

# Behavioral Data Science Week 5 - IRT and GRM

Jonathan Gragg

9/26/2021

## Summary

Based on the IRT model I would say that the distribution looks very similar for 1 IRT model across all questions. The GRM model shows a little disparity however and looks to be 2-3 different distribution shapes. That would lead me to confirm what I discovered last week where I suggested there needed to be 3 factors for these variables.

## Analysis

loading in relevant Libraries

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.0 --

## v ggplot2 3.3.5      v purrr  0.3.4
## v tibble  3.0.3      v dplyr  1.0.2
## v tidyr   1.1.1      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.5.0

## Warning: package 'ggplot2' was built under R version 4.0.5

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(ggplot2)
library(stringr)
library(psych)

## Warning: package 'psych' was built under R version 4.0.5

##
## Attaching package: 'psych'

## The following objects are masked from 'package:ggplot2':
##
## %+%, alpha
```

```
library(GGally)
```

```
## Warning: package 'GGally' was built under R version 4.0.5
```

```
## Registered S3 method overwritten by 'GGally':  
##   method from  
##   +.gg      ggplot2
```

```
library(mirt)
```

```
## Warning: package 'mirt' was built under R version 4.0.5
```

```
## Loading required package: stats4
```

```
## Loading required package: lattice
```

```
library(ltm)
```

```
## Warning: package 'ltm' was built under R version 4.0.5
```

```
## Loading required package: MASS
```

```
##  
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:dplyr':  
##  
##   select
```

```
## Loading required package: msm
```

```
## Warning: package 'msm' was built under R version 4.0.5
```

```
## Loading required package: polycor
```

```
## Warning: package 'polycor' was built under R version 4.0.5
```

```
##  
## Attaching package: 'polycor'
```

```
## The following object is masked from 'package:psych':  
##  
##   polyserial
```

```
##  
## Attaching package: 'ltm'
```

```
## The following object is masked from 'package:mirt':
##
##      Science
```

```
## The following object is masked from 'package:psych':
##
##      factor.scores
```

```
library(lavaan)
```

```
## Warning: package 'lavaan' was built under R version 4.0.5
```

```
## This is lavaan 0.6-9
## lavaan is FREE software! Please report any bugs.
```

```
##
## Attaching package: 'lavaan'
```

```
## The following object is masked from 'package:psych':
##
##      cor2cov
```

```
loading in the data
```

```
data <- read.csv("survey_results_public.csv")

#selecting the columns that we will be working with
df <- data %>% dplyr::select(ProblemSolving:ChangeWorld) %>% drop_na()

#converting all the of the column values into a numerical rating scale
df[df=='Strongly disagree'] <- 1
df[df=='Disagree'] <- 2
df[df=='Somewhat agree'] <- 3
df[df=='Agree'] <- 4
df[df=='Strongly agree'] <- 5

columns = 1:17

for (i in columns){
  df[,i] = as.integer(df[,i])
}

glimpse(df)
```

```
## Rows: 28,892
## Columns: 17
## $ ProblemSolving      <int> 5, 5, 5, 5, 5, 5, 4, 5, 5, 5, 5, 4, 4, 5, 5, 5,...
## $ BuildingThings      <int> 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 4, 4, 5, 5,...
## $ LearningNewTech     <int> 4, 5, 5, 5, 4, 5, 4, 5, 5, 4, 4, 5, 5, 4, 5, 5,...
## $ BoringDetails       <int> 2, 3, 2, 1, 2, 2, 1, 1, 2, 5, 4, 1, 3, 3, 3, 2,...
## $ JobSecurity         <int> 5, 4, 3, 4, 3, 3, 4, 4, 5, 3, 4, 4, 1, 4, 4, 4,...
```

```
## $ DiversityImportant <int> 4, 5, 4, 5, 3, 1, 5, 4, 4, 4, 3, 5, 2, 4, 1, 3,...
## $ AnnoyingUI <int> 4, 4, 4, 5, 5, 4, 5, 4, 3, 4, 4, 5, 5, 4, 5, 5,...
## $ FriendsDevelopers <int> 2, 3, 4, 3, 3, 2, 1, 2, 4, 2, 4, 5, 3, 2, 2, 2,...
## $ RightWrongWay <int> 3, 2, 3, 3, 3, 4, 2, 2, 2, 2, 3, 3, 2, 3, 4, 1,...
## $ UnderstandComputers <int> 2, 2, 1, 1, 2, 2, 5, 5, 2, 3, 1, 4, 3, 4, 2, 3,...
## $ SeriousWork <int> 5, 4, 5, 5, 4, 4, 5, 4, 4, 3, 4, 5, 3, 5, 4, 4,...
## $ InvestTimeTools <int> 5, 3, 4, 2, 4, 4, 4, 4, 4, 4, 4, 3, 3, 4, 3, 5,...
## $ WorkPayCare <int> 1, 2, 2, 1, 2, 3, 1, 1, 3, 1, 2, 2, 2, 2, 5, 1,...
## $ KinshipDevelopers <int> 4, 3, 5, 4, 3, 3, 5, 4, 5, 3, 4, 4, 3, 4, 2, 4,...
## $ ChallengeMyself <int> 4, 4, 5, 5, 4, 4, 4, 4, 5, 4, 4, 4, 5, 5, 4, 4,...
## $ CompetePeers <int> 2, 2, 3, 5, 3, 3, 1, 1, 2, 1, 4, 3, 2, 1, 2, 4,...
## $ ChangeWorld <int> 4, 4, 4, 4, 3, 4, 5, 5, 5, 3, 3, 3, 5, 3, 3, 4,...
```

*#Try 1-factor*

```
fa1 <- fa(df, nfactors = 1, rotate = "none")
print(fa1$loadings, cutoff = 0.001, digits = 3)
```

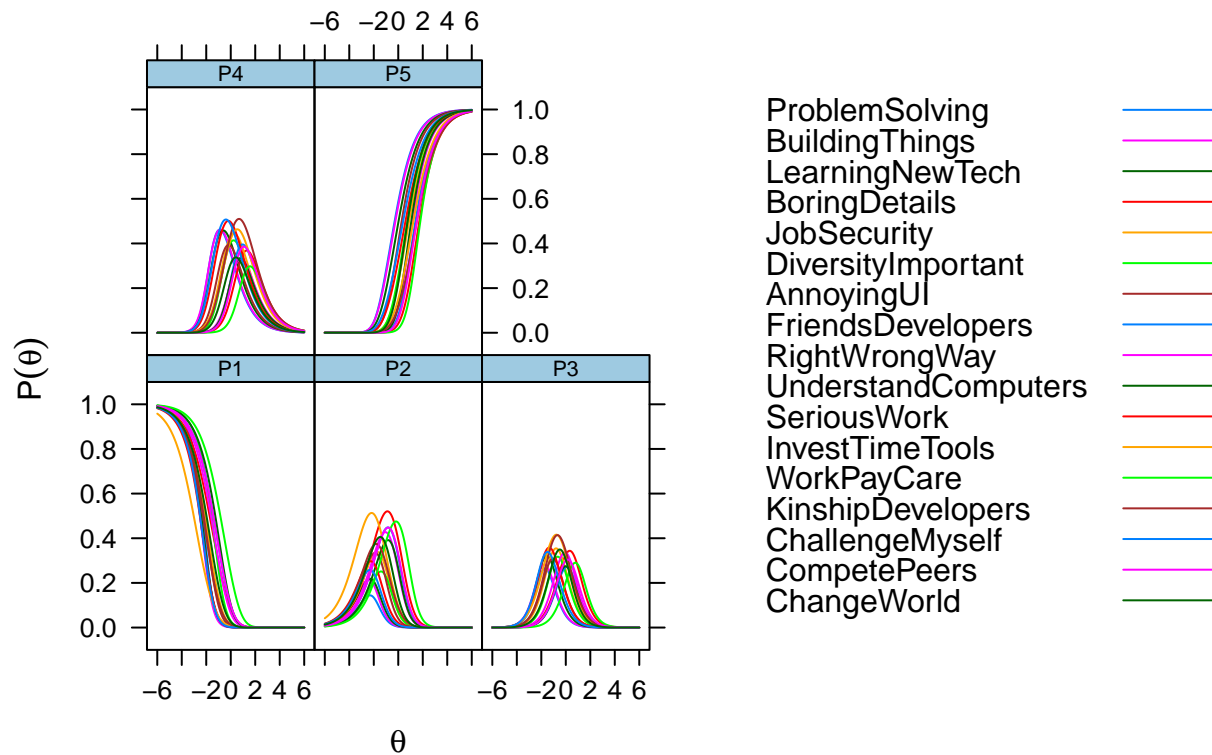
```
##
## Loadings:
##
## ProblemSolving      MR1
## BuildingThings      0.577
## LearningNewTech     0.525
## BoringDetails       -0.099
## JobSecurity         0.121
## DiversityImportant  0.250
## AnnoyingUI          0.206
## FriendsDevelopers   0.120
## RightWrongWay       0.117
## UnderstandComputers -0.055
## SeriousWork         0.458
## InvestTimeTools     0.439
## WorkPayCare         -0.164
## KinshipDevelopers   0.372
## ChallengeMyself     0.631
## CompetePeers        0.166
## ChangeWorld         0.393
##
##
## SS loadings      MR1
## Proportion Var  2.233
## Proportion Var  0.131
```

```
fit1PL <- mirt(df, model = 1, itemtype = "Rasch")
```

```
## Iteration: 1, Log-Lik: -642998.289, Max-Change: 1.00633Iteration: 2, Log-Lik: -624014.072, Max-Change:
```

```
plot(fit1PL, type = 'trace', facet_items = FALSE)
```

## Item trace lines



```
confirm_mod1 <- "Endurance =~ ProblemSolving + BuildingThings + LearningNewTech + BoringDetails + JobSecurity + AnnoyingUI + FriendsDevelopers + RightWrongWay + UnderstandComputers + SeriousWork + InvestTimeTools + WorkPayCare + KinshipDevelopers + ChallengeMyself + CompetePeers + ChangeWorld"
fit1 <- lavaan::cfa(confirm_mod1, data = df, std.lv = TRUE)
summary(fit1, fit.measures = T, standardized = T)
```

```
## lavaan 0.6-9 ended normally after 22 iterations
##
##      Estimator                      ML
##      Optimization method          NLMINB
##      Number of model parameters          34
##
##      Number of observations          28892
##
## Model Test User Model:
##
##      Test statistic          10806.678
##      Degrees of freedom          119
##      P-value (Chi-square)          0.000
##
## Model Test Baseline Model:
##
##      Test statistic          45456.402
##      Degrees of freedom          136
##      P-value          0.000
##
```

```

## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                0.764
##   Tucker-Lewis Index (TLI)                  0.730
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -639382.143
##   Loglikelihood unrestricted model (H1)      -633978.804
##
##   Akaike (AIC)                             1278832.287
##   Bayesian (BIC)                           1279113.512
##   Sample-size adjusted Bayesian (BIC)       1279005.460
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                     0.056
##   90 Percent confidence interval - lower     0.055
##   90 Percent confidence interval - upper     0.057
##   P-value RMSEA <= 0.05                    0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                     0.047
##
## Parameter Estimates:
##
##   Standard errors                          Standard
##   Information                             Expected
##   Information saturated (h1) model         Structured
##
## Latent Variables:
##
##           Estimate  Std.Err  z-value  P(>|z|)  Std.lv  Std.all
##   Endurance =~
##   ProblemSolving    0.388    0.004   96.238   0.000    0.388    0.596
##   BuildingThings    0.353    0.004   84.406   0.000    0.353    0.532
##   LearningNewTch    0.412    0.005   90.722   0.000    0.412    0.566
##   BoringDetails     -0.104    0.007  -15.362   0.000   -0.104   -0.104
##   JobSecurity        0.109    0.006   17.010   0.000    0.109    0.115
##   DivrstyImprtnt    0.254    0.007   36.542   0.000    0.254    0.244
##   AnnoyingUI        0.182    0.006   29.721   0.000    0.182    0.199
##   FriendsDevlprs    0.126    0.008   16.228   0.000    0.126    0.109
##   RightWrongWay     0.118    0.008   15.520   0.000    0.118    0.105
##   UndrstndCmptrs    -0.076    0.008   -9.217   0.000   -0.076   -0.062
##   SeriousWork       0.352    0.005   69.746   0.000    0.352    0.448
##   InvestTimeTols    0.387    0.006   66.735   0.000    0.387    0.431
##   WorkPayCare       -0.164    0.006  -25.325   0.000   -0.164   -0.170
##   KinshipDevlprs    0.322    0.006   55.450   0.000    0.322    0.363
##   ChallengeMyslf    0.473    0.005  103.673   0.000    0.473    0.635
##   CompetePeers      0.177    0.007   23.686   0.000    0.177    0.159
##   ChangeWorld       0.421    0.007   59.143   0.000    0.421    0.385
##
## Variances:
##
##           Estimate  Std.Err  z-value  P(>|z|)  Std.lv  Std.all

```

```
##      .ProblemSolving      0.274      0.003      98.597      0.000      0.274      0.645
##      .BuildingThings      0.317      0.003      104.785      0.000      0.317      0.717
##      .LearningNewTch      0.360      0.004      101.706      0.000      0.360      0.679
##      .BoringDetails      0.996      0.008      119.772      0.000      0.996      0.989
##      .JobSecurity      0.896      0.007      119.676      0.000      0.896      0.987
##      .DivrstyImprtnt      1.020      0.009      117.750      0.000      1.020      0.941
##      .AnnoyingUI      0.800      0.007      118.595      0.000      0.800      0.960
##      .FriendsDevlprs      1.306      0.011      119.723      0.000      1.306      0.988
##      .RightWrongWay      1.265      0.011      119.763      0.000      1.265      0.989
##      .UndrstndCmptrs      1.492      0.012      120.041      0.000      1.492      0.996
##      .SeriousWork      0.492      0.004      110.392      0.000      0.492      0.799
##      .InvestTimeTols      0.657      0.006      111.327      0.000      0.657      0.815
##      .WorkPayCare      0.908      0.008      119.039      0.000      0.908      0.971
##      .KinshipDevlprs      0.683      0.006      114.305      0.000      0.683      0.868
##      .ChallengeMyslf      0.331      0.004      93.661      0.000      0.331      0.597
##      .CompetePeers      1.199      0.010      119.185      0.000      1.199      0.975
##      .ChangeWorld      1.015      0.009      113.416      0.000      1.015      0.852
##      Endurance      1.000      0.000      0.000      0.000      1.000      1.000
```

```
facScores = fscores(fit1PL)
```

```
head(facScores)
```

```
##      F1
## [1,] 0.008897005
## [2,] -0.030545867
## [3,] 0.128181209
## [4,] 0.088247360
## [5,] -0.109027299
## [6,] -0.109027299
```

```
summary(facScores)
```

```
##      F1
## Min.   :-1.6852483
## 1st Qu.: -0.1480877
## Median : 0.0088970
## Mean    :-0.0000013
## 3rd Qu.: 0.1281812
## Max.    : 1.0398703
```

```
faScores = psych::fa(df, 1, rotate = "promax")$scores
```

```
summary(faScores)
```

```
##      MR1
## Min.   :-6.44576
## 1st Qu.: -0.53581
## Median : 0.06505
## Mean    : 0.00000
## 3rd Qu.: 0.63438
## Max.    : 1.92922
```

```
cor(faScores, facScores)
```

```
##           F1  
## MR1 0.766689
```

```
grmMod = mirt(df, model = 1, itemtype = "graded")
```

```
## Iteration: 1, Log-Lik: -622588.525, Max-Change: 1.54736Iteration: 2, Log-Lik: -612307.672, Max-Change:
```

```
coef(grmMod, simplify = TRUE, IRTpars = TRUE)
```

```
## $items  
##           a           b1           b2           b3           b4  
## ProblemSolving      1.719    -4.474    -3.664    -2.140    -0.297  
## BuildingThings      1.417    -5.247    -3.982    -2.310    -0.266  
## LearningNewTech     1.496    -4.680    -3.495    -1.856     0.019  
## BoringDetails      -0.224    10.464     0.425    -5.962   -13.311  
## JobSecurity         0.248   -16.198    -9.168    -2.606     4.811  
## DiversityImportant  0.519    -6.461    -4.166    -1.286     2.100  
## AnnoyingUI         0.440   -10.620    -6.557    -3.017     0.897  
## FriendsDevelopers   0.199   -10.830    -1.747     3.597    11.622  
## RightWrongWay       0.205   -12.402    -3.614     2.900     9.429  
## UnderstandComputers -0.110    18.620     4.629    -5.045   -16.725  
## SeriousWork         1.060    -5.653    -3.756    -1.663     0.808  
## InvestTimeTools     0.974    -5.548    -2.469    -0.354     1.985  
## WorkPayCare        -0.369     3.062    -2.407    -6.194   -10.099  
## KinshipDevelopers   0.792    -5.486    -2.883    -0.338     2.752  
## ChallengeMyself     1.813    -4.115    -2.929    -1.459     0.390  
## CompetePeers        0.292    -7.346    -1.045     3.162     8.745  
## ChangeWorld         0.836    -4.455    -2.090    -0.190     1.605  
##  
## $means  
## F1  
## 0  
##  
## $cov  
## F1  
## F1 1
```

```
plot(grmMod, type = 'trace',  
      facet_items = TRUE, main = "GRM")
```



# GRM

