**Supplementary material**

Table S1

*Means and standard deviations of sample correlation coefficients, and mean absolute difference between sample correlation coefficients and population correlation coefficients (N = 25).*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **ASVAB** | **BFI items** | **BFI scales** | **DBQ items** | **DBQ scales** |
|  | **mean across**  **45**  **correlations** | **mean across**  **946**  **correlations** | **mean across**  **10**  **correlations** | **mean across**  **561**  **correlations** | **mean across**  **10**  **correlations** |
| |Mean *rp*| | .6230 | .1194 | .1739 | .1810 | .4102 |
| |Mean *rs*| | .6104 | .1202 | .1633 | .1749 | .4063 |
| |Mean *rp*|−|*Rp|* | −.0043 | −.0012 | −.0039 | .0096 | −.0096 |
| |Mean *rs*|−|*Rs|* | −.0177 | −.0020 | −.0057 | .0126 | −.0094 |
| *SD* *rp* | .1214 | .2094 | .2122 | .2315 | .1943 |
| *SD* *rs* | .1309 | .2057 | .2053 | .2125 | .1756 |
| Mean |*rp*−*Rp*| | .0954 | .1689 | .1709 | .1895 | .1561 |
| Mean |*rp*−*Rs*| | .0962 | .1690 | .1713 | .1898 | .1572 |
| Mean |*rs*−*Rp*| | .1038 | .1655 | .1652 | .1739 | .1414 |
| Mean |*rs*−*Rs*| | .1031 | .1653 | .1649 | .1742 | .1401 |

*Note*. *rp* = sample Pearson correlation coefficient, *Rp* = population Pearson correlation coefficient, *rs* = sample Spearman correlation coefficient, *Rs* = population Spearman correlation coefficient, ASVAB = Armed Services Vocational Aptitude Battery, BFI = Big Five Inventory, DBQ = Driver Behaviour Questionnaire. The absolute means, standard deviations, and mean absolute differences were calculated for each off-diagonal item of the correlation matrix (45, 946, 10, 561, & 10 correlations for the ASVAB, BFI items, BFI scales, DBQ items, & DBQ scales, respectively) and subsequently averaged. *Rp* and *Rs* were defined as the correlation coefficients for the total sample (*N* = 11,878 for the ASVAB, *N* = 1,895,753 for the BFI, & *N* = 9,077 for the DBQ). The results were based on 50,000 samples of *N* = 25. When the correlation matrix could not be calculated due to the small sample size, the sampling was repeated.

Table S2

*Means and standard deviations of sample correlation coefficients, and mean absolute difference between sample correlation coefficients and population correlation coefficients (N = 1,000).*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **ASVAB** | **BFI items** | **BFI scales** | **DBQ items** | **DBQ scales** |
|  | **mean across**  **45**  **correlations** | **mean across**  **946**  **correlations** | **mean across**  **10**  **correlations** | **mean across**  **561**  **correlations** | **mean across**  **10**  **correlations** |
| |Mean *rp*| | .6273 | .1206 | .1778 | .1709 | .4193 |
| |Mean *rs*| | .6277 | .1222 | .1690 | .1621 | .4154 |
| |Mean *rp*|−|*Rp|* | .0000 | .0000 | .0000 | −.0004 | −.0005 |
| |Mean *rs*|−|*Rs|* | −.0004 | .0000 | .0000 | −.0001 | −.0003 |
| *SD* *rp* | .0182 | .0327 | .0332 | .0401 | .0348 |
| *SD* *rs* | .0193 | .0319 | .0319 | .0331 | .0269 |
| Mean |*rp*−*Rp*| | .0145 | .0261 | .0265 | .0320 | .0277 |
| Mean |*rp*−*Rs*| | .0195 | .0267 | .0284 | .0338 | .0342 |
| Mean |*rs*−*Rp*| | .0201 | .0261 | .0274 | .0291 | .0294 |
| Mean |*rs*−*Rs*| | .0154 | .0255 | .0255 | .0264 | .0214 |

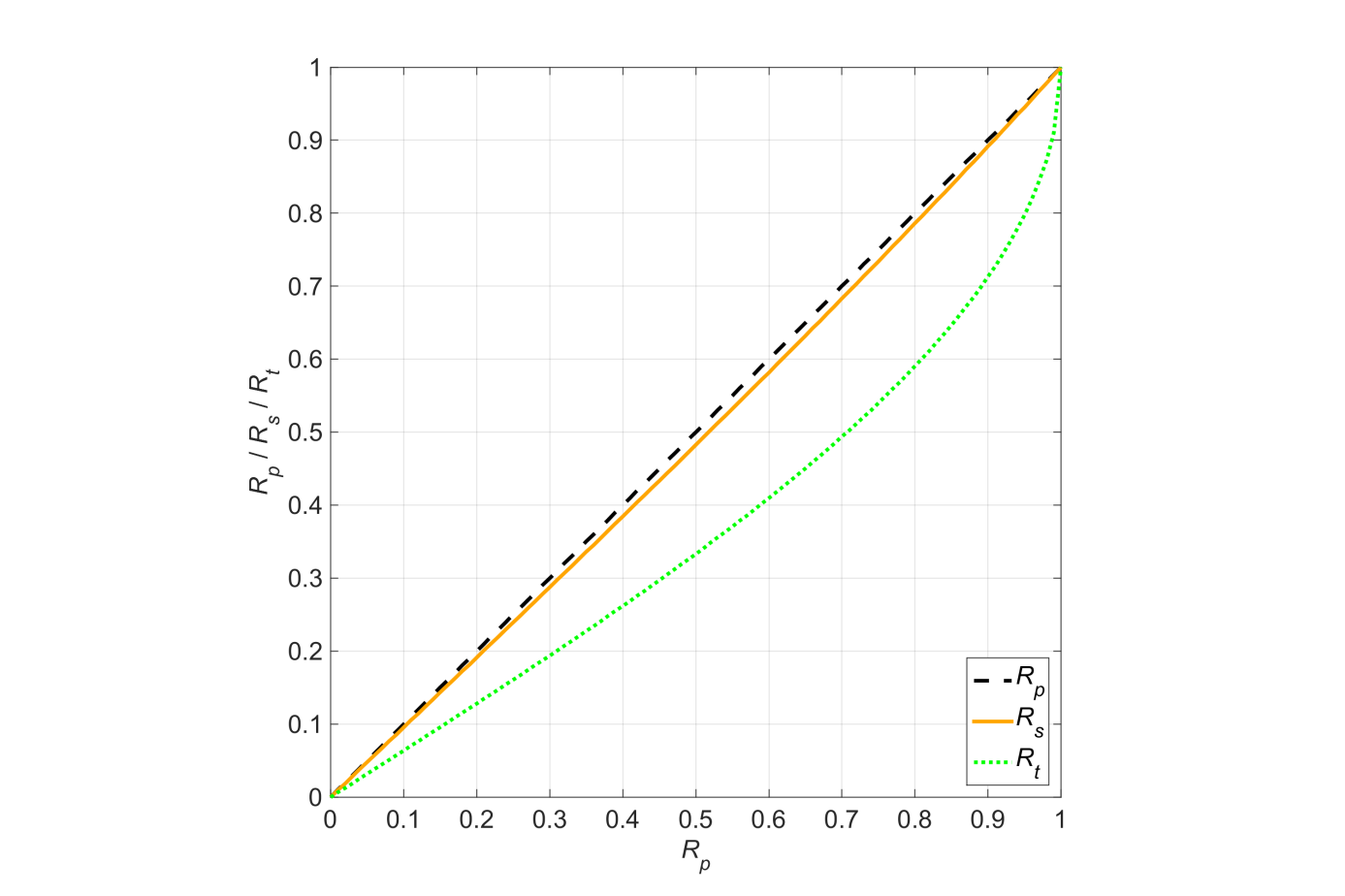
*Note*. *rp* = sample Pearson correlation coefficient, *Rp* = population Pearson correlation coefficient, *rs* = sample Spearman correlation coefficient, *Rs* = population Spearman correlation coefficient, ASVAB = Armed Services Vocational Aptitude Battery, BFI = Big Five Inventory, DBQ = Driver Behaviour Questionnaire. The absolute means, standard deviations, and mean absolute differences were calculated for each off-diagonal item of the correlation matrix (45, 946, 10, 561, & 10 correlations for the ASVAB, BFI items, BFI scales, DBQ items, & DBQ scales, respectively) and subsequently averaged. *Rp* and *Rs* were defined as the correlation coefficients for the total sample (*N* = 11,878 for the ASVAB, *N* = 1,895,753 for the BFI, & *N* = 9,077 for the DBQ). The results were based on 50,000 samples of *N* = 1,000.

Table S3

*Means and standard deviations of the first six eigenvalues of the 34 x 34 correlation matrices of the Driver Behaviour Questionnaire (DBQ).*

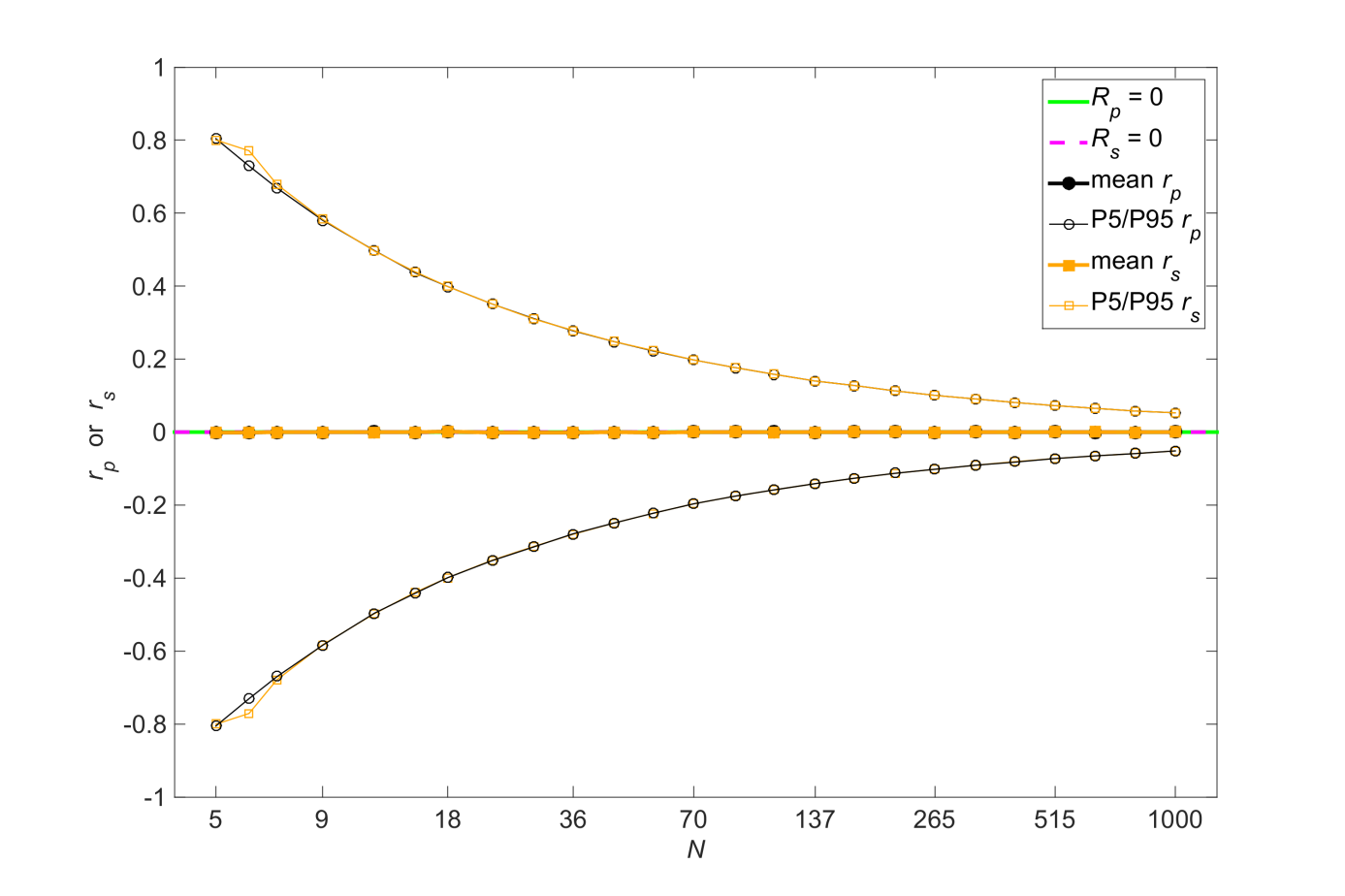
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | ***rp* matrices Mean (*SD*)** | ***rs* matrices**  **Mean (*SD*)** | ***Rp* matrix** | ***Rs* matrix** |
| Eigenvalue 1 | 6.977 (0.896) | 6.666 (0.577) | 6.830 | 6.547 |
| Eigenvalue 2 | 2.910 (0.317) | 2.704 (0.234) | 2.673 | 2.517 |
| Eigenvalue 3 | 1.845 (0.155) | 1.689 (0.092) | 1.274 | 1.238 |
| Eigenvalue 4 | 1.606 (0.092) | 1.530 (0.069) | 1.256 | 1.205 |
| Eigenvalue 5 | 1.459 (0.073) | 1.416 (0.058) | 1.206 | 1.174 |
| Eigenvalue 6 | 1.346 (0.066) | 1.321 (0.052) | 1.049 | 1.061 |

*Note*. *rp* = sample Pearson correlation coefficient, *Rp* = population Pearson correlation coefficient, *rs* = sample Spearman correlation coefficient, *Rs* = population Spearman correlation coefficient. The sample correlation coefficients were based on 50,000 samples of *N* = 200. *Rp* and *Rs* were defined as the correlation coefficients for the total sample (*N* = 9,077).

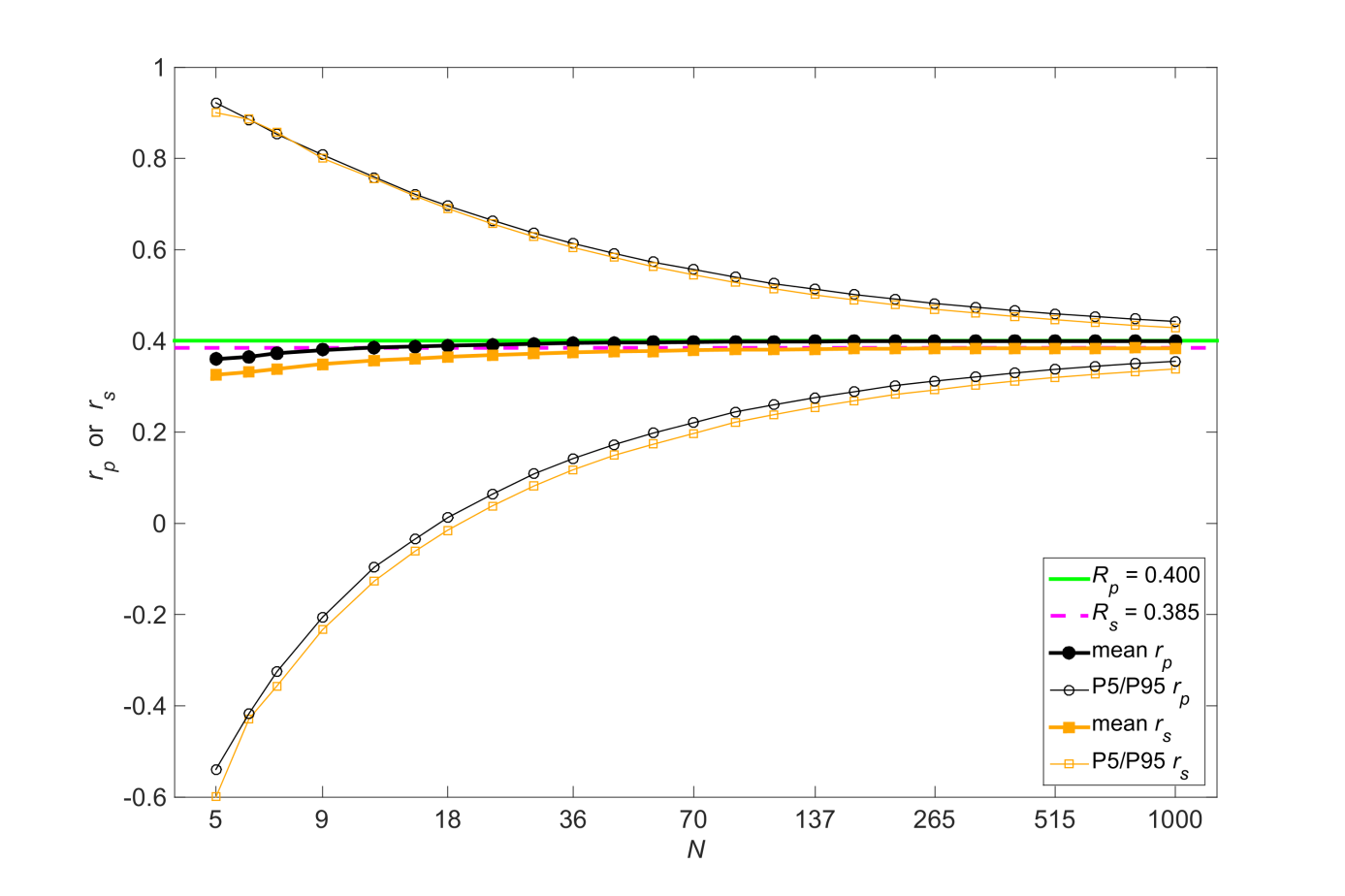


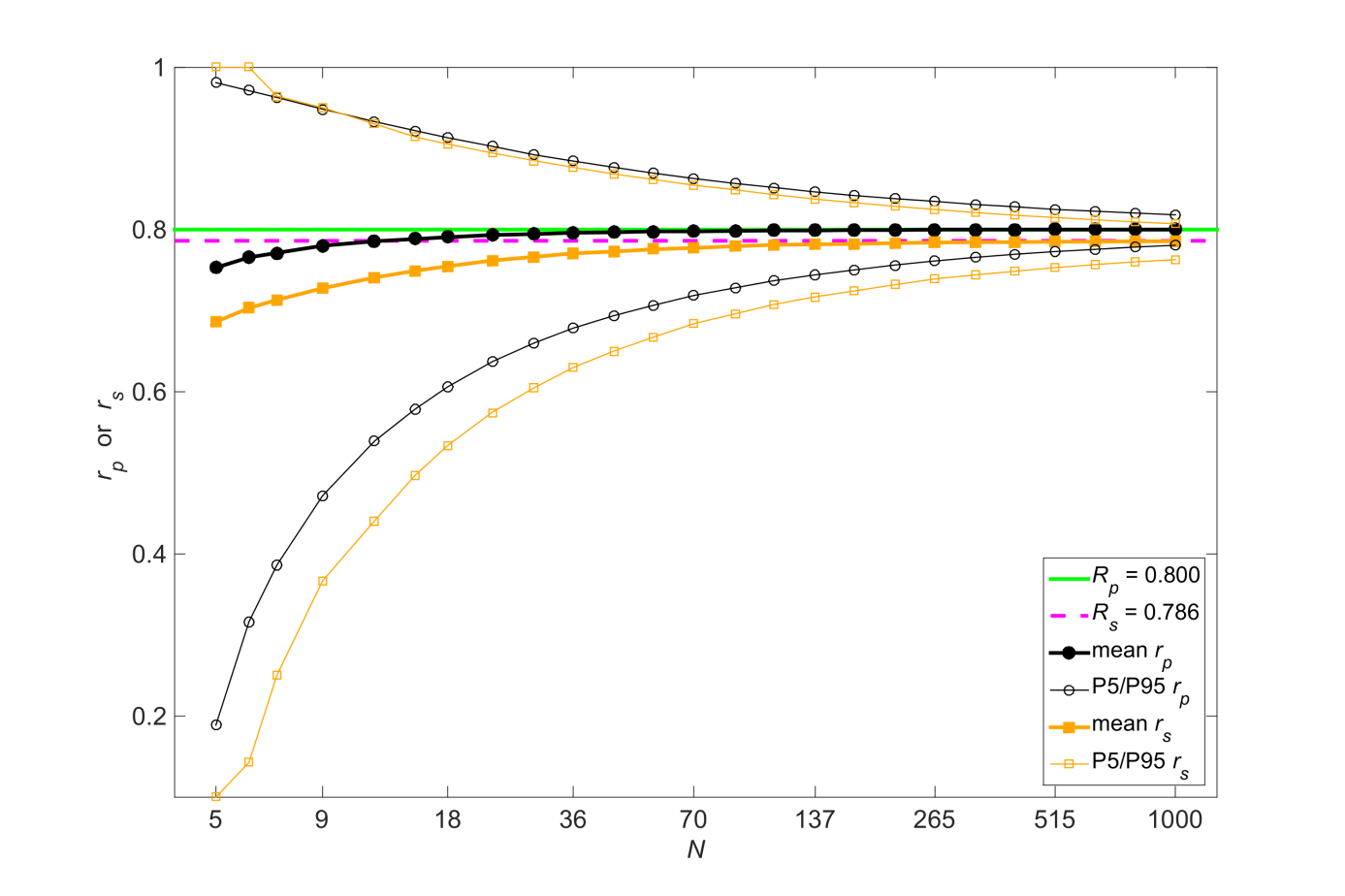
*Figure S1.* The solid line is the relationship between the population Spearman correlation coefficient (*Rs)* and the population Pearson correlation coefficient (*Rp*) in the case of bivariate normality. The dotted line is the relationship between the population Kendall’s tau (*Rt)* and *Rp* in the case of bivariate normality. The dashed line represents *Rp* versus *Rp* and therefore runs diagonally.



*Figure S2.* Simulation results for normally distributed variables having a population Pearson correlation coefficient of .2 (*Rp* = .2). The figure shows the distribution of the Pearson correlation coefficient (*rp*) and the Spearman correlation coefficient (*rs*) for a sample size (*N*) of 5. The distribution was obtained from a simulation of 107 repetitions. The resolution of the distribution was 0.01. The results have been normalized so that the sum of the 201 counts equaled 1. The figure also depicts the exact distribution of *rp* calculated with Equation 4, which lies almost exactly on top of the results of the simulation study. 

*Figure S3.* Simulation results for normally distributed variables having a population Pearson/Spearman correlation coefficient of 0 (*Rp* = *Rs* = 0). The figure shows the mean, 5th percentile (P5), and 95th percentile (P95) of the Pearson correlation coefficient (*rp*) and the Spearman correlation coefficient (*rs*) as a function of sample size (*N*).

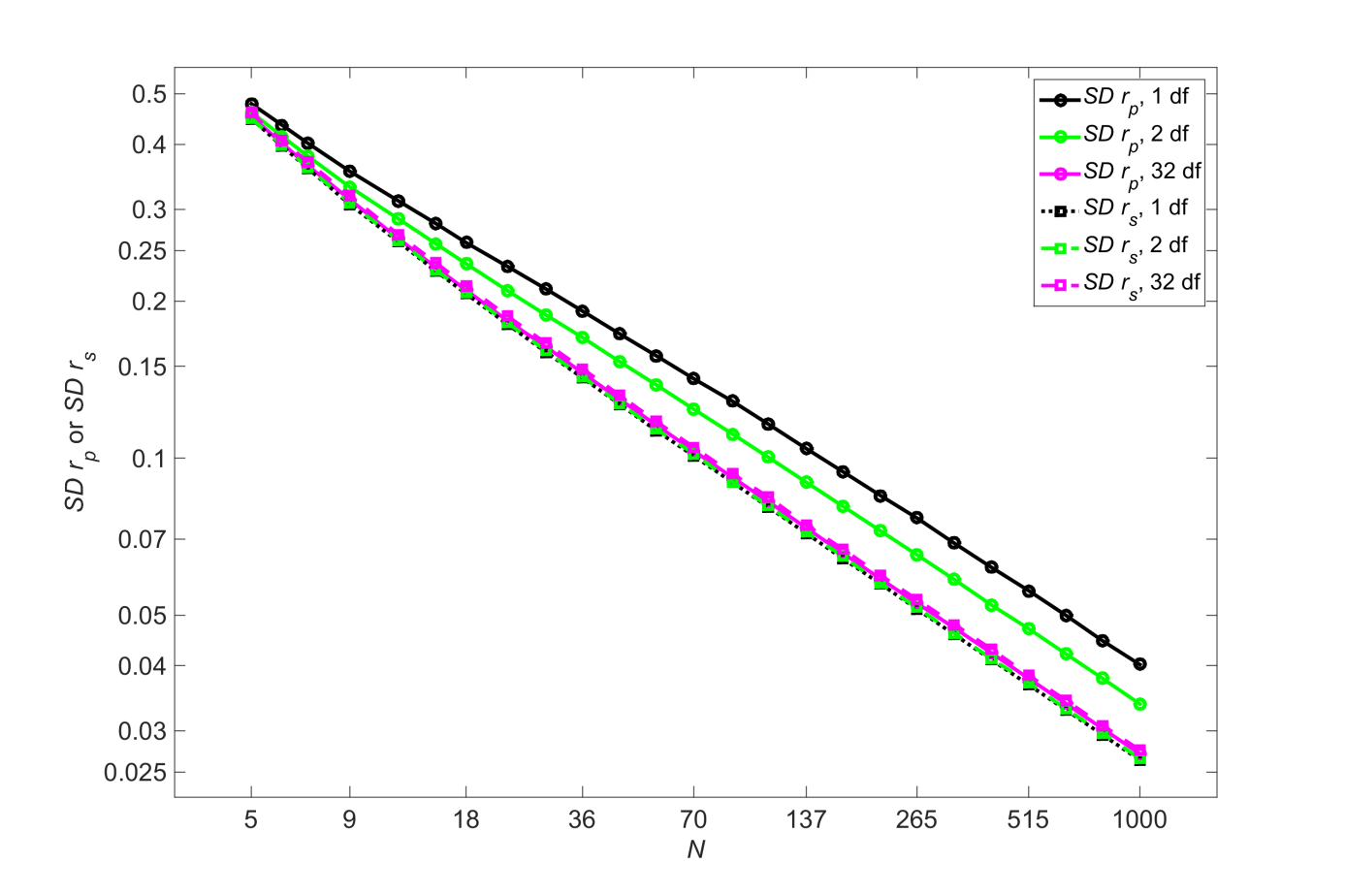
*Figure S4.* Simulation results for normally distributed variables having a population Pearson correlation coefficient of .4 (*Rp* = .4). The population Spearman correlation coefficient (*Rs*) was calculated according to Equation 9. The figure shows the mean, 5th percentile (P5), and 95th percentile (P95) of the Pearson correlation coefficient (*rp*) and the Spearman correlation coefficient (*rs*) as a function of sample size (*N*).



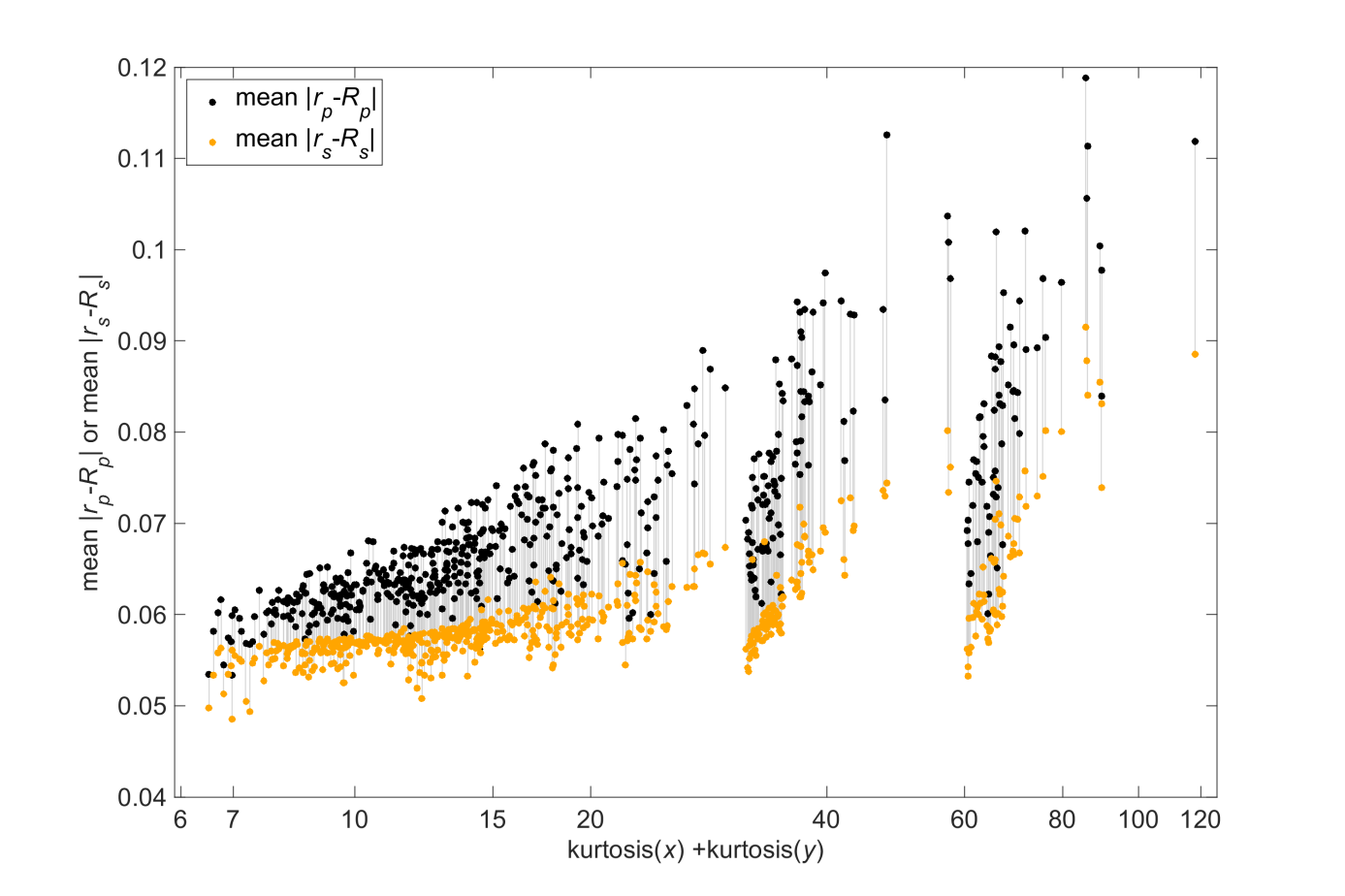
*Figure S5.* Simulation results for normally distributed variables having a population Pearson correlation coefficient of .8 (*Rp* = .8). The population Spearman correlation coefficient (*Rs*) was calculated according to Equation 9. The figure shows the mean, 5th percentile (P5), and 95th percentile (P95) of the Pearson correlation coefficient (*rp*) and the Spearman correlation coefficient (*rs*) as a function of sample size (*N*).

**Supplementary material explaining the behavior of *rp* and *rs* for non-normal distributions and nonlinear associations**

We explored the behavior of *rp* and *rs* for two correlated variables (*Rp* = .4) having a χ2 distribution. Figure S6 shows the standard deviation of *rp* as a function of sample size. It can be seen that the lower the degrees of freedom of the χ2 distributions (and hence the greater the skewness and kurtosis of the two variables), the more variable *rp* is. A χ2 distribution with 32 degrees of freedom closely resembles a normal distribution, which is why the standard deviations of *rp* and *rs* are almost the same in that case.

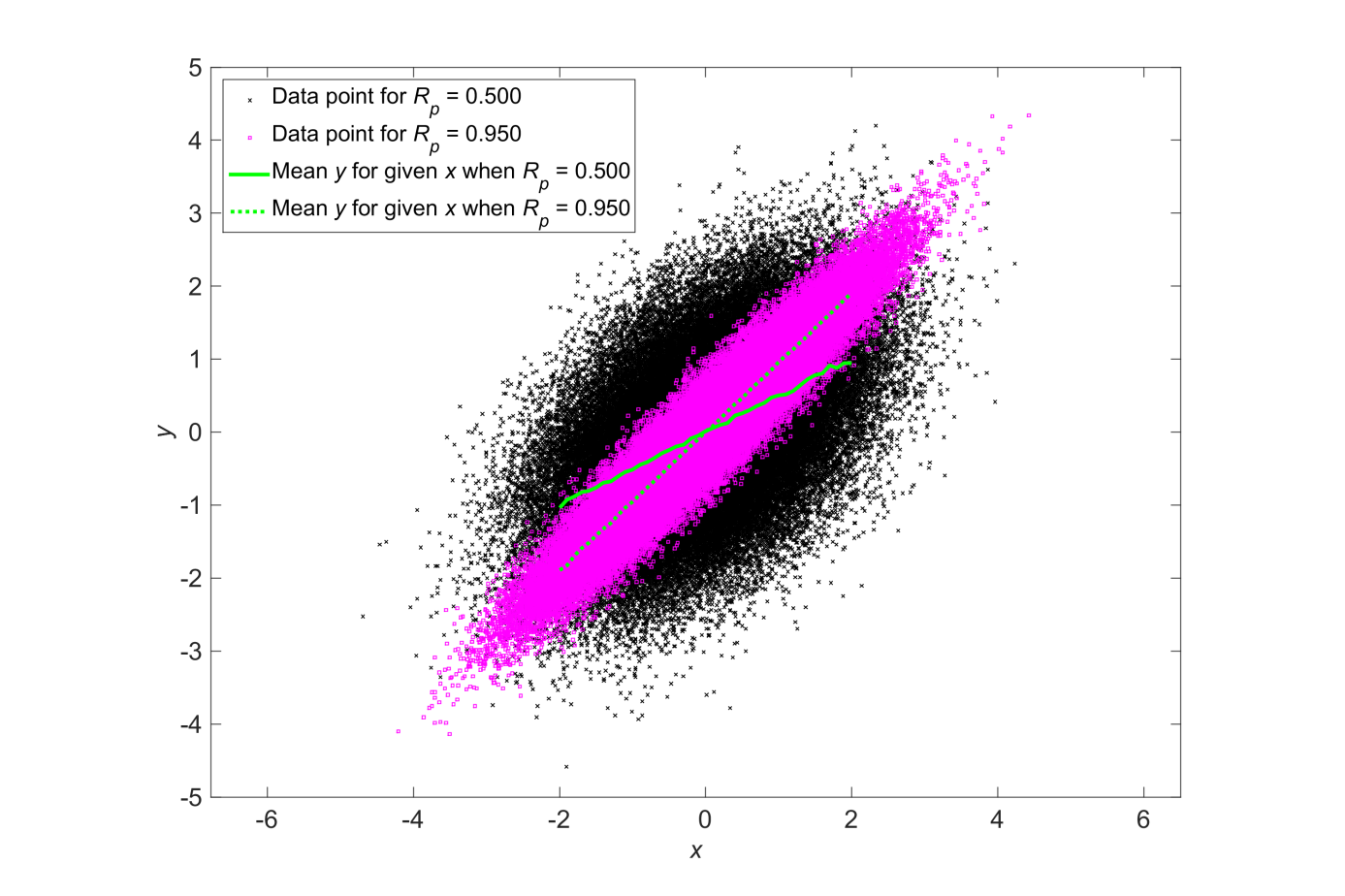


*Figure S6*. Standard deviation of *rp* and standard deviation of *rs* for three approximated χ2 distributions with different degrees of freedom (*df*) and population Pearson correlation coefficient of .4. The population skewness is 2.83, 2.00, and 0.50 for 1 *df*, *2f*, and 32 *df*, respectively. The population kurtosis is 15, 9, and 3.38 for 1 *df*, 2 *df*, and 32 *df*, respectively. A χ2 distribution with 2 *df* is an exponential distribution (see also Figures 3 & 4). The distributions were created using a method by Headrick (2002).



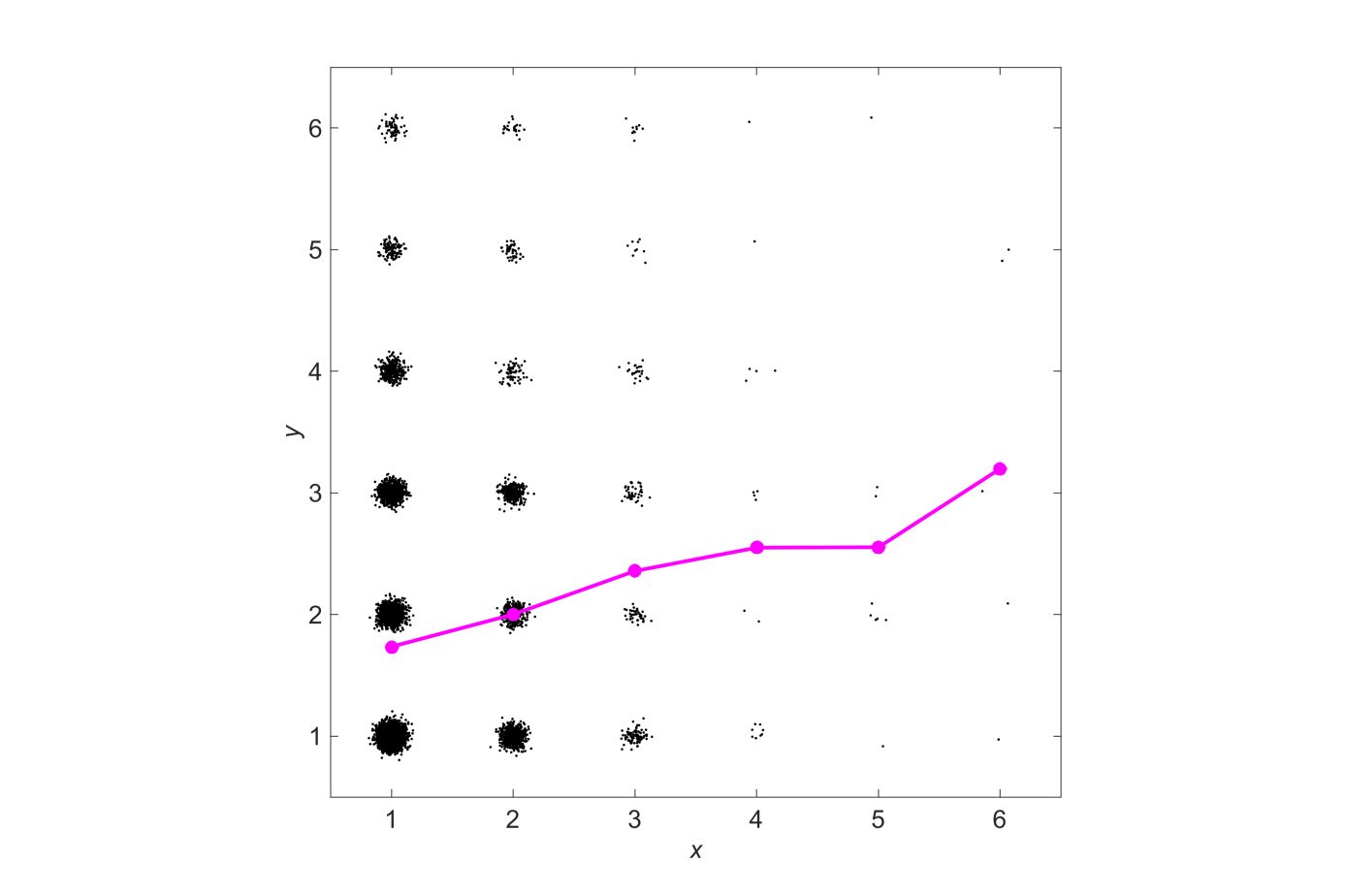
*Figure S7*. Mean absolute differences between the sample Pearson correlation coefficient (*rp*) and the population Pearson correlation coefficient (*Rp*) and mean absolute differences between the sample Spearman correlation coefficient (*rs*) and the population Spearman correlation coefficient (*Rs*), for pairs of variables (*x*, *y*) of the Driver Behaviour Questionnaire (DBQ) dataset as a function of the population kurtosis of *x* plus the population kurtosis of *y* (*N* = 9,077). The mean absolute differences for *rp* and *rs* are connected by a vertical line for each pair of variables. These results were based on 50,000 samples of *N* = 200. The *x*-axis is logarithmic.

When the two variables have a joint normal distribution, then the expected value of *y* for a given *x* is a linearly related to *x* (e.g., Bertsekas & Tsitsiklis, 2014). Figure S8 below illustrates an *Rp* of .95 and an *Rp* of .50 for two variables having a mean of 0 and a standard deviation of 1. The dotted (corresponding to *R* = .95) and solid (corresponding to *Rp* = .5) lines represent the means for a given *x*. The slopes of the lines are equal to the correlation coefficient.



*Figure S8*. Simulation of two normally distributed variables (*x* & *y*) having a population Pearson correlation coefficient *Rp =* .95 (*N* = 100,000) and *Rp* = .5 (*N* = 100,000). The lines represent the mean value of *y* for a given *x*, for *Rp* = .95 and for *Rp* = .5. The means are calculated for bins of *x* that are 0.1 wide.

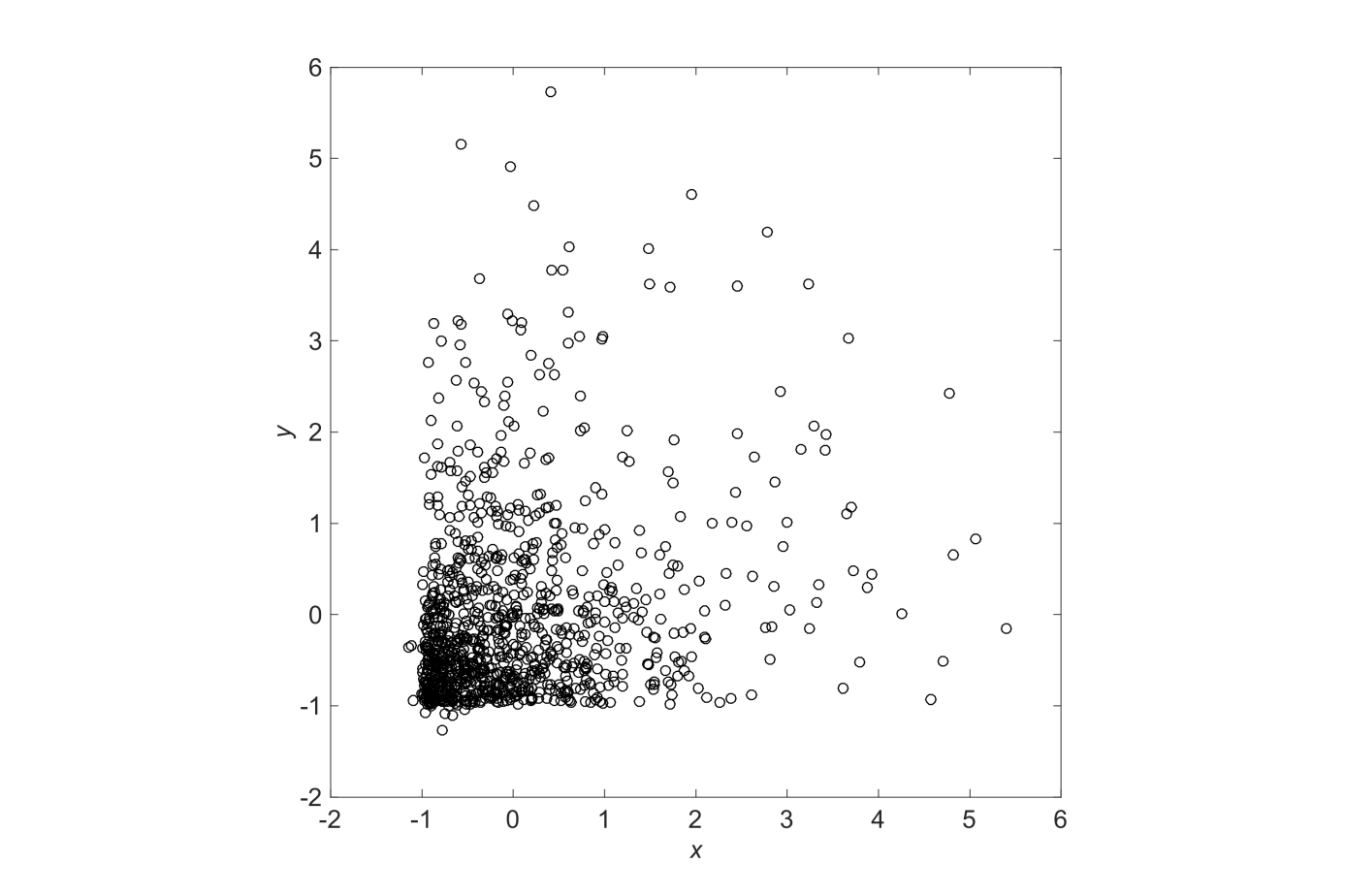
A nonlinear relationship between two variables can only occur when at least one of the two variables is non-normally distributed. However, even two highly skewed variables can be linearly related, if they have the same type of distribution. Figure S9 illustrates the relationship between two items of the Driver Behaviour Questionnaire (DBQ). Both variables are skewed (skewness of *x* = 2.85; skewness of *y* = 1.45; skewness of *x* = 14.5; skewness of *y* = 4.83), yet their relationship is approximately linear. In other words, when two variables are normally distributed, then their relationship is linear. But two highly skewed distributions are not necessarily nonlinearly related.

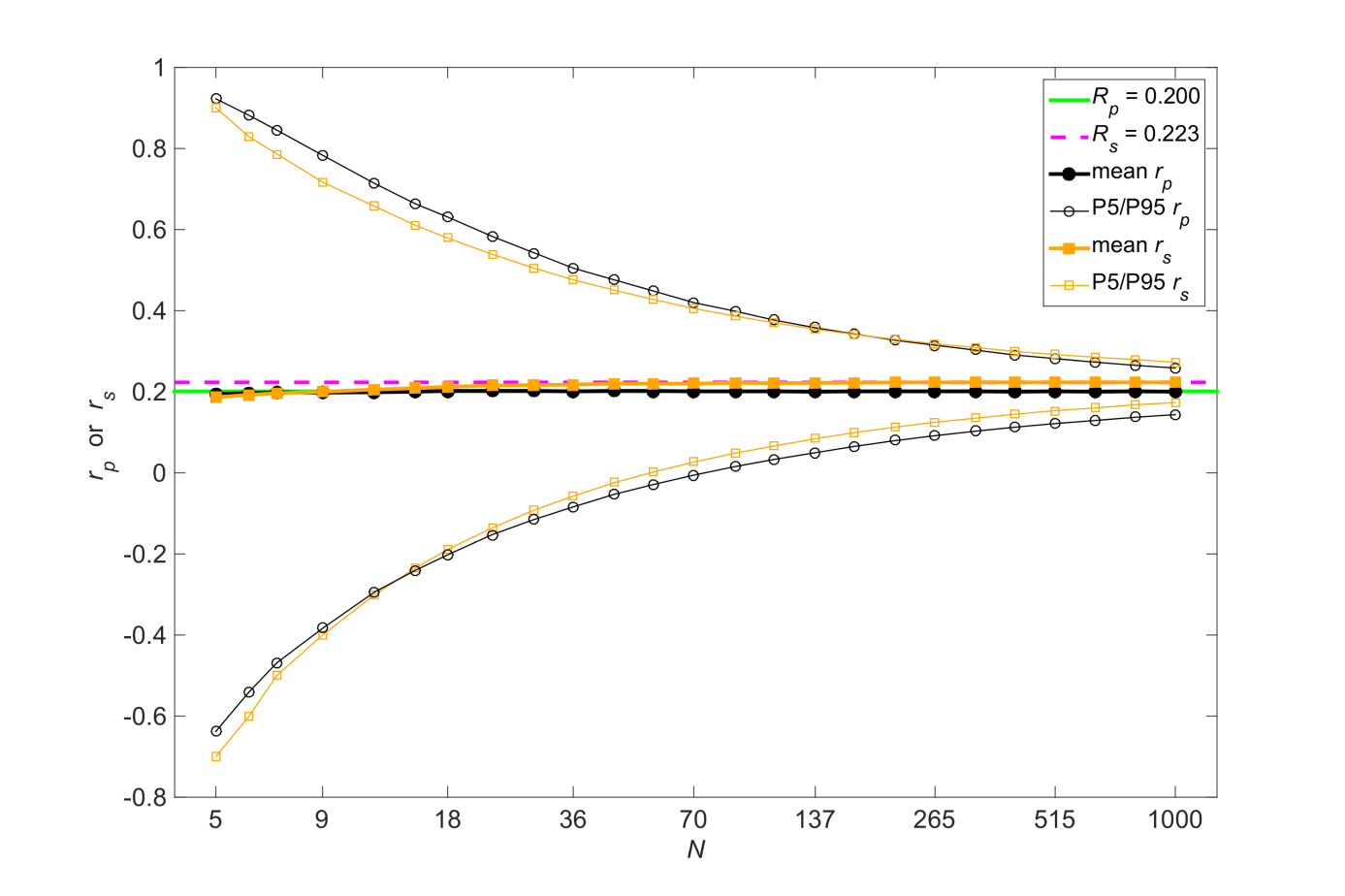


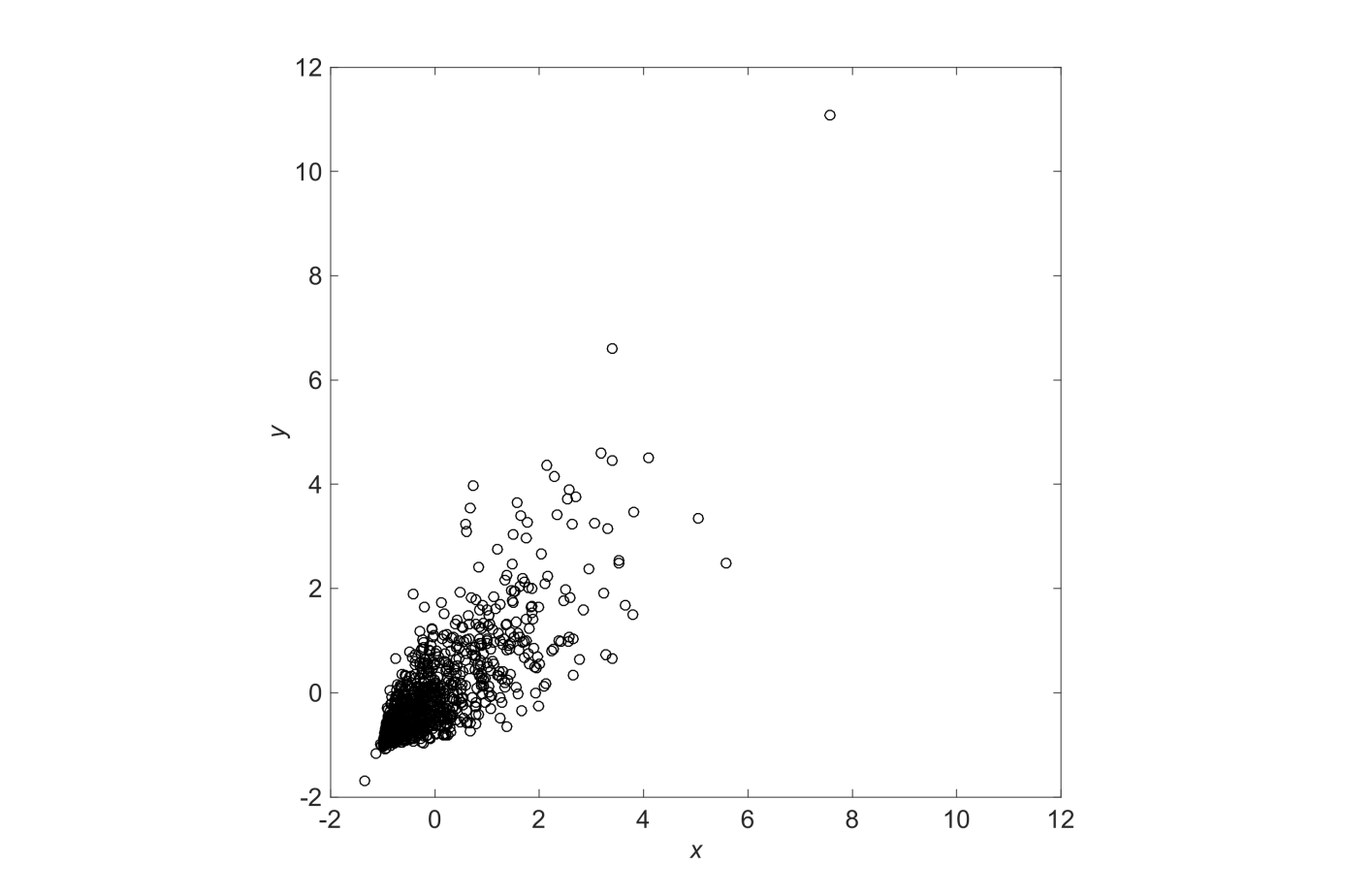
*Figure S9*. Relationship between two items of the Driver Behavior Questionnaire (*x* represents the responses to the item “Brake too quickly on a slippery road, or steer the wrong way into a skid”; *y* represents the responses to the item “Forget where you left your car in a car park”). Noise with a random distribution and a standard deviation of 0.05 is added for each response to prevent overlap of dots (*Rp* = .128; *Rs* = .119; *N* = 9,077). The line with markers represents the mean value of *y* for a given *x*.

We also carried out simulations to explore the effect of the population correlation coefficient (*Rp*) on the behavior of *rp* and *rs*. Figures S10 and S12 illustrate the relationship that we generated, with *x* and *y* being exponentially distributed, and *Rp* = .2 and *Rp* = .8. The simulation results are provided in Figures S11 and S13, respectively.

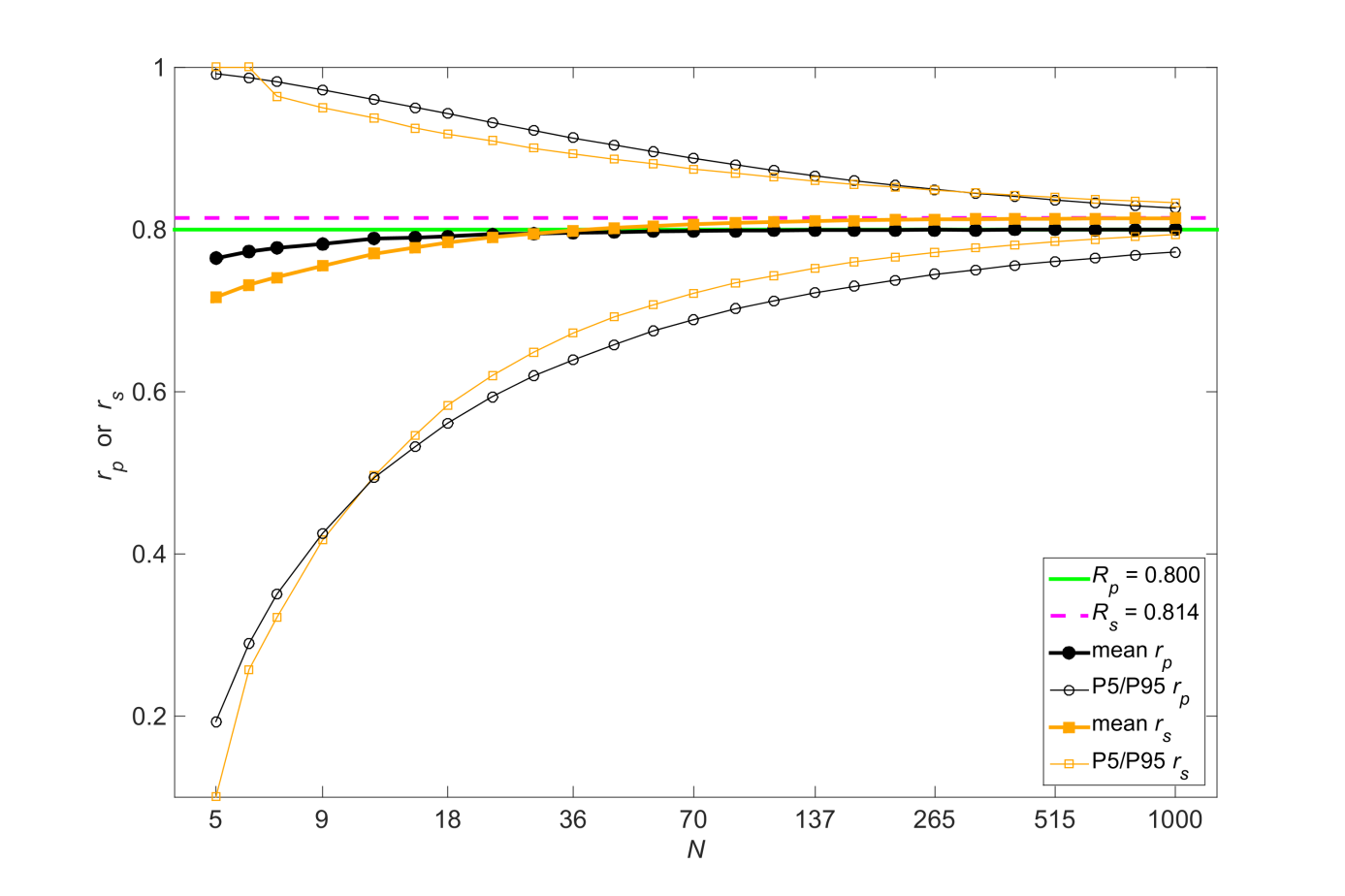
It is possible to devise nonlinear relationships where *rp* is considerably less variable than *rs*. Figures S14 and S15 show results for a nonlinear relationship where an exponential distribution is combined with a beta distribution having a negative skewness (−0.85). So, variables having high skewness or high kurtosis can still yield a stable *rp*.

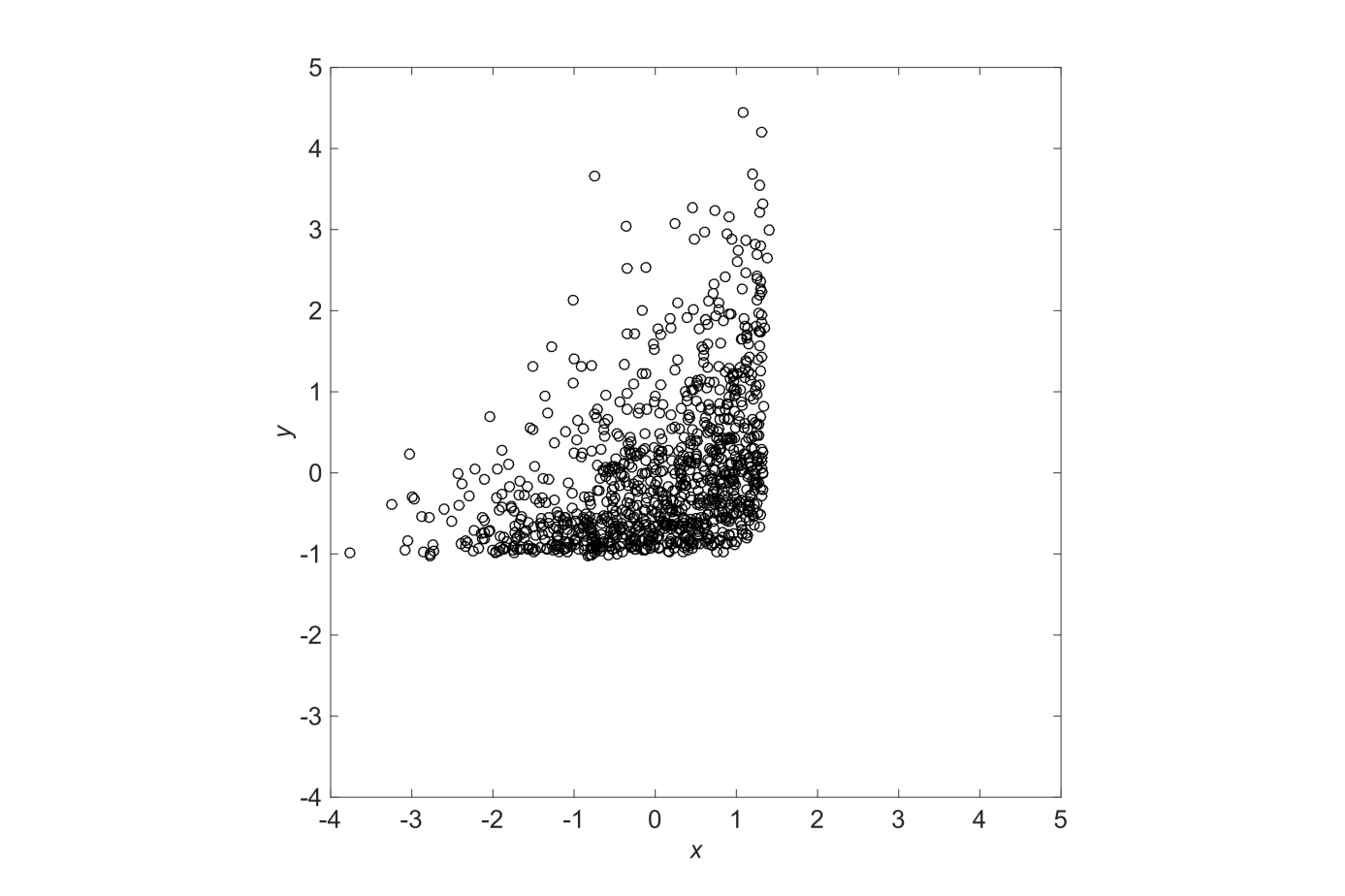
*Figure S10.* Depiction (using *N* = 1,000) of two correlated variables having an exponential distribution with population Pearson correlation coefficient (*Rp*) of .2. *Rp* was obtained by calculating *rp* for a sample of *N* = 107.

*Figure S11.* Simulation results for two correlated variables having an exponential distribution (see Figure S10 for a large-sample illustration of the distribution). The figure shows the mean, 5th percentile (P5), and 95th percentile (P95) of the Pearson correlation coefficient (*rp*) and the Spearman correlation coefficient (*rs*) as a function of sample size (*N*). The population coefficients *Rp* and *Rs* were obtained by calculating *rp* and *rs*, respectively, for a sample of *N* = 107.

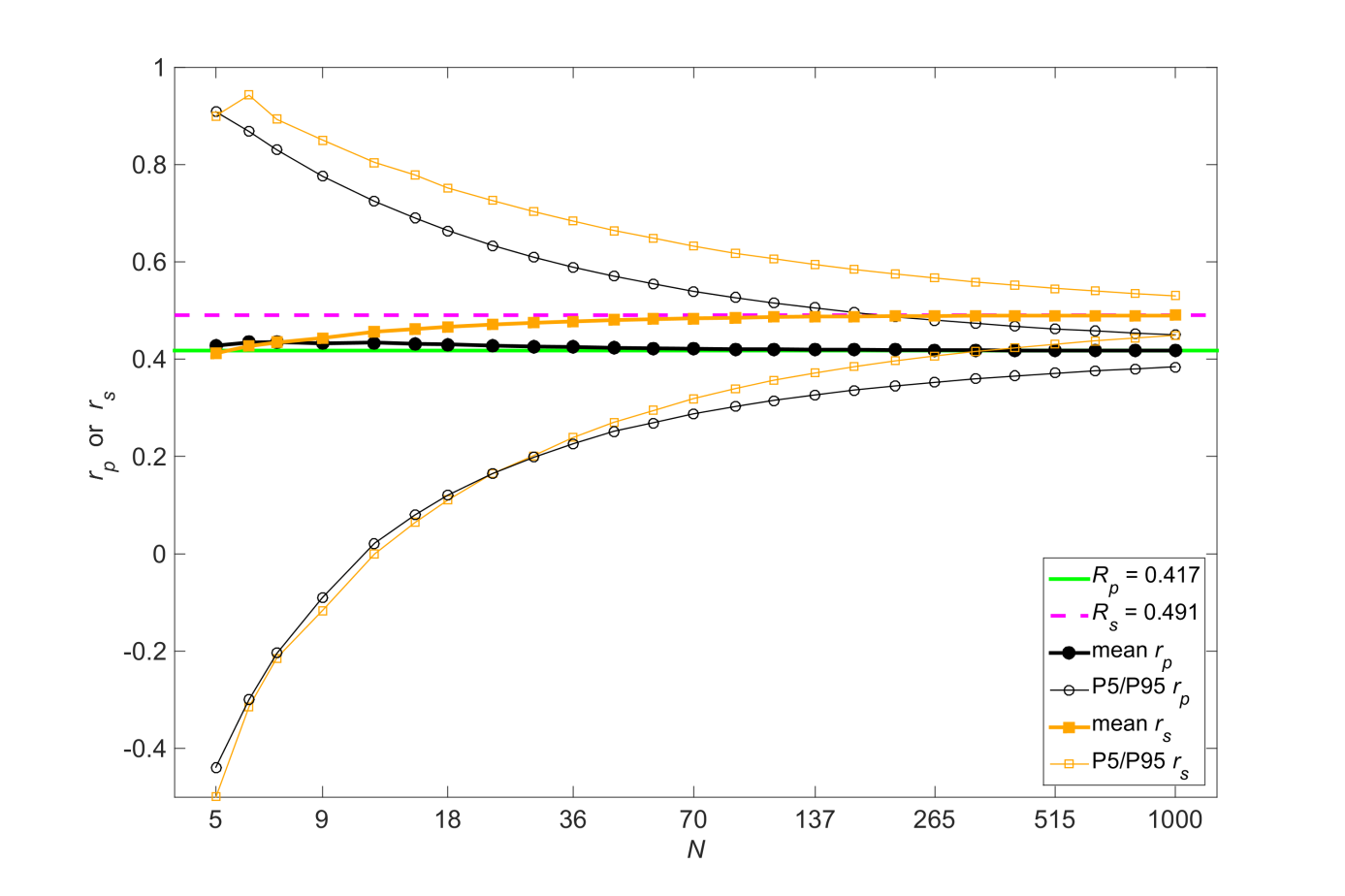


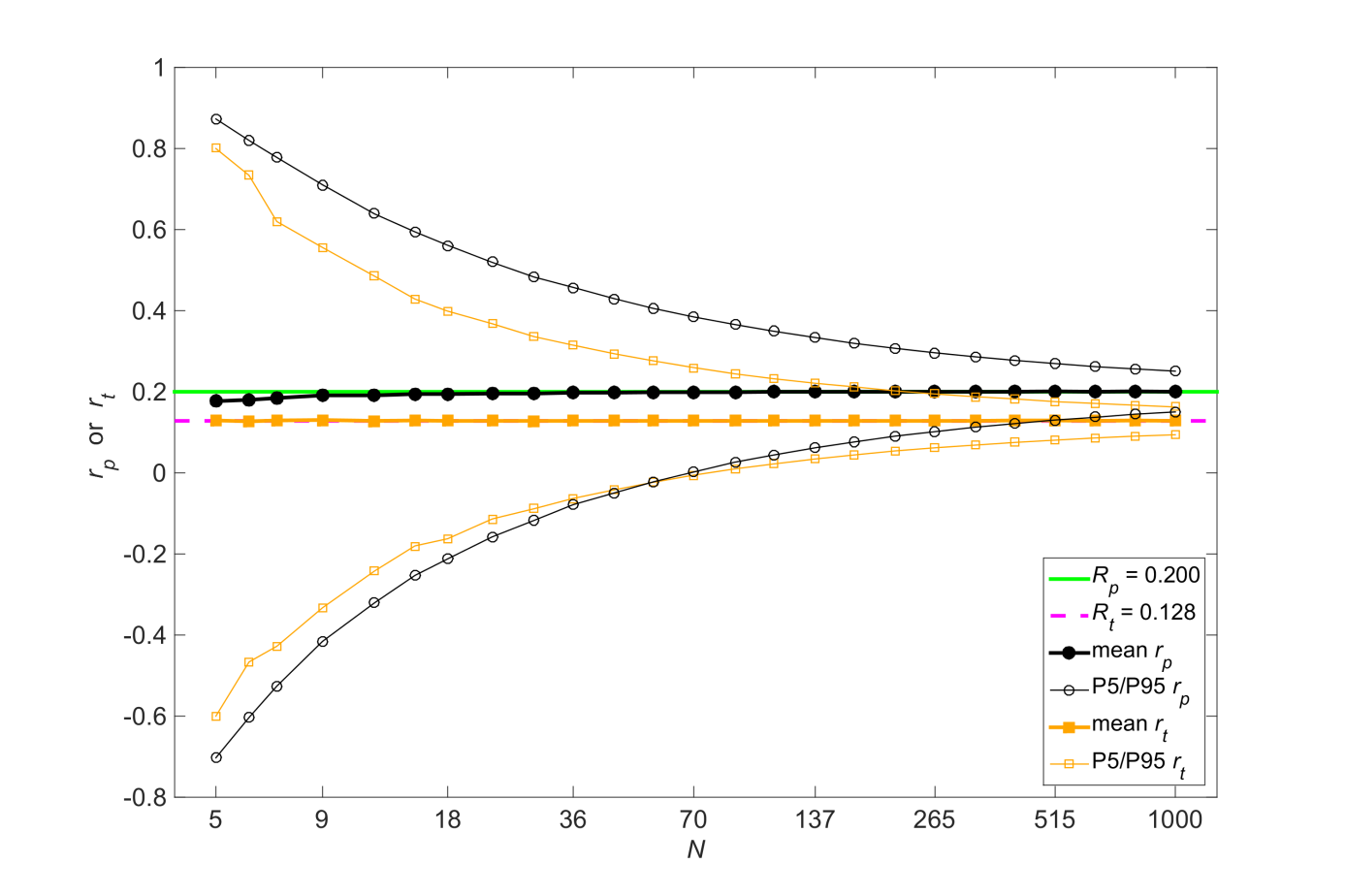
*Figure S12.* Depiction (using *N* = 1,000) of two correlated variables having an exponential distribution with population Pearson correlation coefficient (*Rp*) of .8. *Rp* was obtained by calculating *rp* for a sample of *N* = 107.

*Figure S13.* Simulation results for two correlated variables having an exponential distribution (see Figure S12 for a large-sample illustration of the distribution). The figure shows the mean, 5th percentile (P5), and 95th percentile (P95) of the Pearson correlation coefficient (*rp*) and the Spearman correlation coefficient (*rs*) as a function of sample size (*N*). The population coefficients *Rp* and *Rs* were obtained by calculating *rp* and *rs*, respectively, for a sample of *N* = 107.



*Figure S14.* Depiction (using *N* = 1,000) of a nonlinear relationship between two variables. The variable *x* has a population skewness of −0.85 and a population kurtosis of 3.22, whereas the variable *y* has a population skewness of 2 and a population kurtosis of 9 (*Rp* = .417). These population coefficients were calculated for a sample of *N* = 107.

*Figure S15.* Simulation results for a nonlinear relationship between two variables (see Figure S14 for a large-sample illustration of the distribution). The figure shows the mean, 5th percentile (P5), and 95th percentile (P95) of the Pearson correlation coefficient (*rp*) and the Spearman correlation coefficient (*rs*) as a function of sample size (*N*). The population coefficients *Rp* and *Rs* were obtained by calculating *rp* and *rs*, respectively, for a sample of *N* = 107.

*Figure S16*. Simulation results for normally distributed variables having a population Pearson correlation coefficient of .2 (*Rp* = .2). The figure shows the mean, 5th percentile (P5), and 95th percentile (P95) of the Pearson correlation coefficient (*rp*) and the Kendall tau rank correlation coefficient (*rt*) as a function of sample size (*N*). The population Kendall correlation coefficient (*Rt*) was calculated according to Equation 10.

**Reference**

Bertsekas, D. P., & Tsitsiklis, J. N. (2014). *The bivariate normal distribution*. Retrieved from http://athenasc.com/Bivariate-Normal.pdf