PSYC 7433

Gordon Goodwin

**Problem Set 5**

1. In the file pfs.csv is data derived from *N* = 229 parents enrolled in community-based family support programs in Texas and Kansas.

A preliminary version of the Protective Factors Survey, a scale intended to capture individual differences in constructs thought to inhibit child abuse, was administered to parents (*N* = 229) enrolled in community-based family support programs in Texas and Kansas. Each question was answered on a 7-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree” (see Table 1 for items and variables). As seen in Figure 1, bivariate correlation analyses revealed that there was a relatively high degree of correlation amongst many of the items, indicating that the survey and its associated responses would be a good candidate for factor analysis. In doing so, the goal was to model the interrelationships amongst the observable item responses as a function of fewer underlying latent variables related to the inhibition of child abuse. In order to determine the appropriate number of factors to extract, a parallel analysis, MAP test, and screeplot were all conducted using the *psych* package (Revelle, 2021) in *R* (R Core Team, 2021). Results of the various analyses were ambiguous, with the 1976 MAP test suggesting 1 factor, the 2000 MAP test suggesting three factors, the parallel analysis with principal axis factoring suggesting five factors, the Kaiser criterion suggesting one factors, and the scree plot indicating either one or two factors. As such, five separate models were fit utilizing principal axis factoring and an oblique (Oblimin) rotation to extract one, two, three, four, and five-factor solutions. Pattern matrices for each of the models were then compared to ascertain which solution yielded the “best” simple structure both quantitatively and in terms of interpretability in the context of child abuse inhibition. All loadings less than a threshold of 0.4 were suppressed in order to facilitate simple structure by ensuring that only items with a moderate to strong loading were considered.

After doing so, ultimately a four-factor solution was chosen (see Figure 2 for pattern matrix/loadings) due to interpretability of results in the context of inhibition of child abuse. Specifically, analysis of the simple structure suggested a relatively clear “theme” common to the items that loaded on each of the four factors. As seen in Table 2, the items associated with the first factor (PA1; Support) seemed to all be broadly related to feelings of support, those on PA2 seemed related to the parents’ utilization of behavioral discipline (Behavioral Discipline), those on PA3 seemed related to corporal punishment (Corporal Discipline), and the items associated with PA4 seemed related to feelings of family-related negativity (Negative Outlook). It should also be noted that due to the suppression of loadings less than a threshold of 0.4, there were five items originally considered in the survey that were discarded due to lack of sufficient loading. None of the items that were discarded seem to be glaringly related to any of the factors that were extracted or their constituent items, although this is of course a completely subjective assessment. Regarding model fit, relevant statistics can be seen in Table 3, which shows that the first factor (PA1; Support) was associated with 50% of the total variance explained by the model, with the remainder of the variance explained split roughly equally across the other three factors. In addition, both fit indices consulted (see Table 3 again) suggested an adequate model fit. Lastly, as seen in Table 4, weak to moderate correlations between several of the factors were observed. Specifically, Support had a mild to moderate positive correlation with each of the other three factors, and Corporal Discipline was positively correlated with Negative Outlook. Upon review, the preliminary version of the Protective Factors Survey seems to be relatively successful in its stated goal of measuring constructs related to the inhibition of child abuse. However, there was not a clear “solution” regarding the number of factors to extract, and there were multiple items that did not load strongly on any of the factors. Domain experts should be consulted, and alternative decisions regarding both the number of factors to be extracted and the interpretation of said factors should be considered prior to finalizing the scale.

**Table 1.**

*Protective Factors Survey Variables and Items*

|  |  |
| --- | --- |
| Variable | Item |
| PFS1 | I have neighbors, friends or relatives that help me when I need it |
| PFS2 | My family members feel closer to people outside the family than to our own family members |
| PFS3 | I know where to go in my community to get help with family needs |
| PFS4 | More bad things happen to my family than to other families |
| PFS5 | My family enjoys spending time together |
| PFS6 | When I am worried about my children, I have someone to talk to |
| PFS7 | I don't think my family can survive if another problem hits us |
| PFS8 | I praise my children when they behave well |
| PFS9 | My family shows each other love and affection |
| PFS10 | My family is able to solve our problems |
| PFS11 | When we have disagreements, family members listen to both sides of the story |
| PFS12 | When I discipline my children, I have a hard time keeping my feelings under control |
| PFS13 | I try to comfort my children when something is bothering them |
| PFS14 | My family members discuss problems with each other |
| PFS15 | In my family, we take time to listen to each other |
| PFS16 | I try to take a break when I am frustrated by my children's behavior |
| PFS17 | In my family, we support one another when something goes wrong |
| PFS18 | When my child misbehaves, I... use time-out |
| PFS19 | When my child misbehaves, I... spank |
| PFS20 | When my child misbehaves, I... hit |
| PFS21 | When my child misbehaves, I... ground |
| PFS22 | When my child misbehaves, I... take away privileges |

**Table 2.**

*Assignation of Items Post-Extraction*

|  |  |  |
| --- | --- | --- |
| Variable | Item | Factor Assigned |
| PFS1 | I have neighbors, friends or relatives that help me when I need it | Support |
| PFS3 | I know where to go in my community to get help with family needs | Support |
| PFS5 | My family enjoys spending time together | Support |
| PFS6 | When I am worried about my children, I have someone to talk to | Support |
| PFS9 | My family shows each other love and affection | Support |
| PFS10 | My family is able to solve our problems | Support |
| PFS11 | When we have disagreements, family members listen to both sides of the story | Support |
| PFS14 | My family members discuss problems with each other | Support |
| PFS15 | In my family, we take time to listen to each other | Support |
| PFS17 | In my family, we support one another when something goes wrong | Support |
| PFS18 | When my child misbehaves, I... use time-out | Behavioral Discipline |
| PFS21 | When my child misbehaves, I... ground | Behavioral Discipline |
| PFS22 | When my child misbehaves, I... take away privileges | Behavioral Discipline |
| PFS19 | When my child misbehaves, I... spank | Corporal Discipline |
| PFS20 | When my child misbehaves, I... hit | Corporal Discipline |
| PFS7 | I don't think my family can survive if another problem hits us | Negative Outlook |
| Variable | Item | Factor Assigned |
| PFS4 | More bad things happen to my family than to other families | Negative Outlook |
| PFS2 | My family members feel closer to people outside the family than to our own family members | - |
| PFS8 | I praise my children when they behave well | - |
| PFS12 | When I discipline my children, I have a hard time keeping my feelings under control | - |
| PFS13 | I try to comfort my children when something is bothering them | - |
| PFS16 | I try to take a break when I am frustrated by my children's behavior | - |

*Note*. All loadings less than 0.4 were suppressed.

**Table 3.**

*SS Loadings and Variance Explained*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Support | Behavioral Discipline | Corporal Discipline | Negative Outlook |
| SS Loadings | 4.03 | 1.51 | 1.32 | 1.03 |
| Proportion Var. | 0.18 | 0.07 | 0.06 | 0.05 |
| Cumulative Var. | 0.18 | 0.25 | 0.31 | 0.36 |
| Proportion Exp. | 0.51 | 0.19 | 0.17 | 0.13 |
| Cum. Prop. Exp. | 0.51 | 0.70 | 0.87 | 1.00 |
|  |  |  | *RMSEA* = 0.052 | |
|  |  |  | *TLI* = 0.869 | |

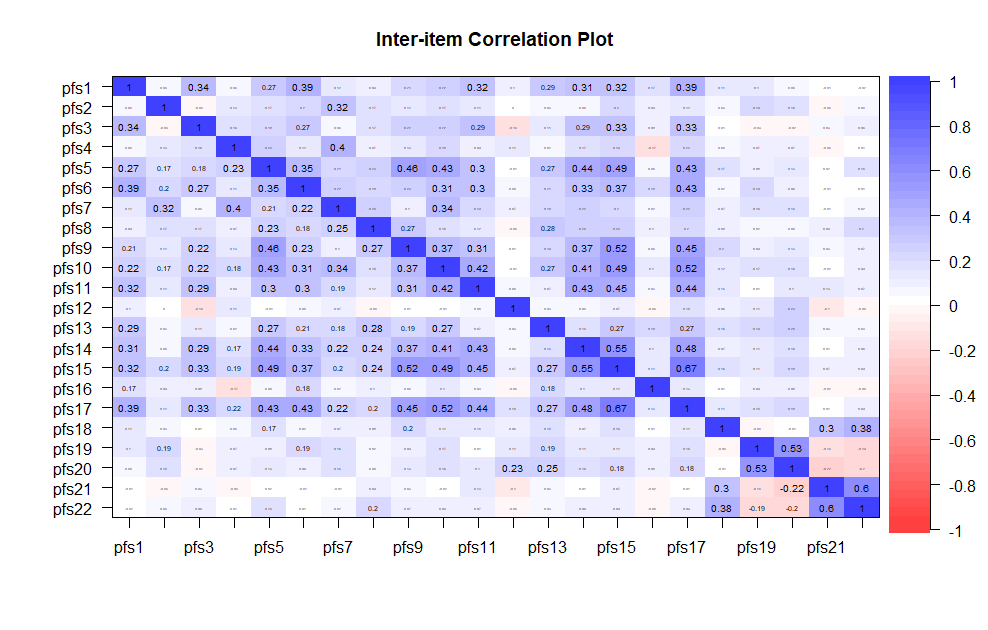
**Table 4.**

*Factor Correlations*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Support | Behavioral Discipline | Corporal Discipline | Negative Outlook |
| Support | 1.00 |  |  |  |
| Behavioral Disc. | 0.14 | 1.00 |  |  |
| Corporal Disc. | 0.26 | -.022 | 1.00 |  |
| Neg. Outlook | 0.33 | 0.06 | 0.21 | 1.00 |

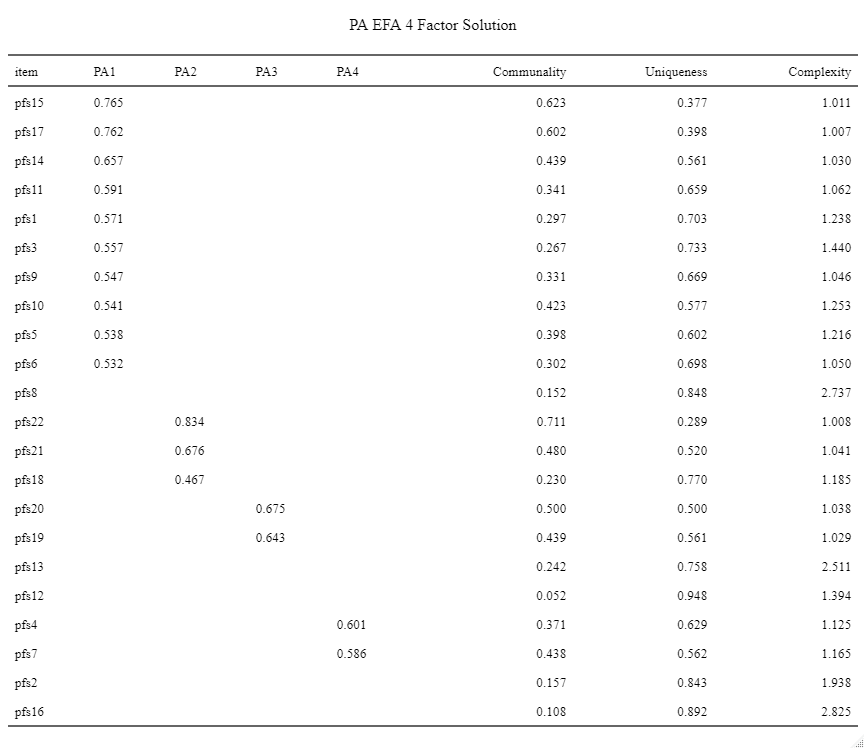
**Figure 1.**

*Zero-order Correlation Matrix Heat Map*



**Figure 2.**

*Pattern Matrix for Four Factor Solution With Principal Axis Factoring and Oblimin Rotation*



*Note*. All loadings less than 0.4 were suppressed.

**References**

R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.<https://www.R-project.org/>.

Revelle, W. (2021). *Psych: Procedure for personality and psychological research*. Northwestern University, Evanston, Illinois, USA. [https://CRAN.R-project.org/package=psych/Version = 2.1.6](https://CRAN.R-project.org/package=psych/Version%20=%202.1.6)

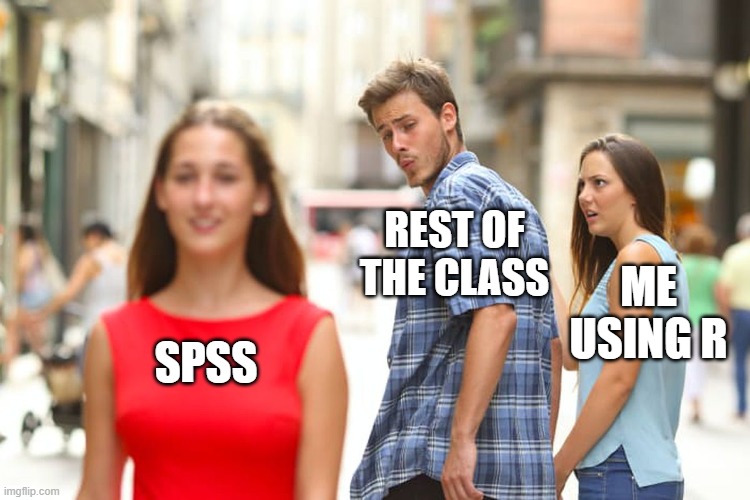
1. **Find an article published in your area of research using cluster analysis/mixture modeling and write a short review**

Wood Jr, F. R., & Graham, R. (2020). “Safe” and “At-Risk”: Cyberbullying victimization and deviant health risk behaviors in youth. *Youth & Society, 52*(3), 449-468. <https://doi.org/10.1177%2F0044118X18810943>

This non-longitudinal study investigated the links between cyberbullying (CB) victimization and a set of health risk behaviors associated with juvenile delinquency (cigarette smoking, marijuana and alcohol usage, and sexual frequency), while controlling for self-control and deviant peer associations. The links between CB victimization and specified deviant health risk behaviors were examined within a theoretical framework aligned with General Strain Theory, in which the presence of unwanted stimuli (i.e. hostile messages intended to inflict harm or discomfort), are hypothesized to produce negative emotions that can lead to deviant behavior (Agnew, 2001). Data were collected from the 2015 *Youth Risk Behavior Survey*, which consisted of a representative sample of 9-12th graders in the U.S. Specifically, this study focused upon the items in which respondents were asked to self-report whether or not they had experienced CB and/or traditional bullying victimization within the previous 12 months, the number of recent sexual partners, and the frequency of recent substance abuse (marijuana and/or alcohol). Using cluster analysis, respondents with complete data observations (*N* = 9,122) were categorized into two groups: “safe” students who reported, on average, no engagement in the behaviors measured, and “at-risk” students who reported, on average, moderate to high levels of engagement in the health risk behaviors outlined above. Following the cluster analysis, the data were also subset into participants who were not physically bullied and those who were bullied.

Logistic regression models were then used to predict group membership (‘at-risk’ vs ‘safe’) from CB victimization, controlling for gender, age, race, and sexual orientation. Two models were fit on each physical bullying subset, one of which included a self-control measure and one of which did not, for a total of four separate models (see figure below). Results indicated that CB victimization significantly increases the odds of a student belonging to the ‘at-risk’ cluster, with the effect remaining consistent when controlling for exposure to physical bullying, a proxy measure of self-control, and various demographic variables. Key limitations of this study include the reliance upon a measure of grades/academic outcomes as an indirect measure of self-control, as well as the reliance upon self-report data (memory, social desirability bias). Despite the other limitations, this was an appropriate use of both cluster analysis and logistic regression given that the cluster analysis was used to form a dichotomous grouping of ‘safe’ vs ‘at-risk’ participants. Future studies should incorporate more precise measures of strain, peer associations, and self-control, as well as utilize a mixed-methods approach to examine qualitative data as well.

**Extra Credit (3 points).** Create a meme about how any topic in multivariate statistics makes you feel. There are lots and lots of meme generators online no need to break out Photoshop skills.



**Syntax**

library(tidyverse)

library(psych)

library(psychTools)

library(readr)

library(EFA.dimensions)

pfs<-read.csv("pfs.csv")

View(pfs)

names(pfs)<-tolower(names(pfs))

# Getting a feel for things

describe(pfs)

apply(pfs,2,hist)

pairs.panels(pfs)

lowerCor(pfs)

corPlot(pfs)

# Determine number of factors

fa.parallel(pfs, fm="pa", fa="both") # Parallel says 5

MAP(pfs) # 1976 MAP says 1, 2000 says 3

pfic<-iclust(pfs)

summary(pfic) # Hierarchical Cluster algo based on reliability says 1

# Perform FA

# PA with 1 Factor & 3 Factors

m1\_1<-fa(pfs, nfactors = 1, rotate = "oblimin", fm = "pa")

m1\_2<-fa(pfs, nfactors = 2, rotate = "oblimin", fm = "pa")

m1\_3<-fa(pfs, nfactors = 3, rotate = "oblimin", fm = "pa")

m1\_4<-fa(pfs, nfactors = 4, rotate = "oblimin", fm = "pa")

m1\_5<-fa(pfs, nfactors = 5, rotate = "oblimin", fm = "pa")

# ML with 1 & 3 Factors

m2\_1<-fa(pfs, nfactors = 1, rotate = "oblimin", fm = "ml")

m2\_2<-fa(pfs, nfactors = 2, rotate = "oblimin", fm = "ml")

m2\_3<-fa(pfs, nfactors = 3, rotate = "oblimin", fm = "ml")

m2\_4<-fa(pfs, nfactors = 4, max.iter=100, rotate = "oblimin", fm = "ml")

# Selected 4 Factor solution (MLE and PA were equivalent)

m1\_4

plot(m1\_4)

fa.diagram(m1\_4)

library(palmerpenguins)

library(tidyverse)

library(psych)

library(flextable)

# table set-up ----

flex <- function(data, title=NULL) {

# this grabs the data and converts it to a flextbale

flextable(data) %>%

# this makes the table fill the page width

set\_table\_properties(layout = "autofit", width = 1) %>%

# font size

fontsize(size=10, part="all") %>%

#this adds a ttitlecreates an automatic table number

set\_caption(title,

autonum = officer::run\_autonum(seq\_id = "tab",

pre\_label = "Table ",

post\_label = "\n",

bkm = "anytable")) %>%

# font type

font(fontname="Times New Roman", part="all")

}

fa\_table <- function(x, cut) {

#get sorted loadings

loadings <- fa.sort(x)$loadings %>% round(3)

#supress loadings

loadings[loadings < cut] <- ""

#get additional info

add\_info <- cbind(x$communality,

x$uniquenesses,

x$complexity) %>%

# make it a data frame

as.data.frame() %>%

# column names

rename("Communality" = V1,

"Uniqueness" = V2,

"Complexity" = V3) %>%

#get the item names from the vector

rownames\_to\_column("item")

#build table

loadings %>%

unclass() %>%

as.data.frame() %>%

rownames\_to\_column("item") %>%

left\_join(add\_info) %>%

mutate(across(where(is.numeric), round, 3))

}

# Tables for SimpStructure

fa\_table(m1\_1, 0.4) %>% flex("PA EFA 1 Factor Solution")

fa\_table(m1\_2, 0.4) %>% flex("PA EFA 2 Factor Solution")

fa\_table(m1\_3, 0.4) %>% flex("PA EFA 3 Factor Solution")

fa\_table(m1\_4, 0.4) %>% flex("PA EFA 4 Factor Solution")

fa\_table(m1\_4, cut = 0) %>% flex("PA EFA 4 Factor Solution")

fa\_table(m1\_5, cut = 0.4) %>% flex("PA EFA 5 Factor Solution")

fa\_table(m2\_1, 0.4) %>% flex("ML EFA 1 Factor Solution")

fa\_table(m2\_2, 0.4) %>% flex("ML EFA 2 Factor Solution")

fa\_table(m2\_3, 0.4) %>% flex("ML EFA 3 Factor Solution")

fa\_table(m2\_4, 0.4) %>% flex("ML EFA 4 Factor Solution")