3 EFA

erin buchanan

6/12/2017

# EFA Example Class Assignment

## Abstract:

System justification theory (Jost & Banaji, 1994; Jost, Banaji, & Nosek, 2004) predicts that ideologies expressed by privileged groups which serve to justify the existence of group inequality can be internalized by members of disadvantaged groups, so that they support and defend the social status quo even at their own direct personal expense and at the expense of other members of their group. Jost and Calogero (2011) have examined self-objectification responses to sexism, which reinforce the status quo of gender inequality; they reported that a low need for cognitive closure as measured by Webster and Kruglanski’s (1994) scale offers protection against self-subjugating system-justifying effects. Women having a low need for closure exhibited a lesser degree of self-objectification upon exposure to sexist ideology. Jost and Calogero, following Glick and Fiske (1996, 2001), differentiated between hostile and benevolent expressions of sexism as well as combining the two into a complementary exposure condition which they observed was most pernicious of all. As there is a scarcity of published work investigating the application of system justification theory to prejudice against individuals of lower socioeconomic status (SES), we will build on Jost and Calogero’s work by searching for system-justifying effects upon exposure to benevolent, hostile, and complementary expressions of SES-based prejudice. System justification theory predicts that in response to statements advocating an ideology prejudiced against them, participants with a lower SES will exhibit stronger self-subjugating responses, which reinforce the status quo of SES inequality.

# Assignment:

master = read.csv("~/OneDrive - Missouri State University/TEACHING/751 SEM/class assignments/R Markdown/3 efa.csv")

## 1. Data screening/clean up (short):

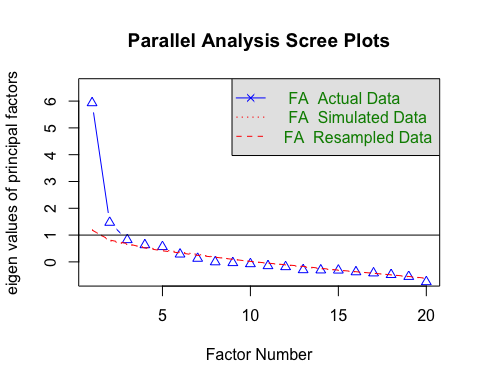
a. Exclude missing cases from the dataset.  
b. Exclude the group variable from the analysis.

nomiss = na.omit(master)  
final = nomiss[ , -ncol(nomiss)]

## 2. Run an EFA analysis on the openness to experience scale, and answer the following questions (you can do multiple R sections here if you want):

a. Number of factors: (As you are watching this video here, note that I added comments to the video as updates)  
 i. Parallel Analysis: 5 factors  
 ii. Scree plot: 2 factors  
 iii. Eigenvalues: 2 or 3 factors

library(psych)  
library(GPArotation)  
  
parallel = fa.parallel(final, fm = "ml", fa = "fa")



## Parallel analysis suggests that the number of factors = 5 and the number of components = NA

parallel$fa.values

## [1] 5.932279963 1.470194637 0.822720380 0.633508663 0.563737302  
## [6] 0.285023513 0.127323562 -0.004958086 -0.034447447 -0.074751171  
## [11] -0.148080832 -0.185564079 -0.298848587 -0.308076122 -0.313017313  
## [16] -0.376815124 -0.418976686 -0.479021550 -0.549202885 -0.746068759

sum(parallel$fa.values > 1) ##old kaiser

## [1] 2

sum(parallel$fa.values > .7) ##new kaiser

## [1] 3

b. Simple structure:  
 i. Set up: run with an oblimin rotation and ML as the type of math.  
 ii. Run the analysis excluding questions as they do not load. 3, 11, 15, 16

twofactor = fa(final, #dataset  
 nfactors = 2, #number of factors  
 rotate = "oblimin", #rotation type  
 fm = "ml") #math   
twofactor

## Factor Analysis using method = ml  
## Call: fa(r = final, nfactors = 2, rotate = "oblimin", fm = "ml")  
## Standardized loadings (pattern matrix) based upon correlation matrix  
## ML2 ML1 h2 u2 com  
## o1 -0.15 -0.91 0.723 0.28 1.1  
## o2 0.17 -0.42 0.279 0.72 1.3  
## o3 -0.09 -0.24 0.044 0.96 1.3  
## o4 0.38 0.05 0.126 0.87 1.0  
## o5 0.71 0.01 0.497 0.50 1.0  
## o6 0.72 0.15 0.438 0.56 1.1  
## o7 0.37 -0.01 0.139 0.86 1.0  
## o8 0.35 -0.12 0.171 0.83 1.2  
## o9 0.65 -0.07 0.461 0.54 1.0  
## o10 0.52 -0.15 0.373 0.63 1.2  
## o11 -0.50 0.40 0.593 0.41 1.9  
## o12 -0.06 0.82 0.720 0.28 1.0  
## o13 -0.58 -0.06 0.304 0.70 1.0  
## o14 -0.17 0.72 0.659 0.34 1.1  
## o15 0.16 0.27 0.058 0.94 1.6  
## o16 -0.30 0.50 0.477 0.52 1.6  
## o17 -0.52 0.16 0.371 0.63 1.2  
## o18 0.00 0.66 0.435 0.56 1.0  
## o19 -0.57 -0.04 0.307 0.69 1.0  
## o20 -0.48 0.21 0.372 0.63 1.4  
##   
## ML2 ML1  
## SS loadings 3.98 3.56  
## Proportion Var 0.20 0.18  
## Cumulative Var 0.20 0.38  
## Proportion Explained 0.53 0.47  
## Cumulative Proportion 0.53 1.00  
##   
## With factor correlations of   
## ML2 ML1  
## ML2 1.00 -0.48  
## ML1 -0.48 1.00  
##   
## Mean item complexity = 1.2  
## Test of the hypothesis that 2 factors are sufficient.  
##   
## The degrees of freedom for the null model are 190 and the objective function was 10.16 with Chi Square of 837.97  
## The degrees of freedom for the model are 151 and the objective function was 3.62   
##   
## The root mean square of the residuals (RMSR) is 0.09   
## The df corrected root mean square of the residuals is 0.1   
##   
## The harmonic number of observations is 91 with the empirical chi square 292.98 with prob < 3.8e-11   
## The total number of observations was 91 with Likelihood Chi Square = 293.51 with prob < 3.3e-11   
##   
## Tucker Lewis Index of factoring reliability = 0.717  
## RMSEA index = 0.113 and the 90 % confidence intervals are 0.085 0.12  
## BIC = -387.63  
## Fit based upon off diagonal values = 0.92  
## Measures of factor score adequacy   
## ML2 ML1  
## Correlation of (regression) scores with factors 0.93 0.95  
## Multiple R square of scores with factors 0.87 0.90  
## Minimum correlation of possible factor scores 0.74 0.81

twofactor2 = fa(final[ , -c(3, 11, 15, 16)],  
 nfactors = 2,  
 rotate = "oblimin",  
 fm = "ml")  
names(final)

## [1] "o1" "o2" "o3" "o4" "o5" "o6" "o7" "o8" "o9" "o10" "o11"  
## [12] "o12" "o13" "o14" "o15" "o16" "o17" "o18" "o19" "o20"

twofactor2

## Factor Analysis using method = ml  
## Call: fa(r = final[, -c(3, 11, 15, 16)], nfactors = 2, rotate = "oblimin",   
## fm = "ml")  
## Standardized loadings (pattern matrix) based upon correlation matrix  
## ML2 ML1 h2 u2 com  
## o1 -0.14 -0.91 0.75 0.25 1.0  
## o2 0.18 -0.43 0.29 0.71 1.3  
## o4 0.37 0.00 0.14 0.86 1.0  
## o5 0.76 0.01 0.58 0.42 1.0  
## o6 0.75 0.13 0.50 0.50 1.1  
## o7 0.35 -0.03 0.13 0.87 1.0  
## o8 0.34 -0.13 0.17 0.83 1.3  
## o9 0.69 -0.05 0.51 0.49 1.0  
## o10 0.46 -0.20 0.34 0.66 1.4  
## o12 -0.07 0.81 0.71 0.29 1.0  
## o13 -0.47 0.02 0.23 0.77 1.0  
## o14 -0.18 0.72 0.67 0.33 1.1  
## o17 -0.49 0.19 0.36 0.64 1.3  
## o18 -0.02 0.64 0.42 0.58 1.0  
## o19 -0.47 0.02 0.23 0.77 1.0  
## o20 -0.44 0.25 0.34 0.66 1.6  
##   
## ML2 ML1  
## SS loadings 3.39 2.97  
## Proportion Var 0.21 0.19  
## Cumulative Var 0.21 0.40  
## Proportion Explained 0.53 0.47  
## Cumulative Proportion 0.53 1.00  
##   
## With factor correlations of   
## ML2 ML1  
## ML2 1.00 -0.43  
## ML1 -0.43 1.00  
##   
## Mean item complexity = 1.1  
## Test of the hypothesis that 2 factors are sufficient.  
##   
## The degrees of freedom for the null model are 120 and the objective function was 6.92 with Chi Square of 579.89  
## The degrees of freedom for the model are 89 and the objective function was 1.78   
##   
## The root mean square of the residuals (RMSR) is 0.08   
## The df corrected root mean square of the residuals is 0.1   
##   
## The harmonic number of observations is 91 with the empirical chi square 147.7 with prob < 9.3e-05   
## The total number of observations was 91 with Likelihood Chi Square = 146.44 with prob < 0.00012   
##   
## Tucker Lewis Index of factoring reliability = 0.828  
## RMSEA index = 0.094 and the 90 % confidence intervals are 0.059 0.109  
## BIC = -255.03  
## Fit based upon off diagonal values = 0.94  
## Measures of factor score adequacy   
## ML2 ML1  
## Correlation of (regression) scores with factors 0.93 0.95  
## Multiple R square of scores with factors 0.86 0.90  
## Minimum correlation of possible factor scores 0.71 0.79

c. Adequate:  
 i. Include fit indices:  
 1. RMSR: .08 good  
 2. RMSEA: RMSEA index = 0.094 and the 90 % confidence intervals are 0.059 0.109, good/adequate   
 3. TLI: 0.828 poor   
 4. CFI: 0.8751034 poor   
 ii. Include reliability: .82, .85 both good  
 iii. Name your factors:  
 one factor is about the arts, the other factor is thinking philosophy

twofactor2

## Factor Analysis using method = ml  
## Call: fa(r = final[, -c(3, 11, 15, 16)], nfactors = 2, rotate = "oblimin",   
## fm = "ml")  
## Standardized loadings (pattern matrix) based upon correlation matrix  
## ML2 ML1 h2 u2 com  
## o1 -0.14 -0.91 0.75 0.25 1.0  
## o2 0.18 -0.43 0.29 0.71 1.3  
## o4 0.37 0.00 0.14 0.86 1.0  
## o5 0.76 0.01 0.58 0.42 1.0  
## o6 0.75 0.13 0.50 0.50 1.1  
## o7 0.35 -0.03 0.13 0.87 1.0  
## o8 0.34 -0.13 0.17 0.83 1.3  
## o9 0.69 -0.05 0.51 0.49 1.0  
## o10 0.46 -0.20 0.34 0.66 1.4  
## o12 -0.07 0.81 0.71 0.29 1.0  
## o13 -0.47 0.02 0.23 0.77 1.0  
## o14 -0.18 0.72 0.67 0.33 1.1  
## o17 -0.49 0.19 0.36 0.64 1.3  
## o18 -0.02 0.64 0.42 0.58 1.0  
## o19 -0.47 0.02 0.23 0.77 1.0  
## o20 -0.44 0.25 0.34 0.66 1.6  
##   
## ML2 ML1  
## SS loadings 3.39 2.97  
## Proportion Var 0.21 0.19  
## Cumulative Var 0.21 0.40  
## Proportion Explained 0.53 0.47  
## Cumulative Proportion 0.53 1.00  
##   
## With factor correlations of   
## ML2 ML1  
## ML2 1.00 -0.43  
## ML1 -0.43 1.00  
##   
## Mean item complexity = 1.1  
## Test of the hypothesis that 2 factors are sufficient.  
##   
## The degrees of freedom for the null model are 120 and the objective function was 6.92 with Chi Square of 579.89  
## The degrees of freedom for the model are 89 and the objective function was 1.78   
##   
## The root mean square of the residuals (RMSR) is 0.08   
## The df corrected root mean square of the residuals is 0.1   
##   
## The harmonic number of observations is 91 with the empirical chi square 147.7 with prob < 9.3e-05   
## The total number of observations was 91 with Likelihood Chi Square = 146.44 with prob < 0.00012   
##   
## Tucker Lewis Index of factoring reliability = 0.828  
## RMSEA index = 0.094 and the 90 % confidence intervals are 0.059 0.109  
## BIC = -255.03  
## Fit based upon off diagonal values = 0.94  
## Measures of factor score adequacy   
## ML2 ML1  
## Correlation of (regression) scores with factors 0.93 0.95  
## Multiple R square of scores with factors 0.86 0.90  
## Minimum correlation of possible factor scores 0.71 0.79

1 - ((twofactor2$STATISTIC - twofactor2$dof) / (twofactor2$null.chisq - twofactor2$null.dof))

## [1] 0.8751034

alpha(final[ , c(4:10, 13, 17, 19, 20)], check.keys = TRUE)

## Warning in alpha(final[, c(4:10, 13, 17, 19, 20)], check.keys = TRUE): Some items were negatively correlated with total scale and were automatically reversed.  
## This is indicated by a negative sign for the variable name.

##   
## Reliability analysis   
## Call: alpha(x = final[, c(4:10, 13, 17, 19, 20)], check.keys = TRUE)  
##   
## raw\_alpha std.alpha G6(smc) average\_r S/N ase mean sd median\_r  
## 0.82 0.83 0.85 0.3 4.8 0.027 3.7 0.58 0.29  
##   
## lower alpha upper 95% confidence boundaries  
## 0.77 0.82 0.88   
##   
## Reliability if an item is dropped:  
## raw\_alpha std.alpha G6(smc) average\_r S/N alpha se var.r med.r  
## o4 0.82 0.83 0.85 0.32 4.7 0.028 0.016 0.33  
## o5 0.80 0.80 0.82 0.29 4.1 0.030 0.013 0.29  
## o6 0.80 0.81 0.83 0.30 4.2 0.030 0.015 0.29  
## o7 0.82 0.83 0.84 0.33 4.8 0.027 0.013 0.30  
## o8 0.81 0.82 0.84 0.31 4.6 0.028 0.018 0.31  
## o9 0.80 0.81 0.83 0.29 4.2 0.030 0.014 0.29  
## o10 0.80 0.81 0.83 0.29 4.2 0.031 0.017 0.28  
## o13- 0.80 0.81 0.82 0.30 4.2 0.031 0.015 0.29  
## o17- 0.80 0.81 0.83 0.30 4.3 0.030 0.016 0.29  
## o19- 0.81 0.81 0.83 0.30 4.4 0.030 0.015 0.29  
## o20- 0.80 0.81 0.83 0.30 4.3 0.030 0.017 0.29  
##   
## Item statistics   
## n raw.r std.r r.cor r.drop mean sd  
## o4 91 0.48 0.49 0.40 0.37 3.8 0.87  
## o5 91 0.65 0.70 0.69 0.58 4.2 0.68  
## o6 91 0.63 0.66 0.62 0.53 4.2 0.90  
## o7 91 0.43 0.45 0.38 0.31 3.7 0.85  
## o8 91 0.55 0.54 0.46 0.42 3.6 1.05  
## o9 91 0.63 0.67 0.64 0.55 4.2 0.72  
## o10 91 0.68 0.66 0.63 0.57 3.4 1.02  
## o13- 91 0.69 0.65 0.62 0.57 3.4 1.15  
## o17- 91 0.65 0.63 0.58 0.53 3.9 1.08  
## o19- 91 0.63 0.60 0.55 0.51 3.4 1.12  
## o20- 91 0.63 0.62 0.56 0.52 3.5 1.00  
##   
## Non missing response frequency for each item  
## 1 2 3 4 5 miss  
## o4 0.00 0.10 0.22 0.49 0.19 0  
## o5 0.00 0.01 0.11 0.52 0.36 0  
## o6 0.01 0.04 0.11 0.37 0.46 0  
## o7 0.02 0.04 0.31 0.48 0.14 0  
## o8 0.04 0.08 0.36 0.31 0.21 0  
## o9 0.01 0.01 0.07 0.56 0.35 0  
## o10 0.02 0.18 0.33 0.32 0.15 0  
## o13 0.20 0.31 0.22 0.24 0.03 0  
## o17 0.35 0.31 0.20 0.13 0.01 0  
## o19 0.21 0.20 0.36 0.20 0.03 0  
## o20 0.13 0.44 0.23 0.18 0.02 0

alpha(final[ , c(1, 2, 12, 14, 18)], check.keys = TRUE)

## Warning in alpha(final[, c(1, 2, 12, 14, 18)], check.keys = TRUE): Some items were negatively correlated with total scale and were automatically reversed.  
## This is indicated by a negative sign for the variable name.

##   
## Reliability analysis   
## Call: alpha(x = final[, c(1, 2, 12, 14, 18)], check.keys = TRUE)  
##   
## raw\_alpha std.alpha G6(smc) average\_r S/N ase mean sd median\_r  
## 0.85 0.85 0.84 0.53 5.6 0.023 2.4 0.9 0.5  
##   
## lower alpha upper 95% confidence boundaries  
## 0.81 0.85 0.9   
##   
## Reliability if an item is dropped:  
## raw\_alpha std.alpha G6(smc) average\_r S/N alpha se var.r med.r  
## o1- 0.79 0.79 0.76 0.49 3.8 0.033 0.0160 0.46  
## o2- 0.86 0.86 0.84 0.61 6.2 0.023 0.0106 0.60  
## o12 0.80 0.80 0.76 0.50 3.9 0.034 0.0086 0.48  
## o14 0.80 0.80 0.77 0.50 4.1 0.032 0.0152 0.47  
## o18 0.84 0.83 0.81 0.55 4.9 0.026 0.0260 0.55  
##   
## Item statistics   
## n raw.r std.r r.cor r.drop mean sd  
## o1- 91 0.86 0.85 0.82 0.76 2.4 1.13  
## o2- 91 0.64 0.67 0.53 0.49 2.0 0.93  
## o12 91 0.86 0.84 0.82 0.75 2.4 1.30  
## o14 91 0.85 0.83 0.79 0.73 2.8 1.27  
## o18 91 0.74 0.75 0.65 0.60 2.7 1.01  
##   
## Non missing response frequency for each item  
## 1 2 3 4 5 miss  
## o1 0.04 0.12 0.24 0.33 0.26 0  
## o2 0.02 0.05 0.15 0.48 0.29 0  
## o12 0.33 0.29 0.15 0.15 0.08 0  
## o14 0.16 0.33 0.19 0.21 0.11 0  
## o18 0.14 0.24 0.43 0.15 0.03 0

# Openness to Experience Scale

Below are some phrases describing people’s behaviors. Please use the rating scale below to describe how accurately each statement describes you. Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself in relation to other people of your gender and of roughly your same age. Please read each statement carefully, and then check the box that corresponds to your response.

Scale: very inaccurate, moderately inaccurate, neither inaccurate nor accurate, moderately accurate, very accurate

1. Believe in the importance of art.
2. Have a vivid imagination.
3. Tend to vote for liberal political candidates.
4. Carry the conversation to a higher level.
5. Enjoy hearing new ideas.
6. Enjoy thinking about things.
7. Can say things beautifully.
8. Enjoy wild flights of fantasy.
9. Get excited by new ideas.
10. Have a rich vocabulary.
11. Am not interested in abstract ideas.
12. Do not like art.
13. Avoid philosophical discussions.
14. Do not enjoy going to art museums.
15. Tend to vote for conservative political candidates.
16. Do not like poetry.
17. Rarely look for a deeper meaning in things.
18. Believe that too much tax money goes to support artists.
19. Am not interested in theoretical discussions.
20. Have difficulty understanding abstract ideas.