The Personality Structures of the 50 States

Anisha Babu, Diana DeWald, Ian Shryock, Dillon Welindt, Futing Zou

Contents

List of Tables

```
SAPA <- SAPA %>%
  # filter in US
  filter(country == "USA") %>%
  # filter 50 states
  filter(!state %in% c('District of Columbia',
                       'Guam',
                       'Palau',
                       'Puerto Rico',
                       'Virgin Islands',
                       'Northern Mariana Islands',
                       'American Samoa',
                       'Marshall Islands',
                       NA))

# filter out rows with all NA
SAPA <- SAPA[rowSums(is.na(SAPA)) < 99, ]

# in list form select only used Q's
usedQ <- colnames(SAPA[8:106])

IPIPkeys <- map(IPIPkeysList, function(x) {
    x[match(usedQ, x)]
    na.omit(x)
})

# select only IPIP 100
IPIPkeys <- IPIPkeys[1:4]
```

```
## $gender
##
## Female   Male
##  56901  24465
##
## $age
##
##   14   15   16   17   18   19   20   21   22   23   24   25   26   27   28   29
##  767 1113 3129 6462 7707 6709 5896 4885 3621 2999 2569 2300 2193 2085 1887 1649
##   30   31   32   33   34   35   36   37   38   39   40   41   42   43   44   45
## 1590 1453 1345 1230 1212 1240 1152 1159 1069  958 1010  840  866  847  815  791
##   46   47   48   49   50   51   52   53   54   55   56   57   58   59   60   61
##  688  691  681  626  665  542  537  500  433  389  309  301  246  217  212  126
##   62   63   64   65   66   67   68   69   70   71   72   73   74   75   76   77
##  132   88   68   82   50   47   29   27   27   22   21    6   12    7    5    4
##   78   79   80   81   82   83   84   85   86   87   88   89   90
##    5    3    4    1    1    2    1    3    2    2    1    2    1
##
## $state
##
##        Alabama         Alaska        Arizona       Arkansas     California
##            643            555            866            577           9709
##       Colorado    Connecticut       Delaware        Florida        Georgia
##           1097            986            592           2936           2414
##         Hawaii          Idaho       Illinois        Indiana           Iowa
```

```
##             292             340            5520            1707             982
##          Kansas        Kentucky       Louisiana           Maine        Maryland
##             808             820            2030             356            1772
##   Massachusetts        Michigan       Minnesota     Mississippi        Missouri
##            1935            2549            2104             604            1611
##         Montana        Nebraska          Nevada  New Hampshire      New Jersey
##             243             580             274             389            2495
##      New Mexico        New York North Carolina    North Dakota            Ohio
##            1199            4942            1454             190            3600
##        Oklahoma          Oregon    Pennsylvania    Rhode Island South Carolina
##             771            1203            4758             422            1010
##    South Dakota       Tennessee           Texas            Utah         Vermont
##             172            1133            4662             487             161
##        Virginia      Washington   West Virginia       Wisconsin         Wyoming
##            2787            1742             384            2377             126
##
## $race
##
## African American          Chinese Indian/Pakistani         Japanese
##             6108             1129              469              257
##           Korean           Latino          Mexican  Native American
##              500             2079             2166              728
##            Other      Other Asian Pacific Islander         Philipino
##             3067              566              305              615
##     Puerto Rican  White/Caucasian
##              512            62859
##
## $education
##
##                College graduate     Currently attending college
##                           12381                           32469
## Graduate or professional degree           High school graduate
##                           10338                            6145
##              Less than 12 years   Some college did not graduate
##                           11759                            8274
```

```
# score items
scores <- psych::scoreItems(keys = IPIPkeys, items = SAPA, min = 1, max = 6,
                            totals = FALSE, impute = 'none')$scores
```

```
# demographic table
demog_tab <- summary(tableby(~ age + gender + race + education,
                             data = SAPA, test = FALSE),
                     title = "Full Sample Demographics")

demog_tab
```

Table 1: (#tab:#3 descriptives statistics)Full Sample Demographics

	Overall (N=81366)
age	

	Overall (N=81366)
Mean (SD)	27.177 (11.343)
Range	14.000 - 90.000
gender	
Female	56901 (69.9%)
Male	24465 (30.1%)
race	
N-Miss	6
African American	6108 (7.5%)
Chinese	1129 (1.4%)
Indian/Pakistani	469 (0.6%)
Japanese	257 (0.3%)
Korean	500 (0.6%)
Latino	2079 (2.6%)
Mexican	2166 (2.7%)
Native American	728 (0.9%)
Other	3067 (3.8%)
Other Asian	566 (0.7%)
Pacific Islander	305 (0.4%)
Philipino	615 (0.8%)
Puerto Rican	512 (0.6%)
White/Caucasian	62859 (77.3%)
education	
College graduate	12381 (15.2%)
Currently attending college	32469 (39.9%)
Graduate or professional degree	10338 (12.7%)
High school graduate	6145 (7.6%)
Less than 12 years	11759 (14.5%)
Some college did not graduate	8274 (10.2%)

```
# PEER REVIEW: To get a more elegant table, perhaps you can create your own function that gives you the

sample <- function(df,x) {
  df %>%
  group_by({{x}}) %>%
  summarize(group_sample = n()) %>%
  ungroup() %>%
  rename(group = {{x}}) %>%
  mutate(total_sample = sum(group_sample),
         percent_sample = group_sample/total_sample) %>%
  select(group,group_sample, percent_sample)
}

sample(SAPA,race)
```

1 A tibble: 15 x 3

group group_sample percent_sample 1 African American 6108 0.0751
2 Chinese 1129 0.0139
3 Indian/Pakistani 469 0.00576
4 Japanese 257 0.00316

5 Korean 500 0.00615
6 Latino 2079 0.0256
7 Mexican 2166 0.0266
8 Native American 728 0.00895
9 Other 3067 0.0377
10 Other Asian 566 0.00696
11 Pacific Islander 305 0.00375
12 Philipino 615 0.00756
13 Puerto Rican 512 0.00629
14 White/Caucasian 62859 0.773
15 6 0.0000737

```
sample(SAPA,education)
```

2 A tibble: 6 x 3

group group_sample percent_sample 1 College graduate 12381 0.152 2 Currently attending college 32469 0.399 3 Graduate or professional degree 10338 0.127 4 High school graduate 6145 0.0755 5 Less than 12 years 11759 0.145 6 Some college did not graduate 8274 0.102

```
# to add for final: demographic table grouped by state

# to add for final:  improve correlation matrix format/names (below), include other personality traits

#RG This table looks great on the pdf, would you consider collapsing the race/ethnicity variables for c

res <- cor(scores, use = "complete.obs")
round(res, 2)
```

IPIP100agreeableness IPIP100conscientiousness

IPIP100agreeableness 1.00 0.21 IPIP100conscientiousness 0.21 1.00 IPIP100extraversion 0.38 0.13 IPIP100intellect 0.16 0.08 IPIP100extraversion IPIP100intellect IPIP100agreeableness 0.38 0.16 IPIP100conscientiousness 0.13 0.08 IPIP100extraversion 1.00 0.22 IPIP100intellect 0.22 1.00

```
apa.cor.table(scores, filename="Corr_table.doc", show.conf.interval=FALSE)
```

The ability to suppress reporting of reporting confidence intervals has been deprecated in this version. The function argument show.conf.interval will be removed in a later version.

Means, standard deviations, and correlations with confidence intervals

Variable M SD 1 2 3
1. IPIP100agreeableness 4.67 0.77

2. IPIP100conscientiousness 4.14 0.92 .21**
   [.21, .22]
3. IPIP100extraversion 3.92 1.02 .38** .13**
   [.37, .38] [.13, .14]
4. IPIP100intellect 4.59 0.73 .16** .08** .22**
   [.15, .16] [.07, .08] [.21, .23]

Note. M and SD are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * indicates p < .05. ** indicates p < .01.

```
apa.cor.table(scores, filename="Corr_table.doc", show.conf.interval=F)
```

The ability to suppress reporting of reporting confidence intervals has been deprecated in this version. The function argument show.conf.interval will be removed in a later version.

Means, standard deviations, and correlations with confidence intervals

Variable M SD 1 2 3
1. IPIP100agreeableness 4.67 0.77

2. IPIP100conscientiousness 4.14 0.92 .21**
   [.21, .22]
3. IPIP100extraversion 3.92 1.02 .38** .13**
   [.37, .38] [.13, .14]
4. IPIP100intellect 4.59 0.73 .16** .08** .22**
   [.15, .16] [.07, .08] [.21, .23]

Note. M and SD are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * indicates p < .05. ** indicates p < .01.

```
# PEER REVIEW: To get a more elegant table, perhaps you can create your own function that gives you the

sample <- function(df,x) {
  df %>%
    group_by({{x}}) %>%
    summarize(group_sample = n()) %>%
    ungroup() %>%
    rename(group = {{x}}) %>%
    mutate(total_sample = sum(group_sample),
           percent_sample = group_sample/total_sample) %>%
    select(group,group_sample, percent_sample)
}

sample(SAPA,race)
```

3 A tibble: 15 x 3

group group_sample percent_sample 1 African American 6108 0.0751
2 Chinese 1129 0.0139
3 Indian/Pakistani 469 0.00576
4 Japanese 257 0.00316
5 Korean 500 0.00615
6 Latino 2079 0.0256
7 Mexican 2166 0.0266
8 Native American 728 0.00895
9 Other 3067 0.0377

10 Other Asian 566 0.00696
11 Pacific Islander 305 0.00375
12 Philipino 615 0.00756
13 Puerto Rican 512 0.00629
14 White/Caucasian 62859 0.773
15 6 0.0000737

```
sample(SAPA,education)
```

4 A tibble: 6 x 3

group group_sample percent_sample 1 College graduate 12381 0.152 2 Currently attending college 32469 0.399 3 Graduate or professional degree 10338 0.127 4 High school graduate 6145 0.0755 5 Less than 12 years 11759 0.145 6 Some college did not graduate 8274 0.102

```
# average scores on included survey questions
average_surveyscores <- SAPA %>%
  summarize_at(vars(q_76:q_1989), mean, na.rm = TRUE)

# average scores on survey questions by state (for final: can group by other variables as well)
average_statescores <- SAPA %>%
  group_by(state) %>%
  summarize_at(vars(q_76:q_1989), mean, na.rm = TRUE)
```

Table 2: Agreeableness Descriptives

State	Mean	SD	N
Alabama	4.610537	4.610537	4.610537
Alaska	4.629332	4.629332	4.629332
Arizona	4.596157	4.596157	4.596157
Arkansas	4.689144	4.689144	4.689144
California	4.663943	4.663943	4.663943
Colorado	4.631178	4.631178	4.631178
Connecticut	4.634948	4.634948	4.634948
Delaware	4.575882	4.575882	4.575882
Florida	4.659417	4.659417	4.659417
Georgia	4.688953	4.688953	4.688953
Hawaii	4.637072	4.637072	4.637072
Idaho	4.672925	4.672925	4.672925
Illinois	4.729347	4.729347	4.729347
Indiana	4.689507	4.689507	4.689507
Iowa	4.667843	4.667843	4.667843
Kansas	4.678785	4.678785	4.678785
Kentucky	4.693537	4.693537	4.693537
Louisiana	4.703087	4.703087	4.703087
Maine	4.667634	4.667634	4.667634
Maryland	4.707332	4.707332	4.707332
Massachusetts	4.654854	4.654854	4.654854
Michigan	4.653881	4.653881	4.653881
Minnesota	4.667296	4.667296	4.667296
Mississippi	4.722192	4.722192	4.722192
Missouri	4.665174	4.665174	4.665174
Montana	4.765261	4.765261	4.765261
Nebraska	4.641667	4.641667	4.641667
Nevada	4.633536	4.633536	4.633536

Table 2: Agreeableness Descriptives *(continued)*

State	Mean	SD	N
New Hampshire	4.629999	4.629999	4.629999
New Jersey	4.669726	4.669726	4.669726
New Mexico	4.727511	4.727511	4.727511
New York	4.663282	4.663282	4.663282
North Carolina	4.647067	4.647067	4.647067
North Dakota	4.615877	4.615877	4.615877
Ohio	4.700273	4.700273	4.700273
Oklahoma	4.652295	4.652295	4.652295
Oregon	4.633837	4.633837	4.633837
Pennsylvania	4.656941	4.656941	4.656941
Rhode Island	4.732280	4.732280	4.732280
South Carolina	4.727500	4.727500	4.727500
South Dakota	4.657946	4.657946	4.657946
Tennessee	4.663913	4.663913	4.663913
Texas	4.652647	4.652647	4.652647
Utah	4.665389	4.665389	4.665389
Vermont	4.776898	4.776898	4.776898
Virginia	4.722891	4.722891	4.722891
Washington	4.653790	4.653790	4.653790
West Virginia	4.618171	4.618171	4.618171
Wisconsin	4.641089	4.641089	4.641089
Wyoming	4.662522	4.662522	4.662522

Table 3: Conscientiousness Descriptives

State	Mean	SD	N
Alabama	4.117571	0.9667964	643
Alaska	4.026173	0.8683283	555
Arizona	4.098011	0.8929967	866
Arkansas	4.147743	0.9248194	577
California	4.095768	0.9086644	9709
Colorado	4.124273	0.9210228	1097
Connecticut	4.120340	0.9307941	986
Delaware	3.997203	0.8632940	592
Florida	4.179669	0.9343133	2936
Georgia	4.129606	0.9180854	2414
Hawaii	4.170548	0.8806935	292
Idaho	4.150588	0.8735417	340
Illinois	4.162300	0.9112091	5520
Indiana	4.218357	0.9211999	1707
Iowa	4.102082	0.9230007	982
Kansas	4.108794	0.9105627	808
Kentucky	4.122991	0.9537198	820
Louisiana	4.209949	0.8955529	2030
Maine	4.205641	0.9188266	356
Maryland	4.128232	0.9095782	1772
Massachusetts	4.118530	0.9365527	1935
Michigan	4.162636	0.9299835	2549
Minnesota	4.096184	0.8982866	2104
Mississippi	4.198448	0.9059218	604
Missouri	4.141600	0.9351437	1611
Montana	4.160722	0.9577293	243
Nebraska	4.156379	0.8777856	580
Nevada	4.082401	0.9495511	274

Table 3: Conscientiousness Descriptives *(continued)*

State	Mean	SD	N
New Hampshire	4.148700	0.8946446	389
New Jersey	4.148890	0.9280449	2495
New Mexico	4.262791	0.8985910	1199
New York	4.148340	0.9339365	4942
North Carolina	4.171695	0.9481568	1454
North Dakota	4.287222	0.8787829	190
Ohio	4.188897	0.9247682	3600
Oklahoma	4.140532	0.9454096	771
Oregon	4.070511	0.9039862	1203
Pennsylvania	4.117532	0.9191805	4758
Rhode Island	4.099572	0.9192980	422
South Carolina	4.136447	0.8851777	1010
South Dakota	4.266537	0.8999202	172
Tennessee	4.208465	0.9470666	1133
Texas	4.131076	0.9285345	4662
Utah	4.102647	0.8660975	487
Vermont	4.090649	0.9496127	161
Virginia	4.141444	0.9057514	2787
Washington	4.127309	0.9222153	1742
West Virginia	4.163824	0.9387757	384
Wisconsin	4.112204	0.9268177	2377
Wyoming	4.188095	0.9480086	126

Table 4: Extraversion Descriptives

State	Mean	SD	N
Alabama	3.744427	1.0737088	643
Alaska	3.802078	1.0298194	555
Arizona	3.826485	1.0391321	866
Arkansas	3.826526	1.1045412	577
California	3.916162	1.0182084	9709
Colorado	3.811444	1.0267798	1097
Connecticut	3.924073	1.0262157	986
Delaware	4.029242	0.9524545	592
Florida	3.888615	1.0586144	2936
Georgia	3.998944	1.0429965	2414
Hawaii	3.835455	1.0309916	292
Idaho	3.752173	1.0263106	340
Illinois	4.012087	0.9837590	5520
Indiana	3.894970	1.0577104	1707
Iowa	3.907930	1.0020006	982
Kansas	3.939745	1.0499232	808
Kentucky	3.938068	1.0360986	820
Louisiana	4.010502	0.9748960	2030
Maine	3.846177	1.0208230	356
Maryland	3.919828	1.0085951	1772
Massachusetts	3.894588	1.0272060	1935
Michigan	3.896950	1.0384003	2549
Minnesota	3.941308	1.0015734	2104
Mississippi	3.920760	1.0388218	604
Missouri	3.912232	0.9806955	1611
Montana	3.847828	1.0609354	243
Nebraska	3.918702	0.9855573	580
Nevada	3.816920	1.0112589	274

Table 4: Extraversion Descriptives *(continued)*

State	Mean	SD	N
New Hampshire	3.862380	0.9731819	389
New Jersey	3.996112	0.9864916	2495
New Mexico	3.898791	1.0508917	1199
New York	3.933406	1.0287508	4942
North Carolina	3.802822	1.0565457	1454
North Dakota	3.803845	1.0412612	190
Ohio	3.920517	1.0329506	3600
Oklahoma	3.804182	1.0792374	771
Oregon	3.886437	1.0061902	1203
Pennsylvania	3.958699	1.0067223	4758
Rhode Island	4.057464	0.9397830	422
South Carolina	4.043584	0.9876283	1010
South Dakota	3.979409	1.0081303	172
Tennessee	3.847732	1.0370815	1133
Texas	3.888641	1.0537326	4662
Utah	3.884845	1.0514536	487
Vermont	3.917118	0.9878257	161
Virginia	3.943700	1.0111944	2787
Washington	3.809801	1.0334914	1742
West Virginia	3.779065	1.0798027	384
Wisconsin	3.920540	1.0092239	2377
Wyoming	3.821847	1.0187586	126

Table 5: Intellect Descriptives

State	Mean	SD	N
Alabama	4.637855	0.7449558	643
Alaska	4.664919	0.7668085	555
Arizona	4.665397	0.7510410	866
Arkansas	4.601727	0.7666522	577
California	4.607159	0.7273981	9709
Colorado	4.670469	0.7275069	1097
Connecticut	4.655686	0.7412064	986
Delaware	4.386655	0.7205952	592
Florida	4.657286	0.7032552	2936
Georgia	4.597719	0.7234032	2414
Hawaii	4.533509	0.7517126	292
Idaho	4.676192	0.7004048	340
Illinois	4.568339	0.7243646	5520
Indiana	4.578118	0.7459488	1707
Iowa	4.533963	0.7319487	982
Kansas	4.600704	0.7605258	808
Kentucky	4.601366	0.7377464	820
Louisiana	4.421264	0.7414032	2030
Maine	4.643924	0.7361798	356
Maryland	4.577738	0.7187010	1772
Massachusetts	4.574818	0.7139949	1935
Michigan	4.656310	0.7285229	2549
Minnesota	4.531498	0.7231444	2104
Mississippi	4.559547	0.7368477	604
Missouri	4.588972	0.7330890	1611
Montana	4.708861	0.7378293	243
Nebraska	4.495270	0.7481413	580
Nevada	4.651490	0.7185546	274

Table 5: Intellect Descriptives *(continued)*

State	Mean	SD	N
New Hampshire	4.630421	0.7565377	389
New Jersey	4.612689	0.7390479	2495
New Mexico	4.579827	0.6900077	1199
New York	4.649041	0.7256438	4942
North Carolina	4.569312	0.7502207	1454
North Dakota	4.570237	0.7177782	190
Ohio	4.556529	0.7364082	3600
Oklahoma	4.605553	0.7834762	771
Oregon	4.619854	0.7436931	1203
Pennsylvania	4.512731	0.7411906	4758
Rhode Island	4.622255	0.6808287	422
South Carolina	4.485802	0.7222830	1010
South Dakota	4.653013	0.6828377	172
Tennessee	4.590234	0.7736854	1133
Texas	4.613907	0.7484548	4662
Utah	4.607547	0.7199137	487
Vermont	4.735498	0.7363306	161
Virginia	4.547663	0.7079233	2787
Washington	4.684024	0.6893191	1742
West Virginia	4.582448	0.7334187	384
Wisconsin	4.504417	0.7397412	2377
Wyoming	4.631488	0.6969593	126

```r
load(here("data", "statefactor.rda"))
stateFactordf
```

```
##   Alabama Alaska Arizona Arkansas California Colorado Connecticut Delaware
## 1       5      6       5        5          5        5           5        6
##   Florida Georgia Hawaii Idaho Illinois Indiana Iowa Kansas Kentucky Louisiana
## 1       5       5      6     6        5       5    5      5        5         5
##   Maine Maryland Massachusetts Michigan Minnesota Mississippi Missouri Montana
## 1     5        5             5        5         5           6        5       7
##   Nebraska Nevada New.Hampshire New.Jersey New.Mexico New.York North.Carolina
## 1        5      7             6          5          5        5              5
##   North.Dakota Ohio Oklahoma Oregon Pennsylvania Rhode.Island South.Carolina
## 1            7    5        5      5            5            7              5
##   South.Dakota Tennessee Texas Utah Vermont Virginia Washington West.Virginia
## 1            7         5     5    5       8        5          5             6
##   Wisconsin Wyoming
## 1         5       8
```

```
## Downloading ingested version of data with readr::read_tsv. To download the original version and remov
```

```
## Rows: 128464 Columns: 390
## -- Column specification --------------------------------------------------------
## Delimiter: "\t"
## chr  (12): gender, relstatus, marstatus, exer, smoke, country, state, ethnic...
## dbl (378): RID, age, height, BMI, weight, q_6, q_22, q_23, q_39, q_40, q_44,...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
## Joining, by = "state"
```

5 Introduction

5.1 Big Five

One of the most widely replicated findings within the field of personality psychology is the Big Five structure of personality. With roots in the 1800's, personality psychology sought to determine the best way to represent the large number of personality traits in a concise structure. This research initially involved researchers providing participants with large numbers of trait descriptive adjectives and asking them to rate the extent to which those adjectives characterize themselves or someone they knew. Dimension reduction analyses were then used to create a simpler structure from those responses.

Multiple research groups began converging on the five factor structure as early as the 1960's, with an increasing consensus by the late 1980's. Most of the recent work on the big five has been conducted through a combination of confirmatory factor analysis and theory driven selection of survey items based on previous findings about the structure.

5.2 Geographical Personality

In recent years, there has been increasing focus on regional variation of personality traits within the United States. Work has examined the extent to which regions of the US differ on the Big Five domains and can be said to have distinct and characteristic combinations of trait levels. For example, Rentfrow and colleagues (2013) show that the south and midwest are best characterized as friendly and conventional, whereas the west is relaxed and creative, and the northeast is temperamental and uninhibited.

A limitation of this work is that it examines the extent to which the five factor structure captures each region and what differences in the levels of each factor are due to regional variation. This research utilizes confirmatory factor analyses that assume that the five factor structure is the ideal level of dimensionality to characterize all regions.

5.3 Cross-Cultural Studies

Much of the cross-cultural work on personality structure has found some support for the notion that the five factor structure has applicability in a number of cultures. However, these studies typically are conducted from an etic perspective that translate the items used in western samples.

However, when studies are conducted from an emic perspective – that is, using trait descriptive adjectives from the language of the culture, rather than translations of items used in the big five framework – different structures emerge. A varying number of factors have been found to best fit different cultures, ranging from one to seven in many cases.

5.4 Geographical Factor Structure within US

Within the US, the regional variation in factor structures has not been an extensively studied topic. Because most research operates within a framework that utilizes confirmatory factor analysis, there is little information on the extent to which regions differ in their factor structure.

In the current study, we use exploratory factor analyses to provide estimates of the optimal factor structures for each of the fifty states.

6 Brief Methods

6.1 Measures

The International Personality Item Pool is an open-source repository of personality trait items that have been researched extensively in the big five tradition. The current study uses ninety nine of one hundred items from the IPIP-100. Participants rated themselves on a number of personality traits from 1- not at all like me to 6- very much like me.

6.2 Data Collection

Data were obtained from the Harvard Dataverse. Data were initially collected using the Synthetic Aperture for Personality Assessment (Revelle et al., 2016; Condon and Revelle, 2014; Wilt et al., 2011) which utilizes a massively missing completely at random design, wherein each participant only provides responses to a fraction of items.

7 Analyses

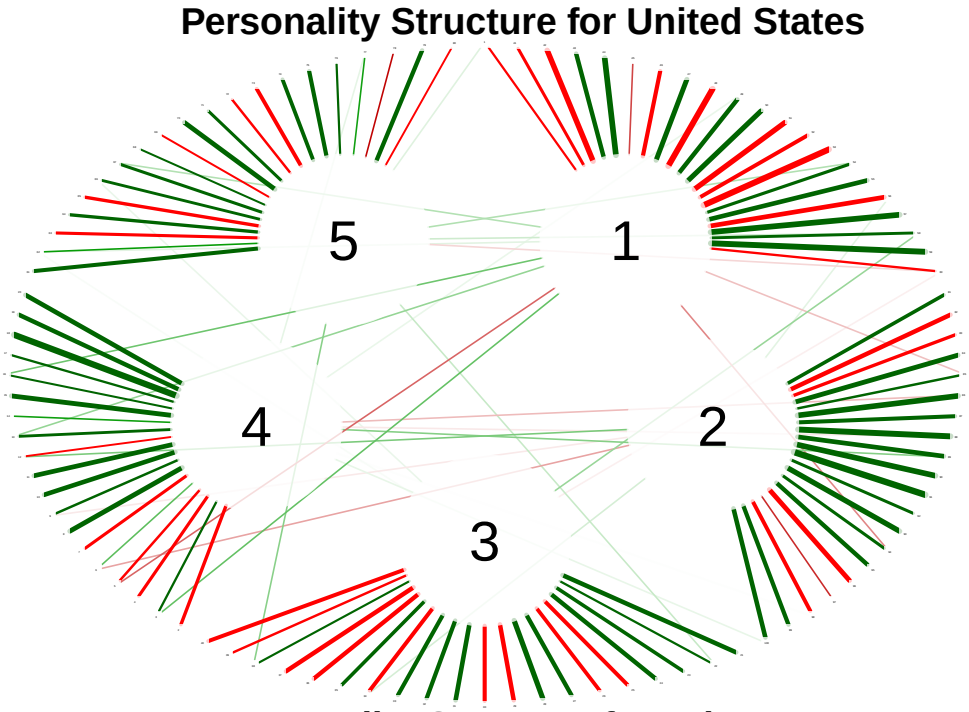
First, we provide descriptive norms for the entire US sample, and then by state.

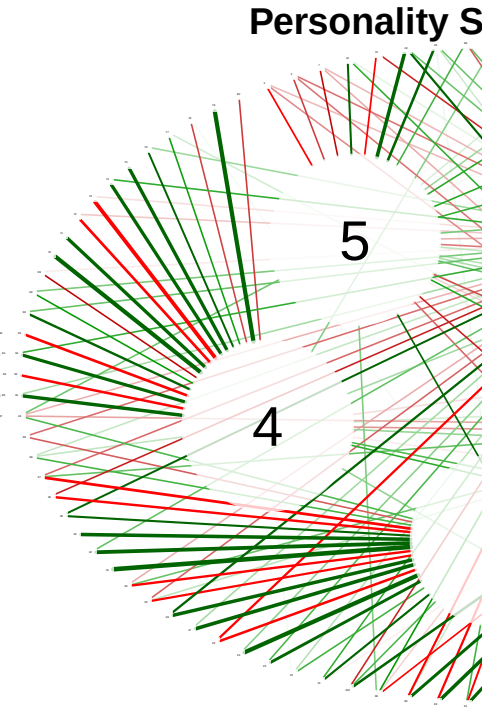
Next, we use parallel analysis to determine the optimal number of factors in the whole sample. Our hypothesis is that five factors will provide an optimal fit.

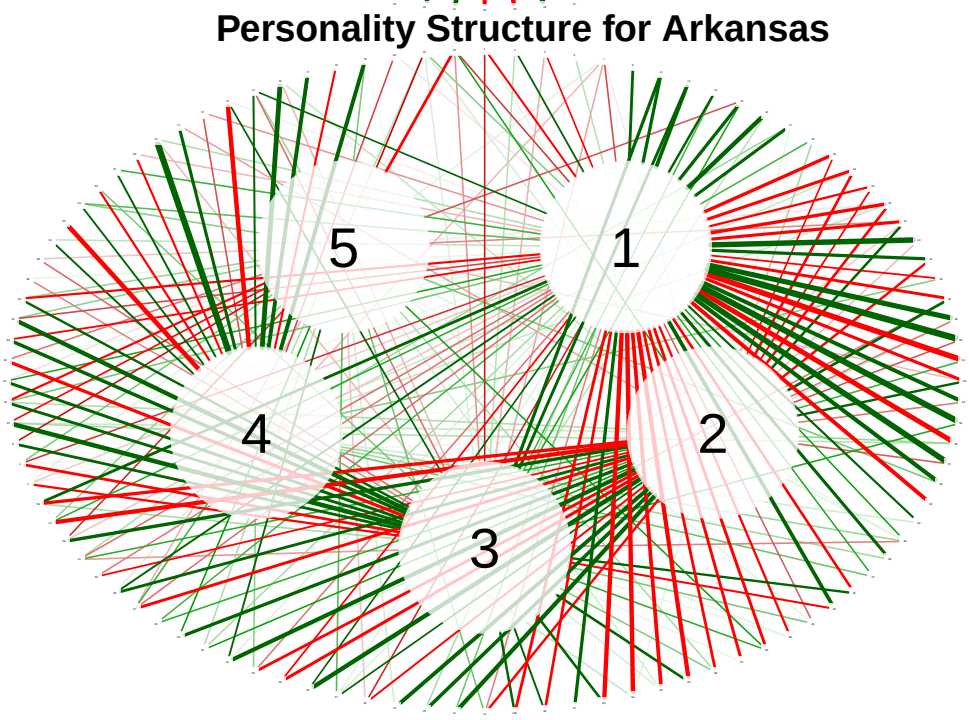
The main analyses are fifty parallel analyses, one for every state, that estimates the optimal number of personality dimensions for each state. We hypothesize that there will be variation in the number of ideal dimensions across states.

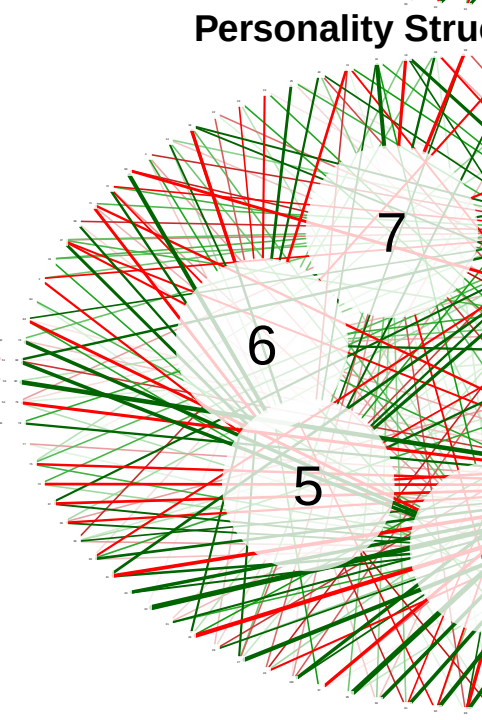
8 Results

The parallel factor analyses indicate that the modal number of factors is 5, as is found in 36 of the 50 states, with 7, 6, and 1 states respectively being better represented by 6, 7, and 8 factors.









8.0.1 EK Peer Review:

9 Areas of strength:

1. Awesome use of branches/git features in general! Made it clear to see who was working on what sections, and how the project was arranged.
2. Code is arranged easily to read, nice use of sections and not doing more than 1~2 things per line of code.
3. Great reproducibility with having data loading working on first try for me, and doesn't require any extra files/folders outside of the "scripts" folder- might be beneficial to cache the data however?

What I learned: Familiarity with the IPIP dataset! I'd really like tos ee some visualizations on how these factors vary across the different states, maybe using a geographic visualization?

10 EK Suggestions:

- Code cleanup: moved library declarations to teh beginning of the script, and added the "needs" package" to simplify package loading