**Common Correlation and Reliability Analysis with SPSS for Windows**

**http://www.nyu.edu/its/statistics/Docs/correlate.html**

***Robert A. Yaffee*** *Statistics and Social Science Group  
Academic Computing Facility  
New York University*

**Keywords:** Correlation, Reliability, Validity, Alpha, Kappa, Intraclass correlation, correction for attenuation, chi-square, gamma, phi, tau, Somer's d, Spearman rho, multiple correlation, partial correlation, semipartial correlation, causation, Godel, coefficient of determination, triangulation.

**Introduction**

This presentation presents the theory of correlation analysis for correlations commonly used in the education of students in the social and natural sciences and then the programming application of it with the SPSS (Version 8 for Windows) statistical package.

In the social and natural sciences, researchers seek to understand and explain the nature of causal relations between phenomena. The phenomena are operationalized into measured relationships that are observed or tested. Hence, correlations serve as empirical indications of possible relationships between variables. Merely because a researcher discovers a correlation does not mean that he has proved the existence of a causal relationship.

Bivariate correlations control for neither antecedent variables nor intervening (mediating) variables. An antecedent variable may cause both of the other variables to change. Without the antecedent variable being operational, the two observed variables, which appear to correlate, may not do so at all. Therefore, it is important to control for the effects of antecedent variables before inferring causation. An intervening variable can also produce the apparent relationship between two observed variables, such that if the intervening variable were absent, the observed relationship would not be apparent. A spurious correlation may arise from failure to control for either of confounding relationships.

It has been argued that because the researcher cannot control for all possible confounding variables for lack of omniscience, that he cannot be sure that the apparent relationship is proven. If the apparent relationship cannot be proven, alas, absolute or the whole truth cannot be known. Nonetheless, we cannot be daunted by the human inability to appreciate the whole or absolute truth. If the subject matter is worth the study, then we need to continue to improve our understanding of it, to model, explain, predict, and possibly even engineer the control of it for the betterment of the human condition.

Correlations are relationships between two or more variables or sets of variables(Cohen and Cohen, 1983). They have three fundamental dimensions: significance, direction, and magnitude. These dimensions will be addressed throughout. The correlations differ with respect to the number and kinds of variables whose relationship is being studied.

The numbers of variables correlated may classify basic kinds of correlations. There are bivariate correlations and multiple correlations. Bivariate Correlations are correlations between two variables, whereas Multiple Correlations are those between one variable and a set of variables. Some bivariate correlations are nondirectional and these are called symmetric correlations. Other bivariate correlations are directional and are called asymmetric correlations. There are multiple correlations that hold part of the set of variables constant: Two of these kinds of correlations are part and partial correlations. The relationship of the partial correlation to the beta weights in regression analysis will be noted. There are also intraclass correlations which can be applied to the computation of reliability in repeated measures analysis of variance. For persons interested in the theory and computation of this, see [Enhancement of Reliability Analysis: Application of Intraclass Correlations with SPSS/Windows v.8](http://www.nyu.edu/its/socsci/Docs/intracls.html) by [Robert A. Yaffee.](mailto:yaffee@nyu.edu)

Less common correlations exist which SPSS does not at this point in time compute. Some of these may have particular applications that are from time to time important, for which reason they are noted here. There are set correlations between two sets of variables also. Right now, SPSS does not handle perform set correlations. Set correlations are defined by Cohen and Cohen (1983) and may be computed with SYSTAT. There are also polyserial and polychoric correlations, which are performed on binary and polichotomous variables and presume an underlying normal distribution. At this time, SPSS does not compute these, although PRELIS can compute such correlations. There are other dynamic or longitudinal correlations --such as the autocorrelation, the partial autocorrelation, and the cross-correlation function--which are beyond the scope of this cross-sectional treatment. Although SPSS Trends can compute the dynamic correlation functions mentioned, SPSS does not compute the other less common correlations. Therefore, we shall reserve such treatment for a more advanced analysis. We will confine ourselves to the kinds of correlations that SPSS can compute.

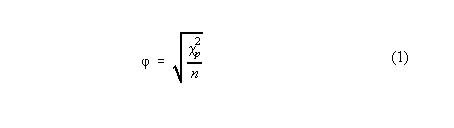
***Kinds of variables (Levels of Measurement):***

When studying correlations, it is necessary to know which correlations are applied to which kinds of variables. Bivariate Correlations are considered first. As for the kinds of variables, the kinds are dependent on the level of measurement of the answer categories to the questions that form the variables in the analysis. By level of measurement, the number of levels in the measurement of the variable is addressed. The variables are either discrete or continuous. Among the discrete variables are the dichotomous variables. These are binary coded. If these variables are coded 0 or 1, they are called dummy variables. There are also discrete variables which have an innate ordering to their measurement. There are ordinal level variables. For example, a Likert type scale may be measured 1 disagree strongly, 2 disagree, 3 unsure, 4 agree, and 5 agree strongly. Answer 1 is less than answer 2 but by how much is not known. Also, there are interval level or ratio level scales. An integer level scale has an ordered set of responses that are whole numbers. An example of an integer level scale may be the age of the respondent. There are also ratio level scales. These scales may have decimals to the right of the decimal point. Body temperature is an example of a ratio level scale. Two is twice one on such a scale. The relative magnitude is precise.

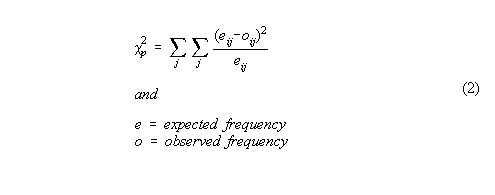
**When ordinal variables may be deemed continuous variables.** Joreskog and Sorbom in a set of Monte Carlo studies have found that ordinal scales that have 15 or more orderings may be considered continuous. For such scales, continuous level correlation coefficients may be used.

***Dichotomous variables*** may be correlated. The kind of correlation that is applied to two binary variables is the phi correlation. A correlation between two

dummy variables is a phi correlation. The phi correlation has a particular formula:

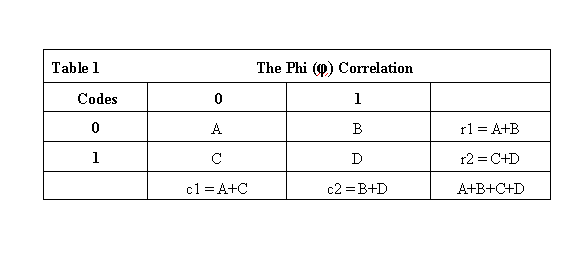


where

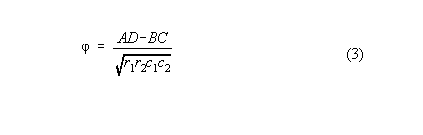


Alternatively, the phi correlation coefficient may be given as a function of the Pearson p 2 where n = sample size. This formula is adapted from SPSS 7.0 Algorithms where n is used instead of the sum of the weights.

If we consider Table 1:



The formula for the phi correlation may be computed from the cells designated accordingly by G. David Garson (1982) Political Science Methods. Boston: Holbrook, 288.



**Significance Test:**

The significance test of the phi is the same as that of a correlation: t = r [ ( n - 2)/ (1 - r 2 )] with df = n - 2. It is helpful to remember that if multiple tests are being conducted, then the procedurewise error rate is .05 for each test. To guard against an experimentwise error rate (which is a sum of the procedurewise errors) climbing to an unacceptably high level, a **Bonferroni correction** for multiple tests can be made. The Bonferroni correction divides the upper limit of the significance level of the individual test by the number of tests (alphaew = [alpha pw/p]){with p = number of coefficients tested in experiment and ew=experimentwise and pw=procedurewise} being conducted in the overall experiment. If the upper significance level of a social science test is .05, and there are three tests (for example, correlation coefficients being tested for signficance), then .05 should be divided by 3. When .05 is divided by three, the adjusted significance level of each test is .017. If each of the three tests are conducted at the .017 level, then the overall experimentwise error rate will be kept within acceptable limits.

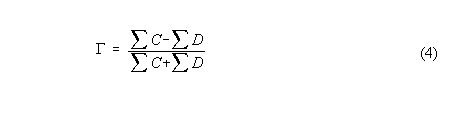
The **Sidak correction** is not quite as conservative as the Bonferroni adjustment. According to Sidak, the experimentwise error may be corrected by   
alpha ew = 1 - (1 - alphapw)1/p where ew=experimentwise, pw=procedurewise, and p = number of procedural tests. The working significance level of the Sidak adjustment tends to be slightly larger than that for the Bonferroni adjustments (Kirk,1982).

***Ordinal Variables***

There are correlations that are applied to two ordinal kinds of variables. These are typically nonparametric correlations. These correlation coefficients are distribution free and are usually applied to the ranks of the two variables. Examples are the Gamma, the Kendal, and the Spearman rank correlation. They measure monotonicity: Whether one variables changes in the same direction as the other variable, when changes from one case to the next is considered. If both variables change in the same direction, a concordance is found. If one variable changes in one direction while the other variable changes in the opposite direction, a discordance is found. The total number of concordances and the total number of discordances for all pairs of observations are counted. The Goodman and Kruskal Gamma and the Kendal's Tau a are nonparametric correlations that measure monotonicity from these counts. In this first section, we will address the symmetric version, for which no directionality is assumed, of these coefficients.

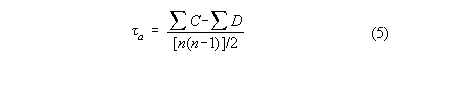
**Goodman and Kruskal Gamma**

The Gamma is a simple symmetric correlation. It tends to be of higher magnitude than the others. It does not correct for tied ranks. It is one of many indicators of monotonicity that may be applied. Monotonicity is measured by the proportion of concordant changes from one value in one variable to paired values in the other variable. When the change in one variable is positive and the corresponding change in the other variable is also positive, this is a concordance. When the change in one variable is positive and the corresponding change in the other variable is negative, this is a discordance. The sum of the concordances minus the sum of the discordances is the numerator. The sum of the concordances and the sum of the discordances is the total number of relations. This is the denominator. Hence, the statistic is the proportion of concordances to the total number of relations.



**Kendall's Tau a**

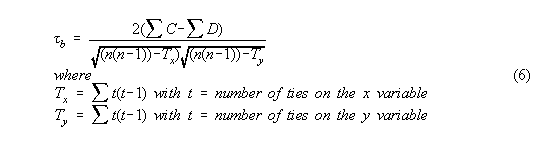
For symmetric tables, Kendall noted that the number of concordances minus the number of discordances is compared to the total number of pairs, n(n-1)/2, this statistic is the Kendall's Tau a:



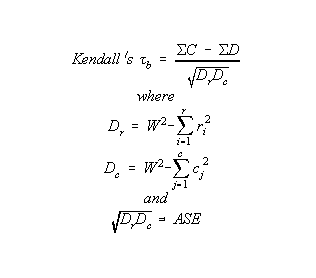
**Correlations that control for tied ranks**

**Kendall's Tau b**

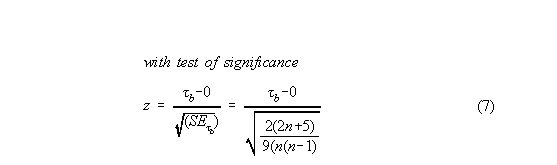
The Kendall's Tau b statistic controls for tied ranks.



It should be, however, noted that SPSS computes this statistic according to the formula on page 684 of the StatXact 4 User's Manual:



even if the usual test of significance comes from

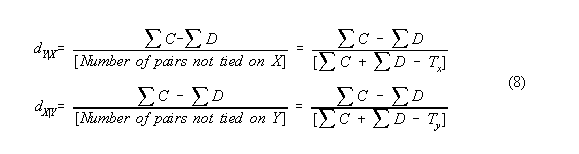
.

**Asymmetric Correlation Analysis**

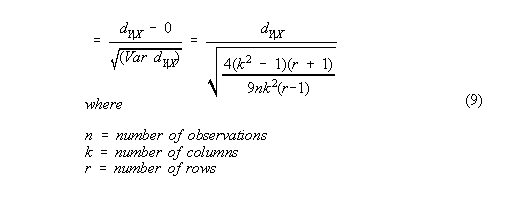
When asymmetric correlations are used, there is a presumed independent variable (iv) and a presumed dependent variable(dv). The notation is such that if K = the coefficient, then KDV|IV reveals the assumed direction of the relationship from the independent variable to the dependent variable.

**Asymmetric Somer's D**

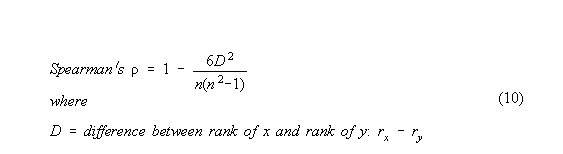
With specific Y and X as dependent variables, respectively:



**Significance Test for Somer's D:**



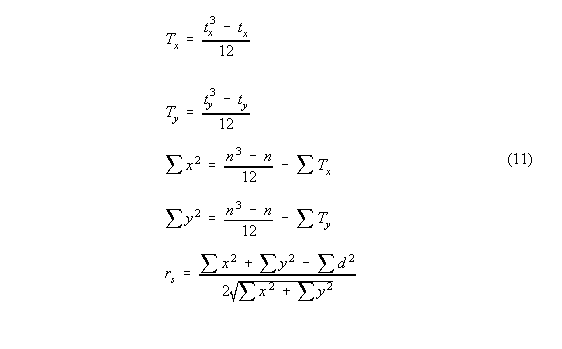
**Spearman rho** : Nonparametric correlation between two ordinal variables. Kendall showed that the Spearman and the Pearson Product Moment Correlation are equivalent.



When the proportion of ties is small, the ties may be ignored. When the proportion of ties is large, they may not be ignored. A correction for ties needs to be made. The correction for ties is performed by giving each tied rank the average score for all of the ranks that are tied.

**Spearman rho corrected for tied ranks**

Daniel, W. (1978) gives the formula for the Spearman correlation corrected for tied ranks.



It should be noted that Kendall discovered that the Spearman rho is equivalent to a Pearson Product Moment Correlation for continuous variables, which is now to be derived.

***Continuous Variables***

**The Pearson Product Moment Correlation**

Measures the extent to which one variable covaries with another. The correlation standardizes the two variables when it computes the covariance. Hence, the correlation is a standardized covariance. This correlation is conventionally a cross-sectional statistic. That is to say, it measures the relationship between two variables within a particular time period. There are longitudinal correlations, called cross-correlations, but they are beyond the scope of this presentation. We confine ourselves to the synchronic correlation coefficient.

Whereas the nonparametric correlations measured monotonicity, the Pearson product moment correlation measures the degree to which variables are related. There is a significance, direction, and magnitude to this correlation as well.

**Assumptions:**

1. interval level data

2. linearity (plot the relationship between the variables with a scattergram

or fit the functional curve formed by the relationship to be sure of linearity).

3. bivariate normality

4. homoskedasticity or equal variances: Truncated variances can attenuate the relationship

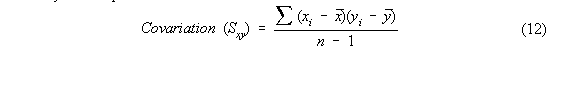
5. independence of observations

6. representative sampling. See (Garson, G.D.,(1999) [Assumptions of correlations per Garson](http://www2.chass.ncsu.edu/garson/pa765/correl.htm#assume)

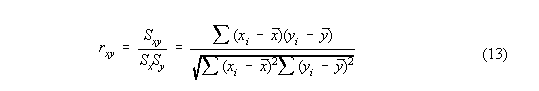
When assumptions are not met, 1) the researcher may transform the variables (with Box-Cox, natural log, Blom, or other transformations) so that the assumptions are met or 2) the researcher may employ nonparametric statistical measures instead.

**Covariance** of two variables: X and Y.

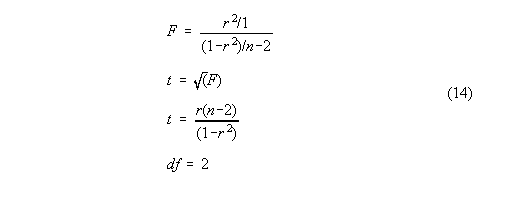
To understand the derivation of the Pearon Product Moment Correlation Coefficient, it is helpful to understand the formula for the covariance of two variables. One variable might be personal disposable income and the other variable might be personal consumption expenditures. The covariation between two variables can be measured by the sum of the product of their mean deviations divided by the sample size.



**Pearson Product Moment Correlation** consists of the covariation divided by the square root of the product of the standard deviations of the two variables.



While the significance test of the correlation can be computed from

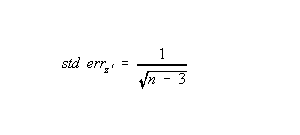


**Test for Difference in Magnitude between Two Independent Correlations**

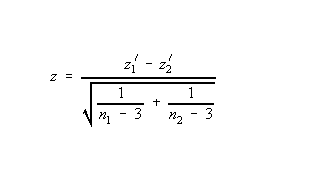
Often, it is necessary to determine whether one correlation is significantly different from another. To peform this test, one has to ascertain a Fisher z' transformation of the correlation. Each of the two correlations needs to be transformed as

z' = 1/2 [ ln(1 + r) - ln (1 -r)]

From these transformations, z1' and z2' are obtained. ln is the natural logarithm. That is the logarithm to the base e. Then it is necessary to compute the standard error for the Fisher's z transformation.



Often the difference is computed between different sized random samples. The differrence between the two transformed correlations is divided by the standard error to yield a normal curve deviate.



If this is greater than 1.96, then the difference between the correlations is significant at the .05 level.

**Coefficient of Determination: R 2** is the proportion of variance explained between two variables. The correlation coefficient that is usually squared can either be a bivariate Pearson Product Moment correlation or a multiple correlation. When squared, the coefficient when applied to a regression model represents the common variance explained by the predicor variables. If the coefficient squared is the semipartial correlation, the coefficient represents the explained variance added by the predictor variable added last. The coefficient of determination is a measure of the strength of the relationship between the predicted variable and model of the predictors in a regression model.

**Correlations between variables with different Levels of Measurement:**

***Dichotomous variable and dichotomous variable with underlying normal distribution***:

**1. Tetrachoric correlation:** Used with two dichotomous variables which assume underlying normal distributions.

***Dichotomous and interval:***

**1. Point biserial correlation** is computed with the formula for the Pearson Product Moment correlation. It has a range from negative unity to positive

unity.

**2. Biserial correlation.** Example according to Garson is a correlation of a dichotomous item in a scale with total scale score of individual. The maximum

value can sometimes exceed unity and there is no established test of significance.  
  
***Ordinal Variables and others***

**1. Polychoric correlation:** Two Ordinal Variables assume underlying normal Distributions. (Available in PRELIS not in SPSS)

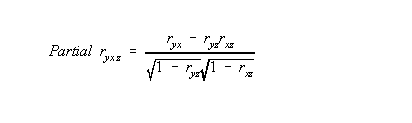
**2. Polyserial correlations:** One ordinal variable and one interval level variable. (Available in PRELIS, not in SPSS).

**Parametric Correlations that control for other variables** : Standard assumptions of Pearson Product Moment Correlations are required.

**1. Partial Correlation:** Shows relationship between x and y while holding z constant.

This correlation is applied to control for potentially confounding variables in correlation and

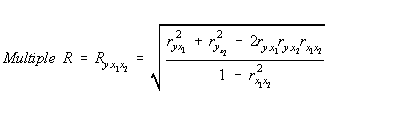
regression analysis. The formula for the partial correlation is



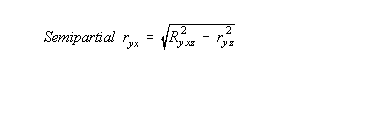
**2. Multiple Correlation:** Indicates the magnitude of the relationship between Y and the set of X variables (X1 and X2 for example)

The squared multiple correlation is called the coefficient of determination and is used to indicate the proportion of variance explained by the model.

The formula for the multiple correlation is

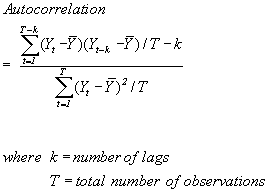


**3. Semipartial correlation:** When squared this correlation is the increment to the variance explained of a model by adding a variable.



**Time Series Analysis**

In Time Series Analysis, a series may be correlated with itself. Under conditions of stationarity (a dynamic equilibrium of mean, variance, and autocovariance), the **Autocorrelation** of a series is constant over the a particular time lag difference. The autocorrelation of a series is simply the correlation of the series with itself at a temporal distance of a particular number of lags of time period.



A **Partial Autocorrelation** is the autocorrelation of a series with itself under conditions of stationarity, while controlling for the effect of intervening lags. A partial autocorrelation reveals the precise autocorrelation of a series with itself without the confounding effects of intervening lagged autocorrelation.

A **Crosscorrelation** is the correlation between two different time series. This correlation is not symmetric. If xt leads yt, then the spikes in the cross-correlation function indicating a cross-correlation will point in one direction. If yt leads xt the spikes in the cross-correlation function will point in the opposite direction. The cross-correlation function is used to identify the direction of relationship between two time series.

An **Extended Autocorrelation Function** is a type of autocorrelation that is used to diagnose the order of an autoregressive moving average process.

**Validity:** The measurement of what is supposed to be measured. It is the extent of unbiasedness of a measure or set of indicators

**Face**: Apparent validity.

**Internal**: Did in fact the experimental treatments make a difference that was accurately and properly tested and measured (Campbell and Stanley, 1963).

**External**: Generalizability of the inferences.

**Concurrent**: Correlation of a measure with established, validated, and known criterion.

**Triangulation**: Multi-method or multi-instrumental analysis of items to reduce systematic error and to corroborate findings.

**Predictive**: correlation of prediction with actual.

**Construct**: A measure of all dimensions of the construct found in the literature. The concept contains all key theoretical components. Cronbach &

Meehl(1955) refer to this "approximate validity of generalizations to higher order constructs from research operations (Cook and Campbell, 1979).

Proportion of common variance explained by factor analysis of items is a good indicator, sometimes measured by omega.

**Discriminant**: Multi-method multi-trait correlations are used show that to that the construct may be differentiated from similiar constructs by one

or several methods.

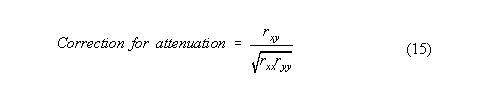
**Convergent**: Multi-method multi-trait correlations are used to show that different methods of measurement of a construct yield similar findings

and confirmation of measurement.

**Statistical conclusion validity**: "inferences about whether it is reasonable to presume covariation given a specified alpha level and the obtained

variances(Cook and Campbell, 1979)." For more on this subject, the reader should refer to the last section of this paper.

**Reliability:** the correlation between the observed variable and the true score when the variable is an inexact or imprecise indicator of the true score (Cohen and Cohen, 1983). Inexact measures may come from random inattentiveness, guessing, differential perception, recording errors, etc. on the part of the observers. These measurement errors are assumed to be random in classical test theory. Under such conditions, the reliability is the ratio of the true score to the observed score variance(Pedhazur and Schmelkin, 1991). In the event of inexact measurement, the correlation between two constructs is often corrected for attenuation(unreliability or imprecise measurement). The correction is computed by dividing the correlation between the measures by of the square root of the product of the reliabilities of the two variables. Reliability is a necessary but not a sufficient condition for validity (Pedhazur & Schmelkin, 1991). The question of measurement of reliability becomes important.



**Types of Reliability**

1. Parallel Measures reliability: Different measures have identical true scores and equal error variances.

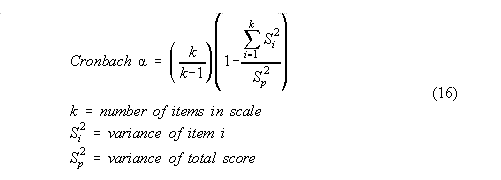
2. Tau Equivalent reliability: Different measures have identical true scores but need not have equal error variances.

3. Congeneric Measures reliability: Different measures have only perfect correlation among their true scores. Therefore, the measures need not have

identical error variances, true score means, nor true score errors.

**Measures of Reliability**

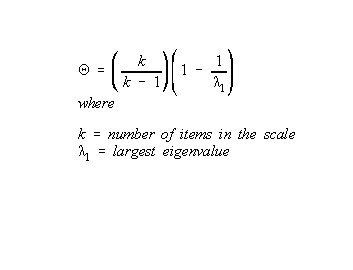
Internal Consistency: (of items in a scale): 1. **Average inter-item correlation** If average inter-item correlation > .6, then standardize items and add them together as an index. 2. **Cronbach's alpha** , which measures " internal consistency of items in a scale" Garson ,G.D.(1999) [Garson on correlations](http://www2.chass.ncsu.edu/garson/pa765/correl.htm)



from which it can be seen that alpha measures true variance over total variance. The range of the alpha is from 0 to 1.0. If the user obtains negative alphas, it means that his items are inconsistently coded. Consistent coding means all items have to be coded so that high values on the items correspond to high values on the total scale scores. If the item-total correlations are negative, then the coding of the items needs to be reviewed and corrected before computation of the alpha. According to J.C. Nunnelly (1998), the alpha of a scale should be greater than .70 for items to be used together as a scale (Nunnelly, 1978). The alpha for the total scale is also computed assuming that the item under examination is deleted. If the alpha increases over the current total scale alpha when an item is deleted, then the rule of thumb is to delete the item unless it is theoretically necessary for the analysis.

If the items are dichotomously coded, then the alpha is called a **Kuder-Richardson 20** coefficient. In this way, the internal consistency of the scale under construction can be enhanced.

If the items in a scale cohere together, then an alpha factor analysis will yield a single factor with loadings that maximize the alpha coefficient. D.J. Armor (1974) derives a **Theta reliability** coefficient that is an alpha analog. Theta reliability is designed to measure the internal consistency of the items (variables) in the first factor scale derived from a factor analysis. Theta is formulated as:

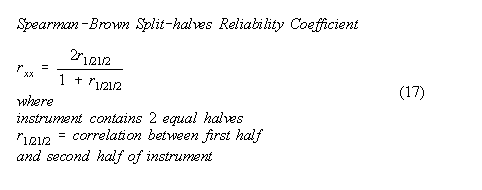


While SPSS does not yet compute the coefficient theta, it can be easily calculated from the factor analysis output produced by SPSS.

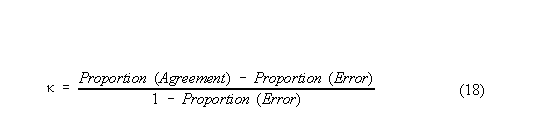
Another measure of reliability in factor analysis is the squared multiple correlation of a variable with that variable has multiple causes or when its error term is related to other error terms (Bollen, 1989).

Stability : This is consistency over time. Test-retest reliability coefficients measure stability

Homogeneity: Split-halves or parallel forms coefficients measure homogeneity. Average inter-item corrections may be employed to asses homogeneity. The number of factors in an alpha factor analysis might be an indicator of homogeneity. Pedhazur and Schmelkin (1991) give the formula for the Spearman Brown Split-halves Reliability coefficient as :



Inter-rater Agreement: **Cohen's kappa** is the ratio of the proportion of agreement (corrected for chance) divided by the maximum number of times they could agree (corrected for chance). SPSS provides a two rater as part of the crosstabs procedure; the generalized kappa for more than two raters is not computed by SPSS. Siegel and Castellan (1988) supply the formula:



**Intraclass correlations** , which are applied when multiple raters judge the same phenomena, are expounded by Shrout and Fleiss, 1978. There is a family of intraclass correlations. The objects being rated are usually considered random, but the raters can be either random or fixed effects. The covariance of the ratings is divided by a form of the total variance to obtain this sort of correlation coefficient. The formula used depends upon whether the ratings are single ones or averages of the raters. It also depends upon whether the raters are deemed fixed or random, and whether the rater variance is included in the denominator of the ratio of the rater variance to the total variance. The principal formulae and programming in SPSS are provided by [Yaffee(1998).](http://www.nyu.edu/its/socsci/Docs/intracls.html) The intraclass correlation are available in versions eight and nine of SPSS for Windows.

**Caveats of Correlation Analysis**

1. Correlation (association) does not prove causation

covariation is necessary but not sufficient for causation; there are other requirements also.  
2. Proof of causation requires temporal sequence of covariation to be demonstrated according to J.S. Mill (Cook and Campbell,1979).

some covariations may be delayed

some covariations may appear only after a threshold is reached

some covariations may require multiple precursors for them to appear

some covariations require interaction (moderation) of other variables

some covariations require intervening (mediating) variables

some covariations are products of antecedent variables  
3. Correlation is nondirectional or bidirectional association

there is no established directionality of the relationship

proof of causation requires directionality be identified

causation presumes unidirectional causal influence from cause to effect though there may

be simultaneity, reciprocal causation, or feedback  
4. Covariation does not demonstrate necessary and sufficient conditions needed to prove causation

There is often only naturalistic observation.

It may be necessary to distinguish necessary precursors from

sufficient conditions for causality to be established.

Without random selection, random assignment to control and

experimental groups, observation before and after the intervention, with proper controls

against factors that corrupt internal and external validity, causation cannot be proven.  
5. There is a need to control for all known plausible alternative predictors before causation can be proven. It may be necessary to eliminate all plausible

alternative explanations before causality is completely established.  
6. Spatial contiguity may be necessary to demonstrate as well causal relationship may require cross-cultural and trans-historical validation before they are

accepted as established.  
7. It is impossible to know for sure that all plausible antecedent or intervening variables are being controlled for, unless or until they are all known. Hence, we

cannot be sure that the relationships observed are other than mere associations. We only successively eliminate these plausible alternative explanations

as we think of them, thereby only incrementally approaching ( by successively falsifying or disconfirming (K. Popper) rather than confirming hypotheses)

an understanding of the whole truth. This is an endless quest rather than an accomplished conquest. Therefore, it is in reality impossible to have complete

proof and hence the absolute or whole truth.   
8. Consistency and completeness are incompatible in the proof

(as that famous Austrian logician, Kurt Godel, has shown). Hence, the complete

truth with certainty may in principle be unknowable.

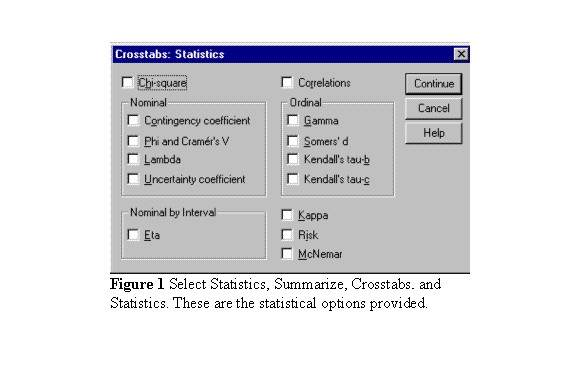
**Causal Modeling**

Causal modeling may be performed with correlations, as standardized regression (correlation) between directly observed variables or scales. Older methods developed employed this approach. Newer methods employ path analysis between such variables with unstandardized regression coefficients. More modern methods using structural equation modeling are regressions among latent variables or factors. These methods seek to model, and hence control for, antecedent and intervening variables. The newer methods can handle reciprocal as well as undirectional relationships as well. They model the paths between variables, whether directly observed or latent. In so doing, they reveal the causal structure of the model. Traditional causal models are estimated with ordinary least squares and can only be properly conducted where the variables are continuous or integer level. More advanced programs-- such as, LISREL and LISCOMP-- can handle dichotomous, ordinal or truncated variables with asymptotic distribution free estimation. Programs that perform these advanced kinds of analysis include LISREL, EQS, AMOS, RAMONA, SAS Proc CALIS, etc.. While modeling interaction terms is cumbersome, it is possible.

**Programming the SPSS Correlations**

With the Crosstabs Procedure

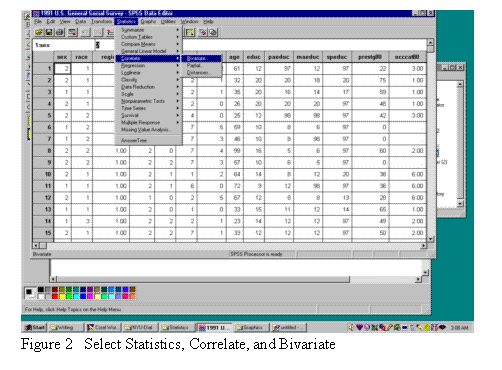
After setting the dependent variable in the column and the independent variable in the row, one clicks on statistics, and selects the appropriate statistical procedure.

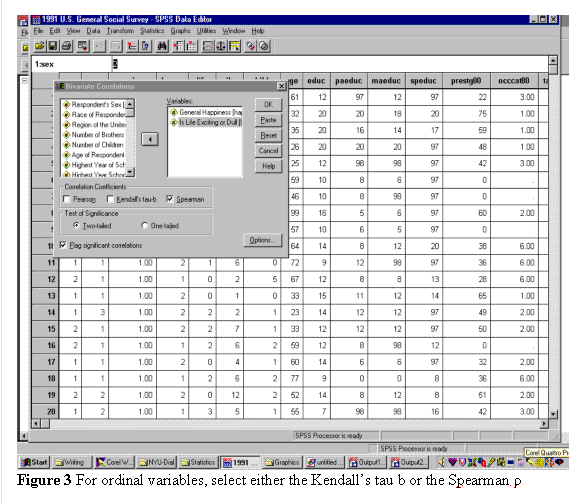


**With the Correlations procedure**

With the correlation procedure, first go to the statistics option in the header bar. Select correlate and then select bivariate:

After selecting the bivariate correlation, proceed to the statistics option and select the appropriate statistical procedure. With 2 ordinal variables, the Spearman is selected.





**Statistical Conclusion Validity**

The statistical analysis should have **adequate statistical power** . Preliminary power and **sample size** analysis should be performed for every statistical test to be conducted in the course of the research. To obtain adequate statistical sensitivity to real differences, sufficient sample size should be determined. Data collection should be designed to acquire that sample size prior to and acquired prior to the performance of the final analysis.

**Proper sampling** should be performed to be sure that the sample representative of the target population. If this cannot be done, then replication of the study in different settings along with a meta-analysis of the results before the study is undertaken should be planned. Probability sampling helps avoid selection bias and random heterogeneity of respondents in small samples. Proper sample weights facilitate generalization to the target population.

Good **reliability of measures** is necessary to be sure that the measures are stable and adequate.

Good **relability of treatment implementation** in experiments also should be monitored and quality assured.

Good **control of the environmental setting** helps exclude the intrusion and adverse effects of random irrelevancies within that milieu.

**Testing for the fulfillment of assumptions** of the statistical tests employed is necessary to be sure of the statistical conclusion validity.

Proper **control for the inflation of the error rate** from repeated comparison testing contributes to the maintenance of statistical conclusion validity as well (Cook and Campbell, 1979).

**References:**

Armor, D.J. (1974). Theta Reliability and Factor Scaling. In H.L.Costner(Ed.), Sociological Methodology 1973-1974. San Francisco: Jossey-Bass, 28.  
Bollen, K. (1989). Structural Equations with Latent Variables. New York: John Wiley, p.221.  
Campbell, D.T.& Stanley, J.C.(1963). Experimental and Quasi-Experimental Designs for Research. Chicago: Rand-McNally, 5.  
Cohen, J. & Cohen, P. (1983) Multiple Regression/Correlation for the Behavioral Sciences 2nd ed. : Hillsdale, NJ: Erlbaum Associates, 67-69, 490-497.  
Cook, T. & Campbell, D.T.(1979). Quasi-Experimentation: Design and Analysis Issues for Field Settings. Boston: Houghton-Mifflin Co. 20, 31, 39-47.  
Daniel, W. W. (1978). Applied Nonparametric Statistics: Boston: PWS-Kent, 262.  
Garson, G.D. (1982) Political Science Methods. Boston: Holbrook, 288.  
Garson, G.D. (1999) World Wide Web URL: http://www2.chass.ncsu.edu/garson/pa765/reliab.htm#intraclass   
Kirk, R. (1982) Experimental Design Procedures for the Behavioral Sciences. 2nd ed.Belmont, Calif.: Brooks/Coles. 101-107.  
Joreskog, K. and Sorbom, D. (1988) PRELIS 2 manual. Scientific Software. Evanston, Ill., 1-2 through 1-11.  
McGraw, K.O. & Wong, S.P.(1996). Forming Inferences About Some Intraclass Correlation Coefficients. Psychological Methods. Vol.1, No.1., 30-46.  
Mehta, C.& Patel, N. (1998-1999) StatXact 4 User's Manual. Boston: Cytel Software, Inc., 684.  
Nunnelly, J.C. (1978) Psychometric Theory 2nd ed. New York: McGraw Hill.  
Pedhazur, E.J. & Schmelkin, L.P.(1991). Measurement, Design, and Analysis. Hillsdale,NJ: Lawrence Erlbaum Publishers, 74-117.  
Shrout, P.E. & Fleiss, J.L. (1979). Intraclass Correlations: Uses in Assessing Rater Reliability, Psychological Bulletin, Vol. 86, 2, 420-428.  
Siegel, S. & Castellan, N.J.Jr. (1988). Nonparametric Statistics for the Behavioral Sciences. 2nd ed. New York: Wiley, 285.  
SPSS, Inc. (1996) Statistical Algorithms version 7. Chicago, Ill: SPSS Inc, 73-74, 100-118.  
Yaffee, R.A. (1998). Enhancement of Reliability Analysis: Applications of the Intraclass Correlation with SPSS v 8. March 11, 1998. World Wide Web URL:

http://www.nyu.edu/acf/socsci/Docs/intracls.html

Date Updated: 6 June 2003   
This article has been accessed http://www.nyu.edu/bin/Count.cgi?df=ray1.dattimes since the counter was installed on 7/30/99.