Paper 331-2012

# Exploratory Factor Analysis with the World Values Survey

Diana Suhr, Ph.D.
University of Northern Colorado

## Abstract

Exploratory factor analysis (EFA) investigates the possible underlying factor structure (dimensions) of a set of interrelated variables without imposing a preconceived structure on the outcome (Child, 1990). The World Values Survey (WVS) measures changes in what people want out of life and what they believe. WVS helps a worldwide network of social scientists study changing values and their impact on social and political life. This presentation will explore dimensions of selected WVS items using exploratory factor analysis techniques with SAS® PROC FACTOR. EFA guidelines and SAS code will be illustrated as well as a discussion of results.

## Introduction

Exploratory factor analysis investigates the possible underlying structure of a set of interrelated variables. This paper discusses goals, assumptions and limitations as well as factor extraction methods, criteria to determine factor structure, and SAS code. Examples of EFA are shown using data collected from the World Values Survey.

The World Values Survey (WVS) has collected data from over 57 countries since 1990. Data has been collected every 5 years from 1990 to 2010 with each data collection known as a wave. Selected items from the 2005 wave will be examined to investigate the factor structure (dimensions) of values that could impact social and political life across countries. The factor structure will be determined for the total group of participants. Then comparisons of the factor structure will be made between gender and between age groups.

## Limitations

Survey questions were changed from wave to wave. Therefore determining the factor structure for common questions across waves and comparisons between waves was not possible. The most recent data available, for 2005, was analyzed to provide examples for this paper.

The examples shown in this paper, include item responses with a 7-point Likert scale, items responses with a 5-point Likert scale, and items responses with a 4-point Likert scale while other item responses are categorical, e.g. gender, age groups. Common practice to determine factor structure examines items with the same response scale. The examples in this paper provide insight into determining factor structure across different response scales.

## Exploratory Factor Analysis

Exploratory factor analysis could be described as orderly simplification of interrelated measures. Traditionally factor analysis has been used to explore the possible underlying structure of a set of interrelated variables without imposing any preconceived structure on the outcome (Child, 1990). By performing exploratory factor analysis (EFA), the number of constructs (dimensions) and the underlying factor structure are identified.

Psychologists searching for a neat and tidy description of human intellectual abilities lead to the development of factor analytic methods. Galton, a scientist during the 19$^{th}$ and 20$^{th}$ centuries, laid the foundations for factor analytic methods by developing quantitative methods to determine the interdependence between 2 variables. Karl Pearson was the first to explicitly define factor analysis. In 1902, Macdonnell was the first to publish an application of factor analysis, a comparison of physical characteristics between 3000 criminals and 1000 Cambridge undergraduates (Child, 1990).

### Exploratory Factor Analysis

- is a variable reduction technique which identifies the number of latent constructs (dimensions) and the underlying factor structure of a set of variables
- hypothesizes underlying constructs, variables not measured directly
- estimates factors which influence responses on observed variables
- allows you to describe and identify the number of factors (dimensions or latent constructs)
- includes unique factors, error due to unreliability in measurement
- traditionally has been used to explore the possible underlying factor structure of a set of measured variables without imposing any preconceived structure on the outcome (Child, 1990).

**Goals** of exploratory factor analysis are

1) to help an investigator determine the number of latent constructs underlying a set of items (variables)
2) to provide a means of explaining variation among variables (items) using a few newly created variables (factors), e.g., condensing information

3) to define the content or meaning of factors, e.g., latent constructs

**Assumptions** underlying exploratory factor analysis are

- Interval or ratio level of measurement
- Random sampling
- Relationship between observed variables is linear
- A normal distribution (each observed variable)
- A bivariate normal distribution (each pair of observed variables)
- Multivariate normality

**Limitations** of exploratory factor analysis are

- The correlations, the basis of factor analysis, describe relationships. No causal inferences can be made from correlations alone.
- the reliability of the measurement instrument (avoid an instrument with low reliability)
- sample size ( larger sample → larger correlation)
  - ◊ minimal number of cases for reliable results is more then 100 observations and 5 times the number of items
  - ◊ since some subjects may not answer every item, a larger sample is desirable. For example, 30 items would require at least 150 cases (5*30), a sample of 200 subjects would allow for missing data
- sample selection
  - ◊ Representative of population
  - ◊ Do not pool populations
- variables could be sample specific, e.g., a unique quality possessed by a group does not generalize to the population
- nonnormal distribution of data

**Statistical Background**

Exploratory factor analysis (EFA) decomposes an adjusted correlation matrix. Variables are standardized in EFA, e.g., mean=0, standard deviation=1, diagonals are adjusted for unique factors, 1-u. Squared multiple correlations (SMC) are used as communality estimates on the diagonals.

The amount of variance explained is equal to the trace of the decomposed adjusted correlation matrix, the sum of the adjusted diagonals or communalities. Observed variables are a linear combination of the underlying and unique factors. Factors are estimated, (X1 = b1F1 + b2F2 + . . . e1 where e1 is a unique factor).

Eigenvalues indicate the amount of variance explained by each factor. Eigenvectors are the weights that could be used to calculate factor scores. In common practice, factor scores are calculated with a mean or sum of measured variables that "load" on a factor.

The EFA Model is $Y = X\beta + E$

where Y is a matrix of measured variables
X is a matrix of common factors
$\beta$ is a matrix of weights (factor loadings)
E is a matrix of unique factors, error variation

Communality is the variance of observed variables accounted for by a common factor. A large communality value indicates a strong influence by an underlying construct. Community is computed by summing squares of factor loadings

$d_1^2$ = 1 – communality = % variance accounted for by the unique factor
$d_1$ = square root (1-community) = unique factor weight (parameter estimate)

Figure 1 below shows 4 factors (circles) each measured by 3 observed variables (rectangles) with unique factors (error). Since measurement is not perfect, error or unreliability is estimated and specified explicitly in the diagram. Factor loadings (parameter estimates) help interpret factors. Loadings are the correlation between observed variables and factors, are standardized regression weights if variables are standardized (weights used to predict variables from factor), and are path coefficients in path analysis. Standardized linear weights represent the effect size of the factor on variability of observed variables.

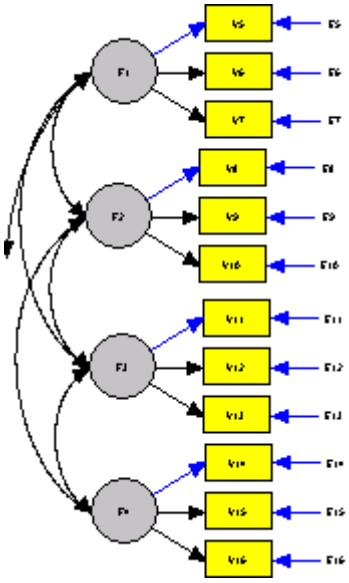


Figure 1 Factor Analytic Model

## Exploratory Factor Analysis Procedures

### Steps in exploratory factor analysis

1) Reliability analysis, reverse item scales if needed
2) Set variance levels a priori, proportion of variance and cumulative variance
3) Initial extraction
   - each factor accounts for a maximum amount of variance that has not previously been accounted for by any of the other factors
   - factors are uncorrelated
   - eigenvalues represent the amount of variance accounted for by each factor
4) Determine number of factors to retain
   - scree test, look for elbow
   - proportion of variance
   - cumulative variance
   - prior communality estimates are not perfectly accurate, cumulative proportion must equal 100% so some eigenvalues will be negative after factors are extracted, e.g., if 5 factors are extracted, cumulative proportion equals 100% and items 6 and above have negative eigenvalues
   - interpretability
     - at least 3 observed variables per factor for significant factors
     - common conceptual meaning
     - each factor measures a different construct or dimension
     - rotated factor pattern has simple structure (no cross loadings)
5) Rotation – a transformation
6) Interpret solution
7) Calculate factor scores
8) Results in a table
9) Prepare results, paper

### Reliability Analysis Prior to EFA Analysis

Reliability refers to the accuracy and precision of a measurement procedure (Thorndike, Cunningham, Thorndike, & Hagen, 1991). Reliability may be viewed as an instrument's relative lack of error. In addition, reliability is a function of properties of the underlying construct being measured, the test itself, the groups being assessed, the testing environment, and the purpose of assessment. Reliability answers the question, "How well does the instrument measure what it purports to measure?".

Some degree of inconsistency is present in all measurement procedures. The variability in a set of item scores is due to the actual variation across individuals in the phenomenon that the scale measures, made up of true score and error. Therefore, each observation of a measurement (X) is equal to true score (T) plus measurement error (e), or X = T + e.

Reliability can be assessed by internal consistency – measured with Cronbach's coefficient alpha. Internal consistency is a procedure to estimate the reliability of a test from a single administration of a single form. Internal consistency depends on the individual's performance from item to item based on the standard deviation of the test and the standard deviations of the items. Cronbach's coefficient alpha is on a scale from zero to one with a value closer to one being a more reliable measurement instrument and showing higher internal consistency.

$$\forall = \frac{(\ n\ )}{(\ n-1\ )} \frac{(\ SD_t^2 - \Sigma SD_i^2\ )}{(\ SD_t^2\ )}$$

where $\forall$ is the estimate of reliability,
n is the number of items in the test,
$SD_t$ is the standard deviation of the test scores
$\Sigma$ means "take the sum" and covers n items,
$SD_i$ is the standard deviation of the scores from a group of individuals on an item.

**Levels of Reliability**
Acceptable levels of reliability depend on the purpose of the instrument. Acceptable reliability of instruments developed for research purposes can be as low as 0.60. An acceptable reliability level of a diagnostic instrument used for making decisions about individuals (e.g., a psychological measure) should be much higher, e.g., 0.95.

**Comparisions**
The reliability coefficient provides a basis for comparison when measurement is expressed in different scales.

**Statistical Power**
An often overlooked benefit of more reliable scales is that they increase statistical power for a given sample size (or allow smaller sample size to yield equivalent power), relative to less reliable measures. A reliable measure, like a larger sample, contributes relatively less error to the statistical analysis.

**Reversing Items**
Examination of reliability analysis and item-to-total correlations reveal which, if any, item scales should be reversed. Items could be worded so that most responses will be in the same "direction". However, if appropriate responses are in the opposite "direction" of the scale, item responses are reserved. Negative item-to-total correlations determine which item response scales to reverse. For example, with a scale of 1=strongly agree, 2=agree, 3=disagree, 4=strongly disagree, responses are reversed to 1=strongly disagree, 2=disagree, 3=agree, 4=strongly agree.

To reverse an item on a 4-point scale the SAS code is

```
x013r = 5 - x013;
```

Substitute the values 1, 2, 3, 4 to verify responses are reversed.

For a 5-point scale the SAS code is

```
x007r = 6 - x007;
```

Substitute the values 1, 2, 3, 4, 5 to verify responses will be reversed.

**Factor Extraction**
Factor analysis seeks to discover common factors. The technique for extracting factors attempts to take out as much common variance as possible in the first factor. Subsequent factors are, in turn, intended to account for the maximum amount of the remaining common variance until, hopefully, no common variance remains.

Direct extraction methods obtain the factor matrix directly from the correlation matrix by application of specified mathematical models. Most factor analysts agree that direct solutions are not sufficient. Adjustment to the frames of reference by rotation methods improves the interpretation of factor loadings by reducing some of the ambiguities which accompany the preliminary analysis (Child, 1990). The process of manipulating the reference axes is known as rotation.

Rotation applied to the reference axes means the axes are turned about the origin until some alternative position has been reached. The simplest case is when the axes are held at $90^o$ to each other, orthogonal rotation. Rotating the axes through different angles gives an oblique rotation (not at $90^o$ to each other).

**Criteria for Extracting Factors**
Determining the number of factors to extract in a factor analytic procedure means keeping the factors that account for the most variance in the data. Criteria for determining the number of factors are:

1) Kaiser's criterion, suggested by Guttman and adapted by Kaiser, considers factors with an eigenvalue greater than one as common factors (Nunnally, 1978)

2) Cattell's (1966) scree test. The name is based on an analogy between the debris, called scree, that collects at the bottom of a hill after a landslide, and the relatively meaningless factors that result from overextraction. On a scree plot, because each factor explains less variance than the preceding factors, an imaginary line connecting the markers for successive factors generally runs from top left of the graph to the bottom right. If there is a point below which factors explain relatively little variance and above which they explain substantially more, this usually appears as an "elbow" in the plot. This plot bears some physical resemblance to the profile of a hillside. The portion beyond the elbow corresponds to the rubble, or scree, that gathers. Cattell's guidelines call for retaining factors above the elbow and rejecting those below it.
3) Proportion of variance accounted for keeps a factor if it accounts for a predetermined amount of the variance (e.g., 5%, 10%).
4) Interpretability criteria
    a. Are there at least 3 items with significant loadings (>0.30)?
    b. Do the variables that load on a factor share some conceptual meaning?
    c. Do the variables that load on different factors seem to measure different constructs?
    d. Does the rotated factor pattern demonstrate simple structure? Are there relatively
        i. high loadings on one factor?
        ii. low loadings on other factors?

**Significant Factor Loadings**

There are several methods to determine significant factor loadings and whether an item should be retained or included when calculating factor scores. A factor loading that is significant means responses on the item are influenced by the underlying construct. One method is to retain items with factor loadings greater than 0.30 or less than -0.30. This method is considered a rigorous level.

Another method treats factor loadings as correlation coefficients in terms of significance levels and takes into account the sample size and the value of the factor loading. Another method, the Burt-Banks formula, adjusts for the sample size, the number of variables, and the number of factors extracted (Child, 1990; Klein, 1994).

## SAS Code - PROC FACTOR and options

| | |
|---|---|
| DATA = | specifies dataset to be analyzed |
| PRIORS =SMC | squared multiple correlations used as adjusted diagonals of the correlation matrix |
| METHOD =ML,ULS | specifies maximum likelihood and unweighted least squares methods |
| ROTATE = | VARIMAX(orthogonal at a 90 degee angle), PROMAX (oblique, not at a 90 degree angle) |
| SCREE | requests a scree plot of the eigenvalues |
| N = | specifies number of factors |
| MINEIGEN=1 | specifies select factors with eigenvalues greater than 1 |
| OUT = | data and estimated factor scores, use raw data and N= |
| FLAG = | include a flag (*) for factor loadings above a specified value |
| REORDER = | sort the loadings from largest to smallest values for each factor |

**Methods**

As an aside, names given to factor extraction methods have some interesting origins.

- Procrustes was a highwayman who tied his victims to a bed and shaped them to its structure either by stretching them or by cutting off their limbs. In factor analysis, the Procrustes technique/method involves testing data to see how close they fit a hypothesized factor structure.
- The plasmode method is taken from well-established areas (e.g., physics, chemistry) so that the factor structure is predictable.

## Exploratory Factor Analysis Example

**Data for Exploratory Factor Analysis**

The World Values Survey (WVS) provided data analyzed for EFA examples in this paper. . WVS data has been collected from over 57 countries since 1990. Each data collection, every 5 years, is called a wave. Items on the survey have been updated or changed from wave to wave. Selecting items across waves resulted in low sample size. Therefore, selected items from the 2005 wave were examined to investigate the factor structure (dimensions) of values that could impact social and political life across countries. A list of items selected can be found in Appendix A.

**Reliability Analysis**

Internal consistency of the measurement instrument can be determined with PROC CORR. Both the raw and standardized Cronbach Alpha will be examined as well as item-to-total correlations. Items that correlate negatively with the total will have item scale responses reversed.

The following SAS code provided Cronbach Alpha and item-to-total correlations.

```
*correlations, cronbach alpha;
```

```
proc corr data=sgf2012w5 nocorr alpha nomiss;
 var a008 b001 b002 b003 c036 c037 c038 c039 c041
     d054 d055 d057 d059 d060 d078 d079 d080 e255
     g006 g019 g020 g021 g023 x001 x003r x007 x025
     x045 x047 x049 x052 x053 x054 x055 y001 y003;
run;
```

where the data set is sgf2012w5, no correlations are printed, Cronbach alpha is requested, observations with missing values are not included in the analysis.

The analysis included 17991 observations and found a Raw Cronbach Alpha = 0.67 and a Standardized Cronbach Alpha = 0.71. A value closer to one means the measurement instrument is more reliable and has higher internal consistency. The analysis found one item, x045,with an item-to-total correlation equal to -0.37. Therefore, the item responses (Would you describe yourself as belonging to) were reversed from a scale of 1=upper class, 2=upper middle class, 3=middle middle class, 4=lower middle class, 5=Working class, 6=Lower class to a scale of 1=Lower class, 2=Working class, 3=Lower middle class, 4=Middle middle class, 5=upper middle class, 6=Upper class. The SAS code for the reverse is

```
   x0045r = 7 - x0045;
```

with a new variable defined and responses of the original item retained.

After reversing item x045, the reliability analysis was run resulting in a Raw Cronach Alpha = 0.70 and a Standardized Cronbach Alpha = 0.74.

**Setting Variance Levels a priori**

Variance levels for factor extraction in this example are set prior to running the factor analysis. The researcher requires the amount of variance explained by each factor to be at least 5% and the cumulative variance explained to be at least 75%.

**SAS Code**

```
proc factor data=sgf2012w5 method=ml priors=smc scree;
  var a008 b001 b002 b003 c036 c037 c038 c039 c041
      d054 d055 d057 d059 d060 d078 d079 d080 e255
      g006 g019 g020 g021 g023 x001 x003r x007 x025
      x045r x047 x049 x052 x053 x054 x055 y001 y003;
```

where the data set is sgf2012w5, method is maximum likelihood, diagonals on the correlation matrix are squared multiple correlations, and factor loadings are listed from largest to smallest value for each factor.

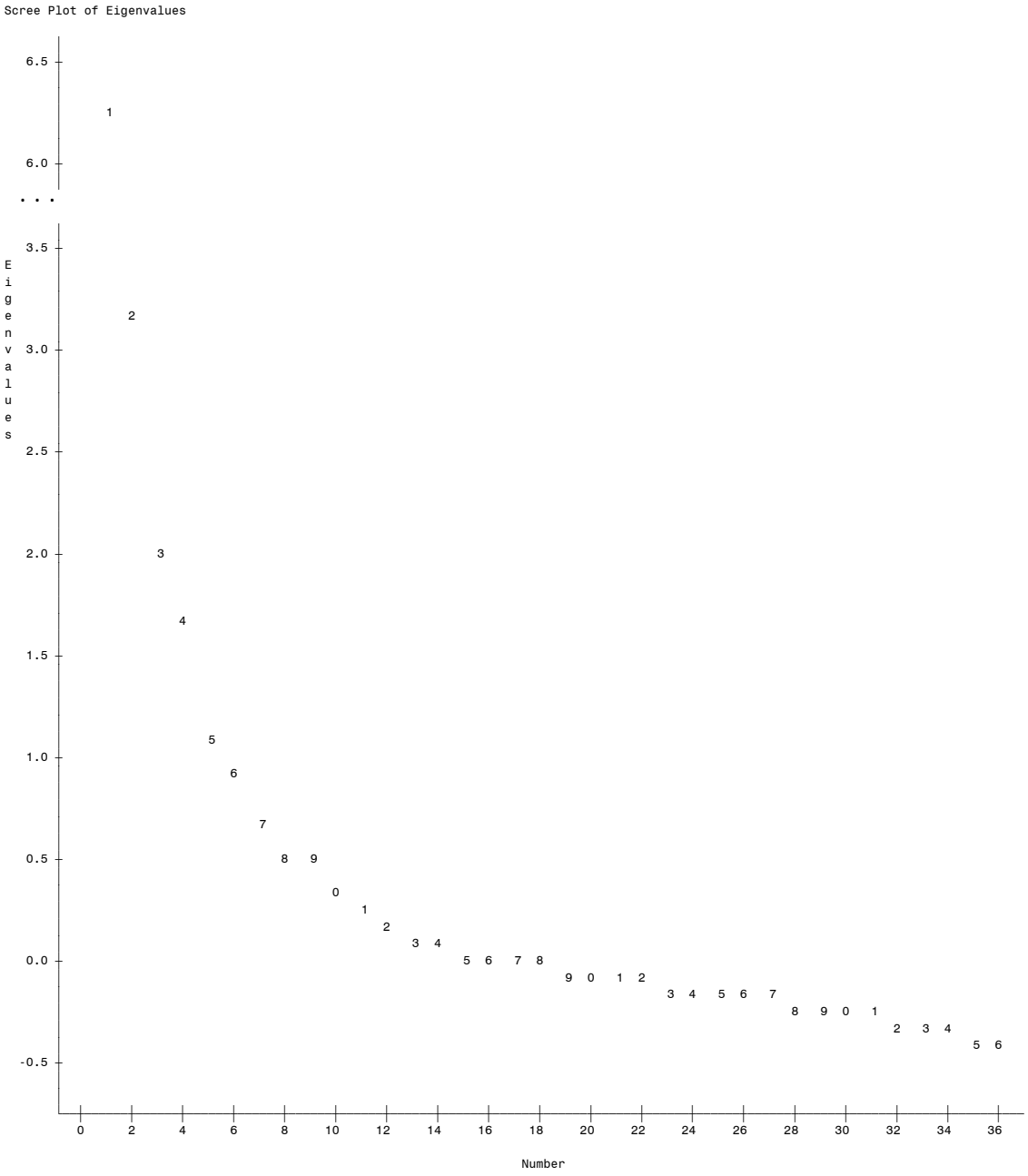
Five factors are retained, cumulative variance is 1.0495.

- Each factor explains 46%, 23%, 15%, 12% and 8% of the variance which meets criteria set a priori.
- Preliminary eigenvalues are shown in the SAS output below,

```
The FACTOR Procedure
Initial Factor Method: Maximum Likelihood

Prior Communality Estimates: SMC  . . .

Preliminary Eigenvalues: Total = 13.5354419  Average = 0.3759845

          Eigenvalue    Difference    Proportion    Cumulative
     1    6.27522773    3.14744413        0.4636        0.4636
     2    3.12778360    1.09312473        0.2311        0.6947
     3    2.03465887    0.36929485        0.1503        0.8450
     4    1.66536402    0.56301741        0.1230        0.9681
     5    1.10234661    0.16312418        0.0814        1.0495
     6    0.93922244    0.30197461        0.0694        1.1189
     7    0.63724782    0.13810555        0.0471        1.1660
 . . .
    35    -.37806406    0.01129996       -0.0279        1.0288
    36    -.38936402                     -0.0288        1.0000

5 factors will be retained by the PROPORTION criterion.
```

Eigenvalues of the weighted reduced correlation matrix are 7.78424513, 4.62240617, 2.74934881, 2.24911005, and 1.43683597. Proportion of variance explained by each factor is 41%, 25%, 15%, 12%, and 8%. Cumulative variance for 5 factors is 100%.

Scree plot of the Eigenvalues shows an elbow at Factor 5.




Hypothesis tests are both rejected, no common factors and 5 factors are sufficient. In practice, we want to reject the first hypotheses and accept the second hypothesis. Tucker and Lewis's Reliability Coefficient indicates good reliability (0.77).

Significance Tests Based on 17991 Observations

| Test | DF | Chi-Square | Pr > ChiSq |
|---|---|---|---|
| H0: No common factors | 630 | 119999.721 | <.0001 |
| HA: At least one common factor | | | |
| H0: 5 Factors are sufficient | 460 | 20124.1900 | <.0001 |
| HA: More factors are needed | | | |

```
Chi-Square without Bartlett's Correction       20142.291
Akaike's Information Criterion                 19222.291
Schwarz's Bayesian Criterion                   15635.382
Tucker and Lewis's Reliability Coefficient         0.774
```

Squared Canonical correlations indicate the amount of variance explained by each factor.

```
Squared Canonical Correlations
Factor1         Factor2         Factor3         Factor4         Factor5
0.88615983      0.82214021      0.73328702      0.69222341      0.58963180
```

### Exploratory Factor Analysis with rotation

After determining the number of factors to retain, 5 factors in this example, the factor analysis is run rotating the factor pattern. Rotation will allow the items to more distinctly group into a factor or dimension.

```
** used flag option, multiple factor loadings by 100 and round to integer;
proc factor data=sgf2012w5 method=ml rotate=v reorder priors=smc FLAG=0.30;
  var a008  b001 b002 b003 c036 c037 c038  c039 c041
      d054  d055 d057 d059 d060 d078 d079  d080 e255
      g006  g019 g020 g021 g023 x001 x003r x007 x025
      x045r x047 x049 x052 x053 x054 x055  y001 y003;
```

Options added to the PROC FACTOR for the example are

- varimax rotation (orthogonal) rather than a promax rotation (oblique)
- flag=0.30 which multiples the factor loadings by 100, rounds loading to an integer, and flags loadings that are greater than or equal to 30.

Other options

- n=5 to keep 5 factors. The number of factors retained may be specified if the default number of factors retained does not meet the levels of variance set a priori.
- out=*dataset name* specifies to save the original data and factor scores.

**Results for the 5 factor model**

Preliminary Eigenvalues, significance tests, squared canonical correlations, Eigenvalues of the weighted reduced correlation matrix are the same values as shown above.

Factor loadings illustrate correlations between items and factors. The higher the value of the factor loading, the higher the value of the item and factor correlation. The REORDER option arranges factors loadings by factor from largest to smallest value for each factor.

```
              Rotated Factor Pattern
                                                                 Factor1 Factor2 Factor3 Factor4 Factor5
x025  Highest educational level attained                          66 *     11       6      -6       1
x045r                                                             61 *      4       0      -3      -1
x053  Nature of tasks: manual vs. Cognitive                       56 *      9       7      -4       2
x047  Scale of incomes                                           55 *      5       1      -5       0
e255  How often use of PC                                        52 *     19      12      -2       4
x054  Nature of tasks: routine vs. Creative                      41 *     11       5      -4       0
y003  Autonomy Index                                             24       21      19      -6       5
x049  Size of town                                              21        4       5       3       6
x055  Nature of tasks: independence                             21        4       7      -5      -9
a008  Feeling of happiness                                      -20        0      -4       9      10
x052  Institution of occupation                                 -21        2      -3       1       5
c041  Work should come first even if it means less spare time    22       61 *    15      -6       9
c039  Work is a duty towards society                            10       60 *     6       3      12
c038  People who don´t work turn lazy                            8       57 *    10      -2       2
c036  To develop talents you need to have a job                  9       49 *     7      -2       2
c037  Humiliating to receive money without having to work for it  3       48 *     5       2       1
d054  One of main goals in life has been to make my parents proud 15       46 *    17       7      10
d055  Make effort to live up to what my friends expect           10       34 *    20       7       3
d080  I decide my goals in life by myself                        -5       24      -1      -1      16
y001  Post-Materialist index 12-item                            18       22      18     -13       5
d079  I seek to be myself rather than to follow others           -7       21      -6       1      17
d078  Men make better business executives than women do          19       17      82 *    -5       2
```

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| d059 | Men make better political leaders than women do | 16 | 22 | 75 * | -5 | 2 | |
| d060 | University is more important for a boy than for a girl | 17 | 11 | 59 * | 2 | -10 | |
| x001 | Sex | 0 | 6 | 27 | 0 | 3 | |
| d057 | Being a housewife just as fulfilling | 1 | 1 | 21 | 0 | 3 | |
| b002 | Increase in taxes if used to prevent environmental pollution | -7 | 7 | 1 | 82 * | 8 | |
| b001 | Would give part of my income for the environment | -8 | 12 | 5 | 76 * | 12 | |
| b003 | Government should reduce environmental pollution | 13 | 12 | 9 | -37 * | -3 | |
| g020 | I see myself as member of my local community | 5 | 14 | 5 | 2 | 67 | * |
| g021 | I see myself as citizen of the [country] nation | 0 | 18 | 6 | -1 | 66 | * |
| g019 | I see myself as a world citizen | -8 | 6 | 0 | 22 | 35 | * |
| g006 | How proud of nationality | 7 | 21 | 9 | 8 | 31 | * |
| g023 | I see myself as an autonomous individual | -13 | 1 | -4 | 7 | 25 | |
| x007 | Marital status | 1 | 2 | 6 | 0 | 8 | |
| x003r | Age group | -8 | 5 | 2 | 4 | -10 | |

Printed values are multiplied by 100 and rounded to the nearest integer.
Values greater than 0.3 are flagged by an '*'.

**Retaining items to define factors**

The researcher retained items with a factor loading of 0.30 or larger. The items retained have a specific correlation, 0.30 or larger. The items that describe the factor (dimension) have a large value for the factor loading (correlation). Some researchers determine that the factor loading must be a specific value, e.g. 0.40 or 0.50 or even 0.70 to be retained and to describe the factor or dimension. See Appendix B for a summary of items, factor loadings and a comparison of 5-factor and 4-factor models.

Items that "load" on more than one factor, e.g. have a factor loading of 0.30 or higher. correlate to more than one factor. That is, the items are not clearly influenced by one dimension (latent construct). Examine the rotated factor pattern to see that no items load on more than one factor.

**Factor scores**

Factor scores could be calculated by weighting each item response by the factor loading. In common practice, factor scores are calculated without weights. A factor is calculated by using the mean or sum of item responses that load, are highly correlated with the factor.

**Interpretability**

Is there some conceptual meaning for each factor? Could the factors be given a name?
Factor 1 could be a dimension of education/task/social class.
Factor 2 could be a dimension of work.
Factor 3 could be a dimension of men vs. women.
Factor 4 could be a dimension of environment.
Factor 5 could be a dimension of citizen.

**4-Factor Model**

For comparison, a 4-factor model was run with n= option included to specify a 4-factor model and flag option changed to flag=0.35. This model closely resembles the 5-factor model. We could examine the 4-factor model if, for example, the proportion of variance was set at 10% and the proportion of variance for factor 5 is 8% which does not meet the criteria.

```
proc factor data=sgf2012w5 method=ml rotate=v reorder priors=smc FLAG=0.35 n=4;
  var a008  b001 b002 b003 c036 c037 c038  c039 c041
      d054  d055 d057 d059 d060 d078 d079  d080 e255
      g006  g019 g020 g021 g023 x001 x003r x007 x025
      x045r x047 x049 x052 x053 x054  x055 y001 y003;
```

Rotated factor pattern for a 4-factor model.

Rotated Factor Pattern

| | | Factor1 | Factor2 | Factor3 | Factor4 |
|---|---|---|---|---|---|
| c039 | Work is a duty towards society | 59 * | 12 | 9 | 2 |
| c041 | Work should come first even if it means less spare time | 57 * | 25 | 18 | -7 |
| c038 | People who don´t work turn lazy | 51 * | 11 | 14 | -4 |
| d054 | One of main goals in life has been to make my parents proud | 46 * | 16 | 20 | 6 |
| c036 | To develop talents you need to have a job | 44 * | 12 | 11 | -4 |
| c037 | Humiliating to receive money without having to work for it | 42 * | 6 | 8 | 0 |

| | | | | | |
|---|---|---|---|---|---|
| g021 | I see myself as citizen of the [country] nation | 41 * | -5 | 4 | 5 |
| g020 | I see myself as member of my local community | 38 * | -1 | 2 | 8 |
| g006 | How proud of nationality | 33 | 4 | 8 | 10 |
| d055 | Make effort to live up to what my friends expect | 31 | 11 | 22 | 5 |
| d080 | I decide my goals in life by myself | 30 | -5 | -1 | -1 |
| d079 | I seek to be myself rather than to follow others | 28 | -8 | -7 | 2 |
| y001 | Post-Materialist index 12-item | 22 | 18 | 19 | -13 |
| x025 | Highest educational level attained | 10 | 66 * | 6 | -5 |
| x045r | | 3 | 61 * | 0 | -2 |
| x053 | Nature of tasks: manual vs. Cognitive | 9 | 56 * | 7 | -4 |
| x047 | Scale of incomes | 4 | 55 * | 2 | -5 |
| e255 | How often use of PC | 18 | 52 * | 13 | -2 |
| x054 | Nature of tasks: routine vs. Creative | 10 | 41 * | 5 | -4 |
| y003 | Autonomy Index | 20 | 24 | 20 | -7 |
| x055 | Nature of tasks: independence | -1 | 22 | 8 | -5 |
| x049 | Size of town | 6 | 20 | 5 | 4 |
| x003r | Age group | 0 | -7 | 3 | 2 |
| g023 | I see myself as an autonomous individual | 13 | -15 | -6 | 9 |
| a008 | Feeling of happiness | 6 | -21 | -4 | 9 |
| x052 | Institution of occupation | 4 | -21 | -3 | 2 |
| d078 | Men make better business executives than women do | 14 | 19 | 82 * | -6 |
| d059 | Men make better political leaders than women do | 18 | 16 | 76 * | -7 |
| d060 | University is more important for a boy than for a girl | 4 | 18 | 60 * | 1 |
| x001 | Sex | 6 | 0 | 27 | -1 |
| d057 | Being a housewife just as fulfilling | 2 | 1 | 21 | 0 |
| x007 | Marital status | 5 | 0 | 5 | 1 |
| b002 | Increase in taxes if used to prevent environmental pollution | 9 | -8 | 2 | 81 * |
| b001 | Would give part of my income for the environment | 15 | -8 | 6 | 77 * |
| g019 | I see myself as a world citizen | 21 | -12 | -2 | 24 |
| b003 | Government should reduce environmental pollution | 9 | 14 | 10 | -38 * |

Printed values are multiplied by 100 and rounded to the nearest integer.
Values greater than 0.35 are flagged by an '*'.

**4-Factor and 5 Factor Model Comparison**

Reliability and interpretability plays a role in your decision of the factor structure. Reliability was determined for each factor using PROC CORR with options ALPHA NOCORR. A comparison between 4- and 5-factor models found similar factors with the 4-factor model combining items for work and citizen factors from the 5-factor model.

5 factor model reliabilities

- Factor 1, educational level/income/tasks/social class/computer use, 0.69
- Factor 2, work/make parents proud/effort to do what friend expect, 0.72
- Factor 3, men vs. women, 0.79
- Factor 4, environment., 0.64
- Factor 5, citizen, 0.59

4 factor model reliabilities

- Factor 1, work/make parents proud/, 0.71
- Factor 2, educational level/income/tasks/social class/computer use, 0.69
- Factor 3, men vs. women, 0.79
- Factor 4, environment, 0.64

Items included in 3 of the factors on each model are the exactly the same items. However, items included on the other factors are slightly different. The 4-factor model combines factors of work and citizen which are 2 distinct factors in the 5-factor model. The 5-factor model includes “effort to do what friends expect” which is not included in the 4-factor model.

Both models exhibit good reliability and have slightly different factor loadings for items on each factor. However, the 5-factor model describes 5 distinct dimensions and it is recommended, in terms of interpretability, that the 5-factor model be retained.

## Further Analysis

To further examine the factor structure of selected items on the World Values Survey, an analysis was conducted by gender and by age group. It is possible that the factor structure for each group could be different, include different items on each factor or measure different dimensions. However, if the measurement instrument is well written and measures what it purports to measure, the factor structure will not differ between groups.

## Analysis by Gender

Exploratory factor analysis investigated the factor structure of selected items on the World Values Survey for males and for females. A data set was created to delete missing values for gender (WVS item x001). Then reliability analysis with PROC CORR found Cronbach Alpha and item-to-total correlations for each group which were examined for negative item-to-total correlations. EFA with PROC FACTOR used maximum likelihood method with a varimax rotation, squared multiple correlations and flagged factor loadings of 0.30 or larger.

```
* analysis by gender;
data gsgf2012w5;
  set sgf2012w5;
  if x001 ne .;
proc sort data=gsgf2012w5;
  by x001;
proc corr data=gsgf2012w5 nocorr alpha nomiss;
 by x001;
  var A008 b001 b002 b003 c036 c037 c038 c039 c041
      d054 d055 d057 d059 d060 d078 d079 d080 e255
         g006 g019 g020 g021 g023      x003r x007 x025
      x045r x047 x049 x052 x053 x054 x055 y001 y003;
title2 'by gender';
run;
proc factor data=gsgf2012w5 method=ml rotate=v reorder priors=smc FLAG=0.30;
  by x001;
  var a008  b001 b002 b003 c036 c037 c038  c039 c041
      d054  d055 d057 d059 d060 d078 d079  d080 e255
      g006  g019 g020 g021 g023      x003r x007 x025
      x045r x047 x049 x052 x053 x054 x055  y001 y003;
run;
```

Factor loadings of the items retained for males and females can be found in Appendix C.

### Results

Exploratory factor analysis for males and females determined 5-factor models for each group. Three of the 5 factors retain exactly the same items although the factor loadings are slightly different. Those factors are 1) environment, 2 ) men vs. women, 3) educational level/ income,/social class/computer use/tasks. There were differences found on the other factors between male and female models.

A fourth factor for females combines items on work, making parents proud, and doing what friends expect while a fifth factor combines items on being a citizen. These factors for females could be compared to 1 factor for males which combines items on work, making parents proud, doing what friends expect, being a citizen, and making your own decisions. It is interesting that this factor, for males, includes items on making your parents proud and doing what friends expect with making your own decisions. The fifth factor for males includes only 2 items, age group and marital status. These 2 items do not have large enough values on the factor loadings to be retained for the female group.

## Analysis by Age Groups

Exploratory factor analysis investigated the factor structure of selected items on the World Values Survey by age groups. A data set was created to delete missing values for age group (WVS item x003r2). Then reliability analysis with PROC CORR found Cronbach Alpha and item-to-total correlations for each group which were examined for negative item-to-total correlation. EFA with PROC FACTOR used maximum likelihood method with a varimax rotation, squared multiple correlations and flagged factor loadings of 0.30 or larger.

```
* analysis by age group;
data asgf2012w5;
  set sgf2012w5;
  if x003r2 ne .;
proc sort data=asgf2012w5;
  by x003r2;
proc corr data=asgf2012w5 nocorr alpha nomiss;
 by x003r2;
```

```
  var s008  b001 b002 b003 c036 c037 c038 c039 c041
      d054  d055 d057 d059 d060 d078 d079 d080 e255
      g006  g019 g020 g021 g023 x001      x007 x025
      x045r x047 x049 x052 x053 x054 x055 y001 y003;
title2 'by age group';
run;
proc factor data=asgf2012w5 method=ml rotate=v reorder priors=smc FLAG=0.30;
  by x003r2;
  var A008 b001 b002 b003 c036 c037 c038 c039 c041
      d054 d055 d057 d059 d060 d078 d079 d080 e255
         g006 g019 g020 g021 g023 x001      x007 x025
      x045r x047 x049 x052 x053 x054 x055 y001 y003;
run;
```

Factor loadings of the items retained by age groups (15-29, 30-49, 50+ years old) can be found in Appendix D.

**Results**

Exploratory factor analysis for age groups found a 4-factor model for 15-29 year olds, and 5-factor models for 30-49 year olds and 50+ year olds. Three of the factors for each age group contained exactly the same items although the factor loadings were slightly different. Those factors are 1) environment, 2 ) men vs. women, 3) educational level/ income,/social class/computer use/tasks. There were differences found between the others factors for age group models.

The remaining factor for the 15-29 year old age group combined items on work, being a citizen and making your parent proud. The other age groups had a being a citizen factor although age group 50+ included an item on "seeing myself as an autonomous individual" in this factor. The fifth factor for the 30-49 and 50+ year old age groups combined items on work, making your parents proud, doing what your friends expect. It is interesting that the item "I make a lot of effort to live up to what my friends expect" did not have a large enough value for the factor loading to be included in the 15-29 year old age group model.

## Discussion

Exploratory factor analysis techniques investigated the underlying factor structure of a set of selected items on the World Values Survey. The examples in this paper have illustrated that the factor structure could be similar but different when comparisons are made between the total group, gender, and age groups. Attitudes and viewpoints differ between males and females. People of different ages respond differently to questions pertaining to their social attitudes and values.

## Conclusion

The dimensions and underlying factor structure of a measurement instrument can be identified. Exploratory factor analysis identifies the possible underlying structure of a set of variables without imposing a preconceived structure on the outcome while SAS® procedures PROC CORR and PROC FACTOR give you the power to maximize your knowledge and answer your research questions.

## References

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245-276.

Child, D. (1990). The essentials of factor analysis, second edition. London: Cassel Educational Limited.

DeVellis, R. F. (1991). Scale Development: Theory and Applications. Newbury Park, California: Sage Publications.

Hatcher, L. (1994). A step-by-step approach to using the SAS® System for factor analysis and structural equation modeling. Cary, NC: SAS Institute Inc.

Kline, P. (1994). An Easy Guide to Factor Analysis. London: Routledge.

Nunnally, J. C. (1978). Psychometric theory, 2nd edition. New York: McGraw-Hill.

SAS® Language and Procedures, Version 6, First Edition. Cary, N.C.: SAS Institute, 1989.

SAS® Online Doc 9. Cary, N.C.: SAS Institute.

SAS® Procedures, Version 6, Third Edition. Cary, N.C.: SAS Institute, 1990.

SAS/STAT® User's Guide, Version 6, Fourth Edition, Volume 1. Cary, N.C.: SAS Institute, 1990.

SAS/STAT® User's Guide, Version 6, Fourth Edition, Volume 2. Cary, N.C.: SAS Institute, 1990.

Thorndike, R. M., Cunningham, G. K., Thorndike, R. L., & Hagen E. P. (1991). Measurement and evaluation in psychology and education. New York: Macmillan Publishing Company.

Truxillo, Catherine. (2003). Multivariate Statistical Methods: Practical Research Applications Course Notes. Cary, N.C.: SAS Institute.

## Contact Information

Diana Suhr is a Statistical Analyst, Institutional Reporting and Analysis Services, University of Northern Colorado. She holds a Ph.D. in Educational Psychology from UNC. She has been a SAS programmer since 1984.
email: diana.suhr@unco.edu

**Appendix A – Selected Items from the World Values Survey**

| Variable | Item | Scale |
|---|---|---|
| a008 | Taking all things together, would you say you are: (feeling of happiness) | 1=Very happy, 2=Quite happy, 3=Not very happy, 4=Not at all happy |
| b001 | Would give part of my income for the environment | 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree |
| b002 | I would agree to an increase in taxes if the extra money were used to prevent environmental pollution. | 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree |
| b003r | The Government should reduce environmental pollution, but it should not cost me any money. | 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree (original scale)<br>Reversed to (1=Strongly disagree, 2=Disagree, 3=Agree, 4=Strongly agree) |
| c036 | To fully develop your talents, you need to have a job. | 1=Strongly agree, 2=Agree, 3=Neither agree or disagree, 4=Disagree, 5=Strongly disagree |
| c037 | It is humiliating to receive money without having to work for it. | 1=Strongly agree, 2=Agree, 3=Neither agree or disagree, 4=Disagree, 5=Strongly disagree |
| c038 | People who don't work turn lazy. | 1=Strongly agree, 2=Agree, 3=Neither agree or disagree, 4=Disagree, 5=Strongly disagree |
| c039 | Work is a duty towards society. | 1=Strongly agree, 2=Agree, 3=Neither agree or disagree, 4=Disagree, 5=Strongly disagree |
| c041 | Work should always come first, even if it means less spare time. | 1=Strongly agree, 2=Agree, 3=Neither agree or disagree, 4=Disagree, 5=Strongly disagree |
| d054 | One of my main goals in life has been to make my parents proud. | 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree |
| d055 | I make a lot of effort to live up to what my friends expect. | 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree |
| d057 | Being a housewife is just as fulfilling as working for pay. | 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree |
| d059 | On the whole, men make better political leaders than women do. | 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree |
| d060 | A university education is more important for a boy than for a girl. | 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree |
| d078 | On the whole, men make better business executives than women do. | 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree |
| d079 | I seek to be myself rather than to follow others. | 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree |
| d080 | I decide my goals in life by myself. | 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree |
| e255 | How often, if ever, do you use a personal computer? | 1=Never, 2-Occasionally, 3=Frequently, 4=Don't know hwat a computer is |
| g006 | How proud are you to be [Nationality]? | 1=Very proud, 2=Quite proud, 3=Not very proud, 4=Not at all proud |
| g019 | I see myself as a world citizen | 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree |
| g020 | I see myself as a member of my local community. | 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree |
| g021 | I see myself as citizen of the [country] nation. | 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree |
| g023 | I see myself as an autonomous individual. | 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree |
| x001 | Sex | 1=Male, 2=Female |
| x003r | Age groups | 1=15-24, 2=25-34, 3=35-44, 4=45-54, 5=55-64, 6=65 and more years |
| X003r2 | Age groups | 1=15-29, 2=30-49, 3=50 and more years |

| | | |
|---|---|---|
| x007 | Marital Status | 1=Married, 2=Living together as married, 3=Divorced, 4=Separated, 5=Widowed, 6=Single/Never married, 7=Divorced, Separated, or Widow, 8=Living apart but steady relation (married, cohabitation) |
| x025 | Highest Educational Level attained | 1=Inadequately completed elementary education, 2=Completed (compulsory) elementary education, 3=Incomplete secondary school (technical/vocational type), 4=Complete secondary school (technical/vocational type), 5=Incomplete secondary (university-prep type), 6=Complete secondary (university-prep type), 7=Some university without degree, 8-University with degree |
| x045r | Would you describe yourself as belonging to the | 1=Upper class, 2=Upper middle class, 3=Middle middle class, 4=Lower middle class, 5=Working class, 6=Lower class (original scale)<br>Reversed to 1=Lower class, 2=Working class, 3=Lower middle class, 4=Middle middle class, 5=Upper middle class, 6=Upper class |
| x047 | Scale of incomes | 1=Lower step, 2=second step, 3=third step, 4=fourth step, 5=fifth step, 6=sixth step, 7=seventh step, 8=eighth step, 9=ninth step, 10=tenth step, 11=Highest step |
| x049 | Size of town | 1=2000 and less, 2=2000-5000, 3=5000-10000, 4=10000-20000, 5=20000-50000, 6=50000-100000, 7=100000-500000, 8=500000 and more |
| x052 | Are you working for, or characterize your work in the past. | 1=Public Institution, 2=Private Business, 3=Private non-profit organization, 4=Self-employed |
| x053 | Nature of Tasks: Manual vs. Cognitive | 1=Mostly manual tasks, 2, 3, 4, 5, 6, 7, 8, 9, 10=Mostly non-manual tasks |
| x054 | Nature of Tasks: Routine vs. Creative | 1=Mostly routine tasks, 2, 3, 4, 5, 6, 7, 8, 9, 10=Mostly not routine tasks |
| x055 | Nature of Tasks: Independence | 1=No independence at all, 2, 3, 4, 5, 6, 7, 8, 9, 10=Complete independence |
| y001 | Post-Materialist Index 12-item | 0=Materialist, 1,2, 3, 4,5=Postmaterialist |
| y003 | Automony Index | 0, 1, 2=Determination, Perseverance/Independence |

**Appendix B**

**Comparison of 5- and 4-factor solutions on selected items from the World Value Survey**

| 5-Factor Solution | Items | Factor Loading | | 4-Factor Solution | Items | Factor Loading |
|---|---|---|---|---|---|---|
| **#1**, alpha = 0.69 | X025 | .66 | | **#1**, alpha = 0.71 | C039 | .59 |
| | X045 (reversed) | .61 | | | C041 | .57 |
| | X053 | .56 | | | C038 | .51 |
| | X047 | .55 | | | D054 | .46 |
| | E255 | .52 | | | C036 | .44 |
| 6 out of 11 items | X054 | .41 | | | C037 | .42 |
| | | | | | G021 * | .41 |
| | | | | 8 out of 13 items | G020 * | .38 |
| | | | | | | |
| **#2**, alpha = 0.72 | C041 | .61 | | **#2**, alpha = 0.69 | X025 | .66 |
| | C039 | .60 | | | X045 (reversed) | .61 |
| | C038 | .57 | | | X053 | .56 |
| | C036 | .49 | | | X047 | .55 |
| | C037 | .48 | | | E255 | .52 |
| | D054 | .46 | | 6 out of 13 items | X054 | .41 |
| 7 out of 10 items | D055 * | .34 | | | | |
| | | | | | | |
| **#3**, alpha = 0.79 | D078 | .82 | | **#3**, alpha = 0.79 | D078 | .82 |
| | D059 | .75 | | | D059 | .76 |
| 3 out of 5 items | D060 | .59 | | 3 out of 6 items | D060 | .60 |
| | | | | | | |
| **#4**, alpha = 0.64 | B0002 | .82 | | #4, alpha = 0.64 | B002 | .81 |
| | B0001 | .76 | | | B001 | .77 |
| 3 out of 3 items | B003 | .37 | | 3 out of 4 items | B003 | .38 |
| | | | | | | |
| **#5**, alpha = 0.59 | G020 | .67 | | | | |
| | G021 | .66 | | | | |
| | G019 * | .35 | | | | |
| 4 out of 7 items | G006 * | .31 | | | | |
| | | | | | | |
| 23 items retained out of 36 items | | | | 20 items retained out of 36 items | | |
| | | | | | | |

Note: Items with factor loadings of 0.30 or greater are included in the 5-factor solution while factor loadings of 0.35 or greater are included in the 4-factor solution. The notation, e.g. 6 out of 11 items, indicates the number of items that were retained (meet the level selected to load on the factor) out of the number of item that loaded on the factor.

*G020, G021 appear in factor #5 in the 5-factor solution but in factor #1 for the 4-factor solution.

*D055,G006, G019 are not included in the 4-factor solution.

**Appendix C**

**Comparison of Exploratory Factor Analysis by Gender on World Values Survey selected items**

| Male | Items | Factor Loading | | Female | Items | Factor Loading |
|---|---|---|---|---|---|---|
| #1 | C041 | .56 | | #1 | C041 | .62 |
| | C039 | .56 | | | C039 | .59 |
| | C038 | .49 | | | C038 | .59 |
| | D054 | .48 | | | C036 | .54 |
| | G021 | .44 | | | C037 | .49 |
| | G020 | .41 | | | D054 | .48 |
| | C036 | .40 | | 7 out of 9 items | D055 | .35 |
| | C037 | .40 | | | | |
| | G006 | .35 | | #2 | X025 | .68 |
| | D055 | .33 | | | X045 (reversed) | .60 |
| | D079* | .31 | | | X053 | .55 |
| 12 out of 14 items | D080* | .31 | | | X047 | .54 |
| | | | | | E255 | .51 |
| #2 | X025 | .64 | | 6 out of 12 items | X054 | .41 |
| | X045 (reversed) | .61 | | | | |
| | X053 | .59 | | #3 | D078 | .78 |
| | X047 | .56 | | | D059 | .71 |
| | E255 | .52 | | 3 out of 5 items | D060 | .57 |
| 6 out of 11 items | X054 | .42 | | | | |
| | | | | #4 | B002 | .82 |
| #3 | D078 | .80 | | | B001 | .74 |
| | D059 | .76 | | 3 out of 3 items | B003 | .37 |
| 3 out of 4 items | D060 | .58 | | | | |
| | | | | #5 | G021** | .68 |
| #4 | B002 | .82 | | | G020** | .64 |
| | B001 | .78 | | | G019* | .34 |
| 3 out of 4 items | B003 | .58 | | 4 out of 6 items | G006** | .32 |
| | | | | | | |
| #5 | X003r | .71 | | | | |
| 2 out of 2 items | X007* | .56 | | | | |
| | | | | | | |
| 26 items retained out of 35 items | | | | 23 items retained out of 35 items | | |

Note: Items with factor loadings of 0.30 or greater are included in the factor solutions. The notation, e.g. 12 out of 14 items, indicates the number of items that were retained (meet the level selected to load on the factor) out of the number of item that loaded on the factor.

*D079, D080, X003r, X007 not included for females.

* G019 not included for males.

** G006, G020, G021 in factor #1 for males, factor #5 for females.

**Appendix D**

**Comparison of Exploratory Factor Analysis by Age Group on World Values Survey selected items**

| 15-29 year olds | Items | Factor Loading | | 30-49 year olds | Items | Factor Loading | | 50+ year olds | Items | Factor Loading |
|---|---|---|---|---|---|---|---|---|---|---|
| #1 | C039 | .57 | | #1 | X025 | .67 | | #1 | X025 | .63 |
| | C041 | .56 | | | X045 (rev) | .63 | | | X053 | .62 |
| | C038 | .49 | | | X053 | .58 | | | X045 (rev) | .61 |
| | D054 | .48 | | | X047 | .58 | | | X047 | .56 |
| | G021* | .45 | | | E255 | .54 | | | X054 | .50 |
| | C036 | .43 | | 6 of 11 | X054 | .42 | | 6 of 12 | E255 | .48 |
| | C037 | .40 | | | | | | | | |
| | G020* | .40 | | #2 | C041 | .62 | | #2 | C041 | .61 |
| 9 of 15 | G006* | .34 | | | C039 | .62 | | | C039 | .60 |
| | | | | | C038 | .57 | | | C038 | .56 |
| #2 | X025 | .67 | | | C036 | .50 | | | C036 | .50 |
| | X045 (rev) | .59 | | | C037 | .50 | | | C037 | .45 |
| | E255 | .55 | | | D054 | .46 | | | D054 | .41 |
| | X053 | .47 | | 7 of 10 | D055 ** | .34 | | 7 of 8 | D055 ** | .32 |
| | X047 | .44 | | | | | | | | |
| 6 of 11 | X054 | .30 | | #3 | D078 | .81 | | #3 | D078 | .82 |
| | | | | | D059 | .76 | | | D059 | .77 |
| #3 | D078 | .80 | | 3 of 6 | D060 | .58 | | 3 of 6 | D060 | .58 |
| | D059 | .73 | | | | | | | | |
| 3 of 5 | D060 | .62 | | #4 | B002 | .81 | | #4 | B002 | .85 |
| | | | | | B001 | .77 | | | B001 | .75 |
| #4 | B002 | .79 | | 3 of 3 | B003 | .35 | | 3 of 3 | B003 | .42 |
| | B001 | .79 | | | | | | | | |
| 3 of 4 | B003 | .36 | | #5 | G020 | .68 | | #5 | G021 | .69 |
| | | | | | G021 | .67 | | | G020 | .63 |
| | | | | | G019 | .40 | | | G006 | .32 |
| | | | | 4 of 5 | G006 | .33 | | | G019 | .31 |
| | | | | | | | | 5 of 6 | G023* | .31 |
| | | | | | | | | | | |
| 21 items retained out of 35 items | | | | 23 items retained out of 35 items | | | | 24 items retained out of 35 items | | |

Note: Items with factor loadings of 0.30 or greater are included in the factor solutions. The notation, e.g. 9 out of 15 items, indicates the number of items that were retained (meet the level selected to load on the factor) out of the number of item that loaded on the factor.

Note: 4-factor solutions for 15-29 year olds compared to a 5-factor solution for 30-49 year olds and for 50+ year olds.

*G006, G020, G021 in factor #1 for 15-29 year olds and in factor #5 for 30-49 year olds and for 50+ year olds.

**D055 not included in the 15-29 year old model.

## Appendix E - Definitions

**Communality** is the variance in observed variables accounted for by common factors.

**Community** is computed by summing squares of factor loadings

$d_1^2$ = 1 – communality = % variance accounted for by the unique factor
$d_1$ = square root (1-community) = unique factor weight (parameter estimate)

**Eigenvalues** indicate the amount of variance explained by each principal component or each factor.

**Eigenvectors** are the weights in a linear transformation when computing principal component scores.

A **factor** or **dimension** could also be referred to as an **underlying construct**, an **unobserved variable**, or a **latent construct**. A **latent construct** can be measured indirectly by determining its influence to responses on measured variables.

**Factor scores** are estimates of underlying latent constructs.

An observed variable "**loads**" on a factors if it is highly correlated with the factor, has an eigenvector of greater magnitude on that factor.

**Obilque** means other than a 90 degree angle.

An **observed variable** can be measured di**rectly, is sometimes called a measured variable or an indicator or a manifest variable.**

**Orthogonal** means at a 90 degree angle, perpendicular.

**Unique factors** refer to unreliability due to measurement error and variation in the data.