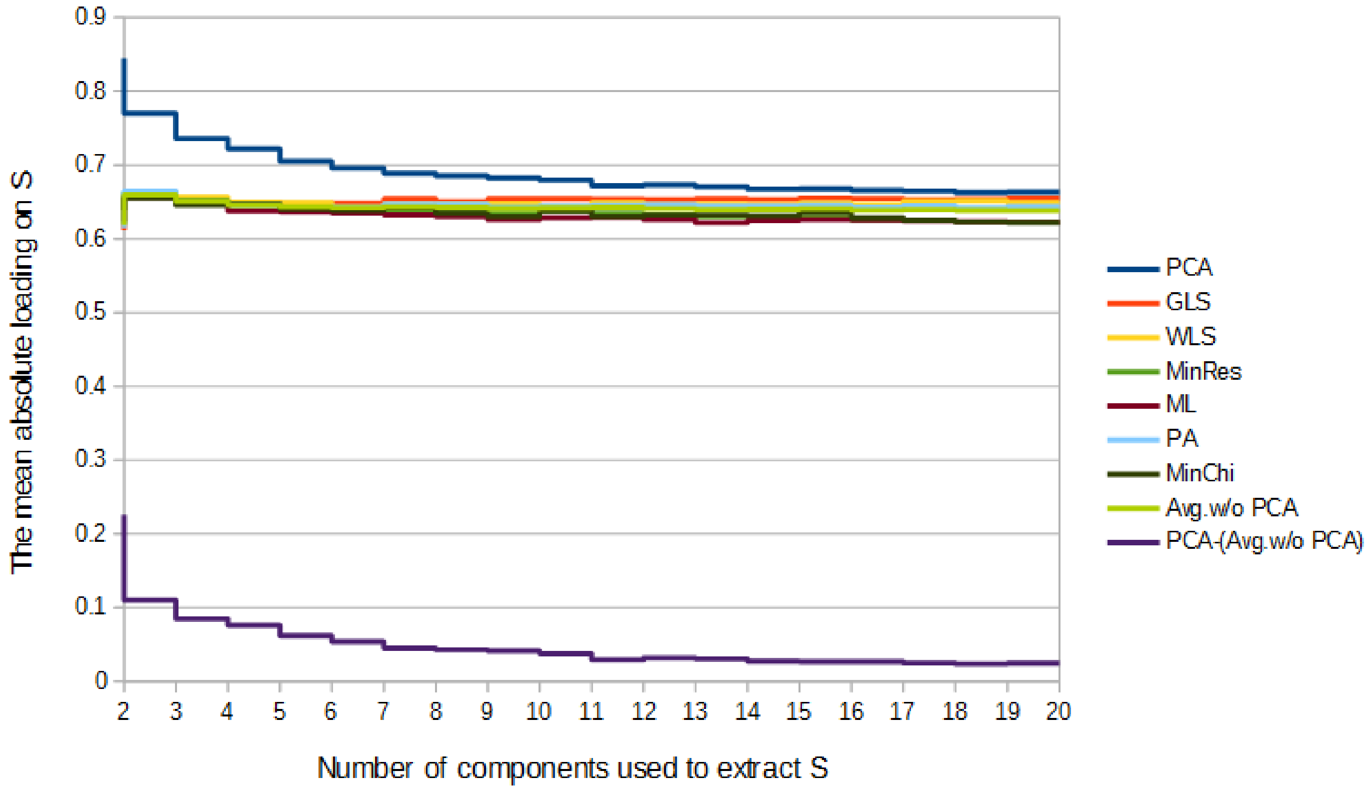

Figure 5: Mean absolute loadings on the first factor with different methods and different sample sizes of variables. Results for DR dataset.

Results are nearly identical for both datasets. The loadings using PCA are higher than with the other methods and this is especially the case for smaller samples of variables. Major et al[16] seem to be right. One should not use estimates of variable loadings from PCA when the number of variables is small.

7 Relationship with national cognitive ability measures

Usually, one finds some interesting country-level variable and then correlates it with national IQ, perhaps with some controls. Often authors will also argue for a causal connection from national IQ/G to country-level variables. The typical example of this is wealth (e.g. [18, 19, 20, 21, 22]). Since it is known that g causes greater wealth at the individual level, and that nations can generally be considered to be a large group of individuals, it would be very surprising, though not impossible, if there was no causation at the group level as well.

I don't want to argue at length for any causal model in this paper, so I merely present the correlations and leave the interpretation to the reader. I choose to look at two cognitive ability measures, the total scores from SPI and DR, as well as the S factors (PCA) from each datasets. For cognitive measures, I use Lynn and Vanhanen's 2012 national IQ estimates[6] and Altinok's educational achievement estimates[23]. Results are shown in Table 7.

I expected the correlations to be quite strong. If population differences in G are the main cause of national differences in many socioeconomic domains, then aggregating measures should increase the correlation with G, since measurement specificity averages out.

r	SPI	DR	IQ	Altinok	SPI_1	DR_1
SPI	1.000	0.928	0.819	0.839	0.981	0.965
DR		1.000	0.712	0.760	0.893	0.909
IQ			1.000	0.910	0.856	0.868
Altinok				1.000	0.870	0.872
SPI_1					1.000	0.975
DR_1						1.000

Table 7: Correlation matrix with Social Progress Index, Democracy Ranking, their S factors, national IQs and Altinok's achievements scores.

As expected, the correlations with the aggregated country-level measures have a very strong correlation with proxies for G. All correlations have P's<0.001 (N's 100-132). Furthermore, the correlations between cognitive measures and the S factors from both datasets are stronger than with the indexes as made by the authors. This indicates either or both of two things: 1) that it is the S factor that drives the relationship, not the remaining variance, or 2) that the summing procedure used by the authors obscures the underlying relationship. Global hereditarians[24] may interpret this result as being in line with predictions, while non-hereditarians may interpret it as showing that national differences in S cause national differences in G, or something else entirely.

I don't report the results using S extracted from the other methods because they are nearly identical.

8 The method of correlated vectors and the S factor

In the study of human cognitive abilities, it is known that different tests have different loadings on the first factor, their g-loading.[3] In other words, some tests measure the underlying factor better than others. Jensen invented the method of correlated vectors (MCV) to test whether it was the g factor that was related to some other variable, or some other part of the variance. I used the same method here in a reverse fashion to test whether it is the S factor found in the SPI and DR datasets that drives the correlation with G proxies or the non-S variance.

I thus calculated the correlation of both G proxies with every variable in both datasets and correlated these values with the S factor loadings. In both datasets, using both national cognitive measures, the correlations are 0.99, confirming the indication from before that it is the S factor that drives the relationships. Nevertheless, it is possible to have a very high correlation and not have a very linear relationship as demonstrated by Anscombe's Quartet.[25] Inspection of all the scatter-plots, however, reveal that the relationships really are very linear. Figure 6 shows a plot of one of these MCV analyses (all the others are similar). Tables 13 and 14 in the Appendix contains both the factor loadings on the S factor and the correlations of each variable with national IQs (96 variables in total).

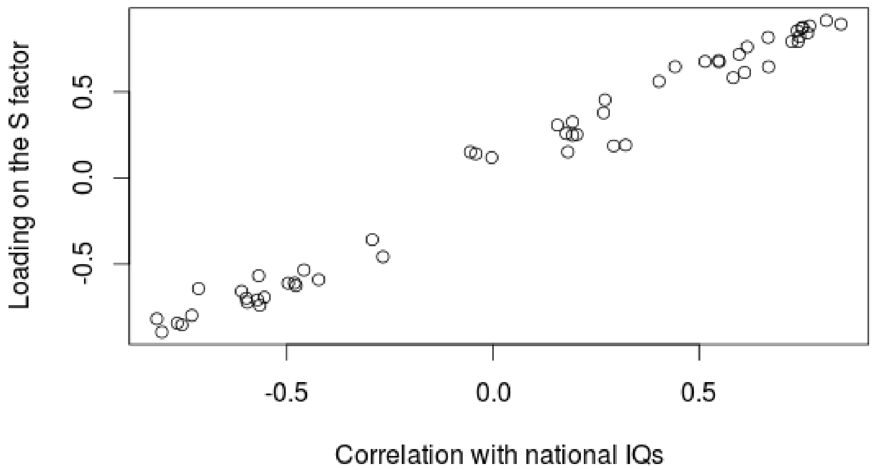

Figure 6: Variable loadings on the S factor (PCA) and their correlation with national IQs in the SPI dataset.

The very strong correlations resulting from the use of MCV in these analyses give indirect support for researchers who have been arguing that the MCV results on cognitive data are both spuriously low and inconsistent due to statistical artifacts. [26, 27, 28, 29] see also [30]. Probably the two primary sources of error are: 1) Restriction of range in factor loadings, and 2) sampling error in the loadings vector due to a small number of variables. Both sources of error seem very small in the present analysis since I used 54 and 42 variables varying in their loading from near -1 to almost 1. Table 8 shows descriptive statistics of variable loadings in both datasets.

Dataset	Min	Max	Mean	SD	N
SPI	-.90	.92	.10	.64	56
DR	-.70	.92	.53	.45	42

Table 8: Descriptive statistics for variable loadings in both datasets using PCA.

There is a question concerning whether it is proper to do the MCV analyses without reversing the variables that have negative loadings on the S factor first. Using the non-reversed variables means the variance is higher which increases the correlation. Reversing them would decrease the correlations. I decided to use the data as they were given by the authors i.e. with no reversal of variables.

If one wanted to reverse them nonetheless, it would not be quite obvious just which ones to reverse. With cognitive data, there is never any disagreement about which direction of a variable is associated with better performance. With social, political and economic variables, this is not always the case. For instance, in the DR dataset, there are two variables (28-29) dealing with expenditure on health: public and private. The public has a strong positive loading (0.658) and the private has a small negative loading (-0.210). There doesn't seem to be much agreement as to whether it is better for health care to be publicly or privately financed, so which direction should they be in? What about electric power consumption per capita (17)? It seems best not to get bogged down into discussions of which of these should be reversed and just use the data as given by authors. Recall that they could not have biased the data to fit my analyses since they put together the datasets for other purposes.

If, however, one insists on reversing the obvious cases, I have done this for those that I considered obvious enough. These are variables 1:6,8,13:18,26,28:33,35,42,48,53 from the SPI dataset. Re-doing the PCA and MCV analysis gives a correlation between S factor loading and national IQ of 0.98. Very little difference.

9 The S loading and the desirability of the variable: Are the first factors really *general* factors?

The mean absolute loading on the S factors is quite high (0.6-0.65). If one inspects the loadings (as shown in the Appendix, Tables 13 and 14), one may note that what loads positively on the S factor is generally considered something desirable, and the reverse for negative loadings. For example, the first variable in the SPI dataset is the percentage of the population that is malnourished. This is clearly something generally considered undesirable and the loading is strongly negative (-0.71).

There are a few variables, however, where the S-loading and the desirability of the variable are opposite. In the SPI dataset, obesity rate has a loading of 0.65 on S, but it is generally thought to be something undesirable, even if it indicates, evolutionarily speaking, that food is plentiful and that there is an absence of non-intentional starvation. The suicide rate from the same dataset shows the same pattern (loading 0.19). Death per capita by air pollution has a loading of 0.19. Finally, use of water resources has a loading of 0.26. There are no variables in the DR dataset like this. All in all, there are 4 variables out of 96 that have opposite signs on their S loading and desirability. Only in the case of obesity is the loading on S > 0.27.

It is worth noting that group-level correlations need not be the same or even in the same direction as individual-level correlations. Both suicide and obesity seem to have the opposite directional relationship at the individual-level.[31, 32]

In cognitive data, if one transforms variables so that a positive value corresponds to better performance (e.g. by reversing response time variables since longer reaction times indicate worse performance), then there is a positive manifold: Every variable has a positive correlation with any other variable. This pattern is seen both in individual-level data with the g factor and in country-level data with the G factor.

In light of the above, one might wonder whether it is fair to call the first factors "general factors" when there are 4 variables that go against the pattern. It comes down to which meaning of "general" is used. Dictionary.com gives these two definitions, among others:

1. of or pertaining to all persons or things belonging to a group or category: a general meeting of the employees.

2. of, pertaining to, or true of such persons or things in the main, with possible exceptions; common to most; prevalent; usual: the general mood of the people.

When I say "general", I mean in the second sense, not the first, and in this sense, the S factor is clearly general since only 4 of 96 variables load in the opposite direction, and only one of them strongly.

10 Number of factors to extract and their size

All previous analyses used only the first factor. However, one might want to know how many factors standard methods indicate that one should extract and if doing so changes results. The R package nFactors includes the nScree() function[5] that calculates the number of factors to extract using 4 different methods: optimal coordinates, acceleration factor (a non-subjective measure based on the Scree plot), parallel analysis and Kaiser's Rule (Eigenvalue >1).

I ran the analysis on both datasets. In the SPI dataset, three of the four methods indicate the number of factors to extract is 9, acceleration factor being the outlier which suggests only to extract 1. The situation is similar in the DR dataset with three methods suggesting 8 factors and acceleration factor again only suggesting 1.

Extracting the first 9/8 unrotated factors using maximum likelihood estimation, both Pearson correlations between scores (SPI: 0.98, DR: 0.97), loadings (SPI: 1.00, DR: 1.00) and congruence

[5], http://artax.karlin.mff.cuni.cz/r-help/library/nFactors/html/nScree.html

coefficients (SPI: 1.00, DR: 1.00) show that the first factors from the 1-factor analyses are very similar to or identical with the first factors from the 8-9 factor analyses.

It is also worth mentioning that the percent of the variance accounted for by each factor quickly decreases from the first factor, and that the first and second factor are not even close in size. Results are shown in Table 9. The variance accounted for by the first factor is slightly lower than in the 1 factor analysis (0.40 vs 0.39 for SPI, and 0.46 vs 0.43 for DR).

Proportion var. for factor N	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
SPI	0.39	0.08	0.07	0.04	0.03	0.03	0.02	0.02	0.01
DR	0.43	0.07	0.07	0.04	0.03	0.02	0.02	0.02	

Table 9: Proportion of variance accounted for each factor. Both datasets.

The relative size of the first factors compared to the second (SPI: 0.39 to 0.08, ratio = 4.9; DR: 0.43 to 0.07, ratio = 6.1) justifies calling them "large". Together with the evidence of their generality in the previous section and in the following two, this justifies calling them "large general factors". This would not have been the case if they were either not general or that the first two factors were about equally large.

11 Correlated factors

Another way to test if there is a general factor, is to use repeated factor analysis while allowing for correlated factors (oblimin rotated). This is sometimes considered the best way to extract g.[14]

The approach taken was this:

1. Determine number of factors to extract with nScree (choosing the number most criteria agree on).
2. Extract that number of factors with oblimin rotation using maximum likelihood estimation.
3. Repeat the above steps.

Using this method on cognitive data usually results in a g (general) factor at the second or third level.[33] The same result was found for both datasets with international data. The 3rd order factor scores correlated 0.97 with the first unrotated factor scores extracted with the same method in both datasets, indicating robust results. It was not possible to use factor congruence or correlations of loadings. In the SPI dataset, there were 9 factors at the first level, 3 at the second and 1 at the third. In the DR dataset, there were 8, 2, and 1, respectively.

11.1 Schmid-Leiman using schmid()

The psych package in R contains the schmid() function[6] which is an implementation of the Schmid-Leiman transformation[34] advocated by e.g. Jensen and Weng. The schmid() function can use three different extraction methods (MR, ML, PA) and three different rotation methods (oblimin, promax, simplimax). I ran the schmid() function on both datasets using all possible combinations. Then I compared the general factor loadings with the loadings from the unrotated first factor extracted using the same method (for technical reasons it was not possible to use factor congruence). Results are shown in Tables 10 and 11.

[6]http://personality-project.org/r/html/schmid.html

Method	Oblimin	Simplimax	Promax
minres	1.00	1.00	1.00
pa	.99	.90	.98
ml	.98	.85	.97

Table 10: Comparison table between general factor extracting using schmid() function and the first unrotated factor using the same extraction method. SPI dataset.

Method	Oblimin	Simplimax	Promax
minres	.97	.97	.99
pa	.97	1.00	.99
ml	.98	.99	.97

Table 11: Comparison table between general factor extracting using schmid() function and the first unrotated factor using the same extraction method. DR dataset.

There doesn't seem to be any overall pattern with regards to factor extraction or rotation method. Simplimax gave somewhat lower results in the SPI dataset, but the results were high in the DR dataset. Results were strongly in line with previous results.

12 Other measures of the strength of a general factor

Recent discussion concerning the existence of a general factor in the personality domain has renewed interest in criteria for considering the strength of a general factor.[35, 36, 37] Revelle and Wilt (RW)[38] have criticized these studies for using inadequate methods for determining whether a factor is plausibly interpreted as a strong general factor. Using simulated datasets, RW show that some of the methods employed by prior studies would indicate a strong general factor even in datasets where there plainly is none (see their Table 2). For instance, the size of the first factor is not a good indication of whether there is a strong general factor or not because the datasets with no general factor can have as large a first factor as those with a strong general factor. They show that a variety of other methods are able to distinguish between datasets with and without a strong general factor.

One of these methods is the hierarchical omega and its asymptotic value (ω_h and ω_{h_∞}). This method was also recently used to estimate the strength of the g factor in different age groups as a test of the Dynamic mutualism theory of g.[39]

A second method is to find the eigenvalue of the general factor in the Schmid-Leiman transformation and divide it by the number of variables ($\Lambda_g(\Lambda/N)$). This is the amount of variance accounted for by the general factor.

A third method is the explained common variance (ECV).

A fourth method is the squared multiple correlation of regressing the original variables on the general factor i.e. the amount of factor variance accounted for by the variables (R^2).

The omega() function[7] can be used to calculate these values.

RW used these four methods (and one more which I did not know how to use on my data) on 8 personality datasets and 5 classic cognitive datasets. I have repeated these analyses on my two datasets. Table 12 shows the average of the personality data, cognitive data and my datasets.

[7] http://personality-project.org/r/html/omega.html

Method	ω_h	ω_{h_∞}	$\Lambda_g(\Lambda/N)$	**ECV**	R^2
Personality data	.37	.48	.16	.34	.41
Cognitive data	.74	.79	.33	.57	.78
International data SPI	.74	.75	.33 (17.7/54)	.48	.81
International data DR	.77	.78	.37 (15.57/42)	.53	.81

Table 12: Comparison of measures of general factor strength on personality data, cognitive data and international data. Values for personality and cognitive data from [38].

The strength of the S factor in the international data is quite similar to, perhaps a little stronger than, the g factor in the classic datasets, while the general factor of personality is clearly weaker.

13 Discussion and conclusion

In the measurement of mental abilities, there has been a long-standing debate on whether g from one battery of tests was the same g from another battery of tests. The issue now seems to have been settled by two studies. The authors had access to datasets where the sample took a couple of different IQ batteries. They then used confirmatory factor analysis to compare latent g factors from different batteries to each other and found them to be at or very close to unity.[40, 41] The only exception was the latent g extracted from the Cattell test, which is probably because it only has one subtest type (non-verbal matrix reasoning) whereas the others had multiple different subtest types.

A second method that has been used is to sample random subsets from the entire battery of tests, factor analyze the subset with and without a probe test, and then compare the g-loading of this probe test in different subsets. Studies of this kind found quite stable g-loadings.[42, 43, 44, 17, 15]

The analyses carried out in this paper suggest that the S factor is quite like g in its stability of measurement. The S factor scores (PCA) extracted from the two international datasets correlated 0.975. Extracting S factors from subsets of variables and comparing them either to each other (subset x subset) or to the S factor from the complete analysis (subset x whole) suggest that S can be quite reliably approximated picking a small collection of random variables. It would be interesting to redo these two analyses on the cognitive data analyzed by Johnson et al.[40, 41]

One difference between cognitive data and international rankings is worth pointing out. The cognitive test variables are actually themselves some sort of sum of (usually unweighted) the actual test items (e.g. number of matrices solved). This is not generally the case for the national variables although some of them, e.g. Freedom House indexes, are some sort of function of subcomponents. It is not clear in which direction this disanalogy biases results, if any.

Results from both datasets examined in this study were generally extremely similar despite the fact that the datasets were created by two different sets of researchers with different goals, none of which was to look for any country-level general socioeconomic factor. This indicates that the results are robust.

A reviewer pointed to a series of studies concerning a national-level K (of r-K life history theory[45]) factor.[46, 47, 48] They employ some of the variables also found in the SPI and DR datasets and find a large general factor. From the perspective of this study, it seems that their K factor is likely just the S factor. Although it is possible that a K factor constitutes a first or second order factor below the S factor. Further research should investigate this.

The paper is an example of the usefulness of programmatic, as opposed to point-and-click/GUI-based, statistical tools applied to psychology. This emerging field has been termed "psychoinformatics" by analogy with "bioinformatics".[49] The benefits of this approach is that the exact analysis methods can be inspected instead of only described vaguely by words. They can also quickly be used on other datasets by other researchers. It is an important step in the push towards open-sourcing science.[50, 51] The main downside is that the learning curve becomes somewhat steeper but this is a very small price to pay.

Finally, should we care whether there is an S factor? Yes. The existence of a strong general socioeconomic factor means that there is a mathematical concept that corresponds to the intuitive concept of country well-being. This is not a necessary state of affairs. Imagine a two factor solution where there are a political freedom factor and an economic freedom factor that don't correlate. That would be very different from what we see when we look at the real world.

14 Datasets and source code

This paper is open source. Everything necessary to reproduce the PDF and results is available in the supplementary material at the journal website.

Table 15 in the Appendix contains the S scores for each country with data. They were computed as the first unrotated factor (PCA). When both datasets covered a country, the average was taken. When only one dataset had data, that value was used. The data can also be found in the supplementary material as well as in the Worldwide Megadataset (http://emilkirkegaard.dk/megadataset).

15 Acknowledgments

Thanks to Laird Shaw for help with R and proofreading. Thanks to reviewers of the *Open Differential Psychology* journal for numerous suggestions that greatly improved the article.

References

[1] Gregory Clark. *The Son Also Rises: Surnames and the History of Social Mobility.* Princeton University Press, 2014.

[2] Emil O. W. Kirkegaard and John Fuerst. Educational attainment, income, use of social benefits, crime rate and the general socioeconomic factor among 71 immmigrant groups in denmark. *Open Differential Psychology*, 2014.

[3] Arthur Robert Jensen. *The g factor: The science of mental ability.* Praeger Westport, CT, 1998.

[4] Heiner Rindermann. The big g-factor of national cognitive ability. *European Journal of Personality*, 21(5):767--787, 2007.

[5] Heitor B. F. Fernandes, Michael A. Woodley, and Jan te Nijenhuis. Differences in cognitive abilities among primates are concentrated on g: Phenotypic and phylogenetic comparisons with two meta-analytical databases. *Intelligence*, 46:311--322, 2014.

[6] Richard Lynn and Tatu Vanhanen. *Intelligence: A unifying construct for the social sciences.* Ulster Institute for Social Research, 2012.

[7] International Monetary Fund. World economic outlook database, 2014. URL `http://www.imf.org/external/pubs/ft/weo/2014/01/weodata/index.aspx`.

[8] Social Progress Imperative. Social progress index 2014, 2014. URL `http://www.socialprogressimperative.org/`.

[9] Scott Stern, Amy Wares, Sarah Orzell, and Patrick O'Sullivan. *SOCIAL PROGRESS INDEX 2014: METHODOLOGICAL REPORT.* Social Progress Imperative, 2014. URL `http://www.socialprogressimperative.org/publications`.

[10] Democracy Ranking. Democracy ranking, 2014. URL `http://democracyranking.org/`.

[11] David F. J. Campbell. *The Basic Concept for the Democracy Ranking of the Quality of Democracy*. Vienna: Democracy Ranking, 2008. URL http://democracyranking.org/?page_id=590.

[12] Nathan Zhao. The minimum sample size in factor analysis, 2009. URL https://www.encorewiki.org/display/~nzhao/The+Minimum+Sample+Size+in+Factor+Analysis.

[13] Graeme D Hutcheson and Nick Sofroniou. *The multivariate social scientist: Introductory statistics using generalized linear models*. Sage, 1999.

[14] Arthur R Jensen and Li-Jen Weng. What is a good g? *Intelligence*, 18(3):231--258, 1994.

[15] Jason Major. The dependability of the general factor of intelligence: Why g is not a first principal component. 2010.

[16] Jason T Major, Wendy Johnson, and Thomas J Bouchard Jr. The dependability of the general factor of intelligence: Why small, single-factor models do not adequately represent g. *Intelligence*, 39(5):418--433, 2011.

[17] Randy G Floyd, Elizabeth I Shands, Fawziya A Rafael, Renee Bergeron, and Kevin S McGrew. The dependability of general-factor loadings: The effects of factor-extraction methods, test battery composition, test battery size, and their interactions. *Intelligence*, 37(5):453--465, 2009.

[18] Richard Lynn and Tatu Vanhanen. *IQ and the wealth of nations*. Greenwood Publishing Group, 2002.

[19] Richard Lynn and Tatu Vanhanen. *IQ and global inequality*. Washington Summit Publishers, 2006.

[20] Gerhard Meisenberg. National iq and economic outcomes. *Personality and Individual Differences*, 53(2):103--107, 2012.

[21] Heiner Rindermann and James Thompson. Cognitive capitalism the effect of cognitive ability on wealth, as mediated through scientific achievement and economic freedom. *Psychological Science*, 22(6):754--763, 2011.

[22] Gregory B Christainsen. Iq and the wealth of nations: How much reverse causality? *Intelligence*, 41(5):688--698, 2013.

[23] Nadir Altinok, Claude Diebolt, and Jean-Luc Demeulemeester. A new international database on education quality: 1965--2010. *Applied Economics*, 46(11):1212--1247, 2014.

[24] John Fuerst and Emil O. W. Kirkegaard. Do national iqs predict u.s. immigrant cognitive ability and outcomes? an analysis of the national longitudinal survey of freshman. *Open Differential Psychology*, 2014.

[25] Wikipedia. Anscombe's quartet --- wikipedia, the free encyclopedia, 2014. URL http://en.wikipedia.org/w/index.php?title=Anscombe%27s_quartet&oldid=613479365. [Online; accessed 4-August-2014].

[26] Jan te Nijenhuis, Annelies EM van Vianen, and Henk van der Flier. Score gains on g-loaded tests: No g. *Intelligence*, 35(3):283--300, 2007.

[27] Jan te Nijenhuis and Henk van der Flier. Is the flynn effect on g?: A meta-analysis. *Intelligence*, 41(6):802--807, 2013.

[28] Jan te Nijenhuis. The flynn effect, group differences, and g-loadings. *Personality and Individual Differences*, 55(3):224--228, 2013.

[29] Jan te Nijenhuis, Birthe Jongeneel-Grimen, and Emil O. W. Kirkegaard. Are headstart gains on the g factor?: A meta-analysis. *Intelligence*, In Press, 2014.

[30] John E Hunter and Frank L Schmidt. *Methods of meta-analysis: Correcting error and bias in research findings*. Sage, 3rd edition, 2014.

[31] Lars Andersson, Peter Allebeck, J-E Gustafsson, and D Gunnell. Association of iq scores and school achievement with suicide in a 40-year follow-up of a swedish cohort. *Acta Psychiatrica Scandinavica*, 118(2):99--105, 2008.

[32] T Chandola, IJ Deary, D Blane, and GD Batty. Childhood iq in relation to obesity and weight gain in adult life: the national child development (1958) study. *International journal of obesity*, 30(9):1422--1432, 2006.

[33] John Bissell Carroll. *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press, 1993.

[34] John Schmid and John M Leiman. The development of hierarchical factor solutions. *Psychometrika*, 22(1):53--61, 1957.

[35] Janek Musek. A general factor of personality: Evidence for the Big One in the five-factor model. *Journal of Research in Personality*, 41(6):1213--1233, 2007.

[36] J Philippe Rushton and Paul Irwing. A General Factor of Personality (GFP) from two meta-analyses of the Big Five: Digman (1997) and Mount, Barrick, Scullen, and Rounds (2005). *Personality and Individual Differences*, 45(7):679--683, 2008.

[37] Dimitri van der Linden, Jan te Nijenhuis, and Arnold B Bakker. The general factor of personality: A meta-analysis of big five intercorrelations and a criterion-related validity study. *Journal of Research in Personality*, 44(3):315--327, 2010.

[38] William Revelle and Joshua Wilt. The general factor of personality: A general critique. *Journal of research in personality*, 47(5):493--504, 2013.

[39] Gilles E Gignac. Dynamic mutualism versus g factor theory: An empirical test. *Intelligence*, 42: 89--97, 2014.

[40] Wendy Johnson, Thomas J Bouchard Jr, Robert F Krueger, Matt McGue, and Irving I Gottesman. Just one g: consistent results from three test batteries. *Intelligence*, 32(1):95--107, 2004.

[41] Wendy Johnson, Jan te Nijenhuis, and Thomas J Bouchard Jr. Still just 1 g: Consistent results from five test batteries. *Intelligence*, 36(1):81--95, 2008.

[42] Robert L Thorndike. Stability of factor loadings. *Personality and Individual Differences*, 8(4): 585--586, 1987.

[43] Malcolm James Ree and James A Earles. The stability of g across different methods of estimation. *Intelligence*, 15(3):271--278, 1991.

[44] Charlie L Reeve and Nikki Blacksmith. Equivalency and reliability of vectors of g-loadings across different methods of estimation and sample sizes. *Personality and Individual Differences*, 47(8): 968--972, 2009.

[45] J Philippe Rushton. Race, Evolution, and Behavior: A life history perspective. 1995.

[46] Donald I Templer. Correlational and factor analytic support for rushton's differential k life history theory. *Personality and Individual Differences*, 45(6):440--444, 2008.

[47] J Philippe Rushton and Donald I Templer. National differences in intelligence, crime, income, and skin color. *Intelligence*, 37(4):341--346, 2009.

[48] Gerhard Meisenberg and Michael A Woodley. Global behavioral variation: A test of differential-K. *Personality and individual differences*, 55(3):273--278, 2013.

[49] Tal Yarkoni. Psychoinformatics new horizons at the interface of the psychological and computing sciences. *Current Directions in Psychological Science*, 21(6):391--397, 2012.

[50] Brian A Nosek and Yoav Bar-Anan. Scientific utopia: I. Opening scientific communication. *Psychological Inquiry*, 23(3):217--243, 2012.

[51] Brian A Nosek, Jeffrey R Spies, and Matt Motyl. Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science*, 7(6): 615--631, 2012.

Loadings for first factor in Social Progress Index		
Subcomponent	*Loading*	*r x IQ*
Malnourishment (% of pop.)	-0.71	-0.57
Depth of food deficit (calories/malnourished person)	-0.69	-0.55
Maternal mortality rate (deaths/100,000 live births)	-0.84	-0.77
Stillbirth rate (deaths/1,000 live births)	-0.86	-0.75
Child mortality rate (deaths/1,000 live births)	-0.90	-0.80
Deaths from infectious diseases (deaths/100,000)	-0.82	-0.82
Access to piped water (% of pop.)	0.88	0.77
Rural vs. urban access to improved water source (absolute difference between % of pop.)	-0.70	-0.60
Access to improved sanitation facilities (% of pop.)	0.86	0.74
Availability of affordable housing (% satisfied)	0.25	0.20
Access to electricity (% of pop.)	0.84	0.76
Quality of electricity supply (1=low; 7=high)	0.79	0.74
Indoor air pollution attributable deaths (deaths/100,000)	-0.74	-0.57
Homicide rate (1= <2/100,000; 5= >20/100,000	-0.64	-0.71
Level of violent crime (1=low; 5=high)	-0.57	-0.57
Perceived criminality (1=low; 5=high)	-0.54	-0.46
Political terror (1=low; 5=high)	-0.61	-0.48
Traffic deaths (deaths/100,000)	-0.66	-0.61
Adult literacy rate (% of pop. aged 15+)	0.82	0.74
Primary school enrollment (% of children)	0.58	0.58
Lower secondary school enrollment (% of children)	0.82	0.67
Upper secondary school enrollment (% of children)	0.88	0.75
Gender parity in secondary enrollment (girls/boys)	0.56	0.40
Mobile telephone subscriptions (subscriptions/100 people)	0.68	0.55
Internet users (% of pop.)	0.92	0.81
Press Freedom Index (0=most free; 100=least free)	-0.46	-0.27
Life expectancy (years)	0.89	0.84
Non-communicable disease deaths between the ages of 30 and 70 (probability of dying)	-0.72	-0.60
Obesity rate (% of pop.)	0.65	0.44
Outdoor air pollution attributable deaths (deaths/100,000)	0.19	0.32
Suicide rate (deaths/100,000)	0.19	0.29
Greenhouse gas emissions (CO2 equivalents per GDP)	-0.36	-0.29
Water withdrawals as a percent of resources	0.26	0.18
Biodiversity and habitat (0=no protection; 100=high protection)	0.25	0.19
Political rights (1=full rights; 7=no rights)	-0.63	-0.48
Freedom of speech (0=low; 2=high)	0.33	0.19
Freedom of assembly/association (0=low; 2=high)	0.45	0.27
Freedom of movement (0=low; 4=high)	0.38	0.27
Private property rights (0=none; 100=full)	0.72	0.60
Freedom over life choices (% satisfied)	0.31	0.16
Freedom of religion (1=low; 4=high)	0.14	-0.04
Modern slavery, human trafficking and child marriage (1=low; 100=high)	-0.61	-0.50
Satisfied demand for contraception (% of women)	0.79	0.73
Corruption (0=high; 100=low)	0.76	0.62
Women treated with respect (0=low; 100=high)	0.15	0.18
Tolerance for immigrants (0=low; 100=high)	0.15	-0.05
Tolerance for homosexuals (0=low; 100=high)	0.68	0.55
Discrimination and violence against minorities (0=low; 10=high)	-0.59	-0.42
Religious tolerance (1=low; 4=high)	0.12	0.00
Community safety net (0=low; 100=high)	0.68	0.51
Years of tertiary schooling	0.65	0.67
Womens' average years in school	0.87	0.75
Inequality in the attainment of education (0=low; 1=high)	-0.80	-0.73
Number of globally ranked universities (0=none; 5= >50)	0.61	0.61

Table 13: Social Progress Index dataset loadings from PCA.

Appendix

Loadings for first factor in Democracy Ranking dataset		
Subcomponent	*Loading*	*r x IQ*
Political rights (aggregated scores): Freedom House	0.67	0.43
Civil liberties (aggregated scores): Freedom House	0.74	0.53
Global Gender Gap Report	0.56	0.39
Press Freedom: Freedom House	0.67	0.44
Corruption Perceptions Index (CPI): Transparency International (TI)	0.79	0.64
Change(s) of the head of government (last 13 years, peaceful)	0.23	0.20
Political party change(s) of the head of government (last 13 years, peaceful)	0.39	0.30
GDP per capita, PPP (constant 2005 international $)	0.87	0.72
GDP per capita, PPP (current international $)	0.87	0.72
Central government debt, total (% of GDP)	0.56	0.51
Inflation, consumer prices (annual %)	0.58	0.49
Unemployment, total (% of total labor force)	0.55	0.55
Unemployment, youth total (% of total labor force ages 15-24)	0.59	0.51
CO2 emissions (kg per 2005 PPP $ of GDP)	-0.17	-0.26
CO2 emissions (metric tons per capita)	-0.59	-0.49
GDP per unit of energy use (constant 2005 PPP $ per kg of oil equivalent)	-0.48	-0.50
Electric power consumption (kWh per capita)	-0.70	-0.58
Electricity production from hydroelectric sources (% of total)	-0.53	-0.54
Labor force, female (% of total labor force)	0.11	0.02
Unemployment, female (% of female labor force)	0.53	0.47
Primary education, pupils (% female)	0.31	0.15
School enrollment, secondary, female (% gross)	0.87	0.76
School enrollment, secondary, female (% net)	0.87	0.73
School enrollment, tertiary, female (% gross)	0.78	0.66
Life expectancy at birth, female (years)	0.86	0.85
Life expectancy at birth, total (years)	0.86	0.84
Health expenditure per capita, PPP (constant 2005 international $)	0.82	0.67
Health expenditure, public (% of GDP)	0.66	0.47
Health expenditure, private (% of GDP)	-0.21	-0.28
Hospital beds (per 1,000 people)	0.52	0.58
Physicians (per 1,000 people)	0.68	0.64
Mortality rate, infant (per 1,000 live births)	0.86	0.82
Mortality rate, under-5 (per 1,000 live births)	0.83	0.81
School enrollment, secondary (% gross)	0.89	0.78
School enrollment, secondary (% net)	0.84	0.74
School enrollment, tertiary (% gross)	0.90	0.83
Pupil-teacher ratio, primary	0.59	0.52
Telephone lines (per 100 people)	0.88	0.81
Internet users (per 100 people)	0.92	0.83
Mobile cellular subscriptions (per 100 people)	0.71	0.61
Research and development expenditure (% of GDP)	0.61	0.52
Scientific and technical journal articles (per 1,000 people)	0.78	0.66

Table 14: Democracy Ranking dataset loadings from PCA.

S score	ID	Names
-1.646905191	AGO	Angola
0.007429029	ALB	Albania
0.738523723	ARE	United Arab Emirates
0.644262065	ARG	Argentina
-0.029950119	ARM	Armenia
1.558365495	AUS	Australia
1.371992315	AUT	Austria

0.042892955	AZE	Azerbaijan
-1.942758154	BDI	Burundi
1.383343184	BEL	Belgium
-1.235078464	BEN	Benin
-1.602305168	BFA	Burkina Faso
-0.873359243	BGD	Bangladesh
0.553371357	BGR	Bulgaria
-0.116325961	BHR	Bahrain
0.066447482	BIH	Bosnia and Herzegovina
0.320160291	BLR	Belarus
-0.265921951	BOL	Bolivia
0.211618534	BRA	Brazil
-0.543886157	BTN	Bhutan
-0.377329857	BWA	Botswana
-2.29869988	CAF	Central African Republic
1.439760092	CAN	Canada
1.513321679	CHE	Switzerland
0.754033701	CHL	Chile
-0.248274788	CHN	China
-1.350950211	CMR	Cameroon
-1.287657542	COG	Congo Rep.
0.043145898	COL	Colombia
0.625901223	CRI	Costa Rica
0.24519397	CUB	Cuba
0.627728605	CYP	Cyprus
0.955500773	CZE	Czech Republic
1.459750885	DEU	Germany
-1.134191132	DJI	Djibouti
1.534062625	DNK	Denmark
-0.114062309	DOM	Dominican Republic
-0.071647669	DZA	Algeria
0.035434359	ECU	Ecuador
-0.332594317	EGY	Egypt Arab Rep.
1.210823896	ESP	Spain
1.070084848	EST	Estonia
1.635287426	FIN	Finland
1.276852479	FRA	France
1.376690022	GBR	United Kingdom
0.040574777	GEO	Georgia
-0.888365187	GHA	Ghana
-1.865749584	GIN	Guinea
0.855402481	GRC	Greece
-0.402289169	GTM	Guatemala
-0.138027257	GUY	Guyana
-0.348985065	HND	Honduras
0.62922268	HRV	Croatia
-1.271672496	HTI	Haiti
0.769771009	HUN	Hungary
-0.302354034	IDN	Indonesia
-0.744089791	IND	India

1.399225147	IRL	Ireland
-0.124524871	IRN	Iran Islamic Rep.
-0.841668176	IRQ	Iraq
1.431614903	ISL	Iceland
1.051434987	ISR	Israel
0.959591737	ITA	Italy
0.199015596	JAM	Jamaica
0.279322831	JOR	Jordan
1.298028434	JPN	Japan
0.2190532	KAZ	Kazakhstan
-1.117062455	KEN	Kenya
-0.291080922	KGZ	Kyrgyz Republic
-0.964596911	KHM	Cambodia
1.145562698	KOR	Korea Rep.
0.665151534	KWT	Kuwait
-0.983513745	LAO	Lao PDR
0.070833285	LBN	Lebanon
-1.282846945	LBR	Liberia
-0.635241651	LBY	Libya
-0.185372171	LKA	Sri Lanka
-1.066267436	LSO	Lesotho
0.777961201	LTU	Lithuania
0.593825046	LVA	Latvia
-0.368161707	MAR	Morocco
0.086260355	MDA	Moldova
-1.36754981	MDG	Madagascar
0.170060159	MEX	Mexico
0.140480755	MKD	Macedonia FYR
-1.412397485	MLI	Mali
0.406047334	MNE	Montenegro
-0.106068729	MNG	Mongolia
-1.560046735	MOZ	Mozambique
-1.578617978	MRT	Mauritania
0.412291654	MUS	Mauritius
-1.431907603	MWI	Malawi
0.262648808	MYS	Malaysia
-0.654597647	NAM	Namibia
-1.752581129	NER	Niger
-1.51473539	NGA	Nigeria
-0.24724988	NIC	Nicaragua
1.555502383	NLD	Netherlands
1.658886469	NOR	Norway
-0.755987594	NPL	Nepal
1.429661061	NZL	New Zealand
-1.344835442	PAK	Pakistan
0.38815429	PAN	Panama
0.101986622	PER	Peru
-0.123768471	PHL	Philippines
-1.040057426	PNG	Papua New Guinea
0.837377811	POL	Poland

1.049452138	PRT	Portugal
-0.291711788	PRY	Paraguay
0.37607494	ROU	Romania
0.264524615	RUS	Russian Federation
-1.147138222	RWA	Rwanda
0.351630848	SAU	Saudi Arabia
-1.484862655	SDN	Sudan
-1.008201142	SEN	Senegal
0.92417306	SGP	Singapore
-1.736454518	SLE	Sierra Leone
-0.052413047	SLV	El Salvador
0.368451917	SRB	Serbia
0.731005733	SVK	Slovak Republic
1.124515502	SVN	Slovenia
1.556969725	SWE	Sweden
-0.932814898	SWZ	Swaziland
-0.826221777	SYR	Syrian Arab Republic
-2.346681181	TCD	Chad
-1.411308631	TGO	Togo
0.094976498	THA	Thailand
-0.451511323	TJK	Tajikistan
-0.780605433	TLS	Timor-Leste
0.513083926	TTO	Trinidad and Tobago
0.006151615	TUN	Tunisia
0.154421183	TUR	Turkey
-1.281883545	TZA	Tanzania
-1.37592458	UGA	Uganda
0.393375955	UKR	Ukraine
0.767864944	URY	Uruguay
1.433892248	USA	United States
-0.01863086	UZB	Uzbekistan
-0.088868721	VEN	Venezuela RB
-1.500981407	YEM	Yemen Rep.
-0.167586574	ZAF	South Africa
-1.199754202	ZMB	Zambia

Table 15: S factor scores for all countries found in either SPI, DR or both.