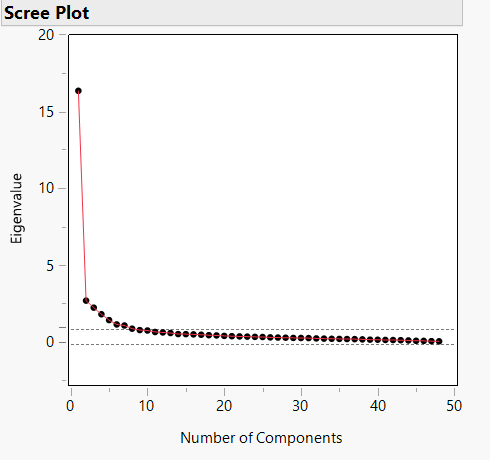
Ayan Sarkar

University of denver

FActor analysis

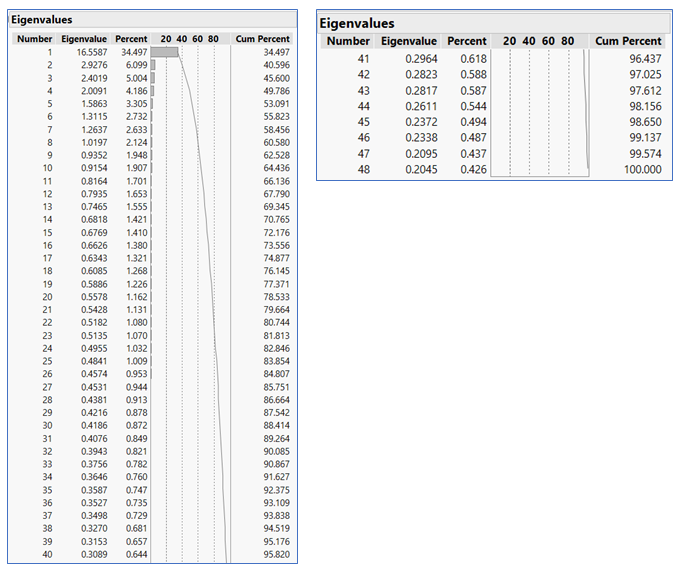
**Factor Analysis Approach**:

* We selected the fields: I1 to I48 as inputs from the source data file: “Team Excellence Data.xls” for our factor analysis.
* The factor analysis was done through the JMP application.
* For imputing the missing values in the data, we opted to use the “Multivariate Nominal Imputation Method “.
* We opted for the “Variance Estimation” as “Default” & “Variation Scaling” as “Correlations” while selecting the inputs for our Factor Analysis.
* Our Scree plot suggested that are 8 factors which have been recognized in the underlying data:



**Fig 1: First iteration: Scree Plot**

* The eigenvalues and percentage of variance accounted for by the latent constructs are as below:

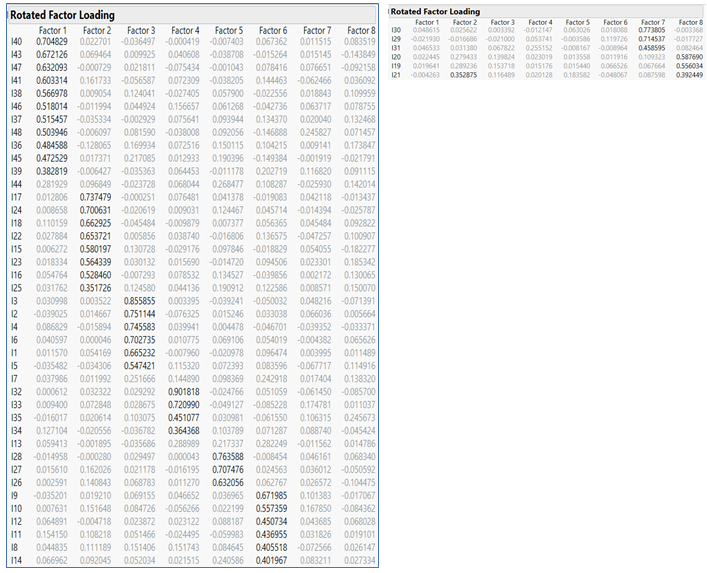


**Fig 2: First iteration: Eigen Values**

As per Kaiser’s Criterion, we usually select the factors with eigen values more than one; in our case, as evident from the table of eigen values, this is the first eight factors who can cumulatively explain 60.580% of the total variance in the data.

The “Percent” field gives us the individual % of variance in the total data explained by the factors on their own.

* For generating the model, we opted for the below options in terms of the following inputs:
  1. Factoring Method: Maximum Likelihood (Due to the presence of missing values in the data.)
  2. Prior Communality: Common Factor Analysis (diagonals=SMC)
  3. Number of factors: 8
  4. Rotation Method: Oblimin ( we opted an oblique rotation)
* The model generated the following factor loadings:

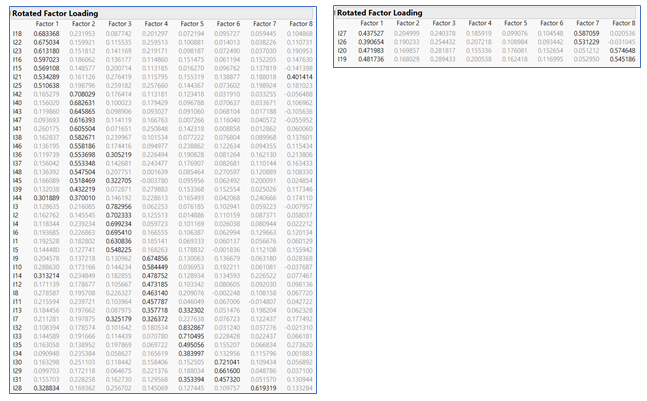


**Fig 3: First iteration: Factor Loadings**

Problematic Loadings: The inputs: I44, I7 & I13 do not have a significant enough loading to belong to any of the 8 factors.

Cross Loading: Input I21 had a cross loading between the factors: 2 & 8 but it is more inclined to Factor 8 in terms of its loading.

* To avoid removing these problematic loadings: I44, I7 & I13 altogether from the data ( to prevent any loss of pattern in the data), we opted for Varimax rotation and re-ran the model to get the below factor loadings.



**Fig 4: Second iteration: Factor Loadings**

Cross Loadings:

Input I21 had a cross loading between the factors: 1 & 8 but it is more inclined to Factor 1 in terms of its loading.

Input I36 had a cross loading between the factors: 2 & 3 but it is more inclined to Factor 2 in terms of its loading.

Input I45 had a cross loading between the factors: 2 & 3 but it is more inclined to Factor 2 in terms of its loading.

Input I44 had a cross loading between the factors: 1 & 2 but it is more inclined to Factor 2 in terms of its loading.

Input I14 had a cross loading between the factors: 1 & 4 but it is more inclined to Factor 4 in terms of its loading.

Input I13 had a cross loading between the factors: 4 & 5 but it is more inclined to Factor 4 in terms of its loading.

Input I7 had a cross loading between the factors: 3 & 4 but it is more inclined to Factor 4 in terms of its loading.

Input I31 had a cross loading between the factors: 5 & 6 but it is more inclined to Factor 6 in terms of its loading.

Input I28 had a cross loading between the factors: 1 & 7 but it is more inclined to Factor 7 in terms of its loading.

Input I27 had a cross loading between the factors: 1 & 7 but it is more inclined to Factor 7 in terms of its loading.

Input I26 had a cross loading between the factors: 1 & 7 but it is more inclined to Factor 7 in terms of its loading.

Input I20 had a cross loading between the factors: 1 & 8 but it is more inclined to Factor 8 in terms of its loading.

Input I19 had a cross loading between the factors: 1 & 8 but it is more inclined to Factor 8 in terms of its loading.

Since we have addressed all the cross loadings now, these are the below factor and element associations we get from the factor loading table:

Factor 1: 15,16,17,18,21,22,23,24,25

Factor 2: 36,37,38,39,40,41,42,43,44,45,46,47,48

Factor 3: 1,2,3,4,5,6

Factor 4: 7,8,9,10,11,12,13,14

Factor 5: 32,33,34,35

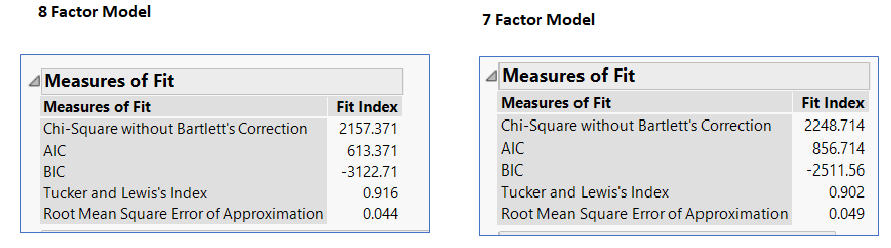
Factor 6: 29,30,31

Factor 7: 26,27,28

Factor 8: 19,20

Justification for Non-Oblique Rotation:

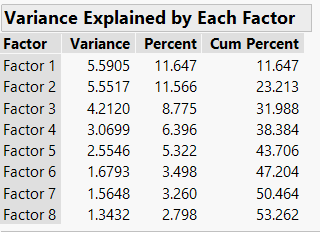
We have opted for Varimax rotation instead of Oblimin rotation because the latter would have required us to remove the problematic loadings: I44, I7 & I13 altogether from the data; this leads to decrease in the number of factors from 8 to 7, which resulted in the Goodness of Fit parameters of the model to degrade from the transition to a model with 8 factors to model with 7 factors and hence our decision to stick with the 8 factor model.



**Fig 5: Measures of Fit for both the models**

**Note: The lesser the AIC, BIC and RMSE & more the Tucker and Lewis’s Index is, the better the model.**

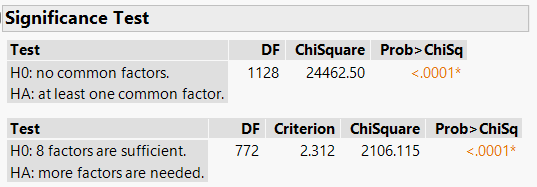
* The variance explained by each of the eight factors are:



**Fig 6: Variance explained by each Factor**

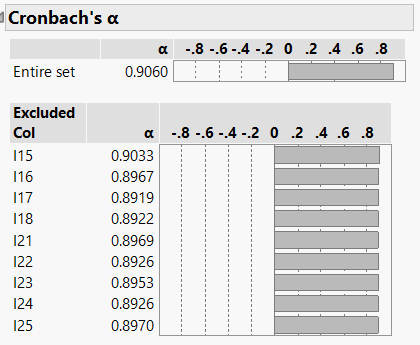
The first three factors on their own represent the most variance in the data: 11.64%, 11.56% and 8.77% respectively.

* All the eight factors are significant in the model and do no have nay commonality among them which has been proved by the Significance tests in the model:



**Fig 7: Significance test for the entire model**

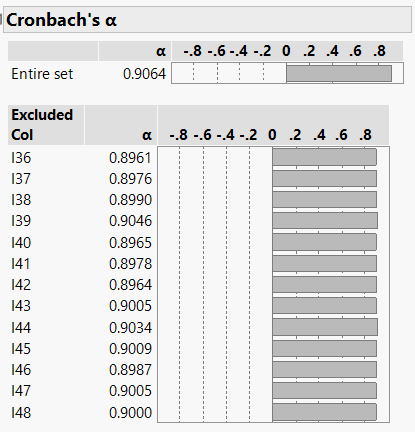
* Item Analysis of Factor 1:



**Fig 8: Item Analysis of Factor 1**

All items in this factor are functioning properly.

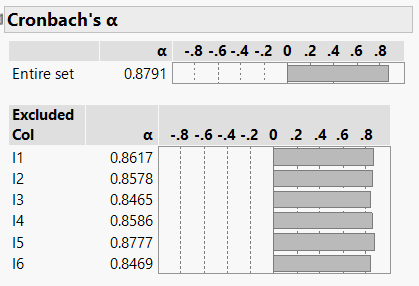
* Item Analysis of Factor 2:



**Fig 9: Item Analysis of Factor 2**

All items in this factor are functioning properly.

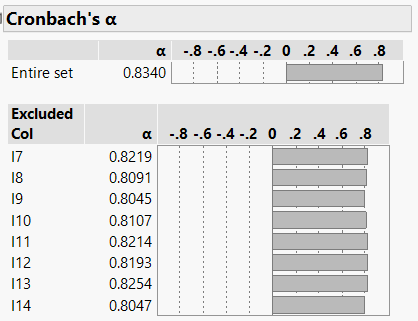
* Item Analysis of Factor 3:



**Fig 10: Item Analysis of Factor 3**

All items in this factor are functioning properly.

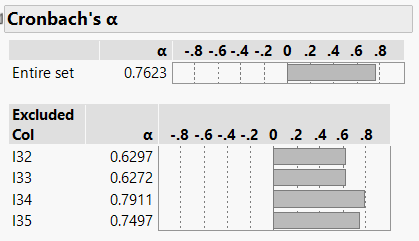
* Item Analysis of Factor 4:



**Fig 11: Item Analysis of Factor 4**

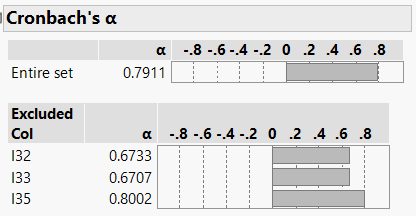
All items in this factor are functioning properly.

* Item Analysis of Factor 5:



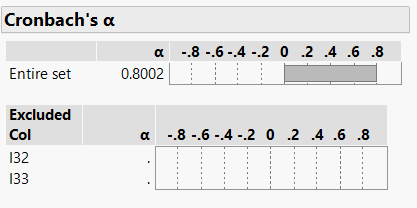
**Fig 12: Item Analysis of Factor 5: 1st Iteration**

The input I34 is not functioning properly. On removing this input from the factor, the reliability improves:



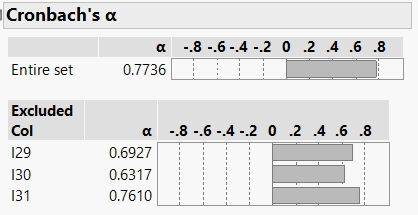
**Fig 13: Item Analysis of Factor 5: 2nd Iteration**

Now we see that the input I35 is not functioning properly. On removing this input from the factor, the reliability improves:



**Fig 14: Item Analysis of Factor 5: 3rd Iteration**

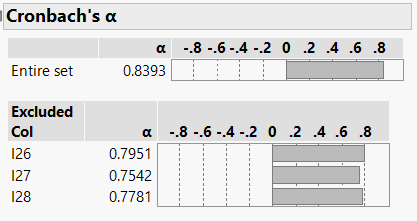
* Item Analysis of Factor 6:



**Fig 15: Item Analysis of Factor 6**

All items in this factor are functioning properly.

* Item Analysis of Factor 7:



**Fig 16: Item Analysis of Factor 7**

All items in this factor are functioning properly.

* Item Analysis of Factor 8:

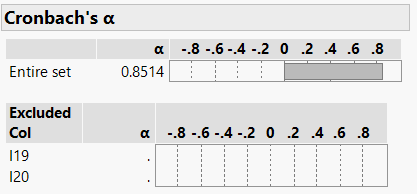
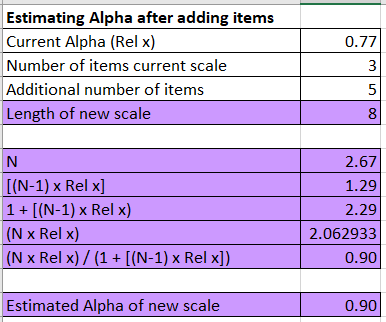


Fig 17: Item Analysis of Factor 8

All items in this factor are functioning properly.

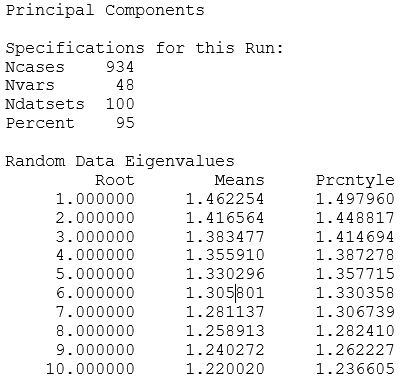
* The factor with the lowest reliability is Factor 6: 0.7736; if we add 5 additional items to this factor, we get a reliability of 0.90



**Fig 18: Spearman Brown Formula**

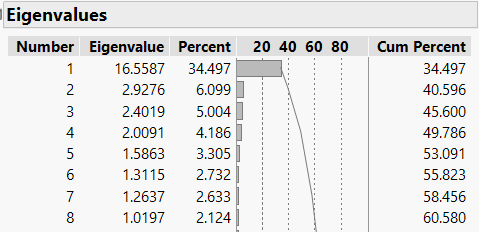
**Appendix**:

The parallel analysis gives us the following average cutoffs that the eigen values of the factors should surpass to be considered as a factor in this data set:



**Fig 19: Given Parallel Analysis Cut-offs**

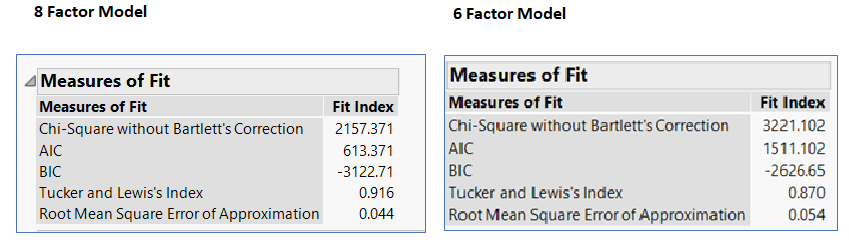
And the eigen values we got for our eight factors were:-



**Fig 20: Eigen Values for our 8 factors**

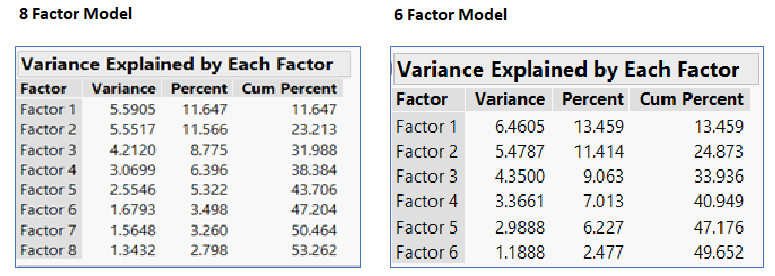
On referring the cutoffs from the parallel analysis, it seems we can only consider first 6 factors in this data set. But if we opt for a 6-factor model, the goodness of fit measures for the model degrades and also the total cumulative percentage of the variance explained by all the factors decreases to a great extent. **Hence, we would prefer to stick with the model having 8 factors.**

Please find below a comparison of the “Measures of Fit” & “Factor Variance” for an 8-Factor & 6-Factor model.



**Fig 21: Measures of Fit for both the models**

**Note: The lesser the AIC, BIC and RMSE & more the Tucker and Lewis’s Index is, the better the model.**



**Fig 22: Factor Variance for both the models**