Measurement, Reliability, Validity, and Latent Factors

Week 3
POLS 8830: Advanced Quantitative Methods

Ryan Carlin
Georgia State University
rcarlin@gsu.edu

Measurement Issues

- Measurement issues fundamental
- Best theory + best model but poor measures . . . meaningless results
- The discipline has not always been terribly interested in measurement
- Perfect measures don't exist

Measurement Issues

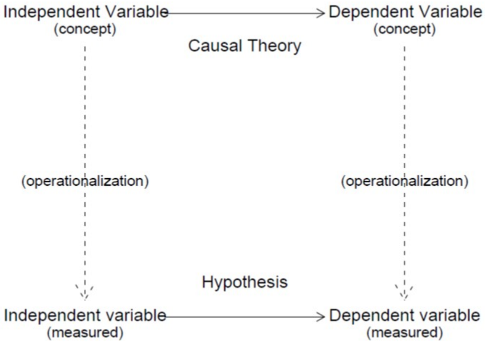
- Poor measures often by-product of the phenomena we theorize about – ideology, political knowledge, political efficacy/alienation
    - Different than hard sciences (e.g., temperature, weight, mass, etc.)
- Frequentist paradigm built on the notion of repeated samples but. . .
- Multiple samples & repeated measures often unfeasible/impractical
- We usually have 1 sample & hope for best $\rightarrow$ crisis of replicability

Attitude

- You care about measurement, embrace it warts and all
- Extant lit full of uncompelling, sometimes downright bad measures... Can we improve them? Can we do better w/o collecting new data?
- What is the underlying construct? How can be operationalized it? What shortcomings does it present? How they matter for inference?
- Focus on
  - Reliability & validity
  - Measurement error
  - Unobservable/latent constructs

The General Goal of Social Science

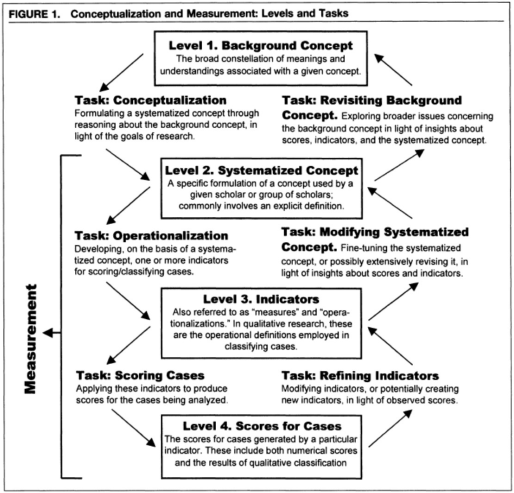
- Some variables observed, others unobserved; all must be measured
- Operationalization connects theory, measure & model
- Measures have varying degrees of reliability & validity, depending on nature of their errors



Clarifying Concepts, i.e. How to Conceptualize

- Inventory the concept's properties
  - Exclude properties that belong to other concepts
  - Exclude impossible to measure conceptual terms
  - Beware of multiple dimensions w/in the concept
    - Dimensions: groups empirically similar, qualitatively distinct
- Properties must:
  - Be concrete
  - Vary, or at least have definable limits
- Operationalize your concept

Adcock & Collier 2001



Reliability

- Definition p. 11 Carmines & Zeller:
- Not a measure of accuracy (validity)
- Boils down to consistency: repeated measures, roughly the same results
- More of a theoretical concept because we don't always re-measure within the same study
- Cronbach's α coefficient

Reliability

- Reliability refers to "the extent to which an experiment, test, or any measuring procedure yields the same result on repeated trials" (Carmines and Zeller 1979, 11)
- Recognition that perfection is not possible
- Rather, focus on consistency across repeated measurements
- More consistency = higher reliability, less consistency = lower reliability

Reliability: Cronbach's Alpha

- Cronbach's Alpha is the measure of the internal consistency of a scale, or a related set of scales

$$\alpha = \frac{N \cdot \bar{c}_{ij}}{\bar{v}_{ij} + (N-1) \cdot \bar{c}_{ij}} \tag{1}$$

  - α: Alpha coefficient (of reliability)
  - N: Number of test items
  - $\bar{c}_{ij}$: Average inter-item covariance among the items
  - $\bar{v}_{ij}$: Average inter-item variance among the items
  - i and j: items

Reliability: Cronbach's Alpha

- This comes with a very simple implementation in the *psych* package
  - `alpha(x, keys=NULL,cumulative=FALSE, title=NULL, max=10,na.rm = TRUE, check.keys=FALSE, n.iter=1,delete=TRUE,use="pairwise", warnings=TRUE, n.obs=NULL,impute=NULL )`
- Key elements:
  - x: data frame or matrix to be analyzed
  - keys: allows for the reverse coding of variables
  - na.rm: Removing NA's (TRUE) or not (FALSE)
  - use: determines the type of correlation calculation to use
  - impute: imputation procedure for missing data

Validity

- Definition p. 12 Carmines & Zeller:
- **Criterion validity:**
  - See Adcock & Collier's "convergent validity", a measure converges strong with an established indicator of the concept

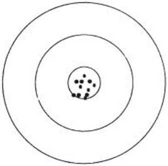
Validity

- Definition p. 12 Carmines & Zeller:
- **Criterion validity:**
  - See Adcock & Collier's "convergent validity", a measure converges strong with an established indicator of the concept
- **Content validity:** does measure capture the entirety of the concept

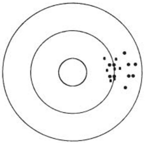
Validity

- Definition p. 12 Carmines & Zeller:
- **Criterion validity:**
  - See Adcock & Collier's "convergent validity", a measure converges strong with an established indicator of the concept
- **Content validity:** does measure capture the entirety of the concept
- **Construct validity:** Is measure correlated with things we expect it to be correlated with? Does it predict what it should? Do things that should predict it, in fact, predict it? Repeated, iterative process
- *Note:* we're talking about measurement validity; internal & external validity are related concepts

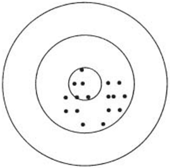
Validity and Reliability Combined



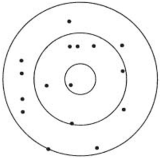
(a) Valid & Reliable



(b) Reliable but not Valid



(c) Fairly Valid but not Reliable



(d) Neither Valid nor Reliable

Validity and Reliability Combined: Correlation Coefficients

- You can use correlation to check for the reliability of measures which use the same approach, or the validity that they are both capturing the same concept if they are measured in different ways
- Pearson's r: for use with continuous data, where relationship is monotonic, linear, and there are no outliers
- Kendall's τ: for use with ordinal data, where comparing comparable groups, and where the relationship is monotonic
- Spearman's ρ: Relaxes the linearity assumption of Pearson's r

Validity and Reliability Combined: Correlation Coefficients

- Pearson's Correlation Coefficient

$$r = \frac{\Sigma(a - \bar{x}_a)(b - \bar{x}_b)}{\sqrt{\Sigma(a - \bar{x}_a)^2 \Sigma(b - \bar{x}_b)^2}} \tag{2}$$

- Where, a and b are vectors of length n, and $\bar{x}_b$ and $\bar{x}_b$ are the sample means of a and b
- Significance determined by:

$$t = \frac{r}{\sqrt{1 - r^2}} \sqrt{n - 2} \tag{3}$$

- $df = n - 2$

Validity and Reliability Combined: Correlation Coefficients

- In R correlation can be tested with:
- `cor()`
- `cor.test()`
- `cor()` only provides the correlation coefficient while `cor.test()` provides additional information including statistical signficance
- Options in `cor.test()` include:
  - `method = c("pearson", "kendall", "spearman")`

Validity and Reliability Combined: Correlation Coefficients

- The output of cor.test will appear as:

```
        Pearson's product-moment correlation

data:  testSet$VCF0210 and testSet$VCF0231
t = 41.5, df = 5403, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.4711603 0.5115997
sample estimates:
     cor
0.491645
```

- The bottom-most value is the correlation coefficient
- The p-value is your indication of statistical significance
- This also provides the 95% confidence interval

Measurement Error

- Random error hurts reliability
- Nonrandom error hurts validity
- X = t + e (true score: avg score after infinite measure of X)

Unobservable Variables

- Latent variable models assume there is some underlying variable that directly measures our concept of interest
- Can't observe or measure the concept directly but...
- We can model the latent variable as a function of a series of items that we believe form part of the concept of interest (based on theory)
- Goal is always to let the data decide – not an arbitrary set of decisions – what should be in an index
- Theory & RQ guide approach (exploratory v. confirmatory)
- Mathematically equivalent, some difference in techniques/mindset/criteria

Concept: Who is a Democrat?

- Is consistently liberal, unconditional, committed
- Prefers democracy to authoritarian alternatives
  - All authoritarian regimes (Communism, military, etc.)
  - Coups d'Etat
- Has positive orientations to democratic norms and principles
  - Public Contestation, Inclusive participation, Checks & Balances/Horizontal Accountability
  - Free & Fair Elections, Rule of Law, Political Participation, Tolerance
- Finds democratic political system & institutions legitimate

Factor Analysis

- What is the min number of common factors needed to satisfactorily reproduce the correlations among observed variables?
- "...assumes that the observed variables are linear combinations of some underlying (hypothetical or unobservable) factors."
- Some factors assumed "common" (correlated) to ≥ 2 variables
- Some factors unique, assumed orthogonal to each other (hence, don't contribute to shared variance)
- Factors extracted based on correlation matrix of all observed variables + rotation

Measures: Democracy vs. Authoritarian Alternatives

- **JC15A:** "When the country is facing difficulties is it justifiable for the president of the republic to close Congress and govern without the Congress?"
- **DEM11:** "Do you believe our country needs a government with an iron fist or that our problems can be resolved with the participation of all?"
- **JC Series** ($\alpha = .83$): Some people say that under some circumstances a military take-over through a coup d'état would be justified. In your opinion would a military coup be justified in the following circumstances?:
    - JC1. When there is high unemployment.
    - JC4. When there are a lot of social protests.
    - JC10. When there is a lot of crime.
    - JC12. When there is high inflation, with excessive prices increases.
    - JC13. When there is a lot of corruption.

Measures: Democracy vs. Authoritarian Alternatives

- **ING4. (Churchillian)** "Democracy may have some problems but it is better than any other form of government. To what extent do you agree?" 1-7 scale → 0-100
- **LINZ.** "For people like me, it doesn't matter whether we have an authoritarian regime or a democratic one"
- "Democracy is preferable to any other forms of government"
- "In some circumstances, an authoritarian government could be preferable to a democratic one"

Measures: Orientations to Democratic Norms & Principles

Political Tolerance: $\alpha = 0.762$

"There are people who speak negatively of the nationality country form of government, not just the incumbent government but the system of government. How strongly do you approve or disapprove of . . .

- **D1:** of such people's right to vote?"
- **D2:** that such people be allowed to conduct peaceful demonstrations in order to express their views?"
- **D3:** such people being permitted to run for public office?"
- **D4:** such people appearing on television to make speeches?"

Measures: System Support

System Support: $\alpha = 0.756$

"To what point do you ...

- **B1:** believe that the justice system of [COUNTRY] guarantees a fair trial?"
- **B2:** have respect for the political institutions of [COUNTRY]?"
- **B3:** believe that the citizens' basic rights are well protected by the [COUNTRY's] political system?"
- **B4:** feel proud to live under the [COUNTRY] political system?"
- **B5:** think that you should support the [COUNTRY] political system?"

Partitioning Variance

- Factor analysis assumes that variance can be partitioned into two types of variance, common and unique
- **Common variance:** amount of variance shared among a set of items. Items that are highly correlated will share a lot of variance.
  - **Communality** (h^2) scales common variance 0-1. Values closer to 1 suggest that extracted factors explain more of the variance of an individual item.
- **Unique variance:** any portion of variance that's not common ($1 - h^2$)
  - **Specific variance:** specific to a particular item
  - **Error variance:** errors of measurement & anything else unexplained

Factor Analysis Calculation

- Covariance matrix & overall variance
- Factor loading: relationship of each variable to underlying factor(s)
  - Can be interpreted as standardized regression coefficients, hence can compare to understand strength
- Eigenvalue = sum of squared component loadings across all items for each component
  - Captures of how much of the variance of the observed variables a factor explains.
  - Kaiser criterion: Any factor with an eigenvalue ≥ 1 explains more variance than a single observed variable and should be "kept".
- Factor score: a regression based method to predict the value of latent variable for any individual

Analysis: Structure of Attitudes

Component	Eigenvalue	Proportion Var.	Cumulative Var.
1	2.10	.23	.23
2	1.42	.16	.39
3	1.04	.12	.51

(a) Unrotated Principal Components Analysis

Variable	1	2	3
Linzian Support	.40	-.22	-.07
Churchillian Support	.40	-.28	-.53
System Support	.11	-.18	.81
Political Tolerance	-.21	.25	.41
Pulse Firme	.35	-.18	.09
Elected Leader > Non-Elected Strongman	.39	-.25	-.20
Reject Pres. Close & Govern w/o Congress	.21	.25	.32
Coup Not Justified when Crime is High	.39	.54	-.03
Coup Not Justified when Corruption is High	.39	.53	-.06

(b) Principal Components

Implementation: Factor Analysis

- One way to accomplish this are the functions in the `psych()` family of packages
- Before performing factor analysis (FA), there are a set of requirements to determine if your data are suitable for FA
- These are generally geared towards identifying if your factors of interest are correlated at a sufficient level such that the FA process will result in meaningful results
  - FA on entirely random data will still produce factors, but these will be meaningless

Implementation: Factor Analysis

- Can perform `cor()` on each variable dyad
- More economical approach to use `lowerCor()`, `pairs.panels()`, and `cor.plot()`

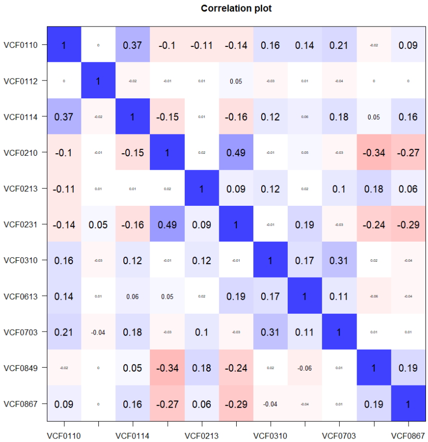
Implementation: Factor Analysis

- Can perform `cor()` on each variable dyad
- More economical approach to use `lowerCor()`, `pairs.panels()`, and `cor.plot()`
- `lowerCor(x)` provides pairwise correlation coefficients for each variable dyad in data.frame **x**

	VCF0110	VCF0112	VCF0114	VCF0210	VCF0213	VCF023	VCF03	VCF06	VCF07	VCF084	VCF086
VCF0110	1.00										
VCF0112	0.00	1.00									
VCF0114	0.37	-0.02	1.00								
VCF0210	-0.10	-0.01	-0.15	1.00							
VCF0213	-0.11	0.01	0.01	0.02	1.00						
VCF0231	-0.14	0.05	-0.16	0.49	0.09	1.00					
VCF0310	0.16	-0.03	0.12	-0.01	0.12	-0.01	1.00				
VCF0613	0.14	0.01	0.06	0.05	0.02	0.19	0.17	1.00			
VCF0703	0.21	-0.04	0.18	-0.03	0.10	-0.03	0.31	0.11	1.00		
VCF0849	-0.02	0.00	0.05	-0.34	0.18	-0.24	0.02	-0.06	0.01	1.00	
VCF0867	0.09	0.00	0.16	-0.27	0.06	-0.29	-0.04	-0.04	0.01	0.19	1.00

Implementation: Factor Analysis

- `cor.plot(x)` provides a graphical presentation of the correlation coefficients for each variable dyad in data.frame **x**



Implementation: Factor Analysis

- These can provide preliminary data to determine what factors are likely to produce meaningful latent factor estimate
- However, there are two ‘Tests’ to determine if FA is appropriate
- These are generally referred to as ‘Tests of Adequacy’
  - Kaiser-Meyer Olkin Measure of Sampling Adequacy
  - Bartlett’s Test of Sphericity

Implementation: Factor Analysis

- Kaiser-Meyer Olkin Measure of Sampling Adequacy
- `KMO(x)`, where x is the data.frame in use
- Looking for values greater than 0.5

```
Kaiser-Meyer-Olkin factor adequacy
Call: KMO(r = faData)
Overall MSA =  0.65
MSA for each item =
VCF0110 VCF0112 VCF0114 VCF0210 VCF0213 VCF0231 VCF0310 VCF0613 VCF0703 VCF0849 VCF0867
   0.62    0.46    0.67    0.66    0.45    0.64    0.62    0.59    0.65    0.68    0.77
```

- Most suited for low-n datasets where variable to observation ratio is around 1:5

Implementation: Factor Analysis

- Bartlett's Test of Sphericity
- `cortest.bartlett(x)`, where `x` is a matrix
  - You can coerce your data.frame to a matrix by replacing `x` with `as.vector(x)` within the test
- You're looking for something like this, where a *p*-value of less than 0.05 is a "Pass"
  - Technically rejecting the null that the data are not suitable

```
> cortest.bartlett(as.matrix(faData, force = TRUE), n=5914)       # Bartlett's Test of Sphericity
R was not square, finding R from data
$chisq
[1] 6573.485

$p.value
[1] 0

$df
[1] 55
```

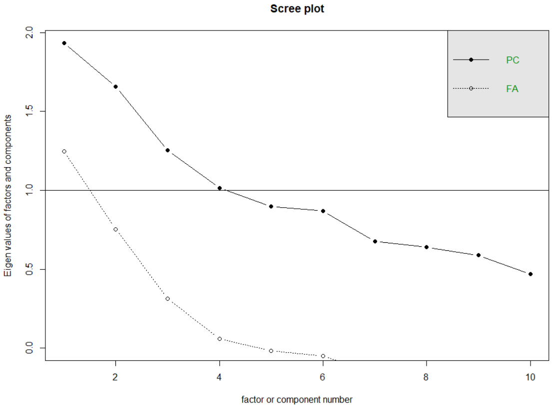
- Limited information for *p*-value; can calculate this manually using the χ^2 and df values

Implementation: Factor Analysis

- Once the preliminary requirements out of the way, the next step in exploratory FA is to determine the appropriate number of factors
- Good way to do this is the `scree(x)` and `fa.parallel(x)` graphs
- Again, `x` is just the data.frame you're using
- Scree plots illustrate the diminishing Eigenvalues for each factor based on the number of factors specified
  - Remember, looking for Eigenvalues in excess of 1.

Implementation: Factor Analysis

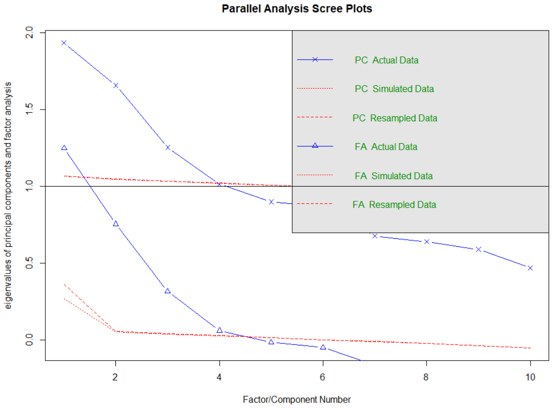
- `scree(x)`



- PC: Principal Component
- FA: Factor Axis Extraction

Implementation: Factor Analysis

- `fa.parallel(x)`



- Using simulated data to provide more information

Implementation: Factor Analysis

- Once you've decided on the appropriate number of factors, use `fa()`
-
```
fa(r,nfactors=1,n.obs = NA,n.iter=1,
rotate="oblimin", scores="regression",
residuals=FALSE, SMC=TRUE,
covar=FALSE,missing=FALSE,impute="median",
min.err = 0.001, max.iter = 50,symmetric=TRUE,
warnings=TRUE, fm="minres",
alpha=.1,p=.05,oblique.scores=FALSE,np.obs=NULL,
use="pairwise",cor="cor",
correct=.5,weight=NULL,...)
```
- Most of this is unnecessary

Implementation: Factor Analysis

- r: dataframe under use
- nfactors: specifying the number of factors to predict
- rotate: options for how to rotate the matrix after the solution is found (preventing correlation amongst factors)
- missing: set to TRUE if you want missing values imputed
- impute: either mean of median values used for imputation (matters if inputs are continuous/ordinal or binary)
- fm: factoring method
- alpha: confidence levels for RMSEA (absolute fit index)
- use: treatment of missing data

Implementation: Factor Analysis

```
> f1<-fa(faData, nfactors = 2, n.obs=n, fm = "pa", rotate = "varimax", impute = "mean") # Creating the factor analysis object;
> f1
Factor Analysis using method =  pa
Call: fa(r = faData, nfactors = 2, n.obs = n, rotate = "varimax", 
    impute = "mean", fm = "pa")
Standardized loadings (pattern matrix) based upon correlation matrix
          PA1   PA2     h2   u2 com
VCF0110 -0.13  0.53 0.2949 0.71 1.1
VCF0112  0.02 -0.04 0.0021 1.00 1.6
VCF0114 -0.22  0.45 0.2515 0.75 1.5
VCF0210  0.69 -0.05 0.4824 0.52 1.0
VCF0213  0.00  0.05 0.0025 1.00 1.0
VCF0231  0.70 -0.04 0.4859 0.51 1.0
VCF0310  0.05  0.44 0.1932 0.81 1.0
VCF0613  0.18  0.28 0.1135 0.89 1.7
VCF0703  0.00  0.48 0.2301 0.77 1.0
VCF0849 -0.41 -0.02 0.1689 0.83 1.0
VCF0867 -0.42  0.06 0.1768 0.82 1.0

                       PA1  PA2
SS loadings           1.41 0.99
Proportion Var        0.13 0.09
Cumulative Var        0.13 0.22
Proportion Explained  0.59 0.41
Cumulative Proportion 0.59 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  55  and the objective function was  1.11 with Chi Square of  6573.48
The degrees of freedom for the model are 34  and the objective function was  0.22

The root mean square of the residuals (RMSR) is  0.06
The df corrected root mean square of the residuals is  0.07

The harmonic number of observations is  4745 with the empirical chi square  1586.06  with prob <  4.7e-312
The total number of observations was  5914  with Likelihood Chi Square =  1299.84  with prob <  2.8e-251

Tucker Lewis Index of factoring reliability =  0.686
RMSEA index =  0.079  and the 90 % confidence intervals are  0.076 0.083
BIC =  1004.55
Fit based upon off diagonal values = 0.88
Measures of factor score adequacy
                                                PA1  PA2
Correlation of (regression) scores with factors   0.84 0.75
Multiple R square of scores with factors          0.70 0.56
Minimum correlation of possible factor scores     0.41 0.12
```

Implementation: Factor Analysis

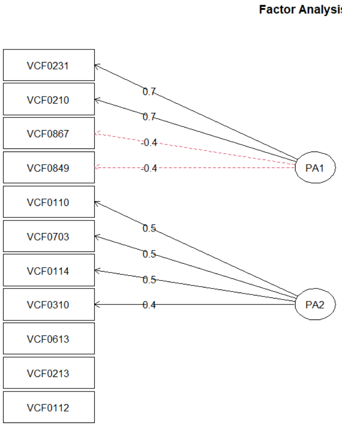
- This provides a lot of information, some useful, all informative
- If you wish to see which variables are loading on which factors:
  - `fa.object<-fa(...)`
  - `fa.object$loadings`

```
Loadings:
          PA1    PA2
VCF0110 -0.134  0.526
VCF0112
VCF0114 -0.220  0.451
VCF0210  0.693
VCF0213
VCF0231  0.696
VCF0310         0.437
VCF0613  0.185  0.282
VCF0703         0.480
VCF0849 -0.411
VCF0867 -0.416

                 PA1   PA2
SS loadings    1.409 0.993
Proportion Var 0.128 0.090
Cumulative Var 0.128 0.218
>
```

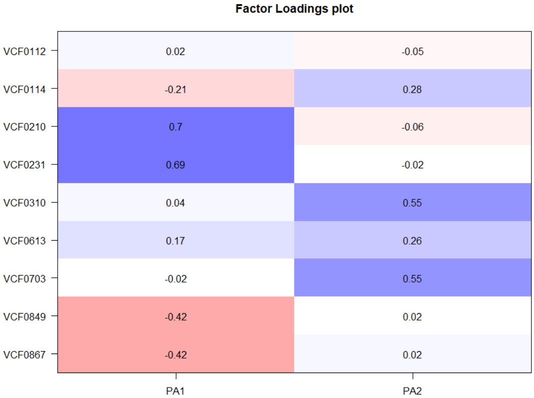
Implementation: Factor Analysis

- An alternative is to graphically present this
- `fa.diagram(fa.object)`



Implementation: Factor Analysis

- You can also graphically present the factors loadings on each factor from each variable
- `cor.plot(fa.object)`



Implementation: Factor Analysis

- If you wish to use these factors for other purposes, you can then combine these with other variables not used in the factor analysis into a data.frame
- `DF2 <- cbind(DF1[vars], fa.object$scores)`
  - `DF1:` Dataframe with other variables
  - `DF2:` Created dataframe with both factors and other variables
  - `fa.object:` The call to your factor analysis results

Implementation: Factor Analysis

- There is a lot of information overlooked in here
- Read the `psych` package documentation for more options, functions, and background information

Implementation: Factor Analysis

- There is a lot of information overlooked in here
- Read the `psych` package documentation for more options, functions, and background information
- Interpreting the effect of the factors on any variable is not straightforward - graphical presentations are often more informative than the information you'll get from a results table

Implementation: Factor Analysis

- There is a lot of information overlooked in here
- Read the `psych` package documentation for more options, functions, and background information
- Interpreting the effect of the factors on any variable is not straightforward - graphical presentations are often more informative than the information you'll get from a results table
- Factor analysis is not appropriate for binary outcomes
  - IRT (Item Response Theory) models are specifically designed for this function
  - With the information gained here, and some research, you should be able to use these if needed