Personality research: an open and shared science

Presented to the Department of Psychology, University of Zurich

William Revelle
Personality, Motivation and Cognition lab
(akaTelemetrics lab)

Northwestern University
Evanston, Illinois USA

Slides available at `personality-project.org/sapa`

Outline

Dominant theme: The importance of open science for psychological research.

1. Open source statistics: The R project
2. Open source materials: IPIP, ICAR, and SAPA data bases
3. Open source methodology: Synthetic Aperture Personality Assessment (SAPA)
   - SAPA: borrowing ideas from radio astronomy
   - Manhattan plots and persome correlations, borrowing ideas from GWAS
4. Open source data: *Journal of Open Psychology Data* and *DataVerse*

Open Science

- Science is an international collaborative endeavor that benefits when more people from more countries participate.
- Scientific societies were started (e.g, the Royal Society in London in 1660) as an “invisible college” to facilitate communication and the sharing of ideas.
- Traditionally we collaborate by publishing our results in scientific journals and by sharing our ideas at national and international conferences or giving guest lectures to our colleagues.
- More recently, there is a trend towards sharing our materials, our methods, and our results, even our data, on the web.
- This makes for better science.

Open Science and the problem of replication

- The last several years has seen a plethora of papers reporting failures to replicate results. This has lead some to worry about the strength of our findings and others to question what does it mean to “replicate” or reproduce a result.
- Others have suggested that we should be more open in our designs, publishing what we plan to do independent of what we actually find.
- This is an important problem that should not be ignored, although pre-registering might inhibit exploratory research.
- But, open science is much more than protecting us from type I errors. It is a philosophy of collaboration. That is what I want to emphasize today.

A Short History of Science: Instrumentation and Modeling

The development of new tools leads to new theories

Instrumentation

1. Telescopes (Galileo)
2. Sailing ships (e.g., Beagle)
3. Depth sounders
4. CO_2 measurement (Keeling)
5. The internet (Al Gore?)
6. WWW (Tim Berners-Lee)
7. Cellphones (e.g., Steve Jobs)

Theories and Models

1. Newton (*Principa*)
2. Darwin/Wallace
3. Plate Tectonics
4. Climate change
5. Open Science
6. Remote assessment
7. Repeated within subject mobile assessment

A Short History of Science: Part 2: Mathematics and Statistics

1. Calculus (Newton/Leibniz)
2. Data visualization (Playfair to Tukey to Cleveland to Tufte)
3. Probability theory (Fermat/Pascal) and the normal curve (Gauss/Quetelet)
4. Correlation (Galton/Pearson/Spearman)
5. Factor analysis (Spearman/Thurstone) and Principal Components analysis (Pearson/Hotelling)
6. Discrete (experimental) conditions and the *t* and *F* (seeds x manure) distributions (Gossett and Fisher)
7. Main frame computation (Ada Lovelace, John von Neumann, Grace Hopper)
8. Randomization and resampling of empirical distributions, not idealized: (Tukey, Efron)
9. Desktop software for us all and open statistical software

Four types of openness:

1. Open source software: The R project (R Core Team, 2018)
2. Open source materials:
   - The International Personality Item Pool (IPIP) (Goldberg, 1999)
   - The International Cognitive Ability Resource (ICAR) (Condon & Revelle, 2014)
   - The Synthetic Aperture Personality Assesment (SAPA) database (Condon, 2017)
3. Open source methodology: The Synthetic Aperture Personality Assessment Project (Revelle, Wilt & Rosenthal, 2010; Revelle, Condon, Wilt, French, Brown & Elleman, 2016)
4. Open source data:
   - Data from the ICAR project (Condon & Revelle, 2016, 2015a)
   - Data from SAPA studies (Condon & Revelle, 2015c,b)

In the process of summarizing the last several years of my students and my research, I will show how we use open source software, items, and methods and then share them with the world.

Four types of openness:

1. Open source software: The R project (R Core Team, 2018)
2. Open source materials:
   - The International Personality Item Pool (IPIP) (Goldberg, 1999)
   - The International Cognitive Ability Resource (ICAR) (Condon & Revelle, 2014)
   - The Synthetic Aperture Personality Assesment (SAPA) database (Condon, 2017)
3. Open source methodology: The Synthetic Aperture Personality Assessment Project (Revelle et al., 2010, 2016)
4. Open source data:
   - Data from the ICAR project (Condon & Revelle, 2016, 2015a)
   - Data from SAPA studies (Condon & Revelle, 2015c,b)

In the process of summarizing the last several years of research, I will show how we use open source software, items, and methods and then share them with the world.

Part I Open Statistics: R

R: What is it?

1. R: An international collaboration for applied statistical research
   - Originally developed in New Zealand in 1991-93
   - Comprehensive R Archive (CRAN) run out of Vienna
   - Core R members in Austria (2), Canada, Denmark, France, Germany (2), India, New Zealand (3), Switzerland, US (6), UK
2. R: The open source - public domain version of S+
3. R: Written by statisticians (and some of us) for statisticians (and the rest of us)
4. R: Not just a statistics system, also an extensible language.
   - This means that as new statistics are developed they tend to appear in R far sooner than elsewhere.
   - R facilitates asking questions that have not already been asked.

Statistical Programs for Psychologists

- General purpose programs
  - R
  - S+
  - SAS
  - SPSS
  - STATA
  - Systat
- Specialized programs
  - Mx
  - EQS
  - AMOS
  - LISREL
  - MPlus
  - Your favorite program

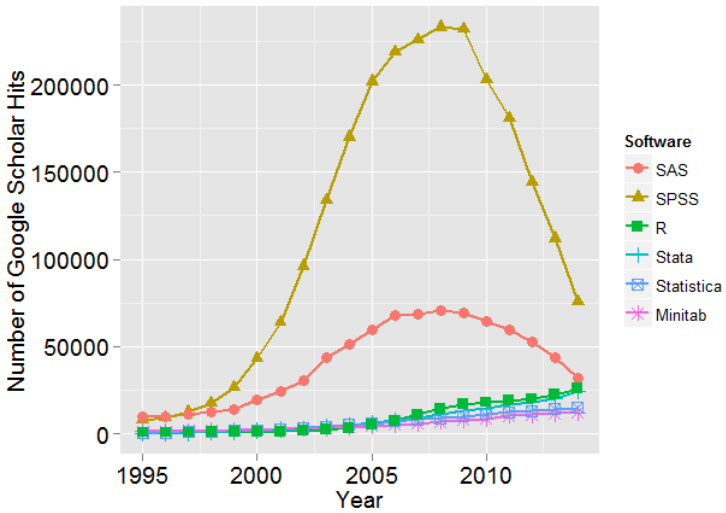
Statistical Programs for Psychologists

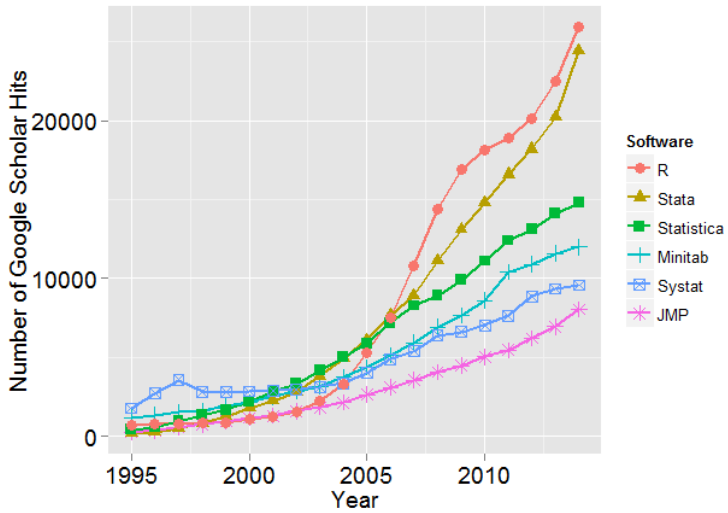
- General purpose programs
  - R
  - $+
  - A
  - $P$$
  - $TATA
  - $y$tat
- Specialized programs
  - Mx (OpenMx is part of R)
  - EQ$
  - AMO$
  - LI$REL
  - MPlu$
  - Your favorite program

R: A brief history

- 1991-93: Ross Dhaka and Robert Gentleman begin work on R project for Macs at U. Auckland (S for Macs).
- 1995: R available by ftp under the General Public License.
- 96-97: mailing list and R core group is formed.
- 2000: John Chambers, designer of S joins the Rcore (wins a prize for best software from ACM for S)
- 2001-2019: Core team continues to improve base package with a new release every 6 months (now more like yearly).
- Many others contribute “packages” to supplement the functionality for particular problems.
  - 2003-04-01: 250 packages
  - 2004-10-01: 500 packages
  - 2007-04-12: 1,000 packages
  - 2009-10-04: 2,000 packages
  - 2011-05-12: 3,000 packages
  - 2012-08-27: 4,000 packages
  - 2014-05-16: 5,547 packages (on CRAN) + 824 bioinformatic packages on BioConductor
  - 2015-05-20 6,678 packages (on CRAN) + 1,024 bioinformatic packages + ?,000s on GitHub/R-Forge
  - 2016-03-21 8,120 packages (on CRAN) + 1,104 bioinformatic packages + ?,000s on GitHub/R-Forge
  - 2019-02-18 13,738 packages (on CRAN) + 1,649 bioinformatic packages + ?,000s on GitHub/R-Forge

Popularity compared to other statistical packages





`http://r4stats.com/articles/popularity/` considers various measures of popularity

1. discussion groups
2. blogs
3. Google Scholar citations ($> 176,300$ citations, $\approx 41,000$ in 2018) for "R foundation for statistical computing"
4. Google Page rank

R as a way of facilitating replicable science

1. R is not just for statisticians, it is for all research oriented psychologists.
2. R scripts are published in psychology journals to show new methods:
   - *Psychological Methods*
   - *Psychological Science*
   - *Journal of Research in Personality*
   - *European Journal of Personality*
   - *Personality and Individual Differences*
3. R based data sets are now accompanying journal articles:
   - The *Journal of Research in Personality* now accepts R code and data sets.
   - JRP special issue in R.
   - The replicability project has released its data and R scripts.
4. By sharing our code and data the field can increase the possibility of doing replicable science.

Reproducible Research: Sweave and KnitR

Sweave is a tool that allows to embed the R code for complete data analyses in LaTeXdocuments. The purpose is to create dynamic reports, which can be updated automatically if data or analysis change. Instead of inserting a prefabricated graph or table into the report, the master document contains the R code necessary to obtain it. When run through R, all data analysis output (tables, graphs, etc.) is created on the fly and inserted into a final LaTeXdocument. The report can be automatically updated if data or analysis change, which allows for truly reproducible research.

Friedrich Leisch (2002). Sweave: Dynamic generation of statistical reports using literate data analysis.

Supplementary material for journals can be written in Sweave/KnitR so that others can redo or extend the analyses. Nicely implemented in RMarkdown and Rstudio which will produce pdf, html or even “word” files.

What is so great about reproducible research?

1. Allows us to share methods with our collaborators.
2. This can be other labs who want to know what you did. It can be your students, it can even be you.
3. David Condon has suggested that your closest collaborator is you, six months ago, but you don't answer your emails.
4. That is, scripted analyses are for you.
5. The Reproducibility Project (`https://osf.io/ezcuj/` has released their 100 replication data set and the R code to analyze it. If any one finds errors or needs more information, they are happy to provide it.
6. See, for instance Dan Gilbert et al. critique (Gilbert, King, Pettigrew & Wilson, 2016) and the response (Anderson, Bahník, Barnett-Cowan, Bosco, Chandler, Chartier & Cheung, 2016).

Misconception: R is hard to use

1. R doesn't have a GUI (Graphical User Interface)
   - Partly true, many use syntax.
   - Partly not true, GUIs exist (e.g., R Commander, R-Studio).
   - Quasi GUIs for Mac and PCs make syntax writing easier.
2. R syntax is hard to use
   - Not really, unless you think an iPhone is hard to use.
   - Easier to give instructions of 1-4 lines of syntax rather than pictures of menu after menu to pull down.
   - Keep a copy of your syntax, modify it for the next analysis.
3. R is not user friendly: A personological description of R
   - R is Conscientious: it wants commands to be correct.
   - R is not Agreeable: its error messages are at best cryptic.
   - R is Stable: it does not break down under stress.
   - R is Open: new ideas about statistics are easily developed.
   - R is Introverted: it will tell you what you want to know if you ask, but not if you don't ask.

Misconceptions: R is hard to learn – some interesting facts

1. With a brief web based tutorial `http://personality-project.org/r`, 2nd and 3rd year undergraduates in psychological methods and personality research courses are using R for descriptive and inferential statistics and producing publication quality graphics.
2. More and more psychology departments are using it for graduate and undergraduate instruction.
3. R is easy to learn, hard to master
   - R-help newsgroup is very supportive (usually)
   - Multiple web based and pdf tutorials see (e.g., `http://www.r-project.org/`)
   - Short courses using R for many applications. (Look at APS program).
4. Books and websites for SPSS and SAS users trying to learn R (e.g., `http://r4stats.com/`) by Bob Muenchen (look for link to free version).

What makes R so powerful are the $> 13,700$ contributed packages

psych A general purpose toolkit for psychological research with a particular emphasis upon

- Basic descriptive statistics and basic graphical tools.
- Basic psychometric procedures including functions for finding α $(= \lambda_3)$, ω_h, and ω_t.
- More advanced data reduction techniques using factor analysis, principal components analysis, mediation, moderation, and cluster analysis.
- Introductory Item Response Theory and Multi-level modeling.

lavaan Basic and advanced structural equation modeling ("The gateway package to R").

sem Structural equation modeling

lme4 Multilevel modeling.

The real power of R

1. Packages and functions can be nested.
2. The output of any function can be used as the input of any other function.
3. By reading the documentation and examining the code, one can add new functions to answer questions you want to answer.
4. The user community helps each other by reporting and fixing bugs so that the next release is better.

Part II: Open Materials

Four types of openness:

1. Open source software: The R project (R Core Team, 2018)
2. Open source materials:
   - The International Personality Item Pool (IPIP) (Goldberg, 1999)
   - The International Cognitive Ability Resource (ICAR) (Condon & Revelle, 2014)
   - The Synthetic Aperture Personality Assesment (SAPA) database (Condon, 2017)
3. Open source methodology: The Synthetic Aperture Personality Assessment Project (Revelle et al., 2010, 2016)
4. Open source data:
   - Data from the ICAR project (Condon & Revelle, 2016, 2015a)
   - Data from SAPA studies (Condon & Revelle, 2015c,b)

In the process of summarizing the last several years of research, I will show how we use open source software, items, and methods and then share them with the world.

Personality, prediction, and life outcomes

1. It has long been known that to predict real world outcomes we need to study more than just ability (Kelly & Fiske, 1950, 1951; Deary, 2008; Roberts, Kuncel, Shiner, Caspi & Goldberg, 2007).
2. Level of education and jobs differ in their intellectual requirements (Gottfredson, 1997).
3. My colleagues and I have shown that there are also temperamental requirements for educational and job choice (Condon & Revelle, 2014; Revelle & Condon, 2012, 2015a; Revelle, Wilt & Condon, 2011; Wilt & Revelle, 2015)
4. We consider individual differences in Temperament, Ability, and Interests (TAI) as they relate to niche selection in choice of college major and in occupational choice (Bouchard, 1997; Hayes, 1962; Johnson, 2010). as well as to the second and third level of personality analysis (between individuals and between groups of individuals (Revelle & Condon, 2015b)).

Measuring individual differences

1. A basic problem in the study of individual differences is that there are so many different constructs that interest us. These include constructs from at least four broad domains
   - Temperament
   - Ability
   - Interests
   - Character
2. Each domain has many constructs:
   - Dimensions of Temperament 2-3-5-6-16-27?
   - Structure of Ability (g - g_f, g_c, V-P-R)?
   - Hierarchical structure of interests people-things, RIASEC .
   - Range of possible measures of character.
3. But many important measures are proprietary.
4. In addition, showing the utility of TAIC measures requires criterion variables, and should include demographics.
5. Our solution: Use and/or develop open source temperament, ability, and interest items.

The International Personality Item Pool (Goldberg, 1999)

1. Perhaps one of the greatest contributions from Lew Goldberg has been his release of the International Personality Item Pool or IPIP (Goldberg, 1999) `http://ipip.ori.org`.
2. The IPIP adapted a short stem item format developed in the doctoral dissertation of Hendriks (1997) and items from the Five Factor Personality Inventory developed in Groningen (Hendriks, Hofstee & De Raad, 1999).
3. Goldberg (1999) used about 750 items from the English version of the Groningen inventory, and has since supplemented them with many more new items in the same format.
4. The IPIP items have been translated into at least 39 languages by at least 65 different research teams. This includes Arabic, German, Farsi, Icelandic, Indonesian, Japanese, Korean, Mandarin, Polish, Portuguese, Russian, Serbian, Spanish, Turkish, Urdu, Slovenian and Swedish.

IPIP and other personality inventories

1. The IPIP was originally meant to be short stems to measure the Abridged Five Factor Circumplex structure of adjectives (Hofstee, de Raad & Goldberg, 1992) but also includes items targeted at most major personality tests.
2. Using a panel of roughly 1000 (very conscientious) residents from Eugene-Springfield, Oregon, Goldberg administered his original IPIP items along with the NEO-PI-R (Costa & McCrae, 1992), the CPI (Gough & Bradley, 1996), the 16PF (Cattell & Stice, 1957), the MPQ (Tellegen & Waller, 2008), the Hogan PI (Hogan & Hogan, 1995), the TCI (Cloninger, Przybeck & Svrakic, 1994), the JPI-R (Jackson, 1983), and the 6FPQ (Jackson, Paunonen & Tremblay, 2000).
3. Goldberg then developed item stems that were highly correlated to the commercial inventories and put these into the public domain with the formation of the IPIP.
4. The items are available at `http://ipip.ori.org` and the Eugene-Springfield data are available from Goldberg.

What are the "Big 5"?: Some representative items

Semantic analysis of many (although primarily European) languages suggest 5 broad factors of the ways in which we describe others.

Conscientiousness Complete my duties as soon as possible. Do things according to a plan. Like order.

Agreeableness Take advantage of others. (R) Am concerned about others. Sympathize with others' feelings.

Neuroticism Get upset easily. Get overwhelmed by emotions. Have frequent mood swings.

Openness/Intellect Am able to come up with new and different ideas. Am full of ideas. Have a rich vocabulary.

Extraversion Like mixing with people. Enjoy meeting new people. Am a talkative person. Am rather lively.

These are sometimes organized as the OCEAN of personality, alternatively, the CANOE of personality.

Extending the IPIP

1. In addition to the basic temperament items at the IPIP site, there are additional items to measure vocational interests (the ORVIS) (Pozzebon, Visser, Ashton, Lee & Goldberg, 2010) as well as avocational interests (Goldberg, 2010) and behavioral measures (Elleman, 2019)
2. David Condon has expanded 2500 IPIP items to include the original IPIP items, the ORVIS, the ORAIS, as well as items from the EPQ (Eysenck, Eysenck & Barrett, 1985), the O*NET interest profile scales (Rounds, Su, Lewis & Rivkin, 2010). These, and other items make a total set of $> 10,000$ items.
3. 7,428 of these are available at `https://sapa-project.org/MasterItemList/`.
4. Condon has also noted that although 18 different inventories (with 168 scales) have what appear to be 1,894 items, their are actually just 696 unique items. In addition, those "magic 696" cover between 57% to 85% of 10 additional inventories with 235 additional scales.

The International Cognitive Ability Resource

ICAR: Extending the IPIP to ability: IPIP:Personality::ICAR:Ability

1. ICAR is an international collaboration to develop open source cognitive ability items.
2. Information at `http://www.icar-project.com/`
3. News letter at `http://www.icar-project.com/ICAR_News_Issue_One.pdf`
4. Key organizers who are coordinating the project:
   - Germany: Phillip Doebler (Münster and Ulm) and Heinz Holling (Münster)
   - U.K.: Luning Sun and John Rust (Cambridge)
   - U.S.A: William Revelle (Northwestern) and David Condon (U. Oregon)
5. Everyone is welcome to join this international collaboration.
6. Supported by Open Research Area (ORA) for the Social Sciences which included participation from national funding agencies (Germany:DFG), (UK:ESRC), (US:NSF).

ICAR: Proof of concept

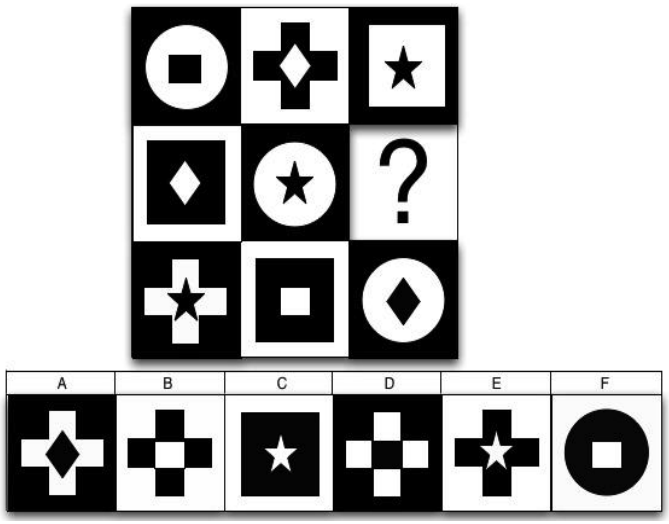
1. About 60 items were developed as part of a honors thesis at Northwestern by Melissa Liebert (Liebert, 2006) and meant to be "Google resistant" (answers are not available on the web).
   - This set was reported at conference in Krakow and in a subsequent book chapter (Revelle et al., 2010).
2. Subsequently David Condon developed some 3 Dimensional rotations items and did some extensive item analysis of the total set.
3. Condon & Revelle (2014) examined the first 60 publicly available items and validated them against self reported SAT exam scores as well as a small sample given the Shipley-2 (Shipley, 2009).
4. The original data set has been released to DataVerse (Condon & Revelle, 2015a) and has been published in the *Journal of Open Psychology Data* (Condon & Revelle, 2015c).
5. An example data set of 16 items with $N = 1,525$ is included as the `ability` dataset in the *psych* package.

The ICAR project extended this set

1. Philipp Doebler (Münster) developed Automatic Item Generation (AIG) functions for reasoning
2. Fiona Chan and Luning Sun (Cambridge) developed “Ravens Like” matrix reasoning
3. Loe & Rust (2017) developed a perceptual maze test.
4. Condon developed 2 dimensional rotation and a forced choice remote associates test
5. Several other item types have been explored.
6. Many of these developments are still be validated.

Sample ICAR items

Matrix Reasoning



Verbal Reasoning

What number is one fifth of one fourth of one ninth of 900?

(1) 2 (2) 3 (3) 4 (4) 5 (5) 6 (6) 7

If the day after tomorrow is two days before Thursday,

then what day is it today?

(1) Friday (2) Monday (3) Wednesday

(4) Saturday (5) Tuesday (6) Sunday

Letter and Number Series

In the following alphanumeric series, what letter comes next?

I J L O S

(1) T (2) U (3) V (4) X (5) Y (6) Z

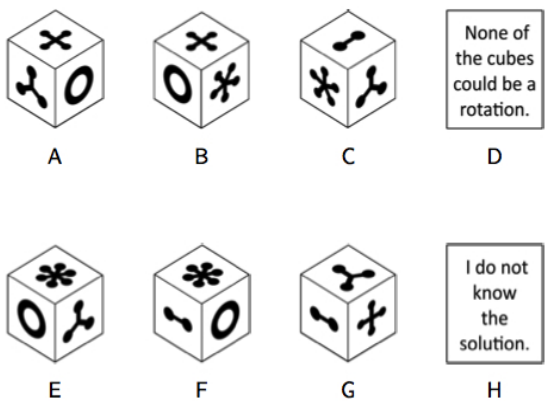
In the following alphanumeric series, what letter comes next?

Q S N P L

(1) J (2) H (3) I (4) N (5) M (6) L

Three-Dimensional Rotation


X



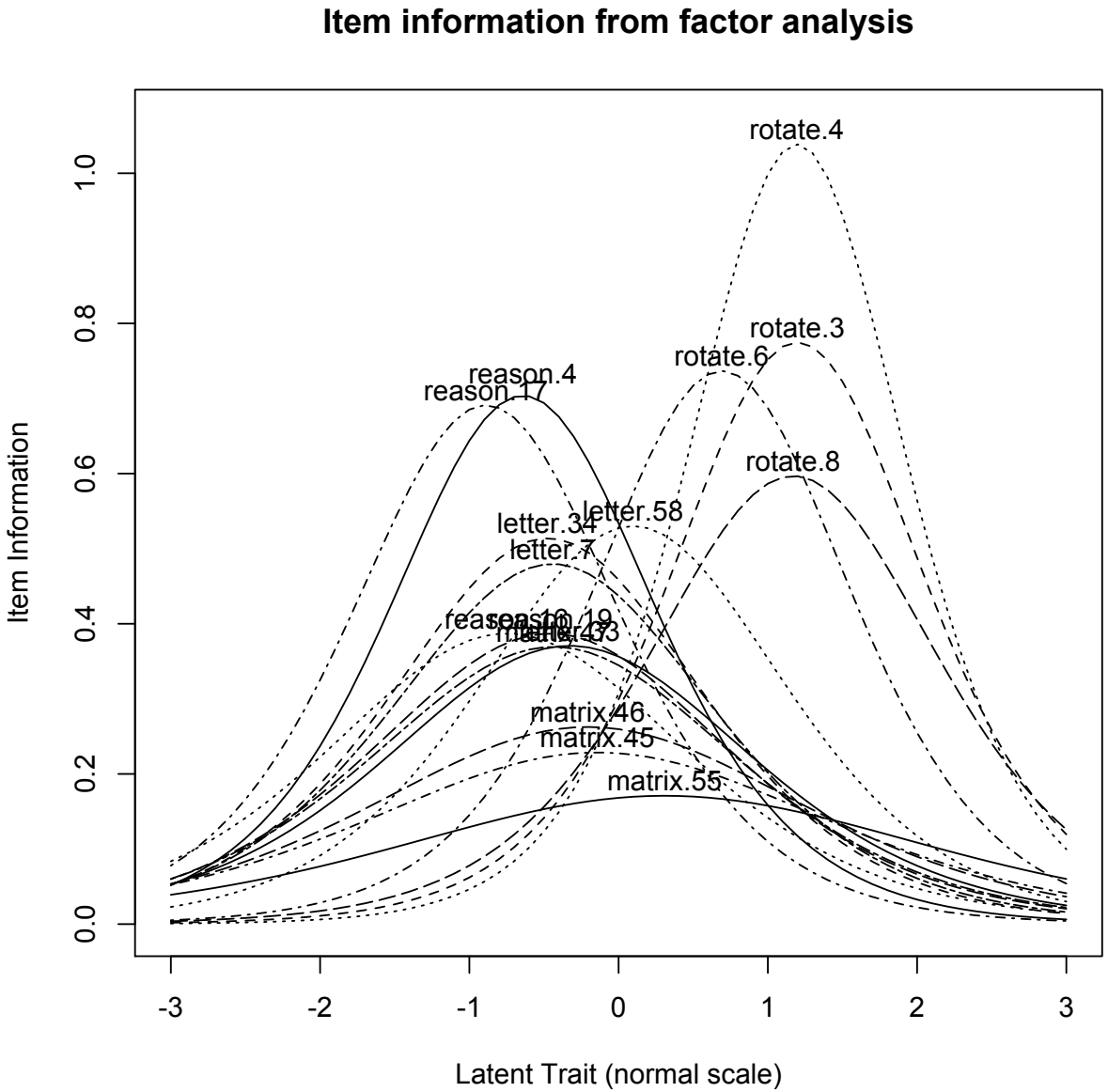
Sample analysis of ICAR items

1. Using basic R functions in the *psych* package (Revelle, 2018) we can evaluate the factor structure of the ICAR items.
2. `irt.fa` will do a factor analysis of the items and report the statistics in terms of those statistics more commonly used in Item Response Theory.
   - The two parameters from factor analysis are item difficulty taken from the τ parameter from the tetrachoric correlation and the item factor loading λ of the matrix of tetrachoric correlations.

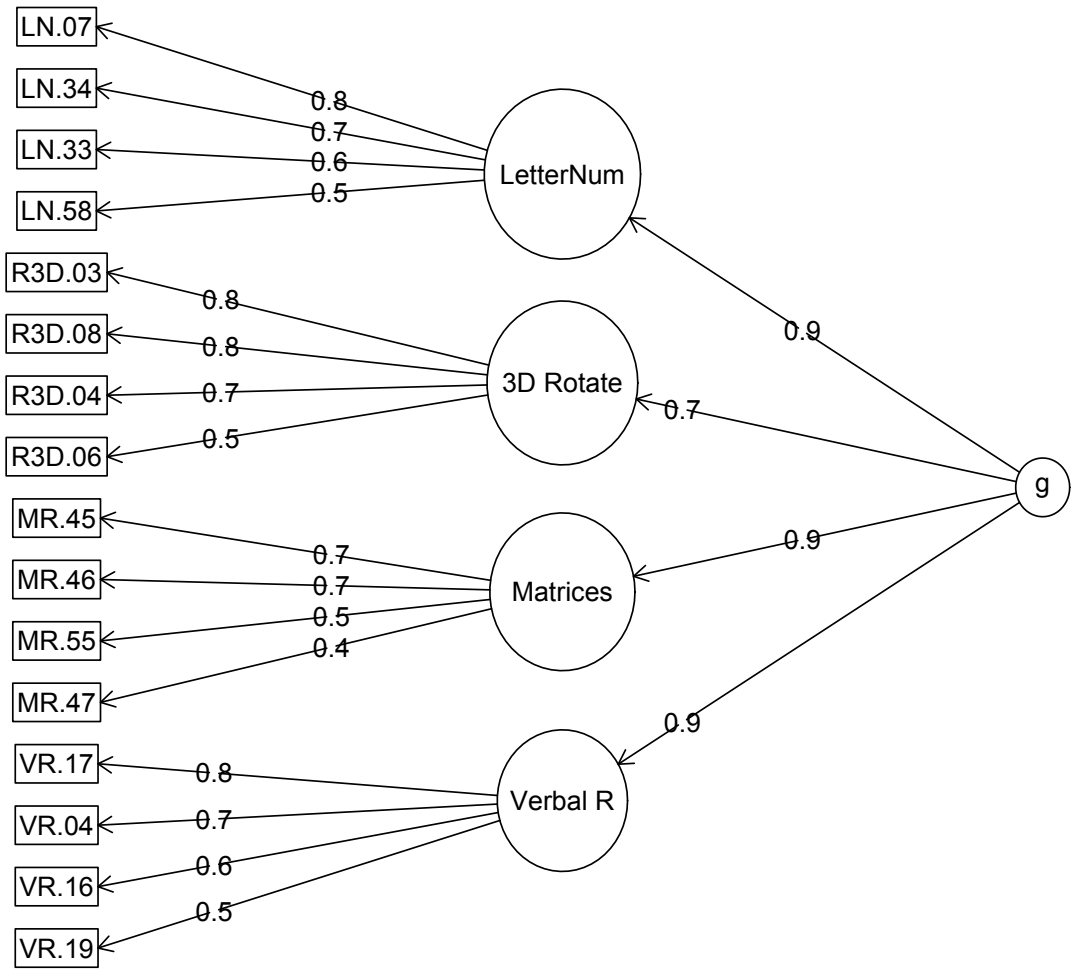
$$a = \frac{\lambda}{\sqrt{1-\lambda^2}} \qquad \delta = \frac{\tau}{\sqrt{1-\lambda^2}}$$

   - The hierarchical structure of the ability items may be shown by factoring the factor intercorrelations.
   - Loadings on a general factor may then be found by using the `omega` function which applies a Schmid Leiman transformation to the resulting higher level solution.

Item information curves for the 16 ICAR sample set

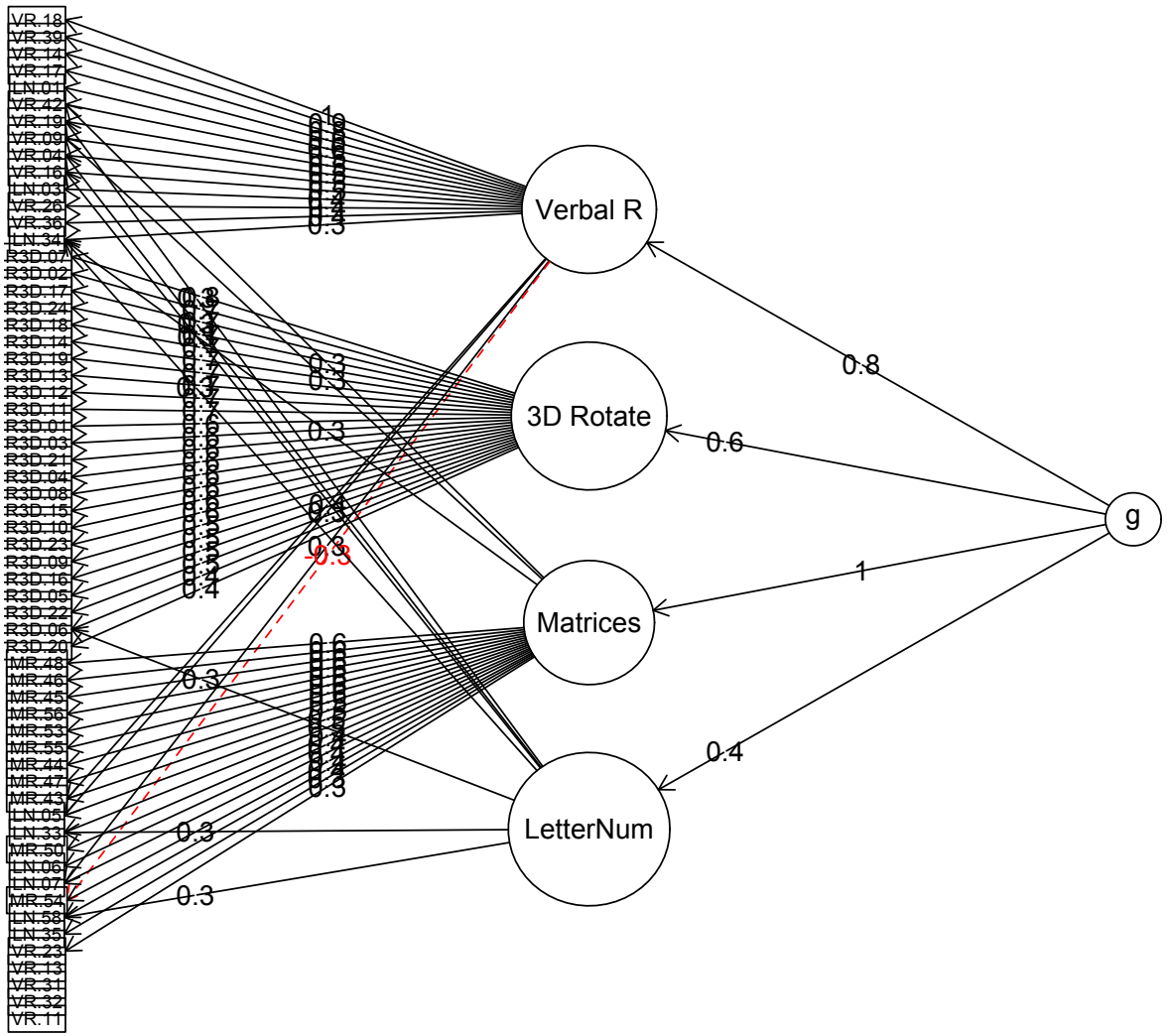


Structure of sample ICAR 16 items shows a clear 4 factor hierarchical solution $\omega_h = .87$



Structure of ICAR 60 items shows a messier 4 factor hierarchical solution $\omega_h = .76$

Hierarchical structure of ICAR60 items



Open materials

1. The International Personality Item Pool items (Goldberg, 1999) as well as the extended IPIP are in the public domain and are available to anyone for free.
2. The items from the International Cognitive Ability resource are also in the public domain and are available to registered users. (We are trying to keep the items relatively secure and do not put all of the actual items up on the web.)
   - We have a basic set of 60 ICAR items (Condon & Revelle, 2014) and the ICAR group is developing and validating item generators to automatically produce hundreds of each of a growing number of item types.
   - Currently have 952 items with $> 220K$ participants.
   - We encourage others to join us in this mission.

Part III: Open Methods

Method: Synthetic Aperture Personality Assessment (SAPA)

Measuring individual differences: the tradeoff between breadth versus depth

- Profile correlations

Synthetic Aperture Personality Assessment

SAPA theory

- Sample items as well as people
- Covariance algebra

SAPA: practice

- Open source software comes to the rescue

Part IV: Open Data

Four types of openness:

1. Open source software: The R project (R Core Team, 2018)
2. Open source materials:
   - The International Personality Item Pool (IPIP) (Goldberg, 1999)
   - The International Cognitive Ability Resource (ICAR) (Condon & Revelle, 2014)
   - The Synthetic Aperture Personality Assesment (SAPA) database (Condon, 2017)
3. Open source methodology: The Synthetic Aperture Personality Assessment Project (Revelle et al., 2010, 2016)
4. Open source data:
   - Data from the ICAR project (Condon & Revelle, 2016, 2015a)
   - Data from SAPA studies (Condon & Revelle, 2015c,b)

In the process of summarizing the last several years of research, I will show how we use open source software, items, and methods and then share them with the world.

Measuring individual differences

1. A basic problem in the study of individual differences is that there are so many different constructs that interest us. These include constructs from at least four broad domains
   - Temperament
   - Ability
   - Interests
   - Character
2. Each domain has many constructs
   - Dimensions of Temperament 2-3-5-6-16-27?
   - Structure of Ability: g - g_f, g_c, V-P-R ?
   - Hierarchical structure of interests people-things RIASEC
   - Range of possible measures of character
3. In addition, showing the utility of TAIC measures requires criterion variables

Breadth vs. depth of measurement

1. Factor structure of domains needs multiple constructs to define structure.
2. Each construct needs multiple items to measure reliably.
3. This leads to an explosion of potential items .
4. But, people are willing to only answer a limited number of items.
5. This leads to the use of short and shorter forms (the NEO-PI-R with 300, the IPIP big 5 with 100, the BFI with 44 items, the TIPI with 10) to include as part of other surveys.

Many items versus many people

1. Not only do want many people, we also want many items.
2. Resolution (fidelity) goes up with sample size, N (standard errors are a function of $\sqrt{N}$)

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N-1}} \qquad \sigma_r = \frac{1-r^2}{\sqrt{N-2}}$$

3. Resolution also increases as number of items, n, measuring each construct (reliability as well as signal/noise ratio varies as number of items and average correlation of the items)

$$\lambda_3 = \alpha = \frac{n\bar{r}}{1+(n-1)\bar{r}} \qquad s/n = \frac{n\bar{r}}{(1-n\bar{r})}$$

4. Thus, we need to increase N as well as n. But how?

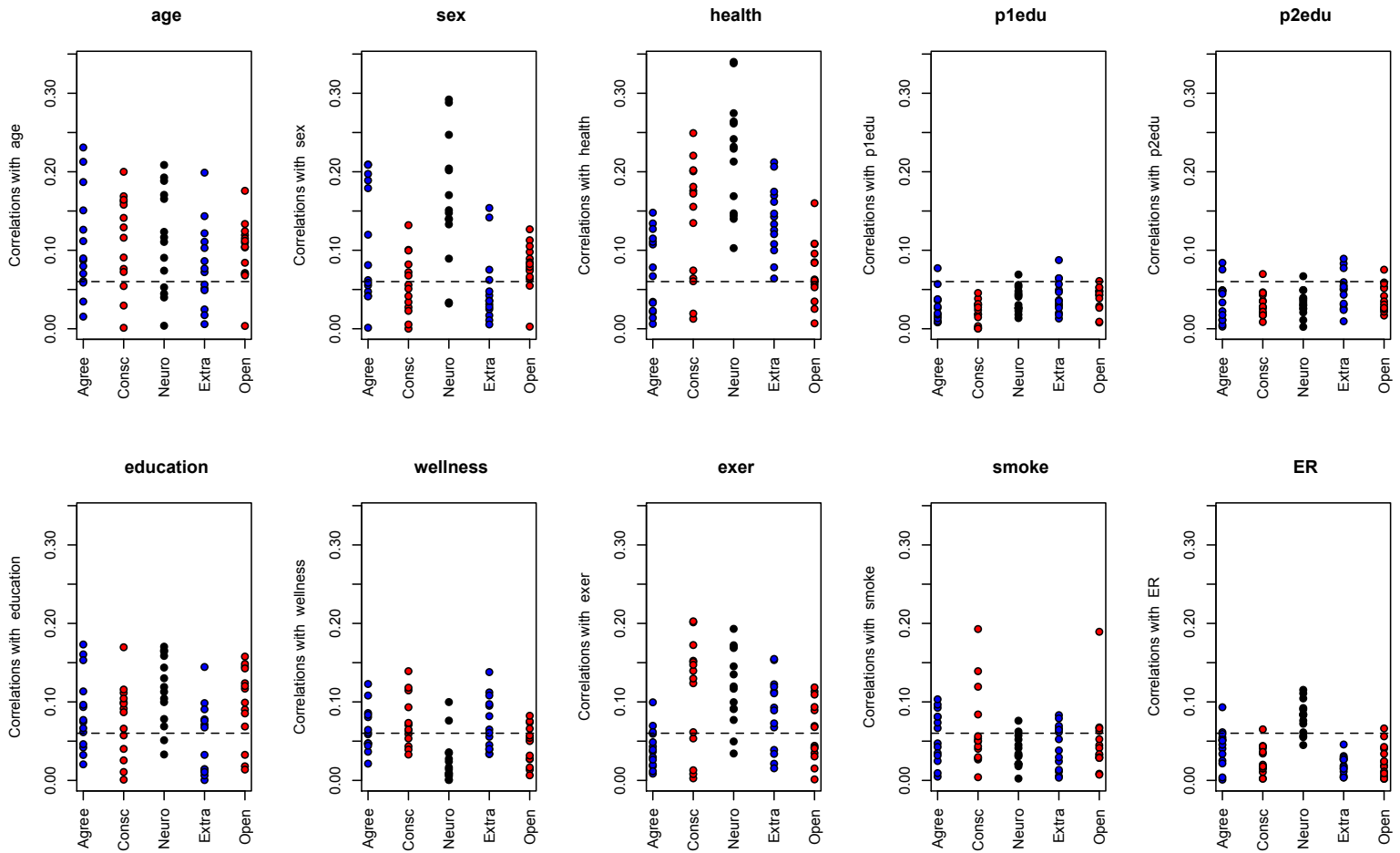
How useful are items?

1. Common observation is that items have low correlations with other items.
2. From a classical reliability perspective: Item variance = general + group + specific + error.
3. The “gospel” is that items are mainly error variance.
4. This is true from a latent variable perspective, but less true if we actually examine item variance.
5. Perhaps 20% of an item is general factor variance, another 10-20% group variance but about 40% is specific and reliable variance.
6. We can see this by doing a variance decomposition of items that are repeated across time.
7. So what?
8. Lets look at the correlates of items.

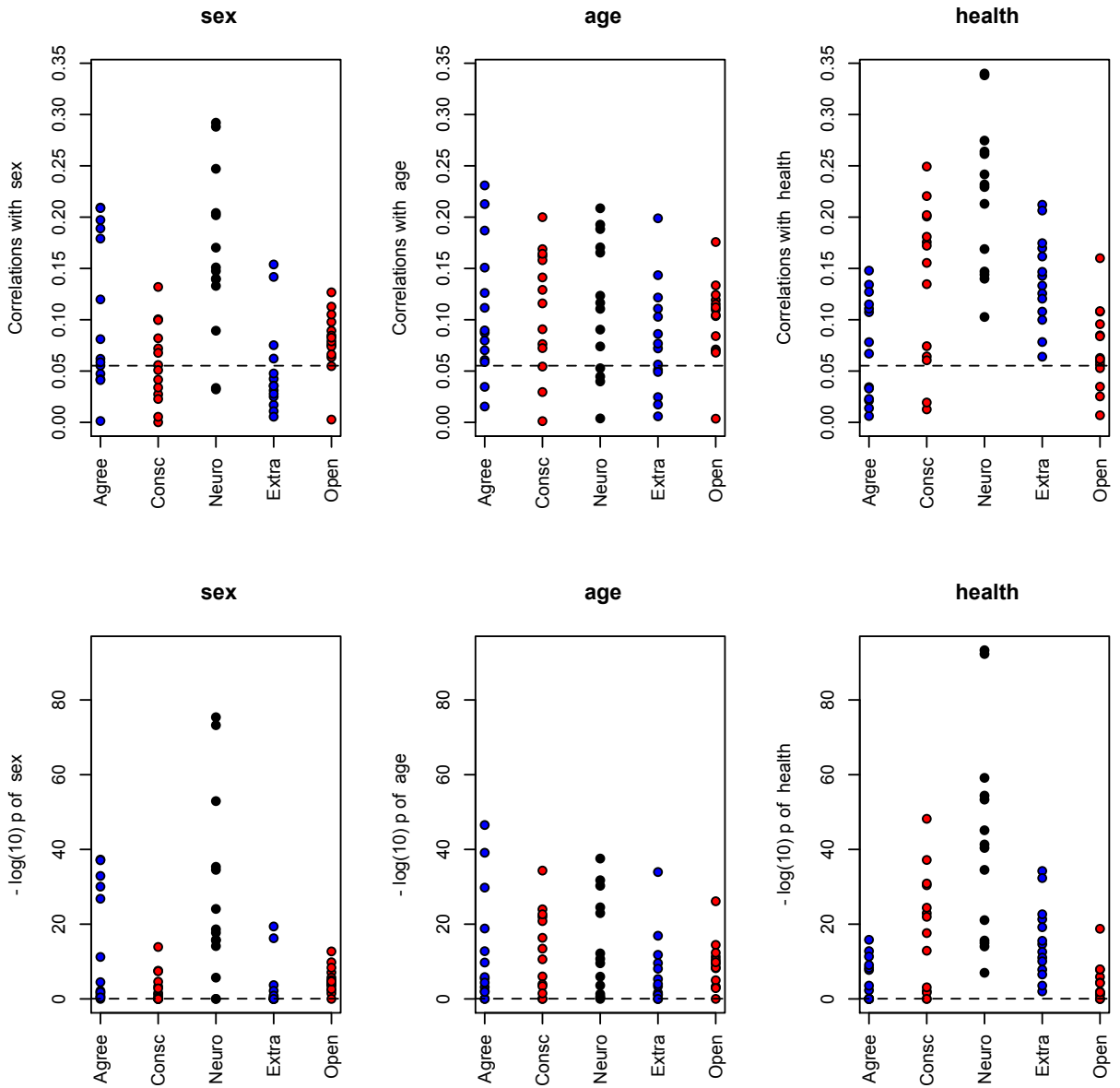
Items as analogous to SNPs in GWAS studies

1. In Genome Wide Association Studies one examines phenotypic variation as it correlates with differences in SNP frequencies across the genome.
2. Do the same by examining phenotypic variation and correlation across the persome (Möttus, Sinick, A.Terracciano, Hřebíckova, Kandler & Jang, 2018)
3. A typical approach is to show the correlations and their probability values (corrected for multiple tests)
   - Typically displayed in "Manhattan Plots" across the genome. We do this across the "Persome".
4. First show plots for an open source data set (spi) available in the *psych* package.
   - This is a set of 135 temperament items with 10 criteria for 4,000 subjects.
5. Then do the same for items from the Big 5, then an extend set (the little 27), then for a bigger data set with even more items.

A "Manhattan plot" of the spi items on the big 5 for 10 criteria



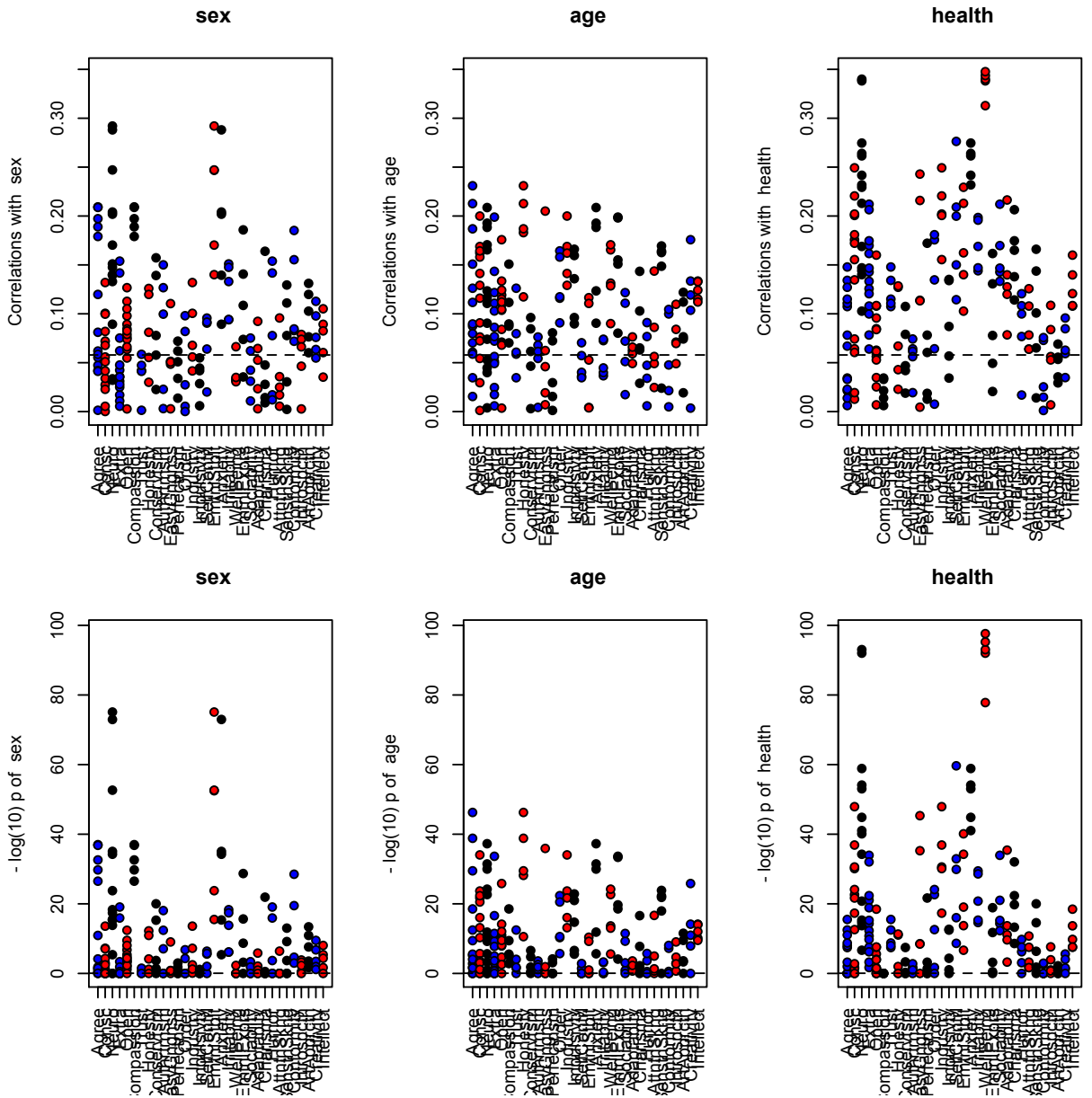
A "Manhattan plot" of the spi items for 3 criteria big 5



Correlations (absolute values)

Log p values (Holm corrected for multiple tests)

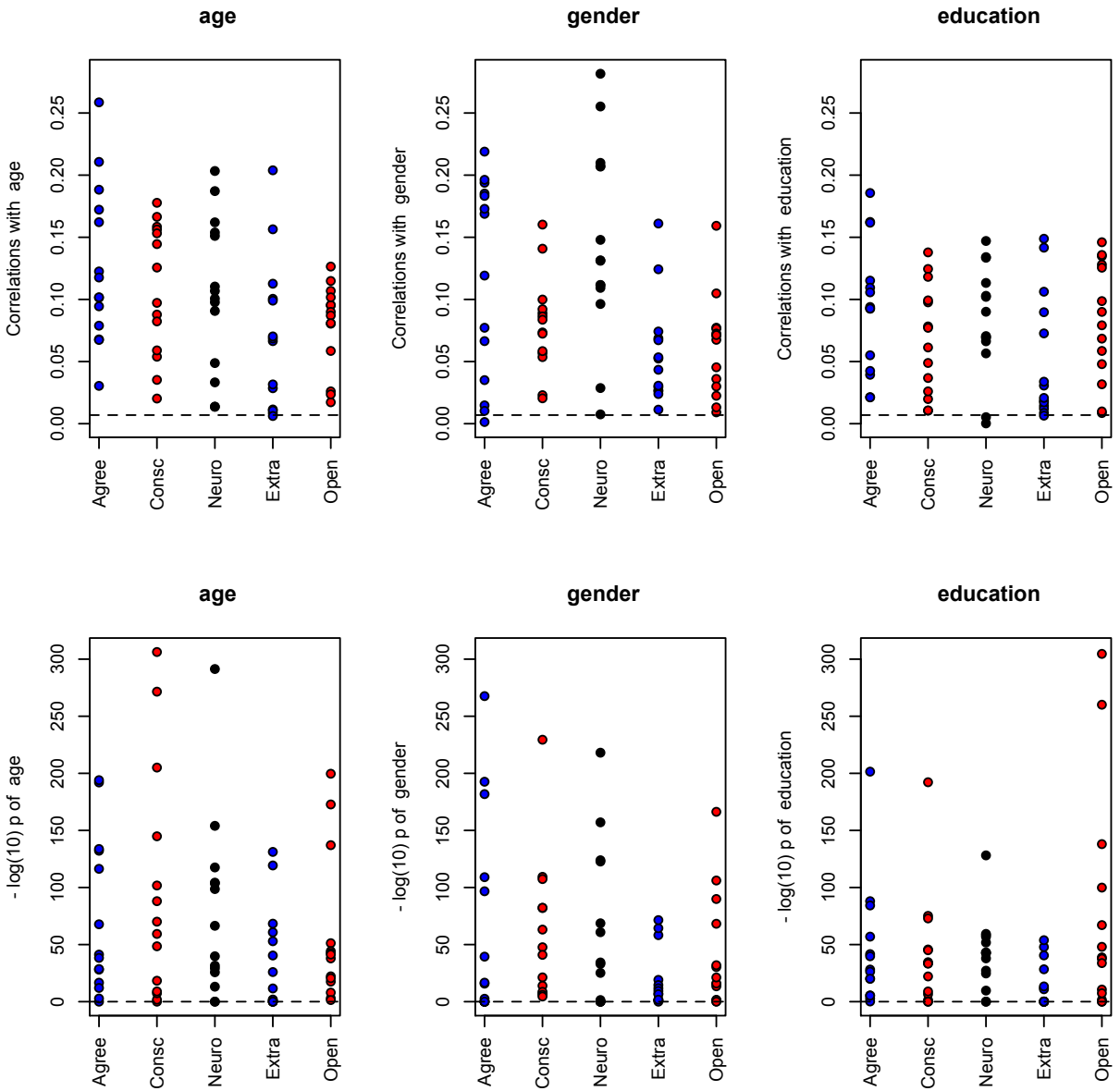
More predictors: 3 criteria big 5 + spi 27, N =4000



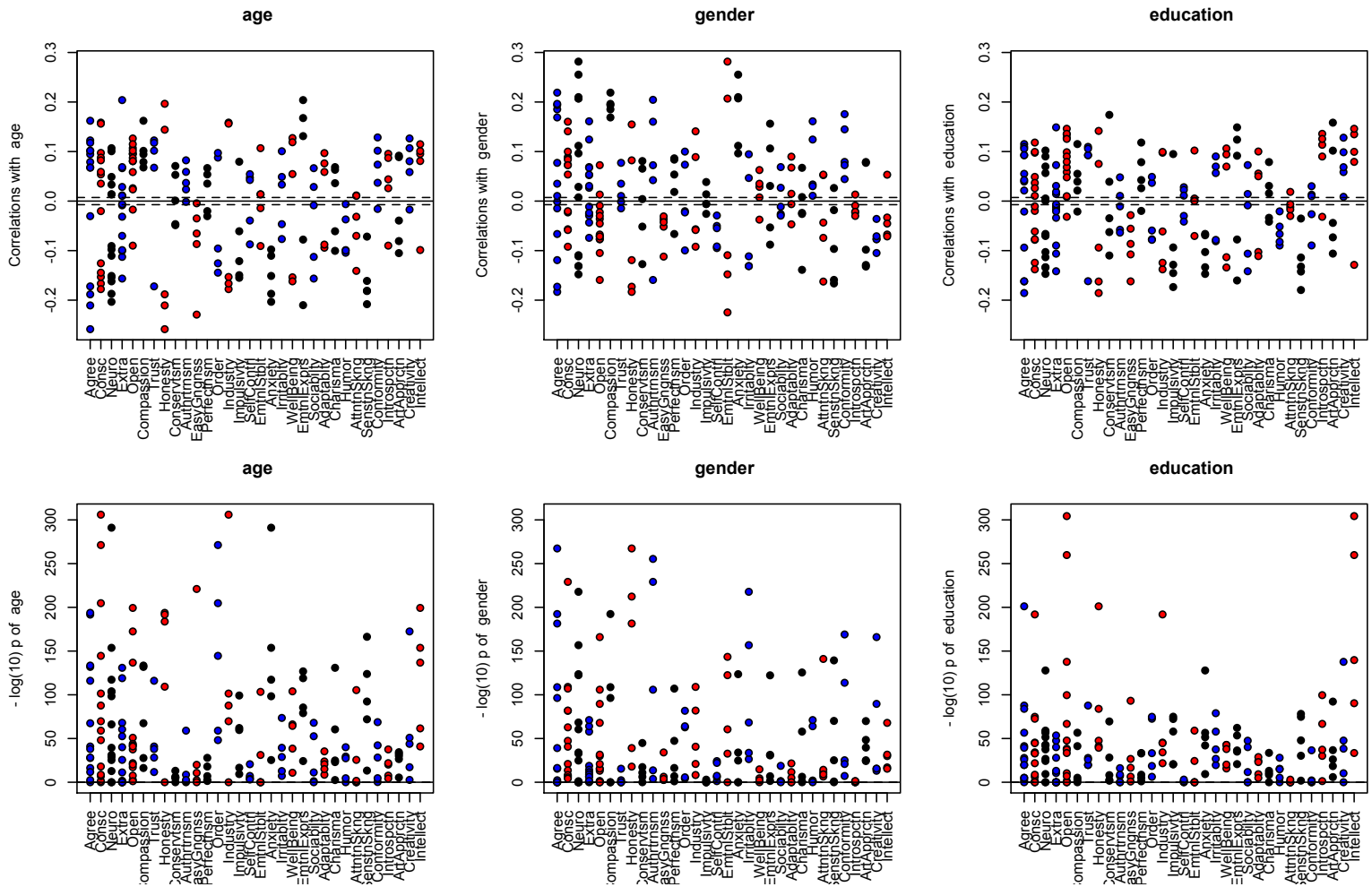
Correlations (absolute values)

Log p values (Holm corrected for multiple tests)

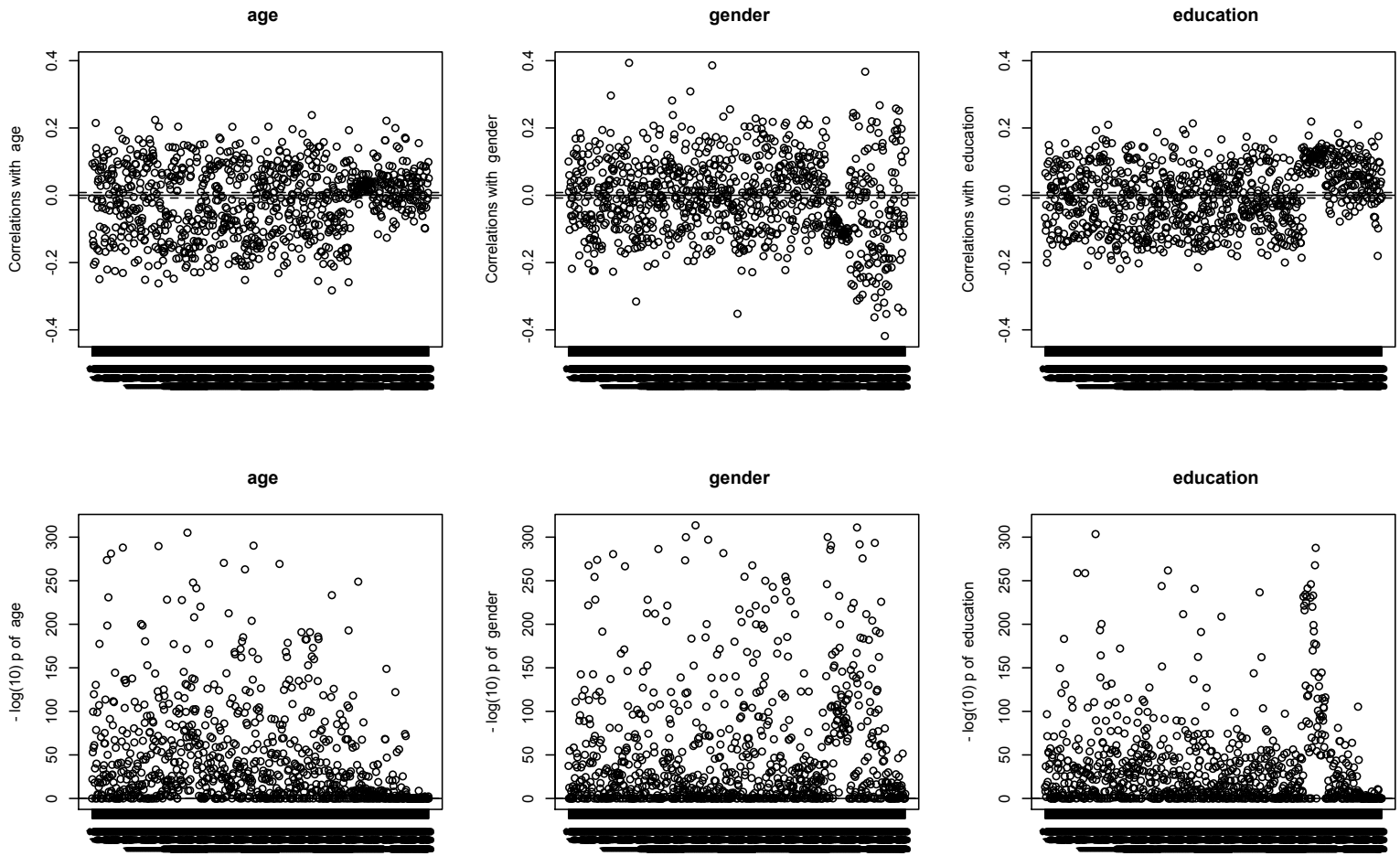
More subjects: 3 criteria big 5, N = 255,000



More subjects: 3 criteria - Big 5 + little 27 items, N = 255,000



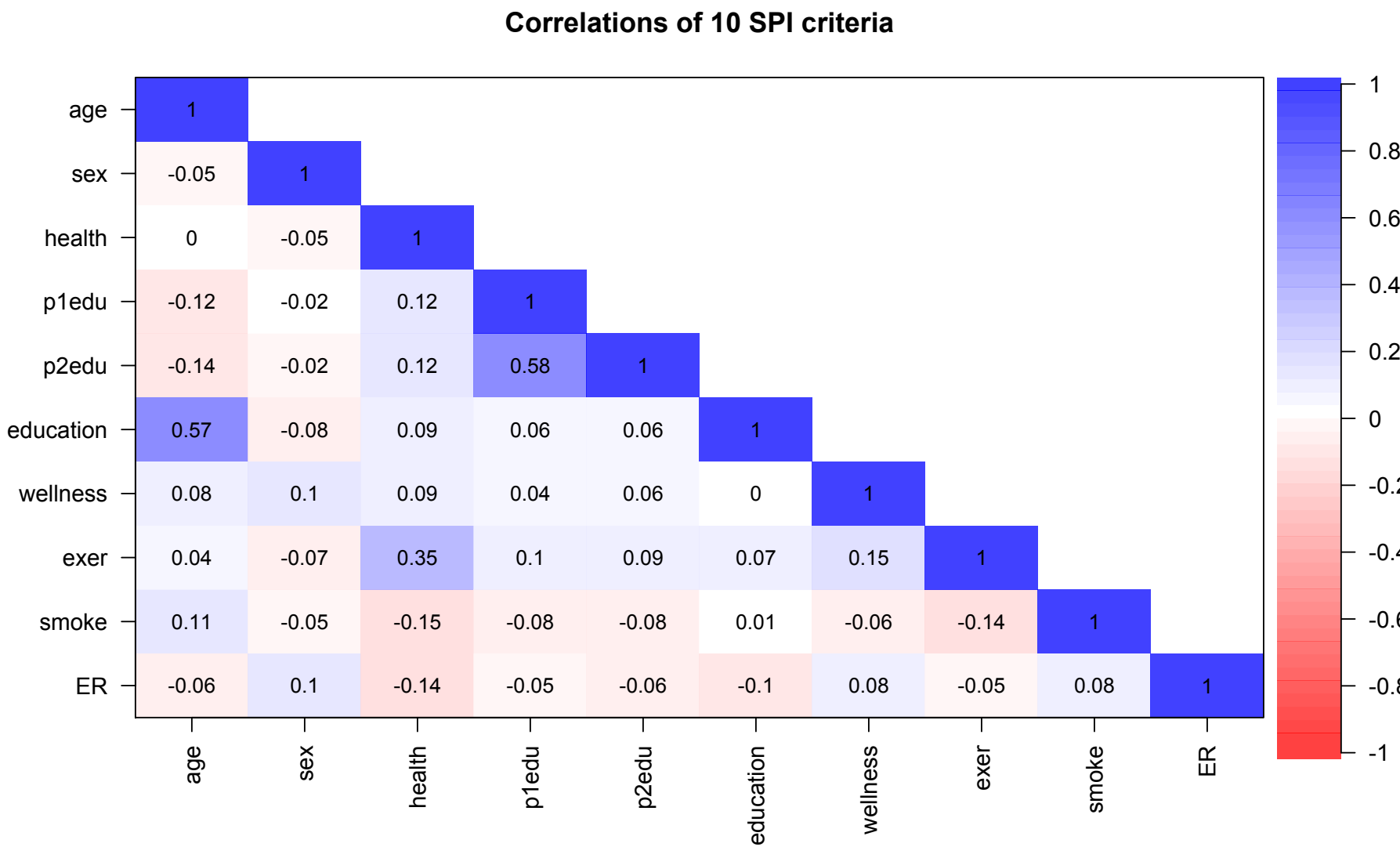
More subjects: 3 criteria - 904 items (temperament, abilities, interests)



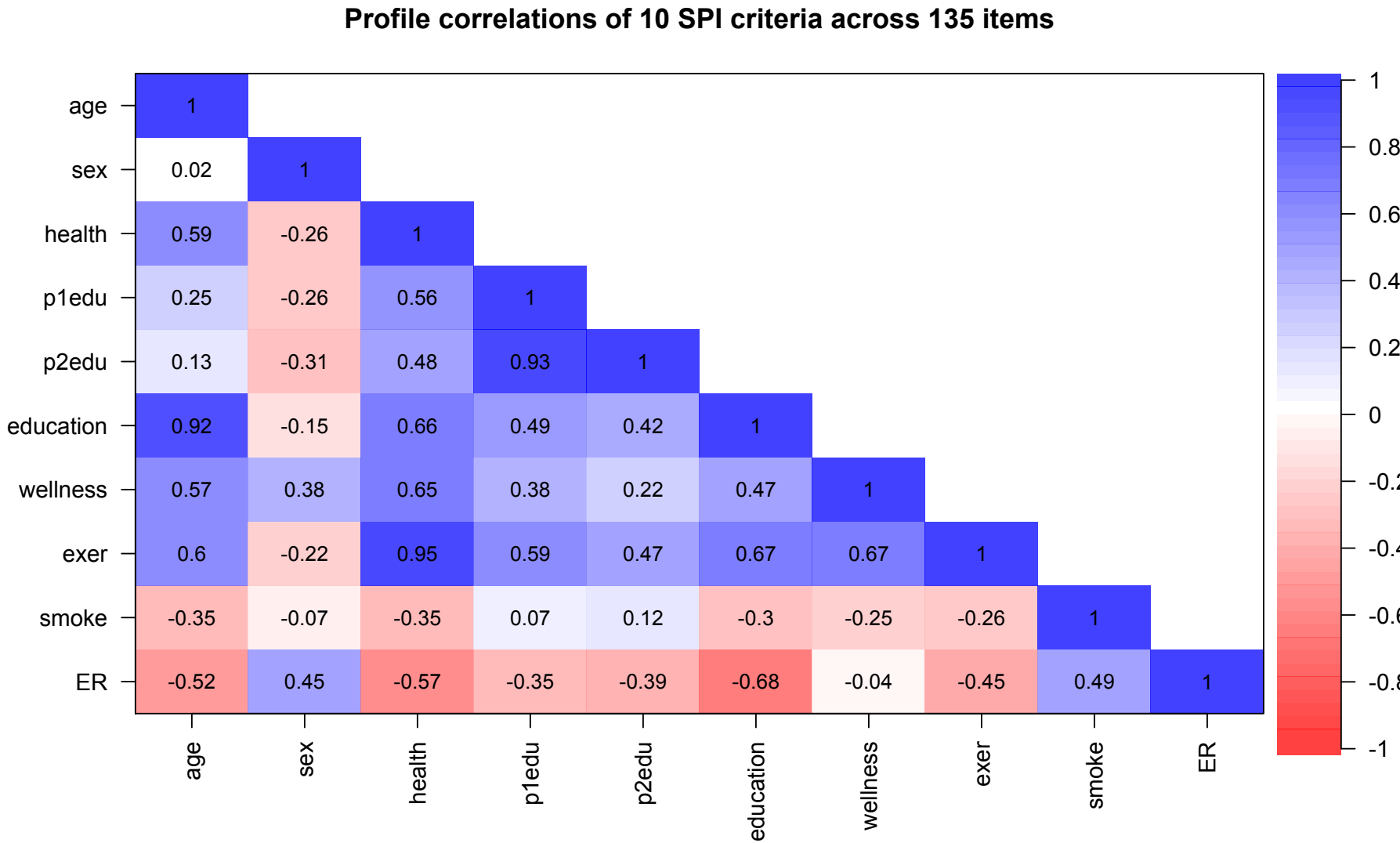
Profile correlations are analogous to the "genetic correlation"

1. For any set of criteria or grouping variables we can find a vector of validity correlations across our predictor set.
2. We can then correlate these vectors. This is analogous to the genetic correlation across SNPs.
3. Basically, we are correlating the profiles of the Manhattan plots
4. I show this using the 10 criteria in the `spi` data set
5. First the raw correlations, then the profile correlations

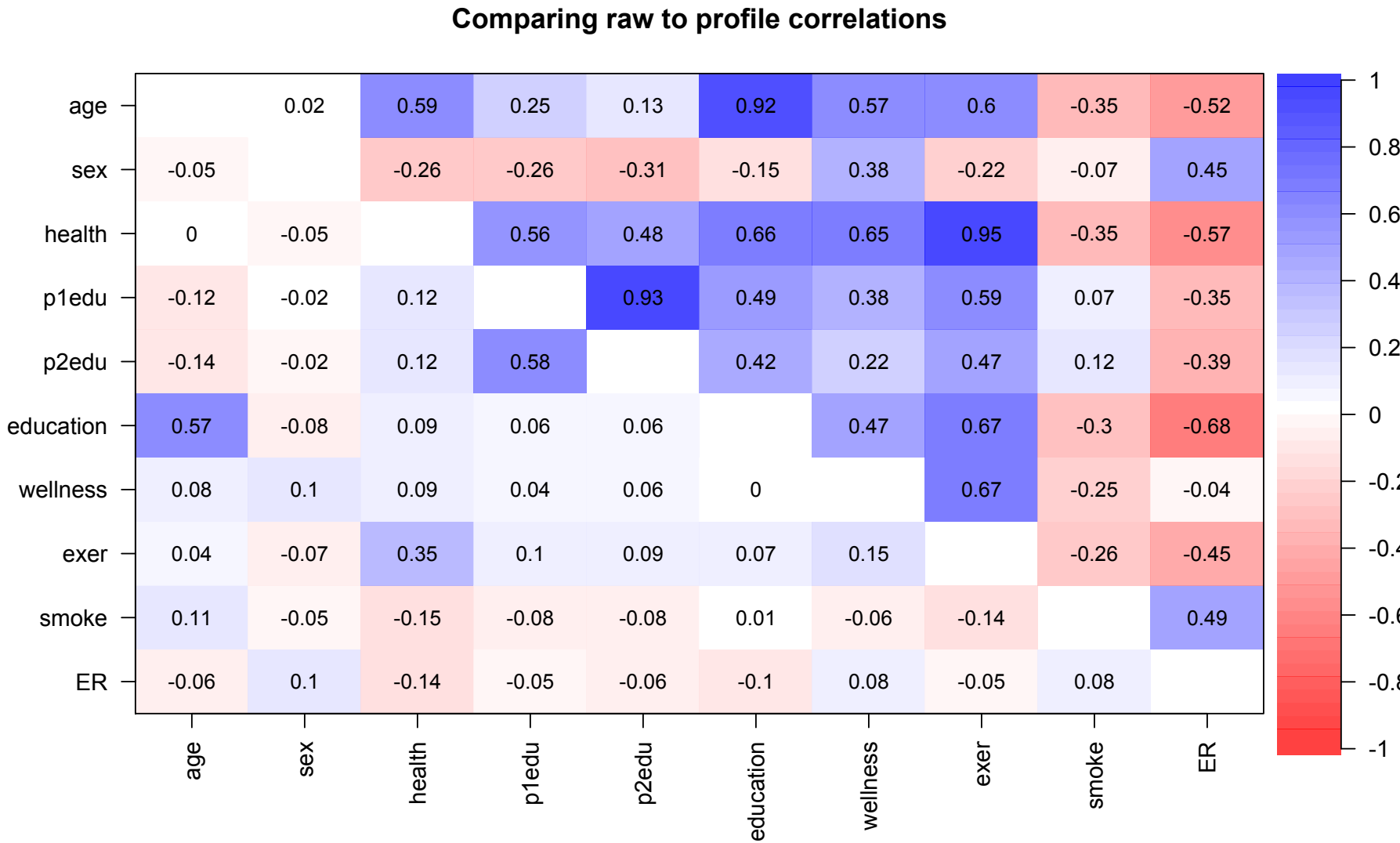
10 criteria from the SPI data set, raw correlations



10 criteria from the SPI data set, profile correlations



Comparing raw and profile correlations from the SPI dataset

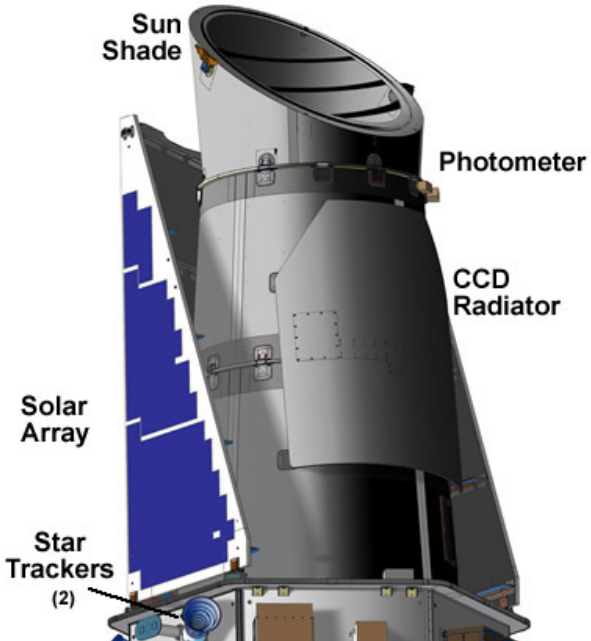


How do we get lots of items and lots of people?

1. Use the web and SAPA
2. SAPA: Synthetic Aperture Personality Assessment
3. Analogous to synthetic aperture radio telescopes
4. Just a fancy name for Massively Missing Completely at Random data

A short diversion: the history of optical telescopes

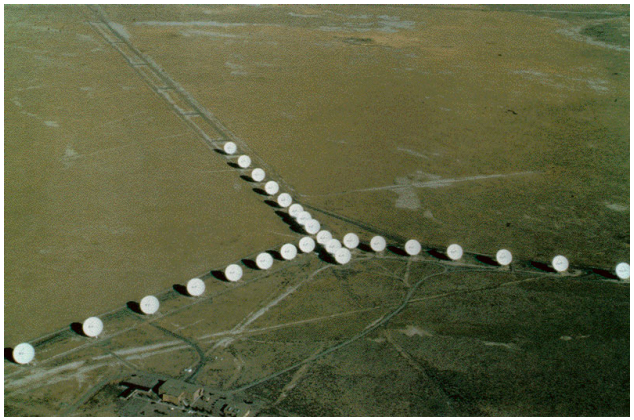
Resolution varies by aperture diameter (bigger is better)

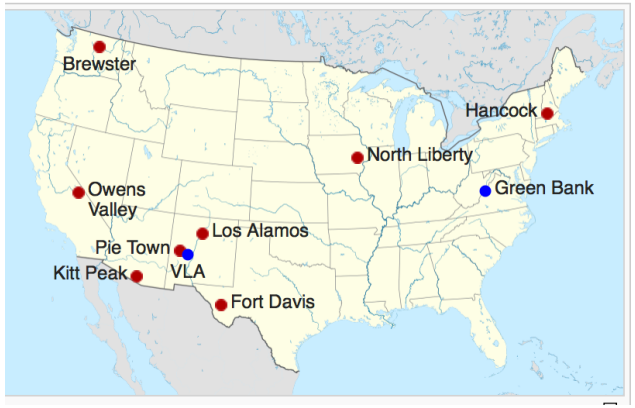


A short diversion: history of radio telescopes

Resolution varies by aperture diameter (bigger is better)

Aperture can be synthetically increased across multiple telescopes or even multiple observatories





Can we increase N and n at the same time?

1. Frederic Lord (1955) introduced the concept of sampling people as well as items.
2. Apply basic sampling theory to include not just people (well known) but also to sample items within a domain (less well known).
3. Basic principle of Item Response Theory and tailored tests.
4. Used by Educational Testing Service (ETS) to pilot items.
5. Used by Programme for International Student Assessment (PISA) in incomplete block design (Anderson, Lin, Treagust, Ross & Yore, 2007).
6. Can we use this procedure for the study of individual differences without being a large company?
7. Yes, apply the techniques of radio astronomy to combine measures synthetically and take advantage of the web.

Subjects are expensive, so are items

1. In a survey such as Amazon's Mechanical Turk (MTURK), we need to pay by the person and by the item.
2. Why give each person the same items? Sample items, as we sample people.
3. Increasing the number of items allows analysis analogous to what is done in genetics with the GWAS analysis of SNPs.
   - Can find the profile correlations of groups across many items.
4. Synthetically combine data across subjects and across items. This will imply a missing data structure which is
   - Missing Completely At Random (MCAR), or even more descriptively:
   - Massively Missing Completely at Random (MMCAR)
5. This is the essence of Synthetic Aperture Personality Assessment (SAPA).

3 Methods of collecting 256 subject * items data

a) 8 x 32 complete

```
46213634521143453443645331212414
21243623166421516154432261516513
51661351155165463622224435623344
11141343362332215612152135614522
25353121264561433433232246526411
61335154566424114612641225353516
24634342151536242425413513435116
11554654453123111162423325516334
```

b) 32 x 8 complete

```
46323114
25443314
43315423
26314145
41435614
42236153
62421344
35234443
34514166
63415154
44441342
13514321
66365663
12264546
31466135
32645514
66151251
14411441
62443636
33316236
63325425
11531126
61155546
33245361
52241654
63212356
24414663
63661414
45555223
14364433
21461416
33232365
```

c) 32 x 32 MCAR p=.25

```
..3..2..6.....4.55.......44.....
..........4..6..45..3.4..6....1
6..3.......6.1.....6.2.......5.6
....3522......5.3...3......5....
....3.2.2.......3..2......65..5.
.....51....324.........23......5
....552............25...54.5....
...44.4.5....3..6...6........3..
....61.523.2....2...........3...
5.............42.4..6.5......61.
....3....3.6..1.4...1..5......5.
1....54..........2.4.33..6......
4.....52..6.....44.3...........2
..44...1........1..42....5..1...
..1..3.......2..3.521.......6...
.......3.142.........22......12.
.4...2..........3..162...4.....4
..4..6..3.4...1....5.33.........
5..........243..5....41......1..
..5..3..4...4.4..5..1.........4.
.....4.......3..5.2.....64.4..4.
...1.1.2...6....4......55....2..
.....3..2..53.....2..2.3.3......
......1...2..43...3.13........5.
...2.........4..54...2.3..62....
22.......332..1.....5......6....
...5..3.4.....3....5.241........
......63.1.......6...5..4..2...5
..2.4..5..........52.4.....44...
2.55.....2.....6.....6.....55...
..5..........4....6341.4..2.....
....55......5.......45....3..32.
```

Synthetic Aperture Personality Assessment

1. Give each participant a random sample of *pn* items taken from a larger pool of n items.
2. Find covariances based upon "pairwise complete data".
3. Find scales based upon basic covariance algebra.
   - Let the raw data be the matrix $\boldsymbol{X}$ with N observations converted to deviation scores.
   - Then the item variance covariance matrix is $\boldsymbol{C} = \boldsymbol{X}\boldsymbol{X}'N^{-1}$
   - and scale scores, $\boldsymbol{S}$ are found by $\boldsymbol{S} = \boldsymbol{K}'\boldsymbol{X}$.
   - $\boldsymbol{K}$ is a keying matrix, with $\boldsymbol{K}_{ij} = 1$ if $item_i$ is to be scored in the positive direction for scale j, 0 if it is not to be scored, and -1 if it is to be scored in the negative direction.
   - In this case, the covariance between scales, $\boldsymbol{C}_s$, is

$$\boldsymbol{C}_s = \boldsymbol{K}'\boldsymbol{X}(\boldsymbol{K}'\boldsymbol{X})'N^{-1} = \boldsymbol{K}'\boldsymbol{X}\boldsymbol{X}'\boldsymbol{K}N^{-1} = \boldsymbol{K}'\boldsymbol{C}\boldsymbol{K}. \qquad (1)$$

4. That is, we can find the correlations/covariances between scales from the item covariances, not the raw items.

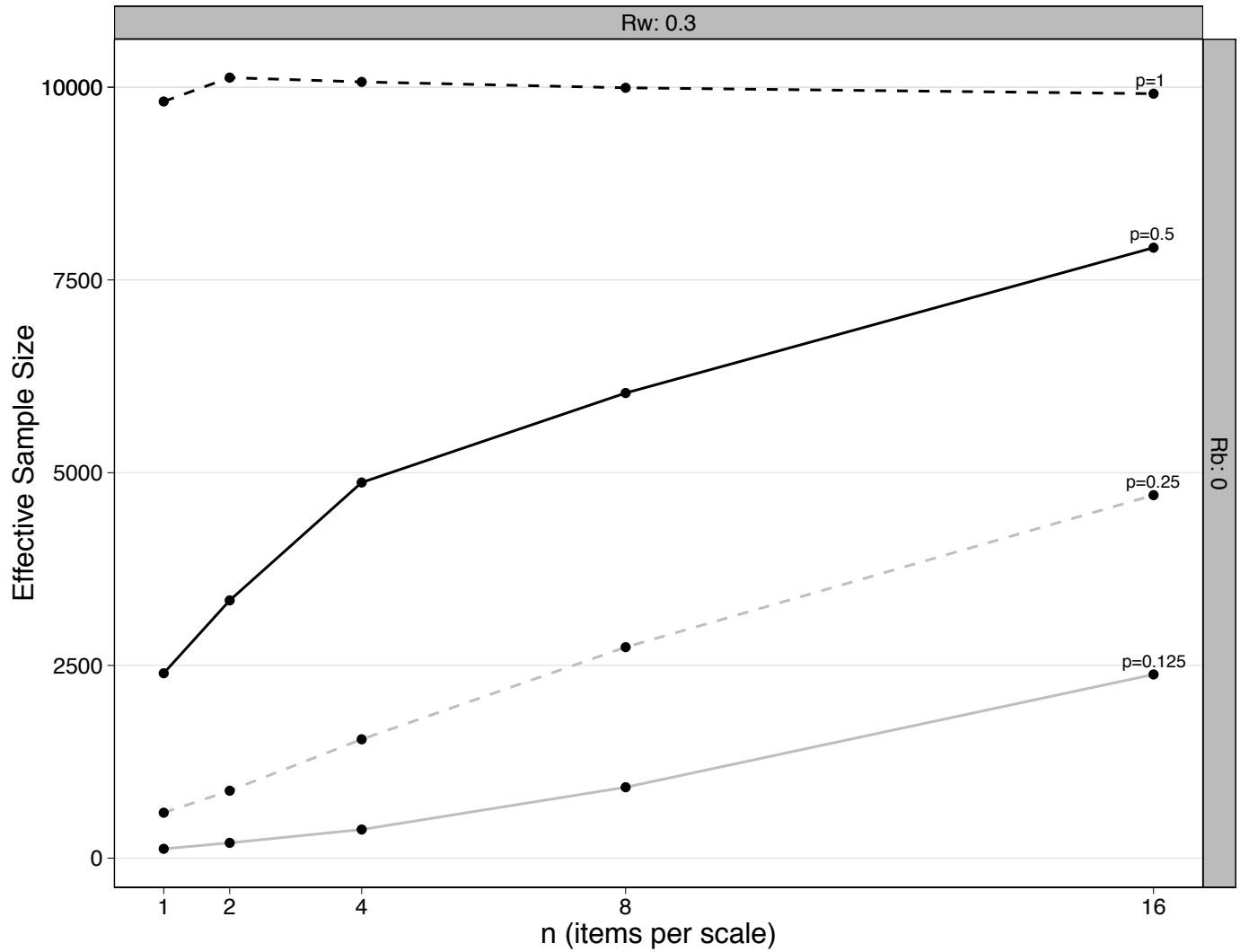
SAPA standard errors and effective sample size

1. When forming synthetic scales from MMCAR based items, the standard error of correlations decreases as a function of the Total number of subjects (N), the percentage of items samples (p), *and* the number of items forming the scale (n).
2. Ashley Brown has shown this quite clearly by simulation (Brown, 2014).
3. A good way to visualize this is to examine the standard error of correlations as a function of N, p, and n.
4. An even more dramatic way is to plot the *Effective Sample Size* (N_{eff}) which because

$$\sigma_r = \frac{1 - r^2}{\sqrt{N - 2}} \text{ is merely } N_{eff} = \frac{(1 - r^2)^2}{\sigma_r^2} + 2$$

Effective sample size varies by the size of the composite scale.

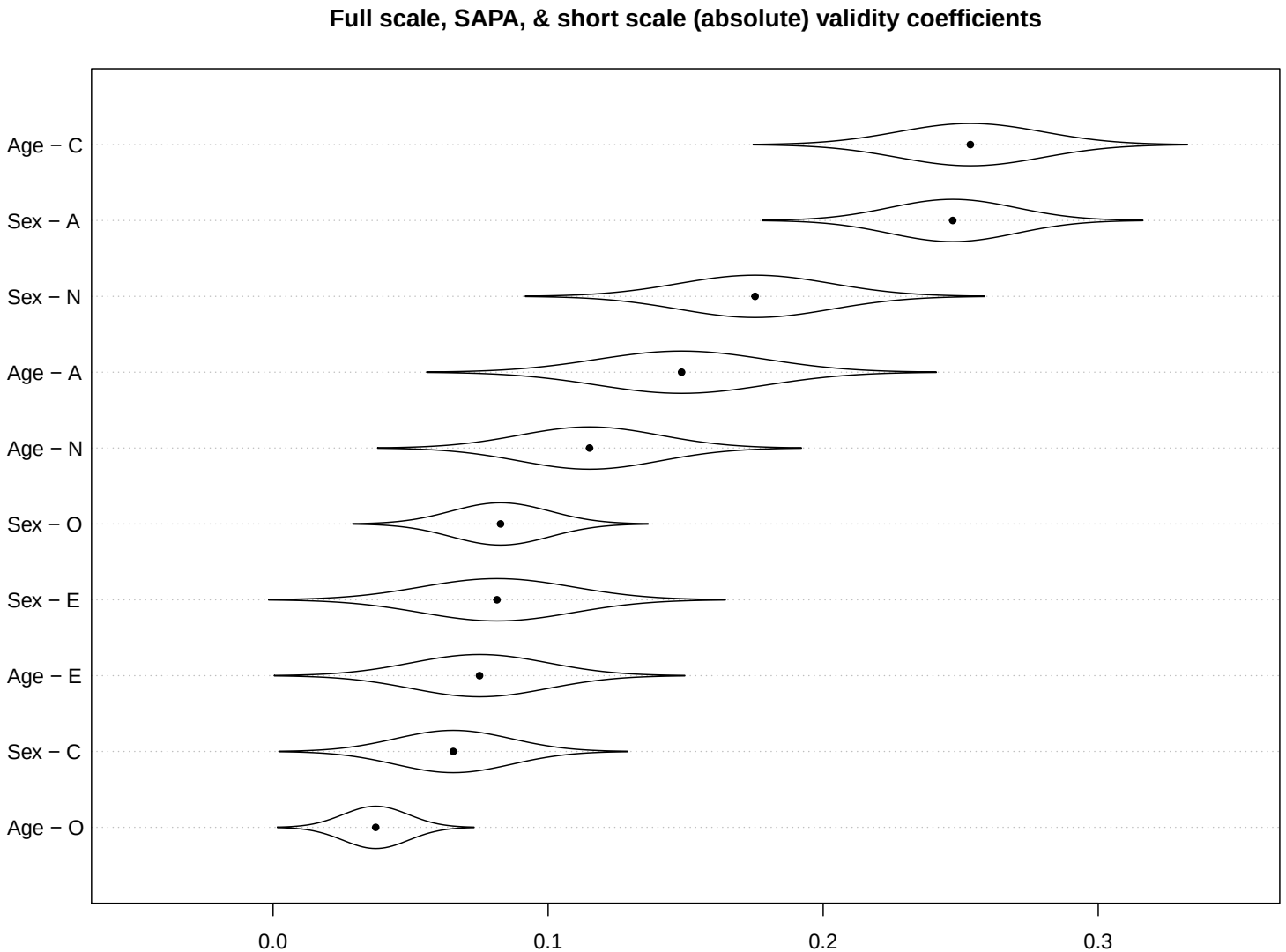
Simulating N= 10,000 with probability of any item ($p = .125, .25, .5$, or 1) and items in the composite 1 , 2, 4, 8, 16.



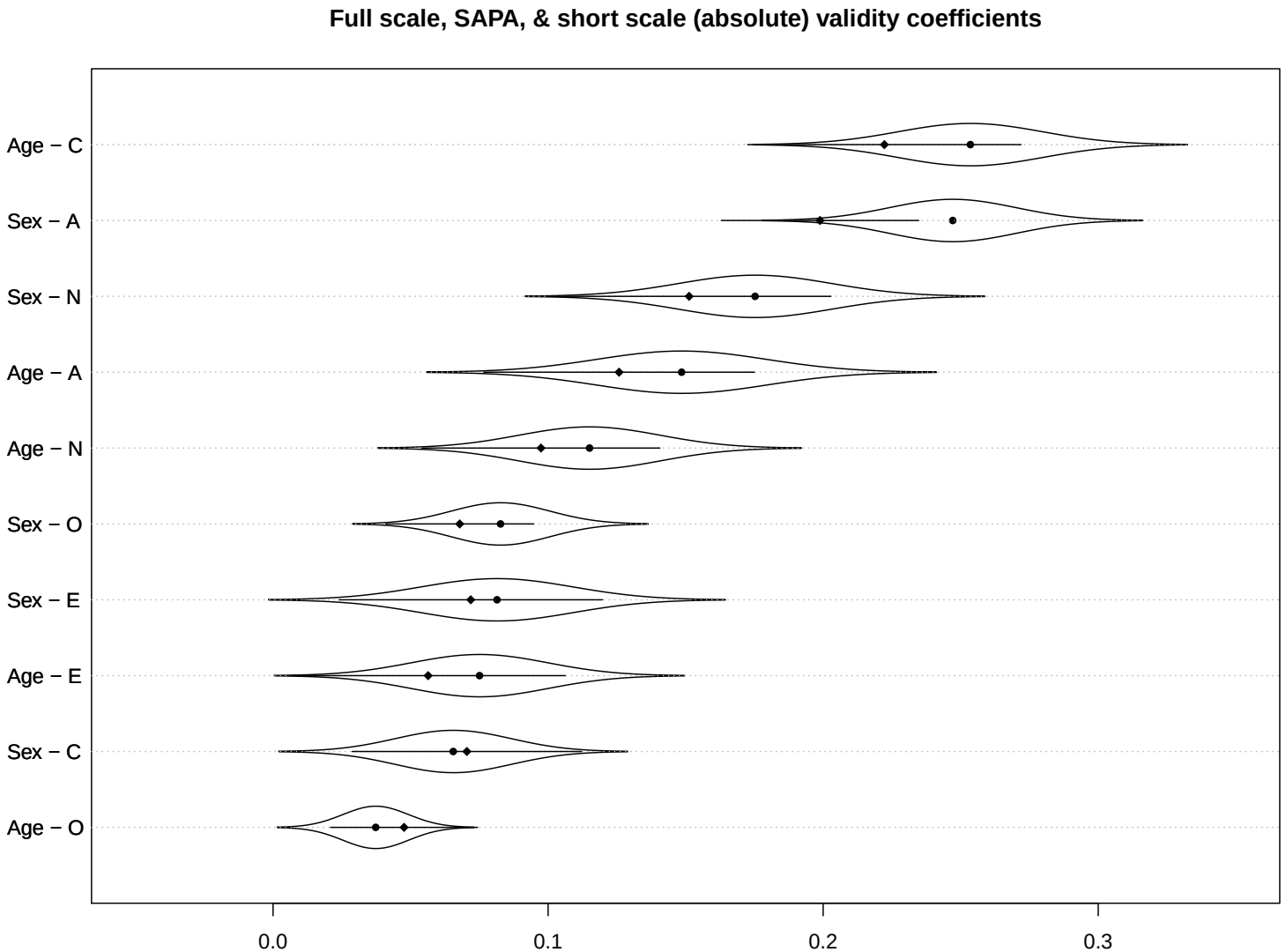
Factor structures can be recovered from SAPA sampling

1. With Sonja Heintz (of Zurich) and David Condon (of the University of Oregon) we have shown that the SAPA sampling techniques can recover factor structures for as low as 200-400 people.
2. In addressing the question of "how low can you go" for sample sizes, Sonja showed through simulated sampling of real data that 25% sampling of 120 items recovers the structure as well as does complete sampling with as few as 400 subjects.
3. Liz Dworak and Sonja also showed this worked when simulating mood data from a Ecological Momentary Sampling data set.
4. I show the sampling variation of validities for 20 samples from a complete data set versus 20 samples from a "SAPAized" data set showing the two agree almost perfectly, and are far superior to the alternative of using just a short form.

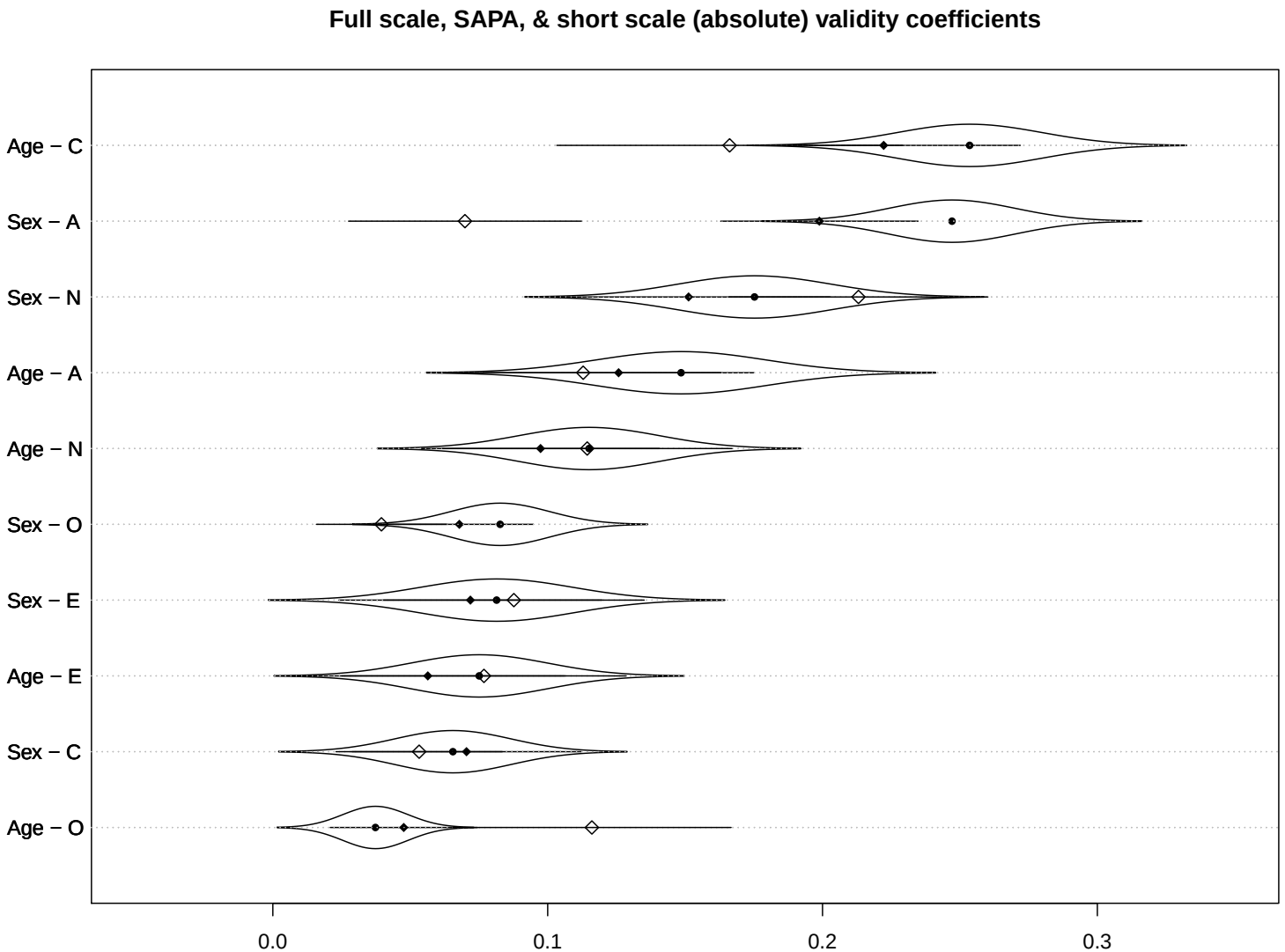
N=400 from Johnson 120 item IPIP, full scales: Validity



N=400 Johnson 120 item IPIP, sapa sampling 30 items: Validity



N=400 120 item IPIP, sapa 30 items + Short scales: Validity



SAPA is not magic: We can obtain high accuracy at the structure level but accuracy is much lower at the single subject level

1. Reliability of composite scales is high when formed from synthetic matrices $\boldsymbol{C_s} = \boldsymbol{K'CK}$ because the number of items per scale/per subject is the nominal amount.
2. Reliability of single scores is much less because very few items measuring a single trait are given to a single subject $\boldsymbol{S} = \boldsymbol{K'X}$.
3. However, the precision of the estimate of subject means ($\bar{x}$) is high because $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{Np-1}}$ and Np is large.
4. SAPA technique is very powerful for research of structure, but less powerful for research based upon single subjects.

How does it work?

1. Give our basic belief in open science, we use public domain items, open source software:
   - Apache webserver, MySQL data bases, PHP and HTML5 web tools, R for statistics.
   - Extensive coding in PHP and MySQL to present item sets in random fashion (Joshua Wilt, David Condon, Jason French)
   - Code written for psychometric measurement and scale construction as implemented in the *psych* package (Revelle, 2018) using R (R Core Team, 2018)
2. Domains measured and item sources
   - Temperament items taken from International Personality Item Pool (IPIP) (Goldberg, 1999) (`ipip.ori.org`) and supplemented with other items.
   - Ability items have been validated (Condon & Revelle, 2014) as part of the International Cognitive Ability Resource Project (`ICAR-project.org`). (ICAR:Ability::IPIP:Temperament)
   - Interest items taken from Oregon Vocational Interest Survey (ORVIS) (Pozzebon et al., 2010)

SAPA overview

1. A "Personality Test" is included as a resource at the `http://personality-project.org` and gives feedback to all participants.
2. Some participants then link their feedback to their social media sites which then appeals to yet more to take it.
3. Some professors assign it to their students in various classes.
4. About 200-600 people per day from around the world visit the `personality-project.org` or `sapa-project.org` websites. This does not sound like much, but over a year, we get around 16,000 participants/month and 452,000 this year. (50K in one day due to news coverage in Washington Post). 838,5132 since 8/18/2010 with another $\approx$100K from 2006-2010.

The SAPA Project

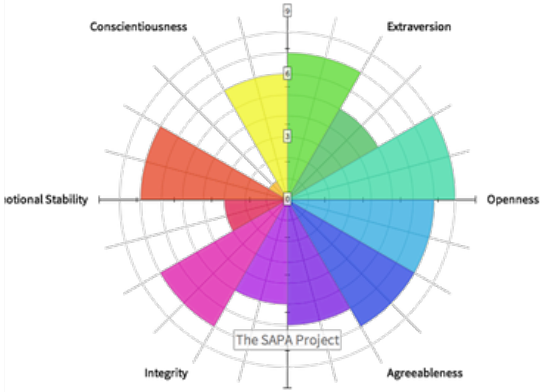
Take the test.

Explore your personality.

Advance the study of individual differences.

Start the test

More info

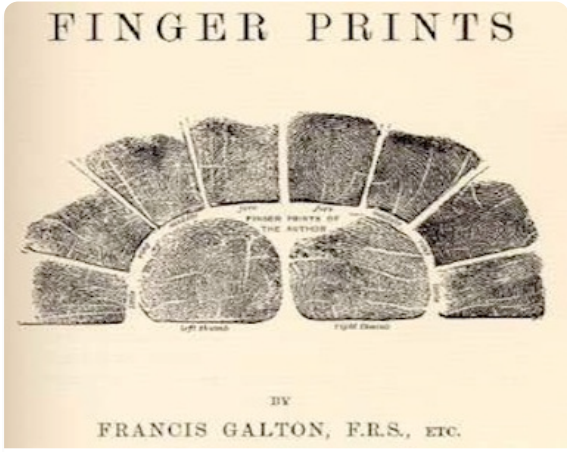


FAQ about the test

Is it long? (not really) Is it free? (yes)

The research behind SAPA

How was the test developed?



Individual Differences

Learn more about differential psychology.

How does it work?: part II

1. Participants find us by searching web for “personality tests”, etc. and find `personality-project.org` or `sapa-project.org`
2. Each participant is given a number of web pages
   - **Consent Form** Basic description of project and question whether they have taken test before.
   - **Demographics** Age, sex, height, weight, education, parental education, country, state, ZipCode (if US), ...
   - **TAIC questions** Temperament/Ability/Interest questions (25 per page, 21 T/I, 4 Ability per page
   - **Continuation pages** After each page, told that feedback will be more accurate if they keep going.
   - **Optional modules** Creativity, Peer ratings, interests, ...
   - **Feedback** Personality feedback based upon scores on temperament items.
3. Results are stored (page by page) on the MySQL server.

How does it work: part III

1. Various data cleaning scripts run using the *SAPA-tools* package (French & Condon, 2015) in R.
   - Screen for duplicate responses based upon a Random Identification Number issued when subjects start the page. We drop all subsequent pages.
   - Screen for subjects < 14 or > 90.
2. Subsequent analyses are done primarily using functions in the *psych* package (Revelle, 2018) for R.
3. Analyses are done at multiple levels:
   - 3.1 At the item and scale covariance level to examine the structure of items
   - 3.2 At the multiple levels of aggregation: zip code, state, college major, occupation. This requires finding individual level scores and then examining the structure of group means through basic multi-level techniques.

Part IV: Open Data

Four types of openness:

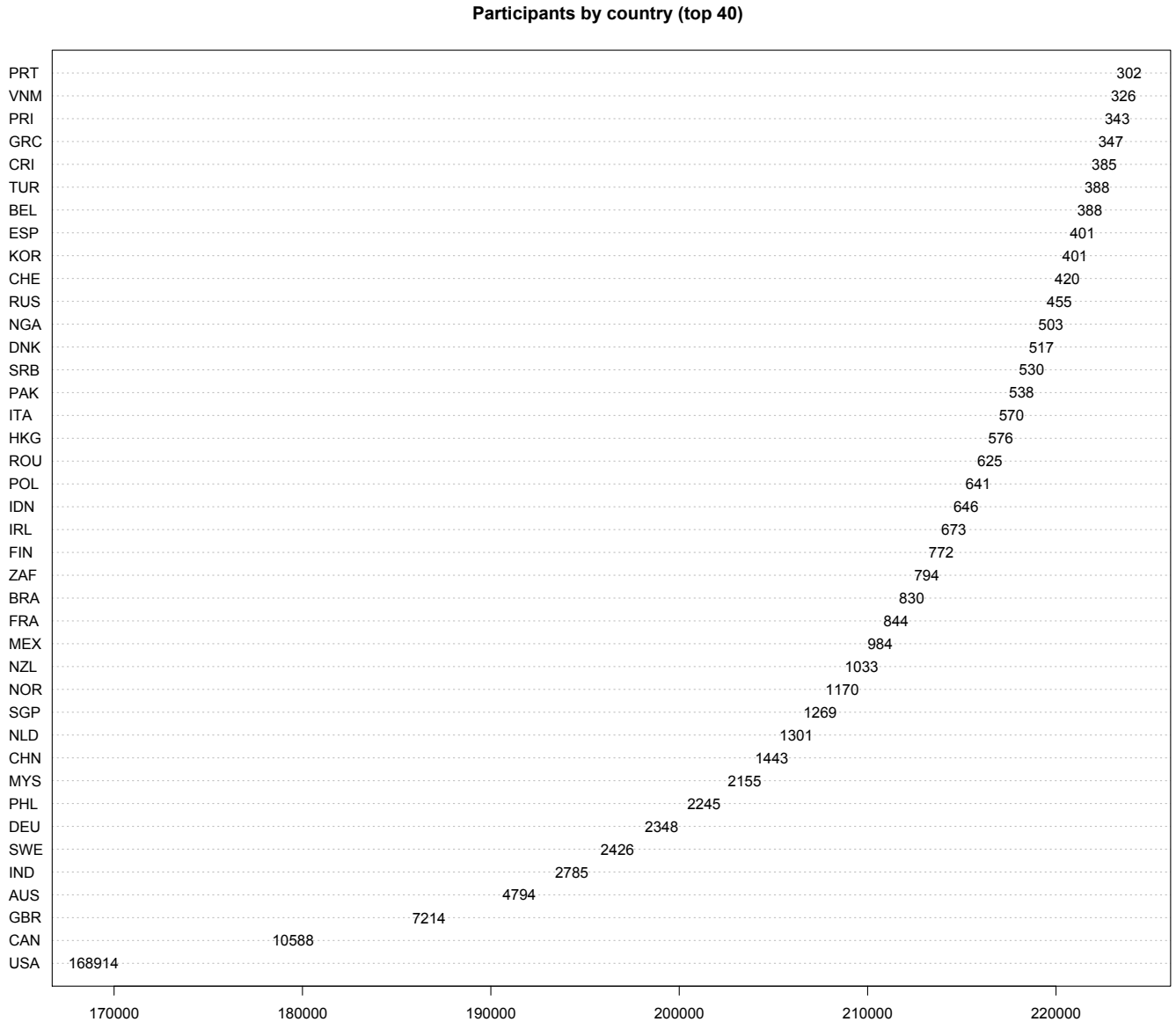
1. Open source software: The R project (R Core Team, 2018)
2. Open source materials:
   - The International Personality Item Pool (IPIP) (Goldberg, 1999)
   - The International Cognitive Ability Resource (ICAR) (Condon & Revelle, 2014)
   - The Synthetic Aperture Personality Assesment (SAPA) database (Condon, 2017)
3. Open source methodology: The Synthetic Aperture Personality Assessment Project (Revelle et al., 2010, 2016)
4. Open source data:
   - Data from the ICAR project (Condon & Revelle, 2016, 2015a)
   - Data from SAPA studies (Condon & Revelle, 2015c,b)

In the process of summarizing the last several years of research, I will show how we use open source software, items, and methods and then share them with the world.

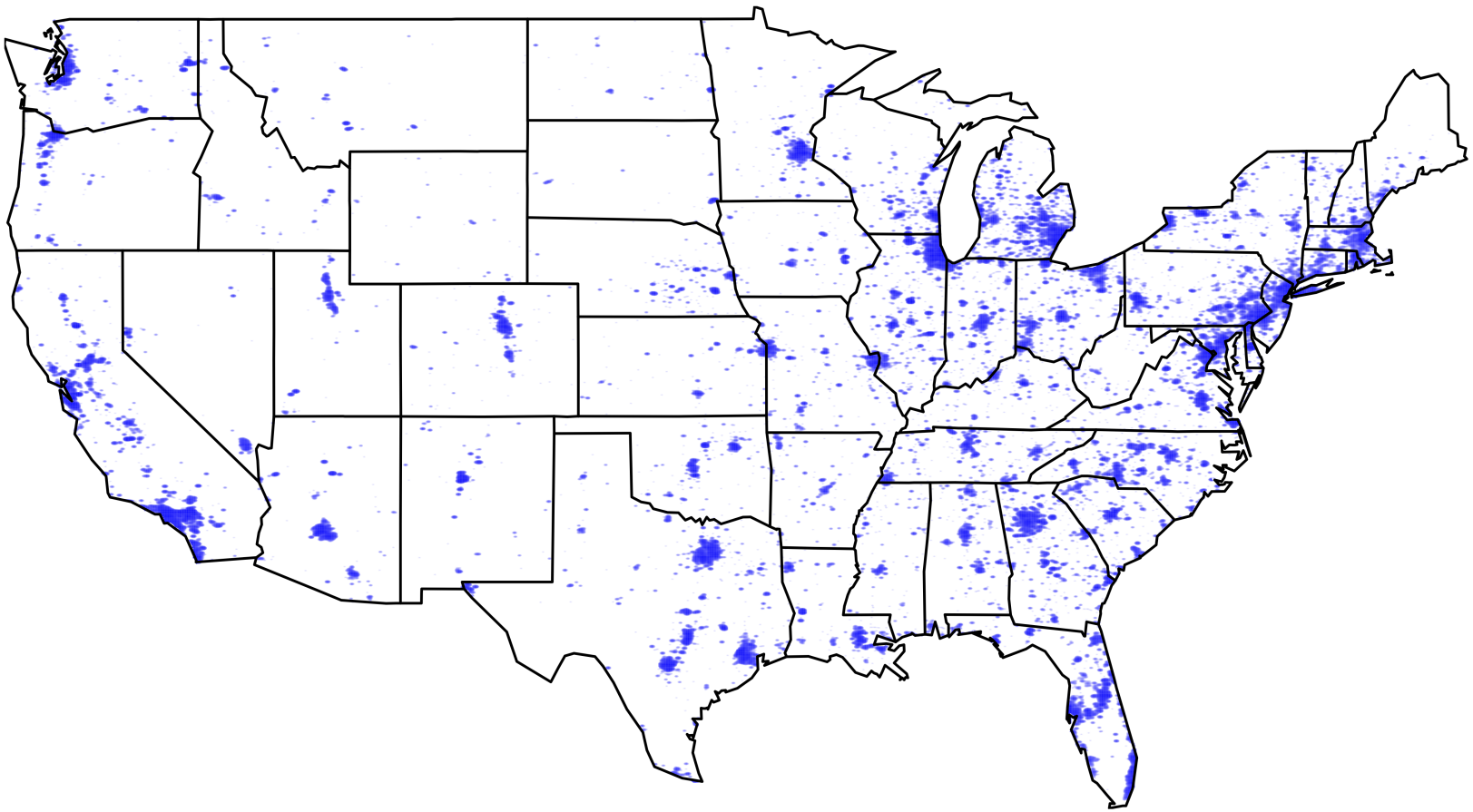
Our data to be discussed today

Time Frame	Data collected at `personality-project.org` and `sapa-project.org` from August 18, 2010 to February, 2017
Subjects	N = 255,348 (77,550 males, 129,451 females)
Materials	953 items (696 temperament, 60 ability, 152 interests, 45 demographic)
Scales used	27 Temperament, 4 Ability, 6 Interests
N in workforce	N = 97,782
N students	N = 102,638
Occupations	973 separate occupations, following a Pareto distribution with $\approx$ 80% represented by the top 20% of occupations
$N \geq 100$	192 occupations for 71,298 participants
$N \geq 100$	120 college majors for 130,584
$N \geq 300$	40 countries for 223,884

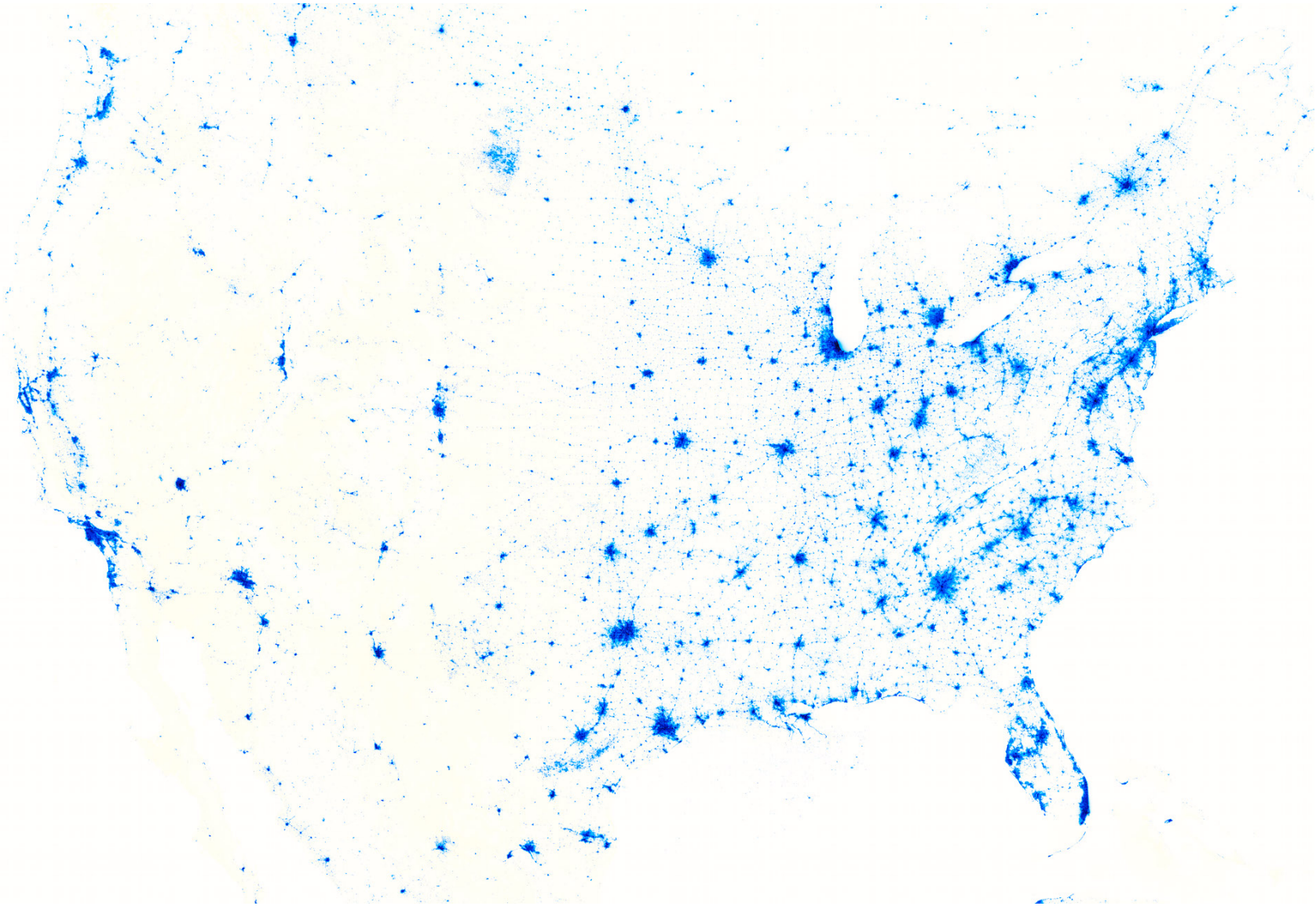
The top 40 countries account for 88% of the sample



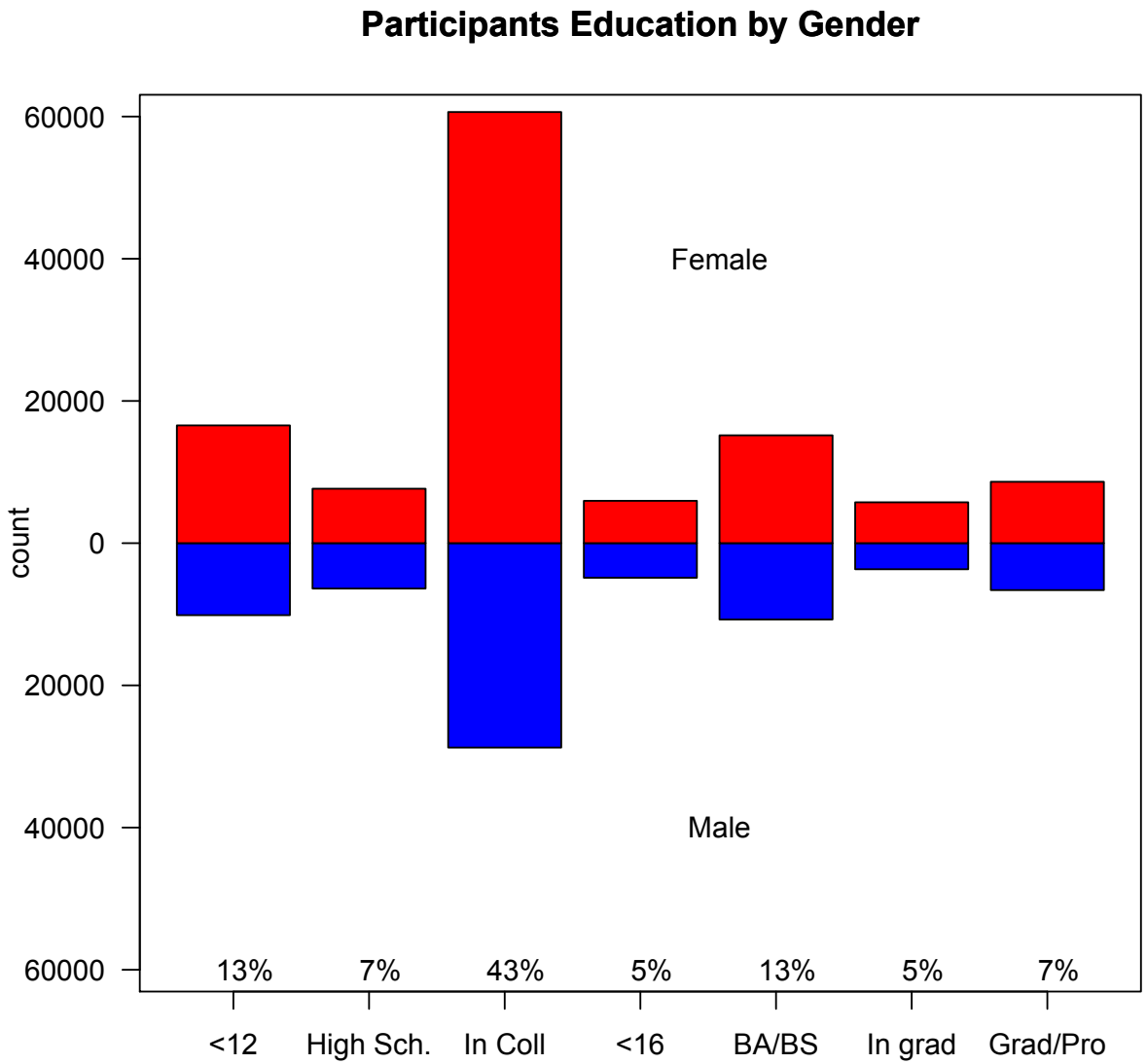
Where do they come from? US SAPA data by zipcodes



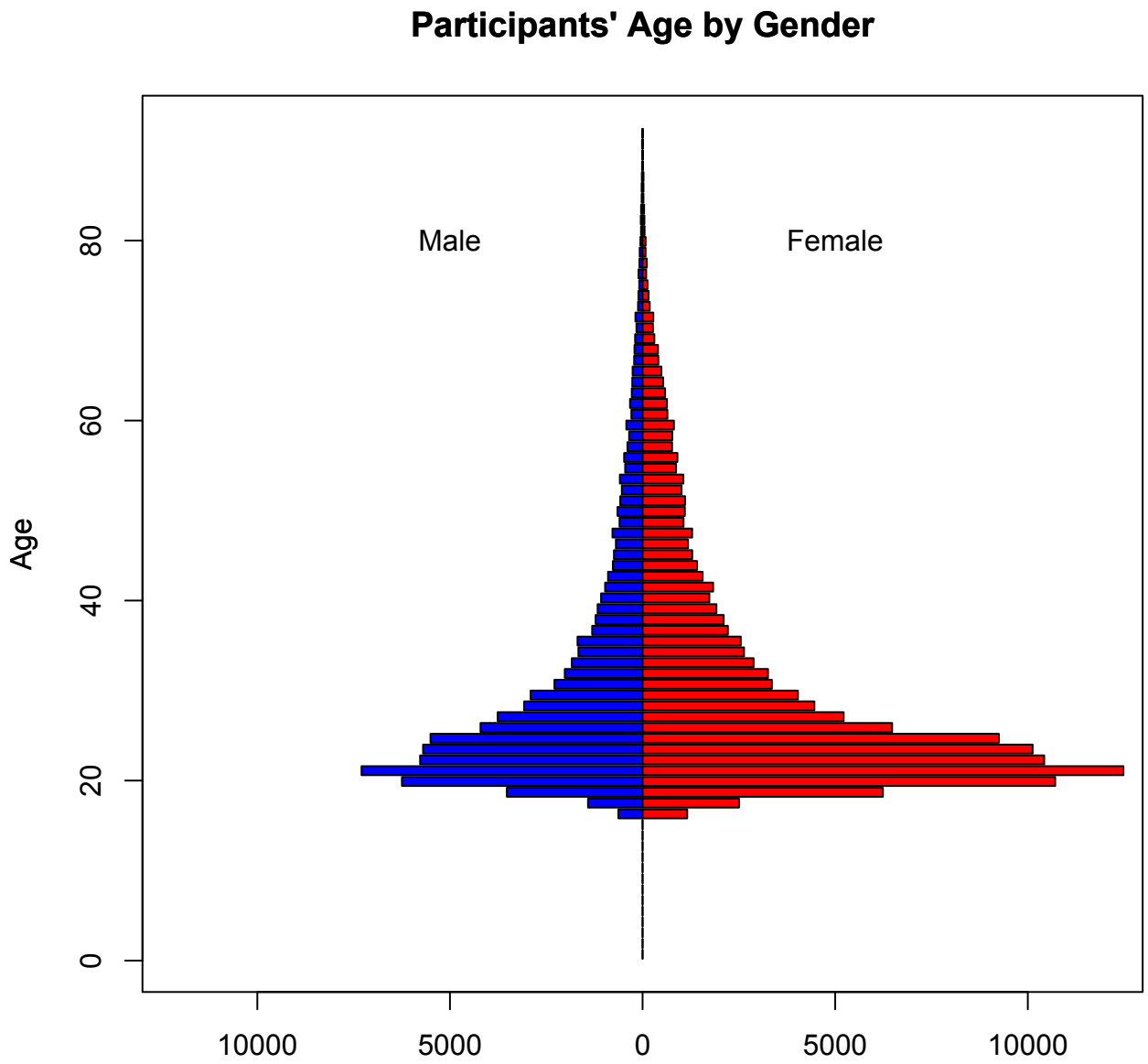
US lights (from NASA)



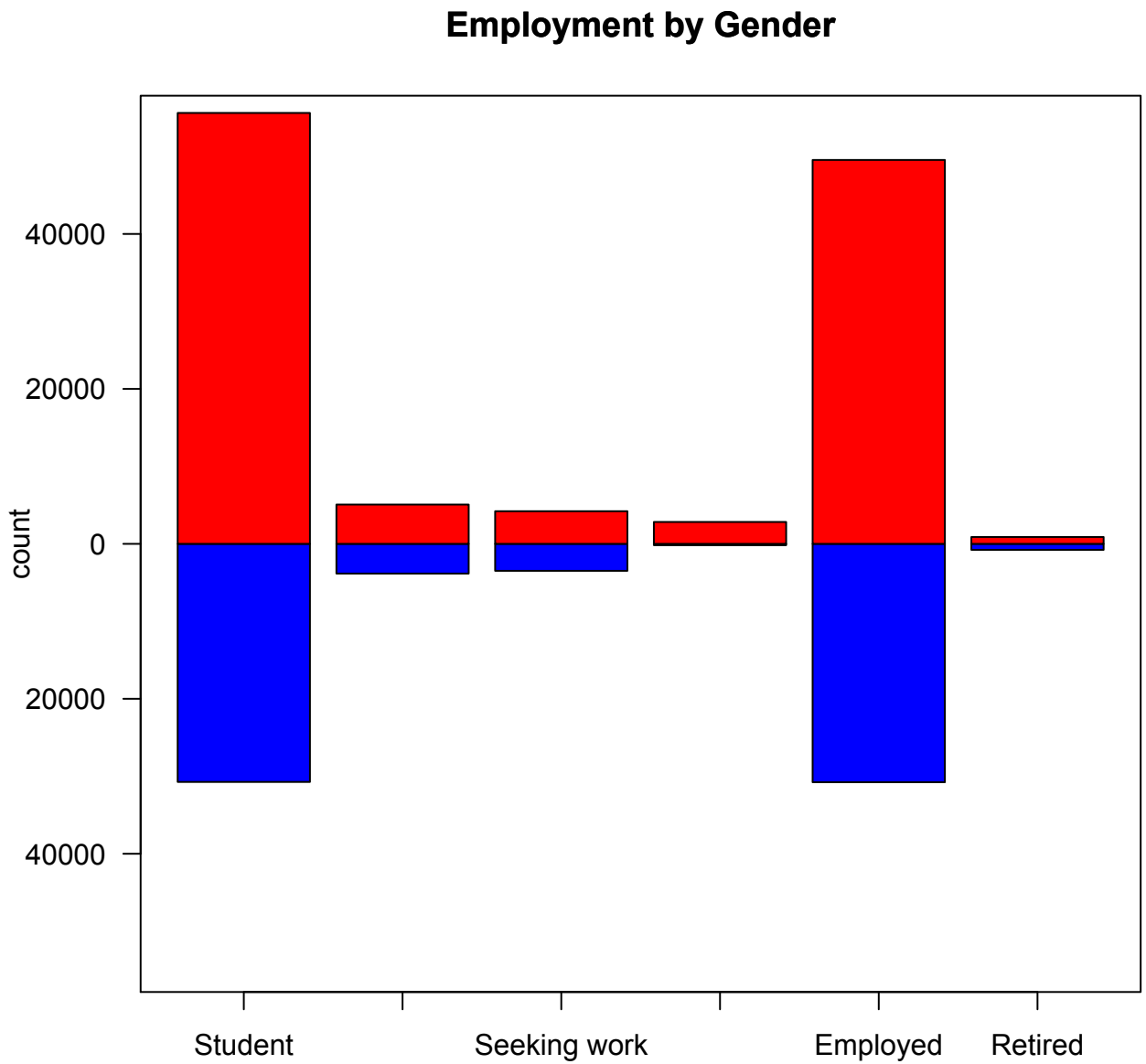
Not a random sample of either education or gender (62% female)



Not a random sample of either age or gender (62% female)



102,638 students, 97,782 employed)



Are the "Big Five" really big?

1. There is a "consensus" about the proper number of factors/components of personality (Goldberg, 1990, 1992; Hofstee et al., 1992).
2. This seems to match life challenges of Getting Along and Getting Ahead

   Conscientiousness Work
   Agreeableness Love
   Neuroticism Effective functioning in many domains
   Openness/Intellect Play
   Extraversion Leadership
3. Additional work has been done on the same 800-1000 person Eugene-Springfield sample and suggests a hierarchical structure (DeYoung, Quilty & Peterson, 2007).
4. But what happens with a larger and different sample?

Multiple solutions to the dimensionality of temperament

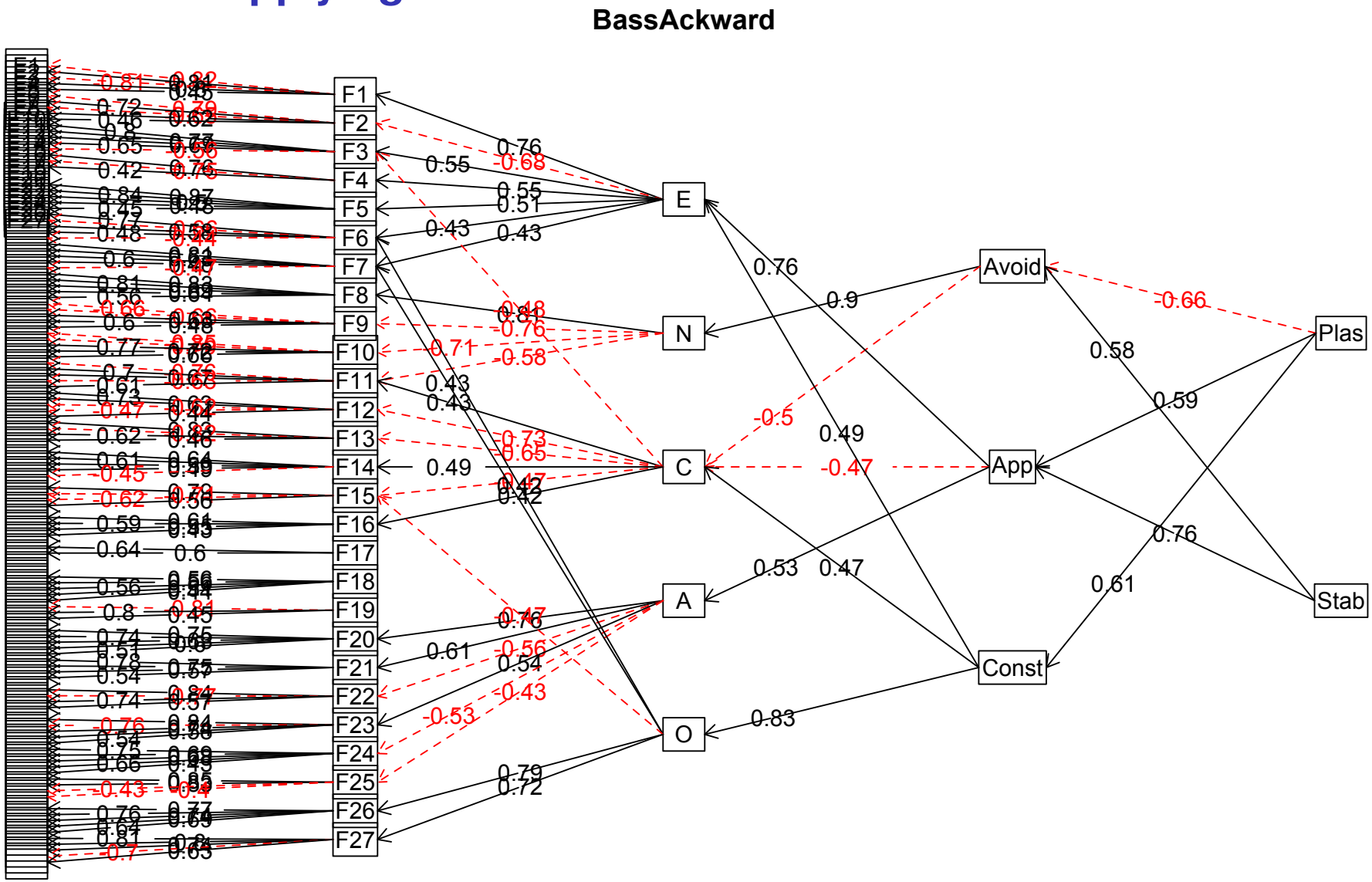
1. Digman alpha and beta (Digman, 1997), DeYoung stability and plasticity (DeYoung, Peterson & Higgins, 2002)
2. Eysenck "Giant 3" (Eysenck, 1994)
3. The "Big 5" (Digman, 1990; Goldberg, 1990)
4. The HEXACO 6 (Lee & Ashton, 2004; Ashton, Lee & Goldberg, 2007)
5. Tellegen 7-9 (Tellegen & Waller, 2008)
6. Comrey 8-9 (Comrey, 2008)
7. Cattell 16 Personality Factors (Cattell, 1957)
8. Condon (2014, 2015) examined 696 non-overlapping items from IPIP:100, IPIP:NEO, IPIP:MSQ, BFAS, EPQ, etc. (Goldberg, 1999; DeYoung et al., 2007; Eysenck et al., 1985)

   Found meaningful 3, 5, and 27 factor solutions.
9. The Condon 3/5/27 form a heterarchical and non hierarchical structure (i.e., lower levels are not cleanly nested in higher levels.)

Personality shows a heterarchical even fractal structure

1. David Condon (2014, 2015) and in prep has shown:
2. The structure of 696 personality items given to 100-200,000 participants does not show a clean organization.
3. The number of factors problems (and its non-solutions) will break the heart of most investigators.
4. No clear structure at any level.
5. Possible to show relationships across different number of factors (note, this is not hierarchical factoring, but merely showing the correlations across different solutions.)

Applying the 'Bass Ackward' function

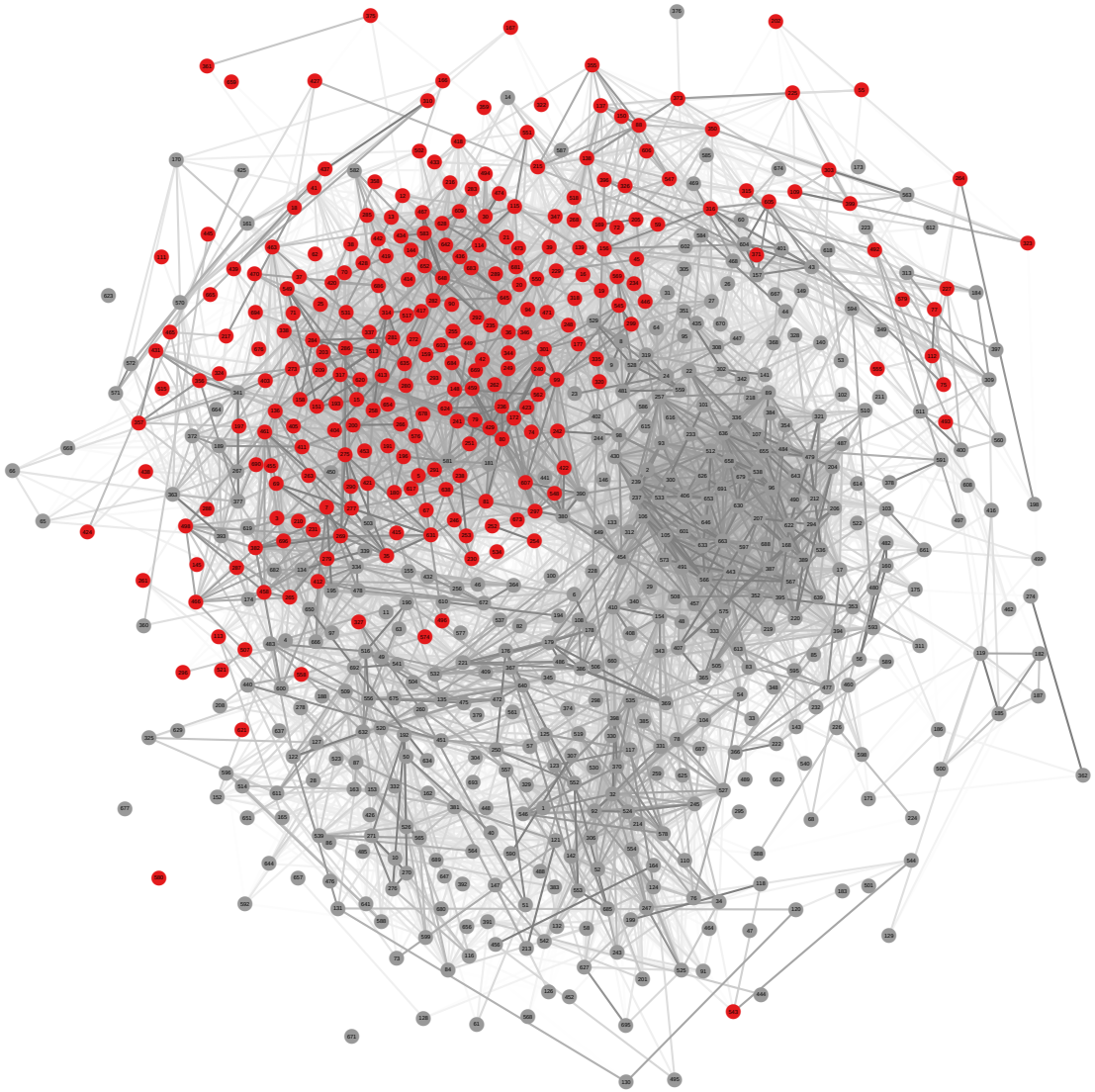


Sample items from each of the SPI 27

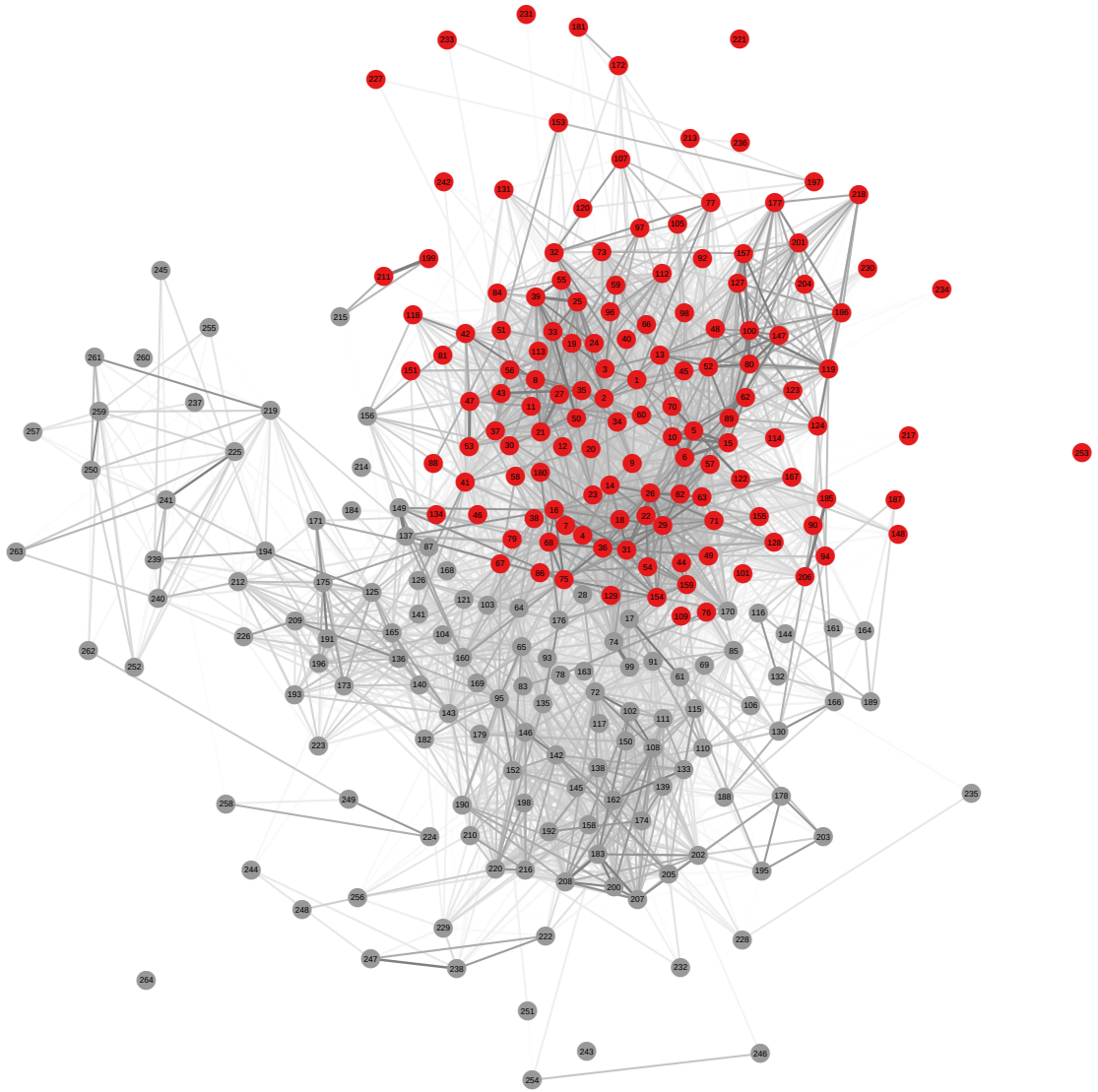
Each scale has 5 items

SPI	Item	Item
Compassion	Am sensitive to the needs of others.	Am concerned about others.
Irritability	Get angry easily.	Lose my temper.
Sociability	Usually like to spend my free time with people.	Avoid company.
WellBeing	Dislike myself.	Feel a sense of worthlessness or hopelessnes
SensationSeeking	Love dangerous situations.	Seek danger.
Anxiety	Worry about things.	Would call myself a nervous person.
Honesty	Tell a lot of lies.	Tell the truth.
Industry	Find it difficult to get down to work.	Start tasks right away.
Intellect	Learn things slowly.	Am quick to understand things.
Creativity	Am full of ideas.	Am able to come up with new and different ide
Impulsivity	Act without thinking.	Make rash decisions.
AttentionSeeking	Make myself the center of attention.	Like to attract attention.
Order	Keep things tidy.	Leave a mess in my room.
Authoritarianism	Believe that laws should be strictly enforced.	Respect authority.
Charisma	Am skilled in handling social situations.	Find it difficult to approach others.
Trust	Trust what people say.	Trust people to mainly tell the truth.
Humor	Laugh a lot.	Laugh aloud.
EmotionalExpressiveness	Am open about my feelings.	Have difficulty expressing my feelings.
ArtAppreciation	Do not enjoy going to art museums.	Believe in the importance of art.
Introspection	Love to reflect on things.	Spend time reflecting on things.
Perfectionism	Dislike imperfect work.	Want every detail taken care of.
SelfControl	Never splurge.	Rarely overindulge.
Conformity	Like to be thought of as a normal kind of person.	Would hate to be considered odd or strange.
Adaptability	Dislike changes.	Dont like the idea of change.
EasyGoingness	Like to take it easy.	Like a leisurely lifestyle.
EmotionalStability	My moods don't change more than most people	Experience very few emotional highs and lows
Conservatism	Tend to vote for conservative political candidates.	Don't consider myself religious.

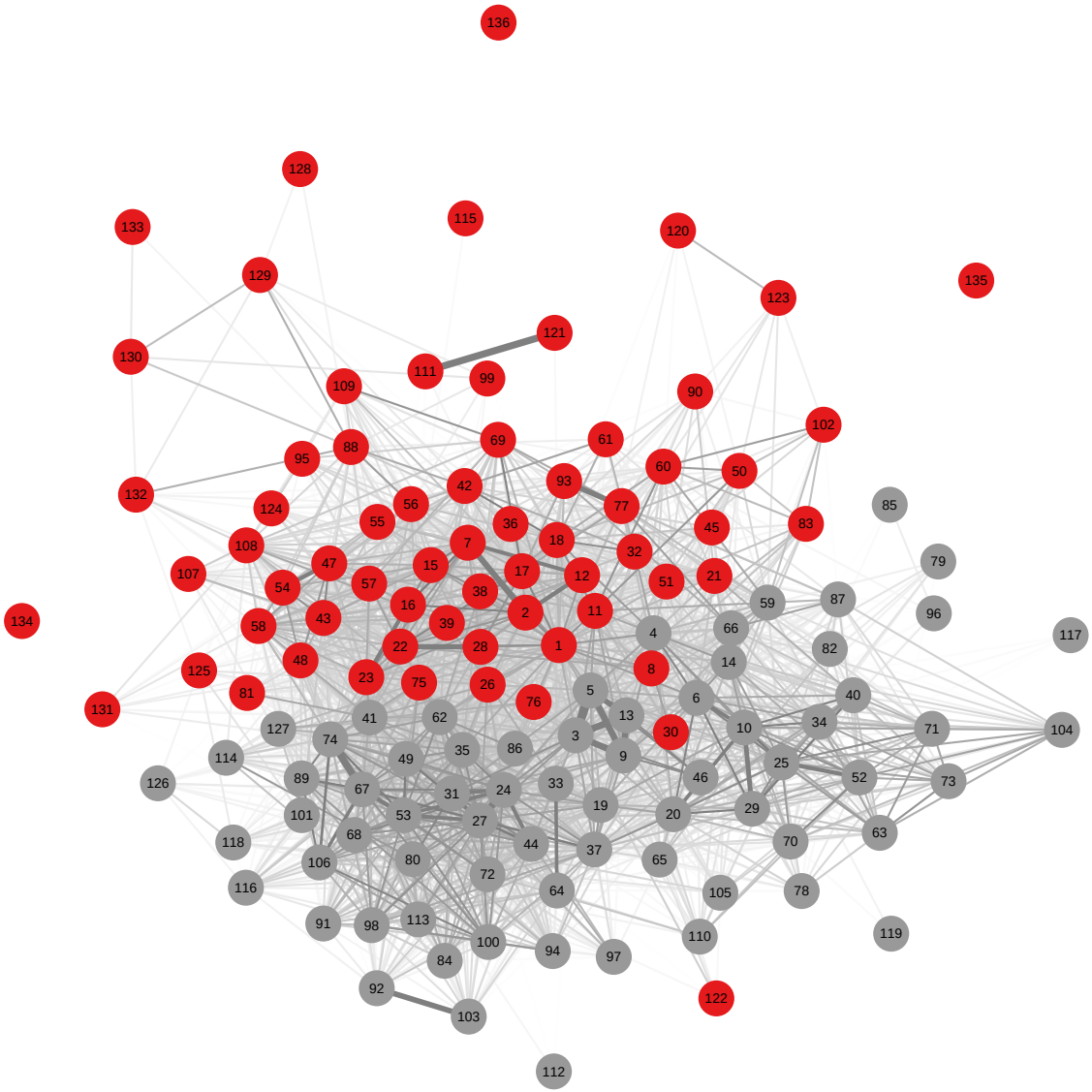
The correlations of 696 personality items



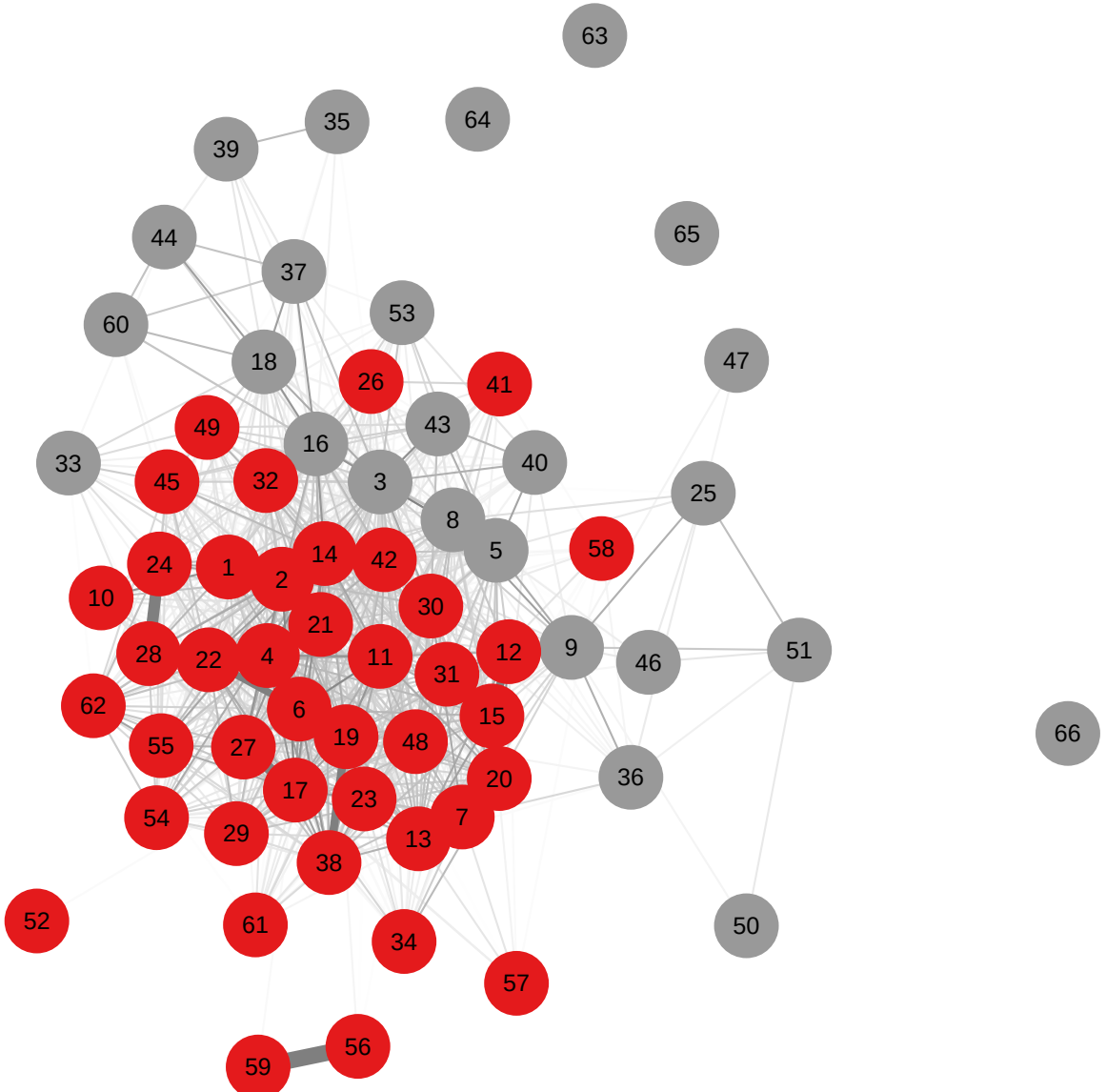
Factoring the items on the first factor of the 696



And doing it again

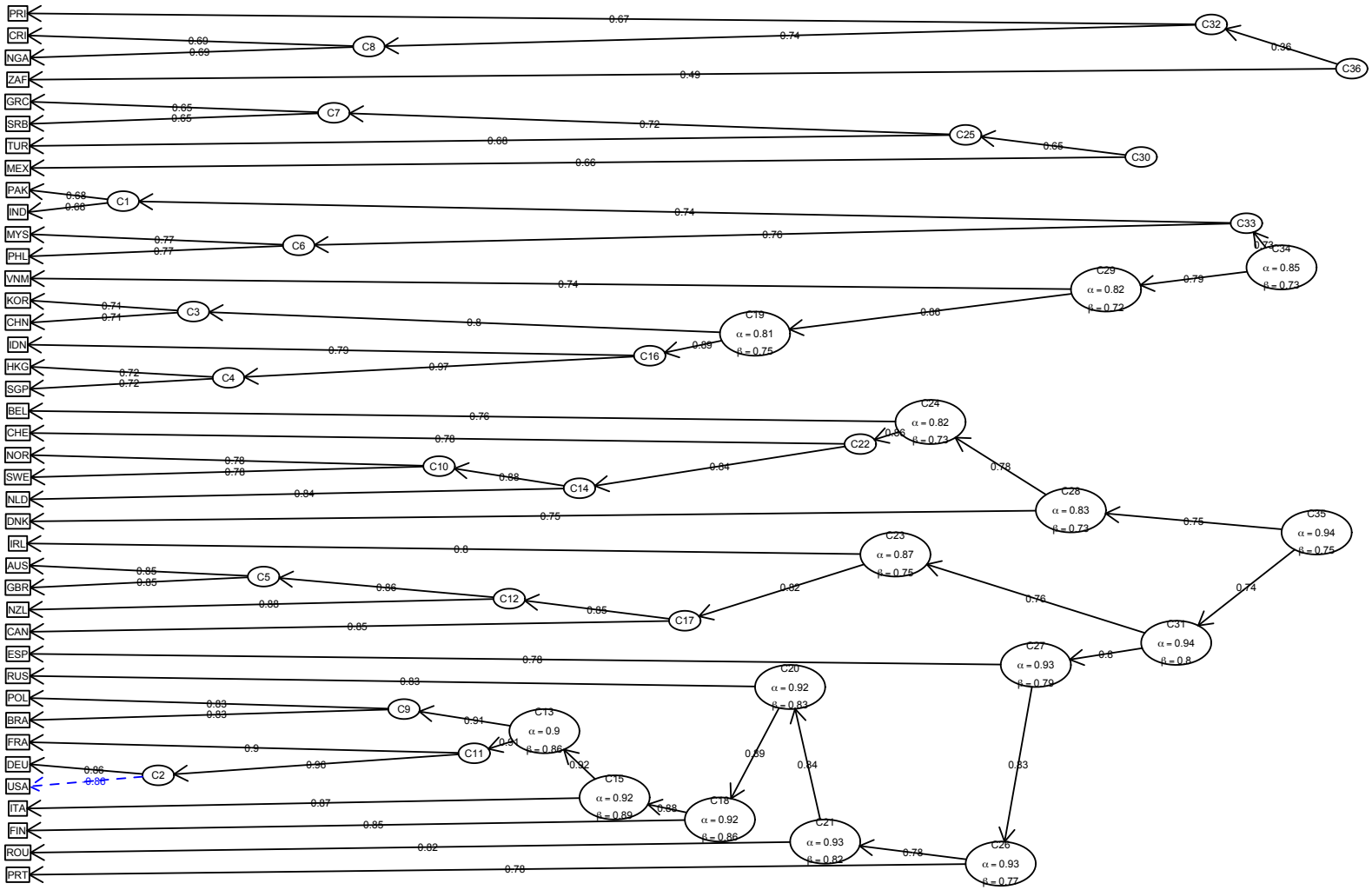


And again


63
35
64
39
44
65
37
53
60
18
47
26
41
49
43
16
33
40
45
32
3
25
8
5
58
24
1
14
42
10
2
30
21
9
46
51
28
22
4
11
12
31
62
6
15
66
55
27
19
48
36
20
17
23
54
29
7
13
38
52
61
50
34
57
56
59

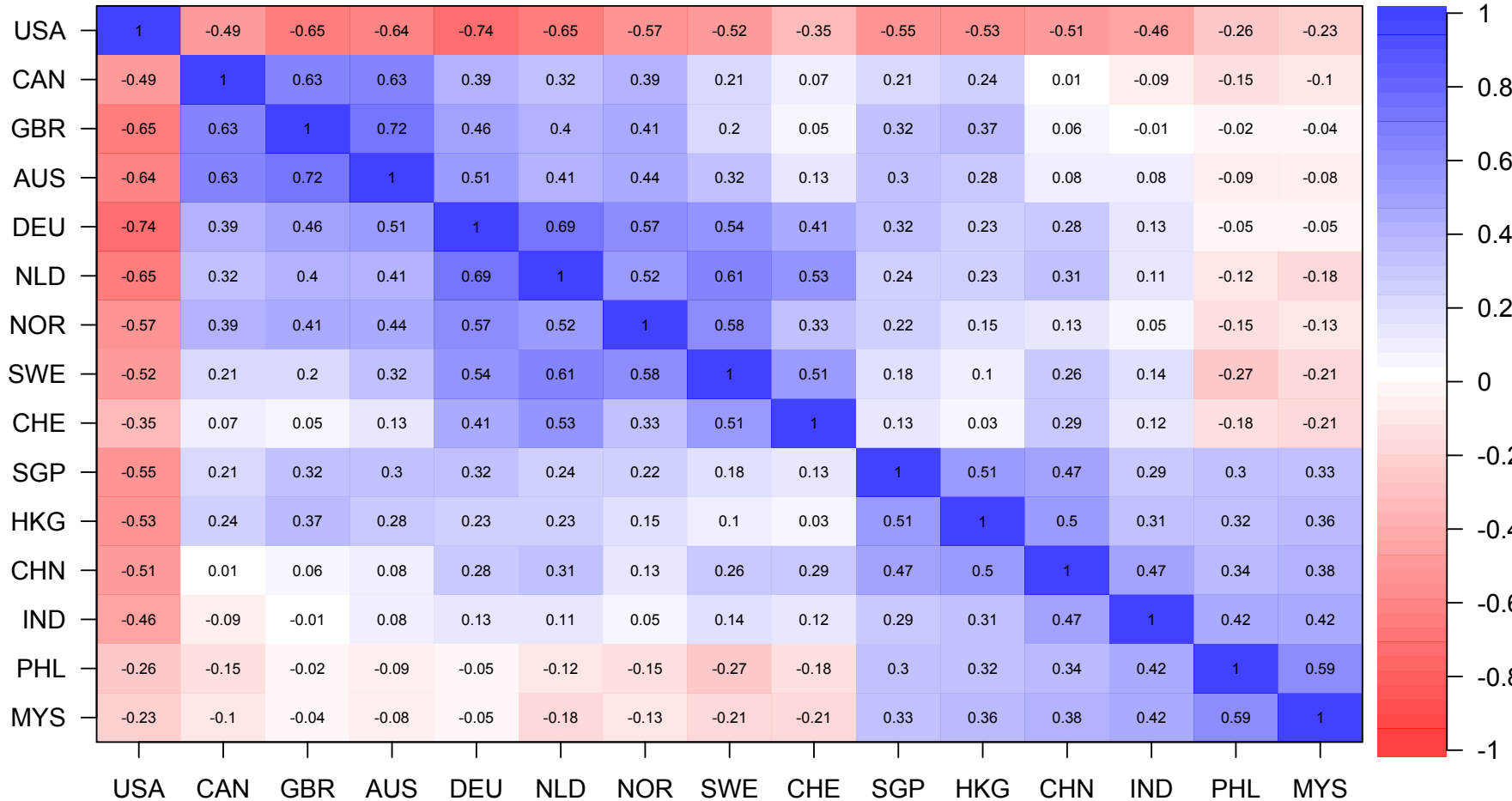
40 countries across 908 items

ICLUST of 40 country profiles across 908 items



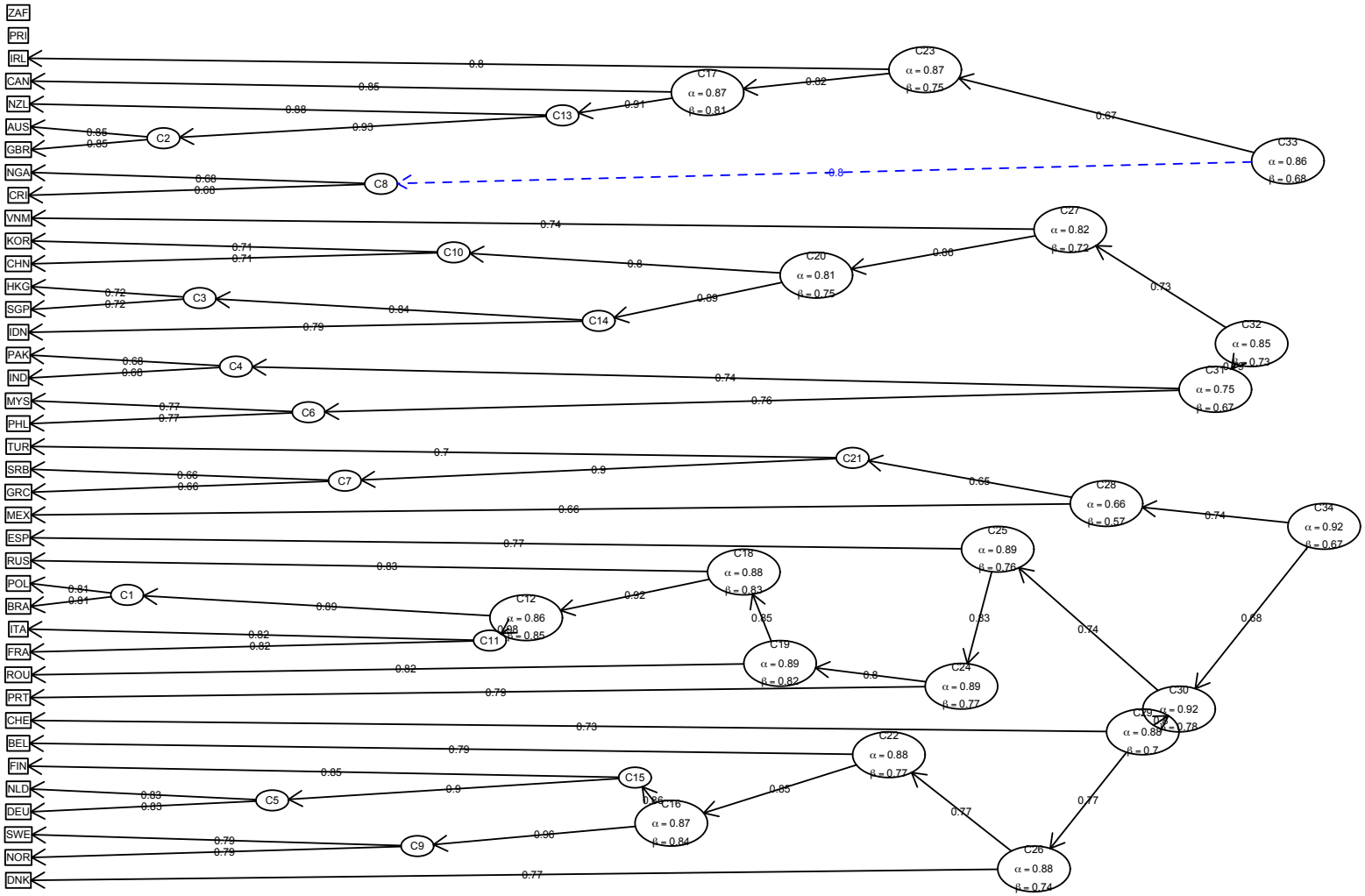
40 countries across 908 items

Profile correlations across 908 items for selected countries



Drop the USA

ICLUST of 39 country profiles across 908 items



Best Items correlating with being from Switzerland

Table: Top correlations with being from CHE

Top items		
CHE	item	itm_s
0.05	ICAR	VRiq10
0.05	Would like to play a musical instrument.	Artistic - ONETshort
0.05	Would like to keep inventory records.	Conventional - ONETshort
-0.05	Would like to paint sets for plays.	Artistic - ONETshort
0.05	Would like to teach a high-school class.	Social - ONETshort
-0.04	Would like to operate a calculator.	Conventional - ONETshort
-0.04	Like to stand during the national anthem.	IPIP
0.04	Would like to buy and sell stocks and bonds.	Enterprising - ONETshort
0.04	ICAR	R3Diq7
0.04	Would like to develop a spreadsheet using computer .	Conventional - ONETshort
0.04	ICAR	VRiq1
0.04	Would like to do laboratory tests to identify diseases.	Investigative - ONETshort
0.04	Would like to take care of children at a day-care center.	Social - ONETshort
-0.03	Suffer from sleeplessness.	EPQ:N
0.03	ICAR	MRiq5

Best Items correlating with being from the UK

Table: Top correlations with being from GBR

GBR	item	itm_s
-0.18	Like to stand during the national anthem.	IPIP
-0.10	Just know that I will be a success.	IPIP
-0.09	Believe in one true religion.	IPIP
0.08	ICAR	VRiq10
-0.08	Like to compete in athletic events.	ORVIS - Adventure
-0.08	Am an extraordinary person.	IPIP
0.08	Dont consider myself religious.	IPIP
0.08	Dislike myself.	IPIP
-0.07	Go straight for the goal.	IPIP
0.07	Have a low opinion of myself.	IPIP
0.07	ICAR	VRiq14
0.07	Would like to put out forest fires.	Realistic - ONETshort
-0.07	Like to make important things happen.	ORVIS - Leadership
0.07	Do too little work.	IPIP
0.07	Waste my time.	IPIP

Best Items correlating with being from USA

Table: Top correlations with being from USA

USA	item
0.24	Like to stand during the national anthem.
-0.20	People spend too much time safeguarding their future with savings and insurance.
-0.19	ICAR
-0.18	Think marriage is old-fashioned and should be done away with.
0.18	Work hard.
-0.18	Get even with others.
-0.17	Believe that there is no absolute right and wrong.
0.17	Will do anything for others.
0.16	Laugh aloud.
-0.16	Believe that I am better than others.
0.15	Push myself very hard to succeed.
-0.15	Dislike routine.
-0.15	ICAR
-0.15	Dont consider myself religious.
-0.15	Admire a really clever scam.
-0.15	Would like to be a foreign correspondent.
-0.15	Never splurge.
0.15	Laugh a lot

Profiles across 908 items of countries correlated with demographic profiles suggest sampling differences across countries

Table: Profile correlations of demographics by countries

Variable	gendr	age	BMI	exer	smoke	edctn	p1edu	p2edu
USA	0.57	0.18	0.52	0.32	-0.13	-0.12	-0.40	-0.38
CAN	-0.27	-0.23	-0.28	-0.37	0.24	-0.01	0.47	0.46
GBR	-0.30	-0.37	-0.32	-0.55	0.29	-0.14	0.36	0.36
AUS	-0.36	-0.10	-0.24	-0.31	0.31	0.16	0.48	0.47
DEU	-0.38	-0.02	-0.44	-0.14	0.06	0.27	0.52	0.50
NLD	-0.52	0.00	-0.45	0.08	0.06	0.26	0.50	0.48
NOR	-0.30	0.02	-0.34	-0.07	0.05	0.23	0.41	0.39
SWE	-0.39	0.37	-0.21	0.24	-0.11	0.56	0.36	0.34
CHE	-0.26	0.20	-0.26	0.32	-0.07	0.36	0.31	0.27
SGP	-0.17	-0.12	-0.25	-0.32	-0.17	0.01	0.01	0.02
HKG	-0.24	-0.32	-0.33	-0.36	-0.02	-0.18	0.05	0.07
CHN	-0.26	-0.04	-0.31	0.00	-0.15	0.08	0.02	0.02
IND	-0.32	0.04	-0.17	0.06	0.03	0.13	0.02	0.01
PHL	-0.03	-0.34	-0.14	-0.29	-0.07	-0.35	-0.23	-0.23
MYS	0.04	-0.23	-0.07	-0.28	-0.03	-0.30	-0.32	-0.32

6 factors of interests

1. 6 factors from the O*NET interest profiler scales (60 items; Rounds et al., 2010)
2. 8 factor Oregon Vocational Interest Scales (92 items; Pozzebon et al., 2010)
3. Oregon Avocational Interest Scales (199 items; Goldberg, 2010)
4. Formed into 6 scales fitting a "RIASEC" structure (60 items)

Realistic	"Like to work with tools and machinery."
Investigative	"Would like to do laboratory tests to identify diseases."
Artistic	"Would like to write short stories or novels."
Social	"Would like to help conduct a group therapy session."
Enterprising	"Would like to be the chief executive of a large company."
Clerical	" Would like to keep inventory records"

The correlational structure of scales are found from the SAPA item correlations

1. Given the raw data matrix, we can find the covariances (using pairwise complete data) and the find the scale intercorrelations.
2. The correlations with scales with overlapping items can be corrected for overlap `scoreOverlap` using a correction by Bashaw & Anderson Jr (1967); Cureton (1966)

Personality at 3 levels of analysis (Revelle & Condon, 2015b)

Personality can be examined at three levels of analysis

1. Personality as a unique temporal signature of one's Affect, Behavior, Cognition and Desires (ABCDs) as they change over time and space within a single individual.
   - Measuring within person patterning requires repeated measures on single subjects over time. We do this with open source text messaging procedures e.g., (Wilt, Funkhouser & Revelle, 2011; Wilt, 2014).
2. Personality is also how people differ in their patterning of the ABCDs between people.
   - This can be multilevel modeling of data collected within subjects showing that the correlational structure within subjects differs across subjects (Wilt et al., 2011; Revelle & Wilt, 2016).
   - It is also the more conventional structure of personality items as collected from the SAPA project.
3. But people choose groups such as college major or occupation based upon their unique aptitudes and appetites.
   - We can analyze this niche selection in terms of the covariance of the mean personality of the group.

TAI for groups is not the same as TAI for individuals

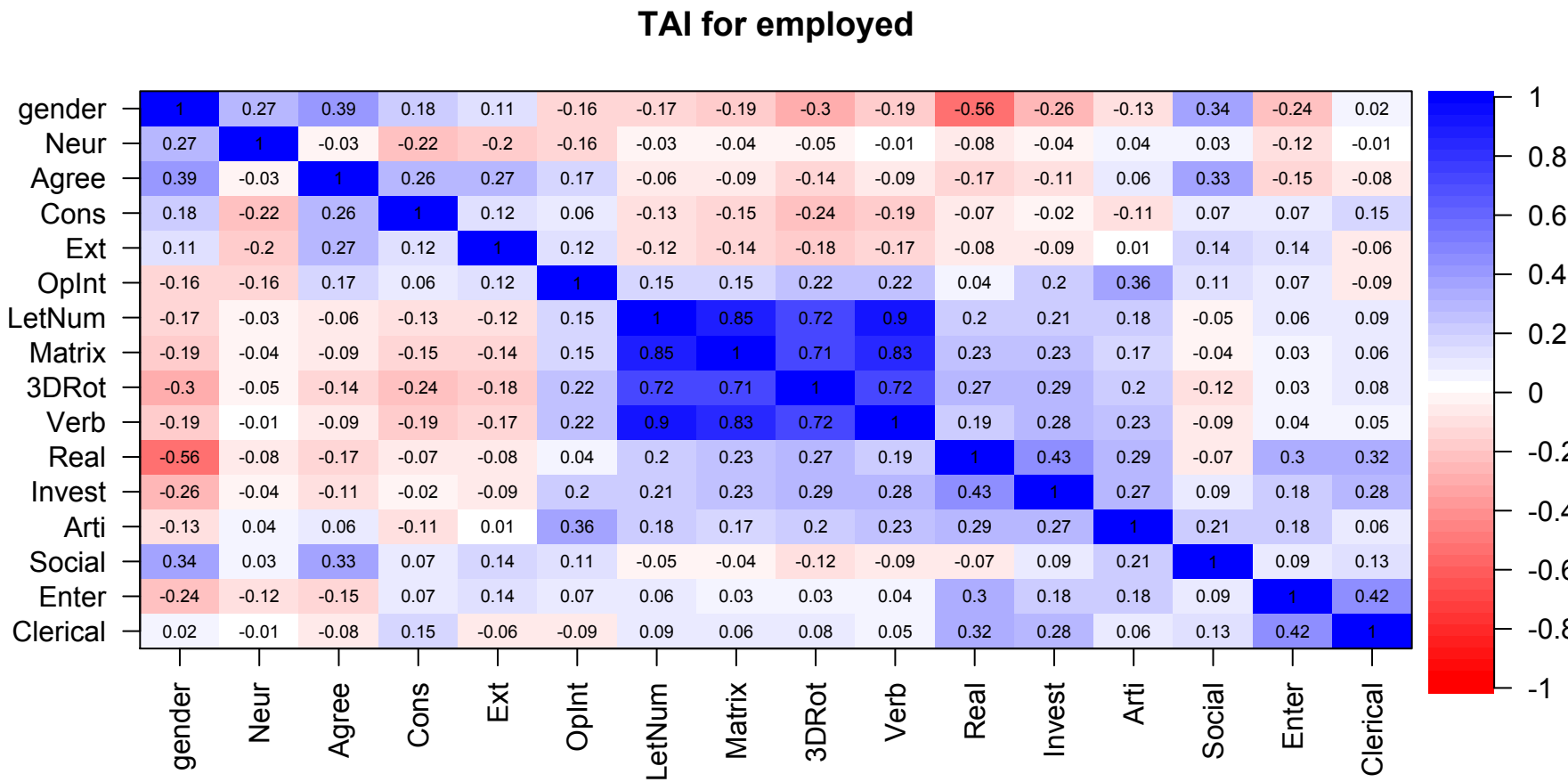
1. How do occupational groups or college majors differ on TAI?
   - The mean scores for groups allow us to compare the groups
   - But it is the structure of these group means that are particularly interesting for they allow us to examine niche selection.
2. Overall correlation is a function of within group correlations and between group correlations.
3. Correlations of aggregate scores $r_{xy_{bg}}$ (between groups) $\neq$ aggregate of correlations $r_{xy_{wg}}$ (within groups)
4. The overall correlation r_{xy} is a function of the within and the between correlations

   $r_{xy} = eta_{x_{wg}} * eta_{y_{wg}} * r_{xy_{wg}} + eta_{x_{bg}} * eta_{y_{bg}} * r_{xy_{bg}}$
5. These multi level correlations sometimes lead to what is known as the Yule-Simpson paradox (Kievit, Frankenhuis, Waldorp & Borsboom, 2013; Simpson, 1951; Yule, 1903)
   - These are independent and useful information.

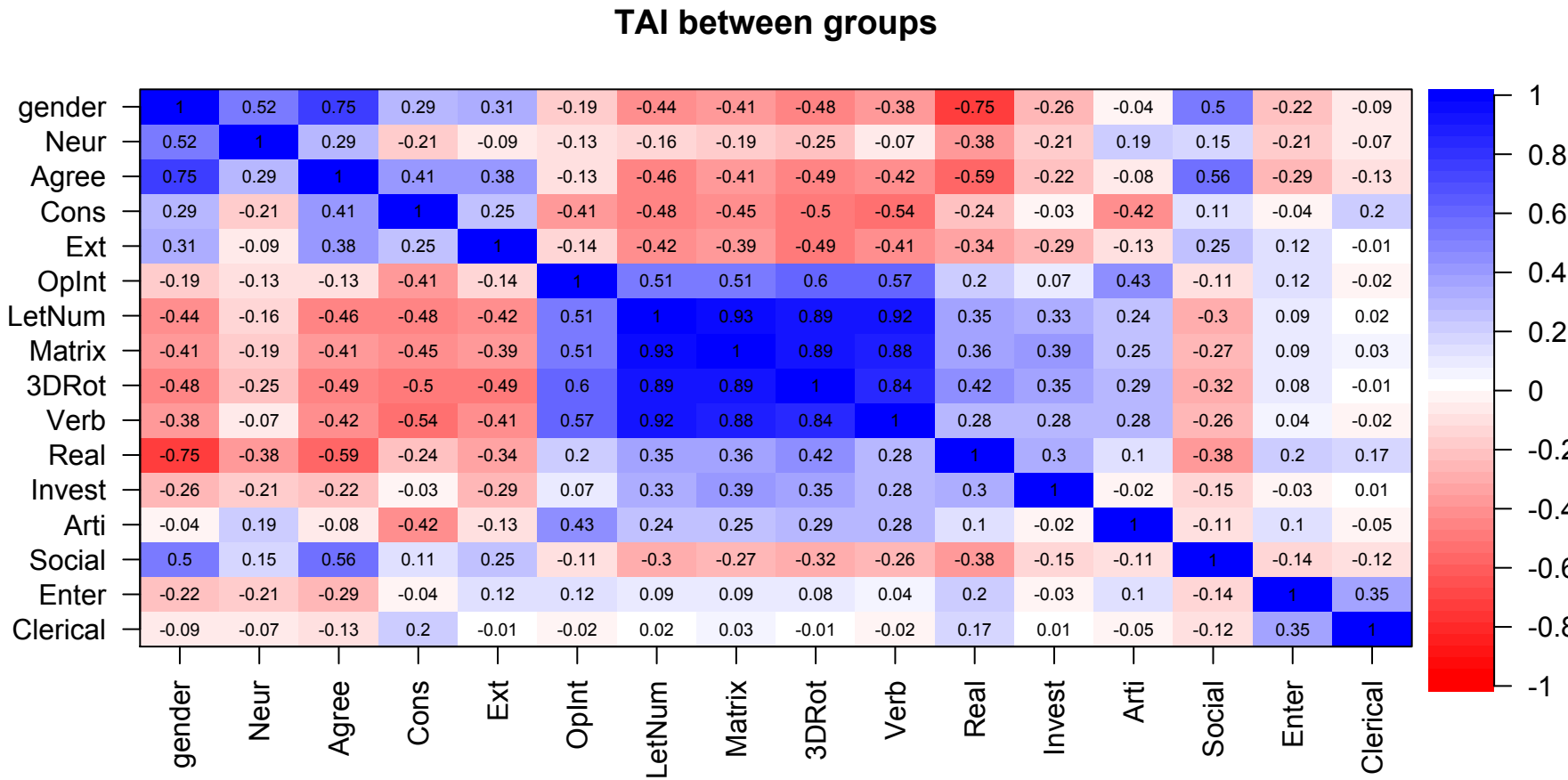
Temperament, Ability, and Interests – within and between groups

1. Examined the factor structure of the TAI scales at the normal, between subjects (across groups) level.
   - This produces the normal factor structure of temperament, of ability and of interests
   - Can show these correlations as a “heatmap”
2. But when analyzing the structure of the mean scores for each of 196 occupational groups (minimum size of 75 members), the structure is drastically different.
   - Several dimensions of temperament and interests are now negatively correlated with ability, others are orthogonal
   - Can also show these correlations as a “heatmap”

Subject Level data of 5 personality scales, 6 interests, 4 ability



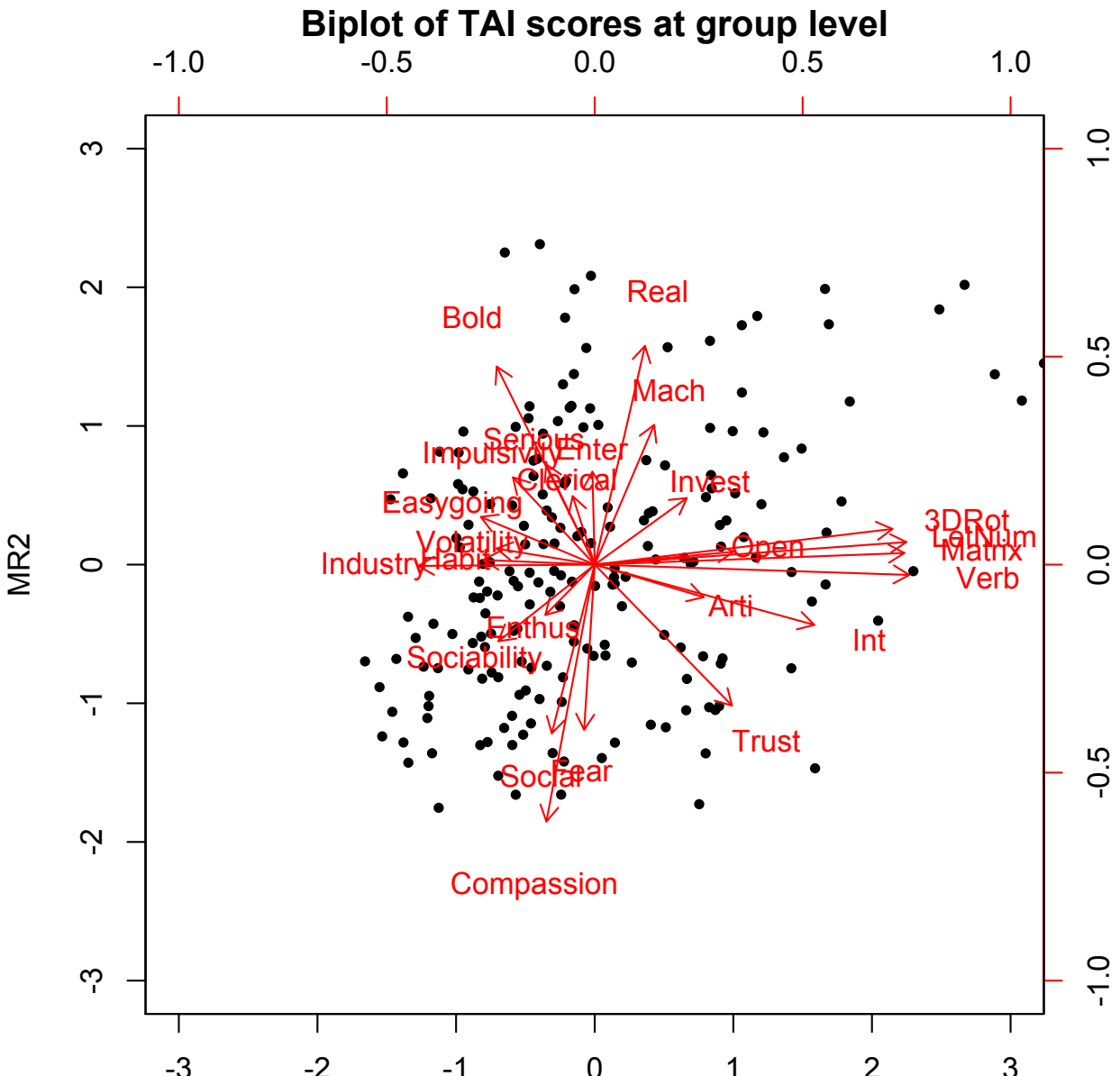
Group Level data of 15 personality scales, 6 interests, 4 ability



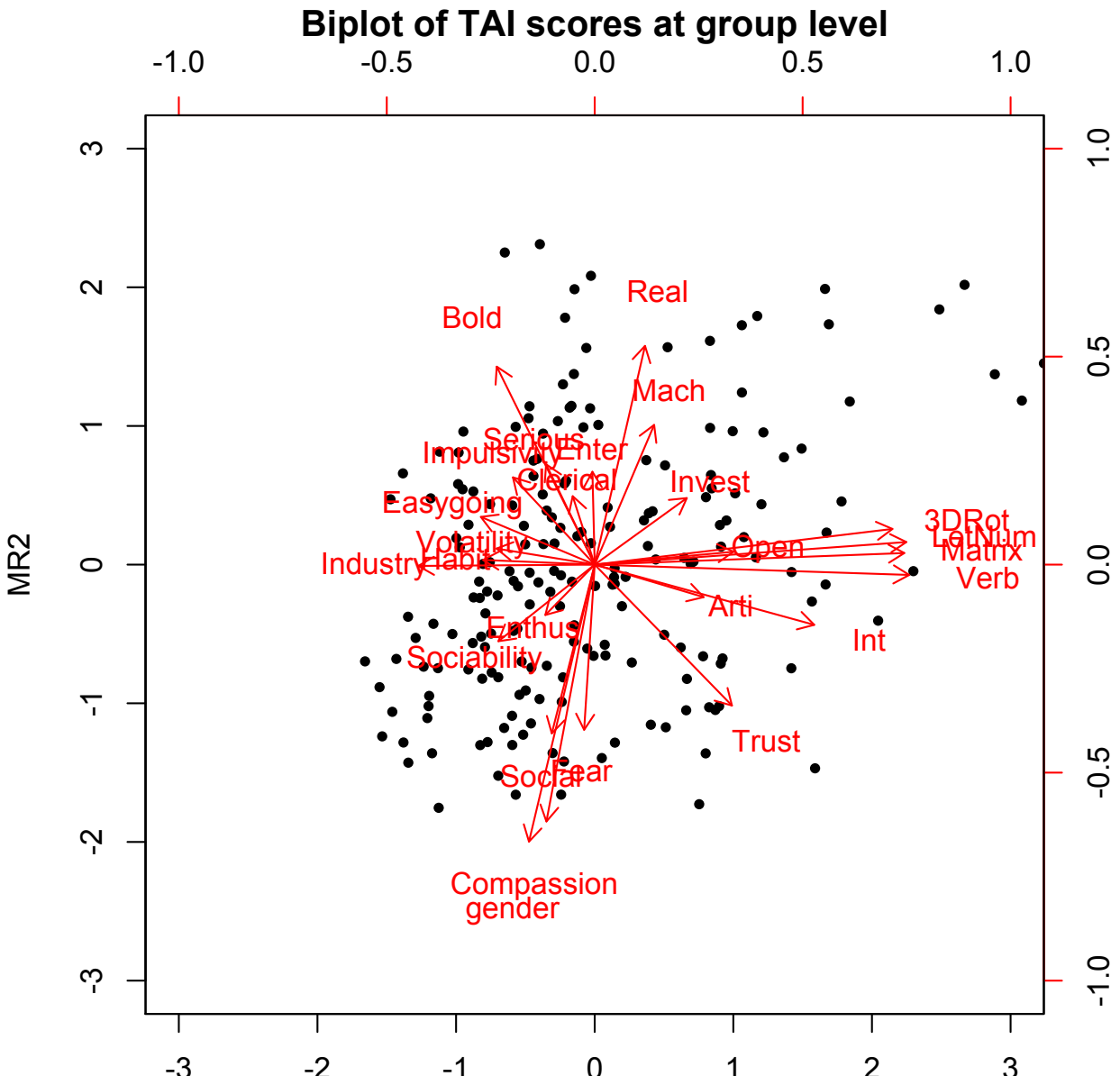
Niche selection

1. Occupations differ systematically in the intellectual Ability they require.
2. But they also differ in the Interests and Temperament they require.
3. A simple two factor solution shows that high ability can trade off for low Industry or Conscientiousness and that Boldness (low Anxiety) and Realistic interests differs from high Anxiety and Social interests.
4. We can examine the extent to which this second dimension a difference of gender using factor extension.

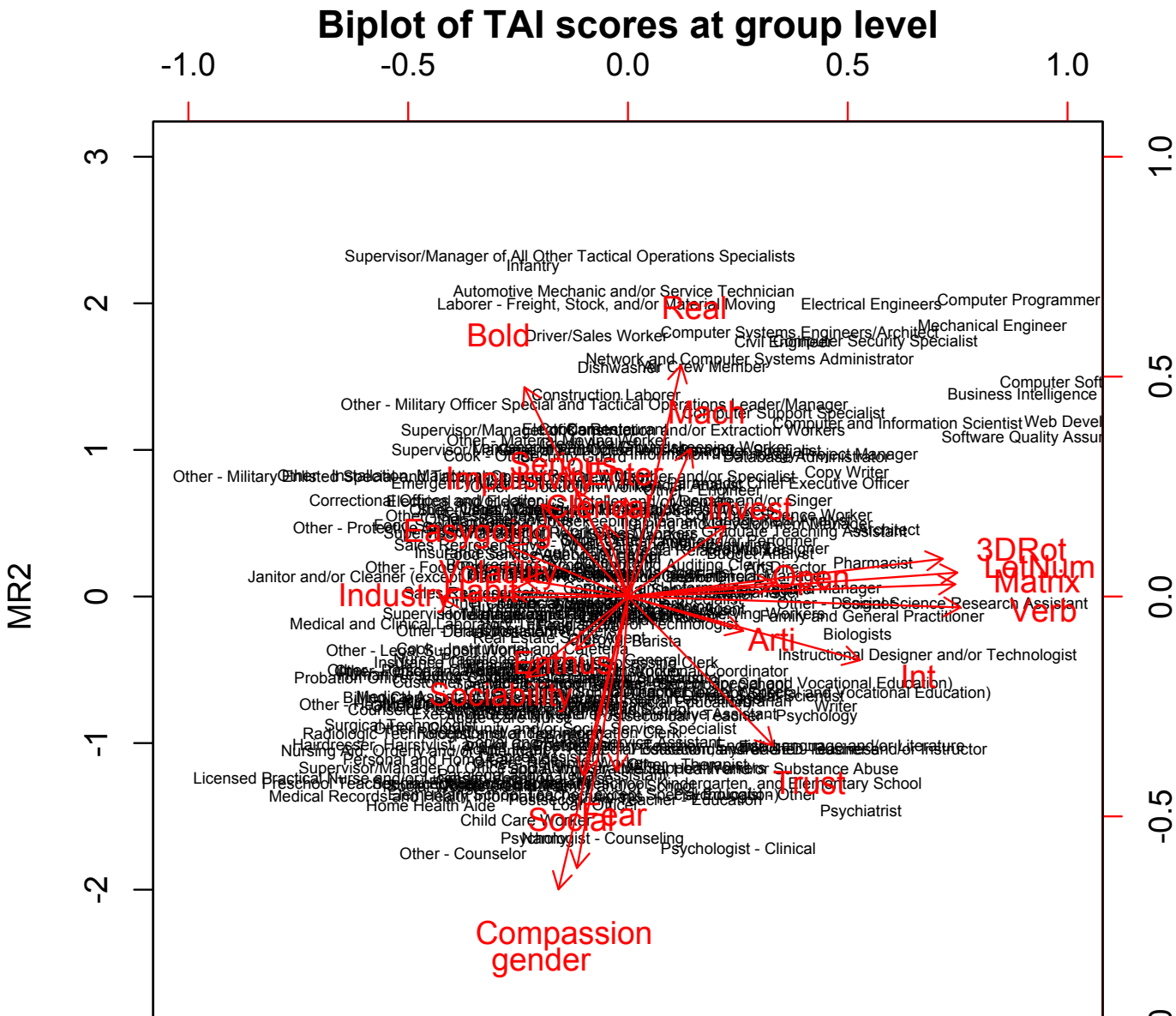
Biplot of a two factor solution to the group level data



Add gender to the extended factor solution of the group data



Biplot of a two factor solution to the group level data



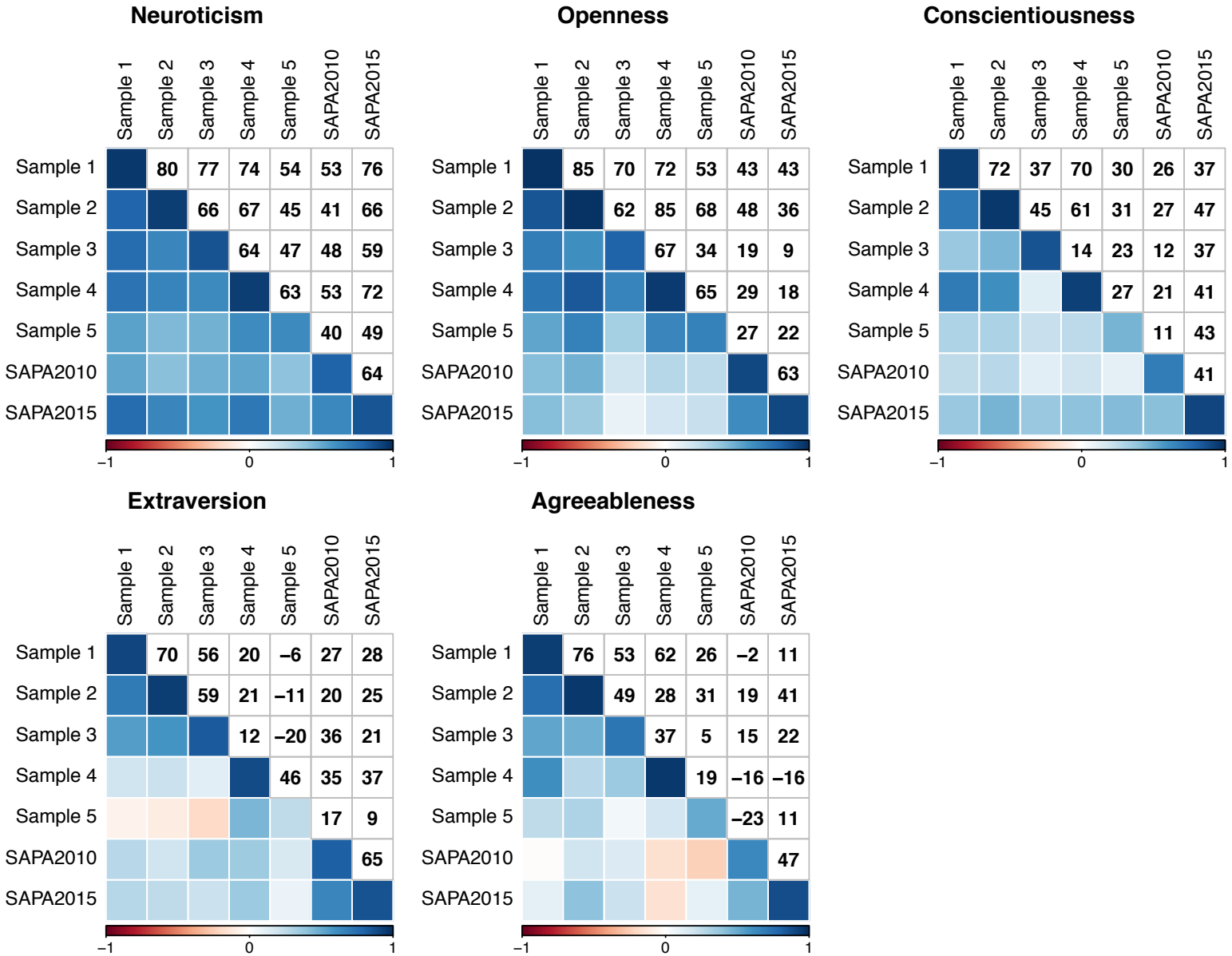
Part V

Open Scientific Research

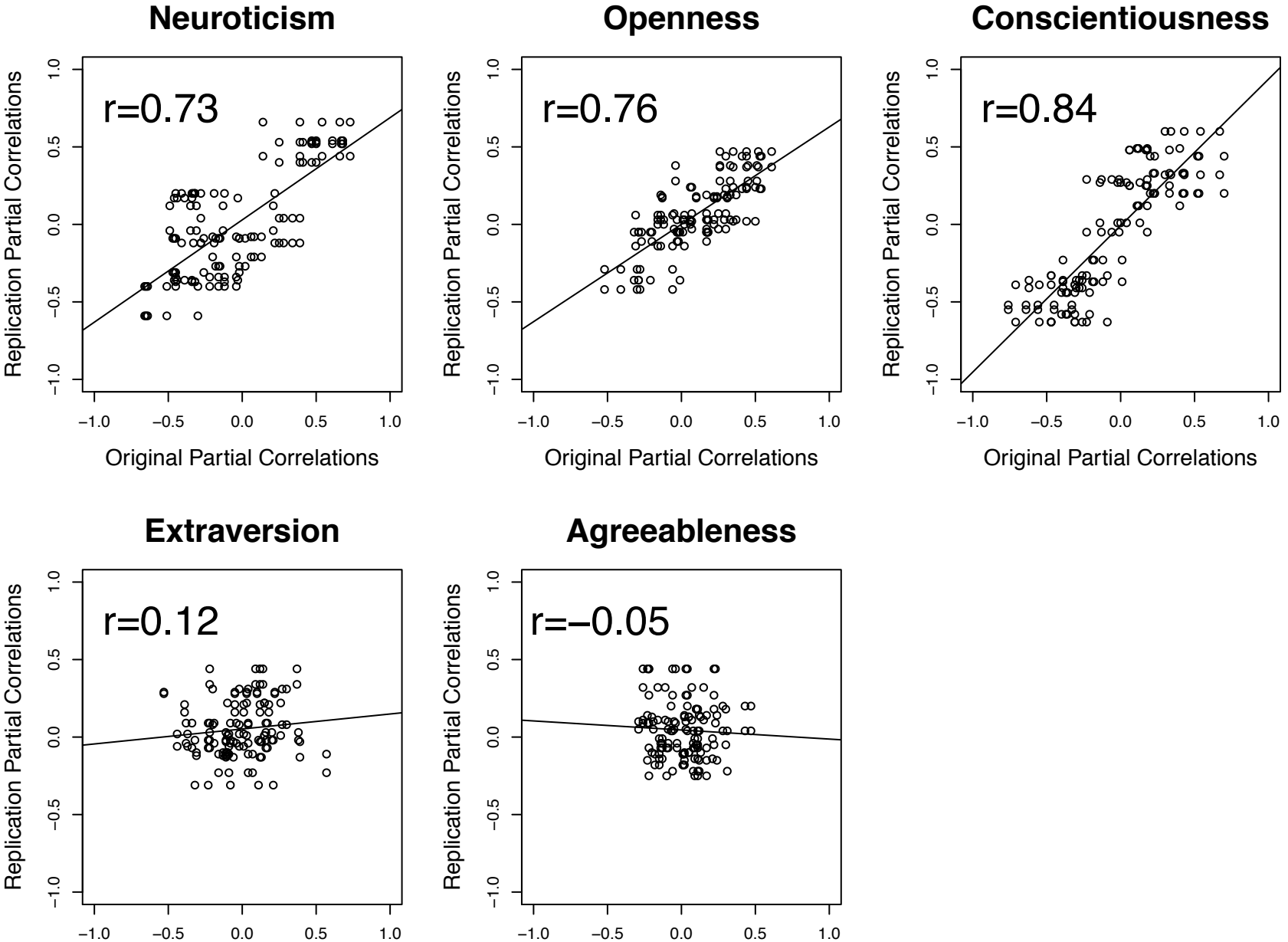
State level personality differences: Are they replicable?

1. Rentfrow and Gosling have reported data from several large internet surveys (much larger than ours).
2. Are they replicable?
3. State differences are small, but reliable for some measures, not all

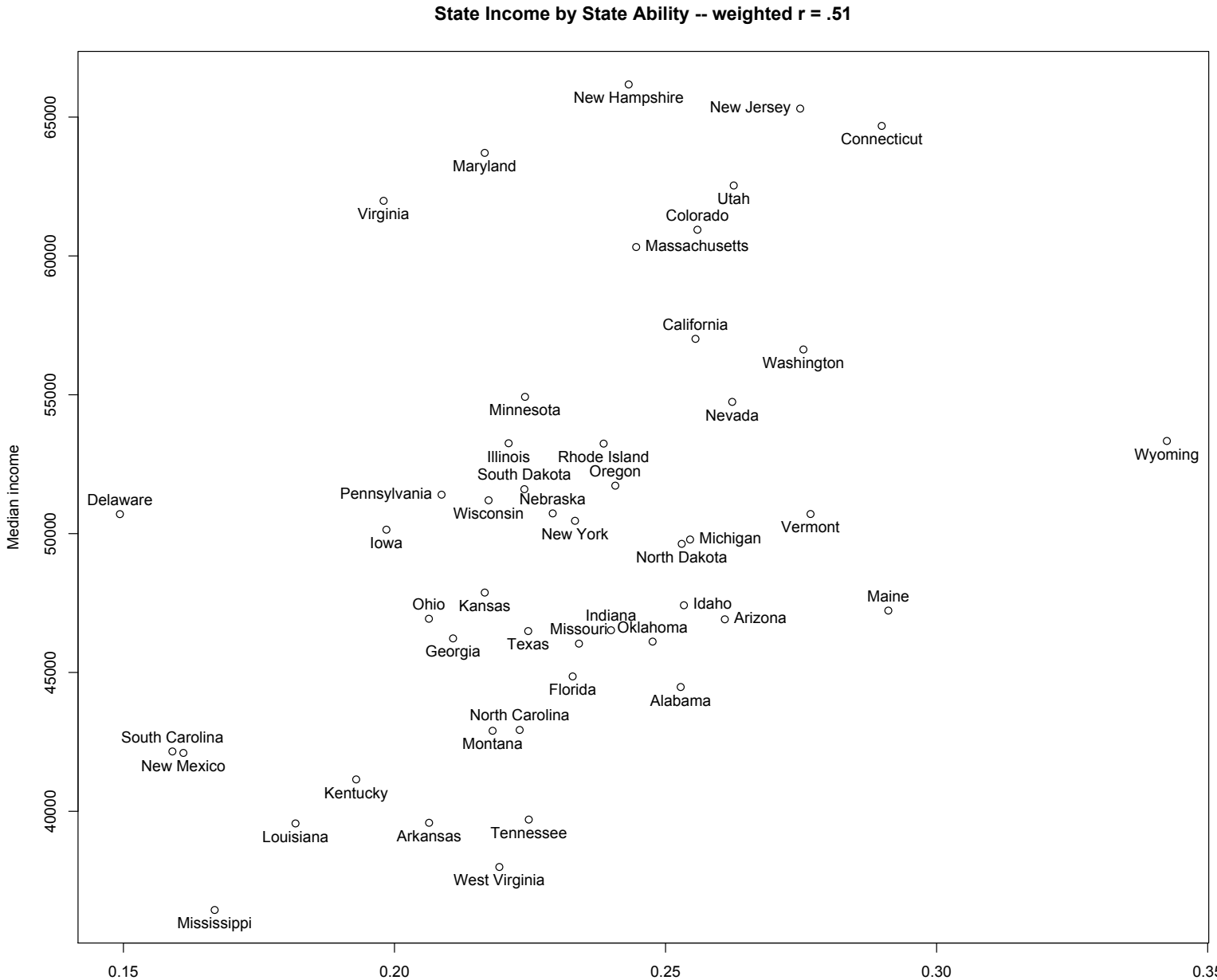
Replicability of state differences depends upon the trait being measured



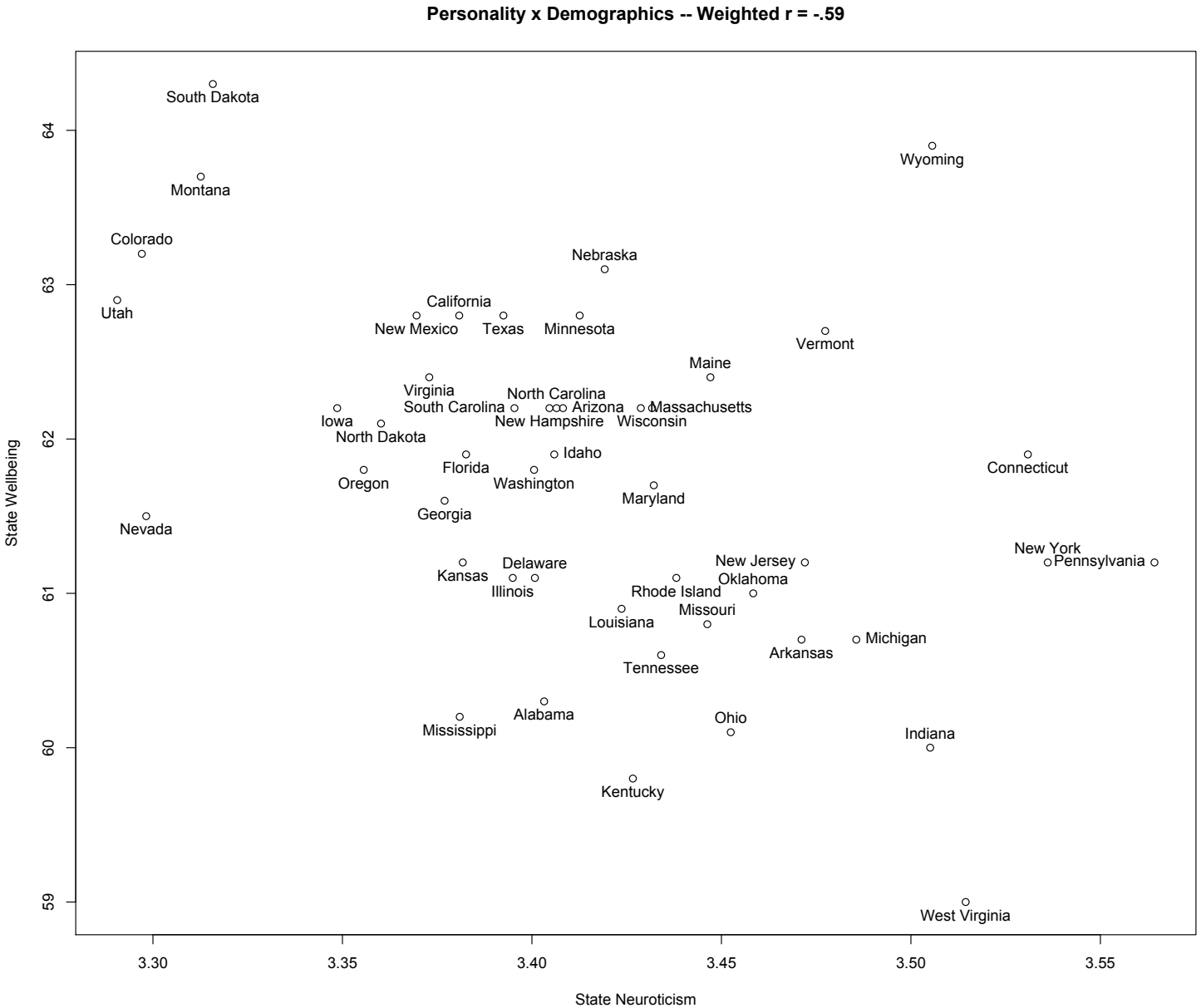
Replicability of personality by state demographics depends upon trait



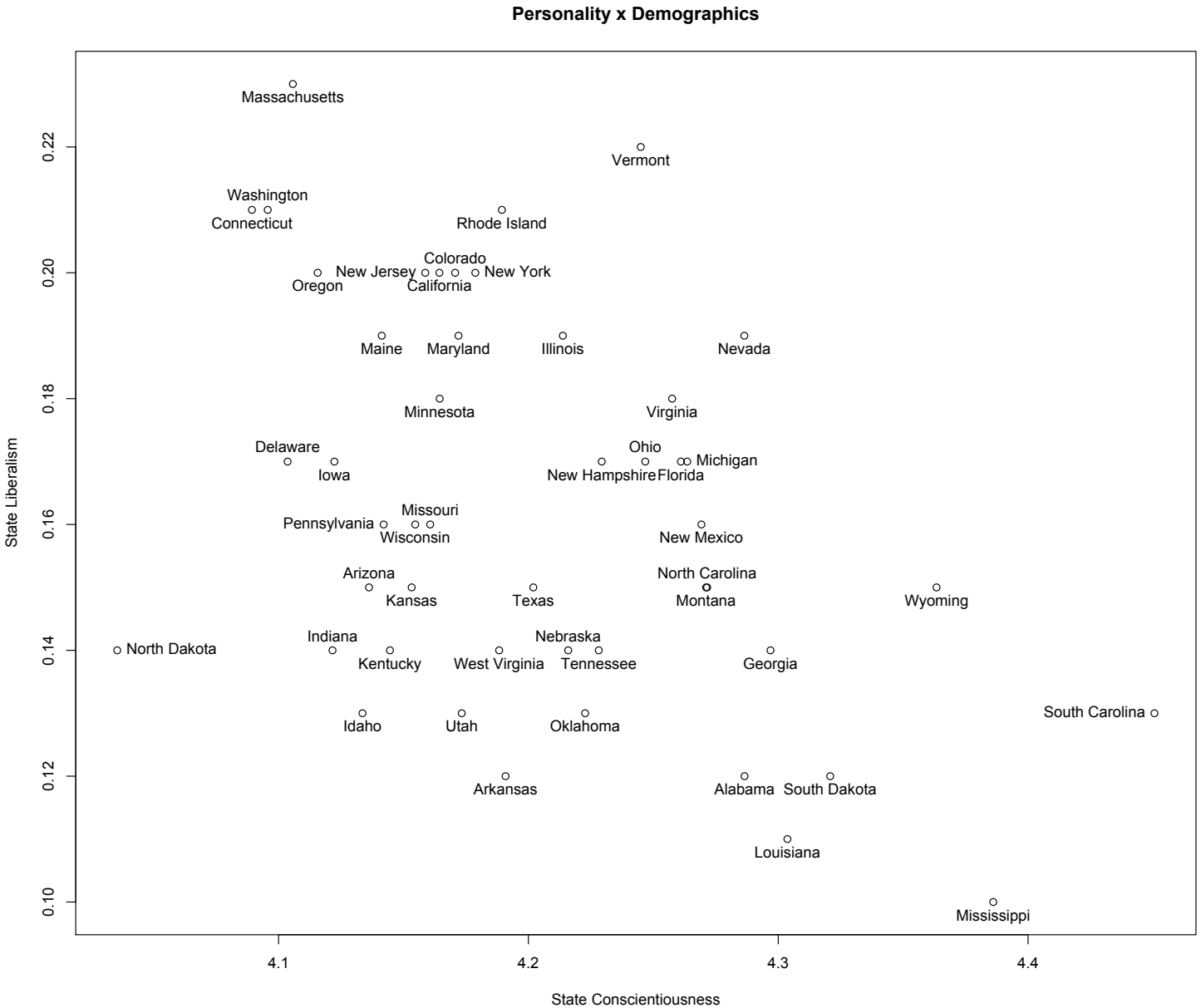
State level: Income varies by iq (weighted r = .51)



State level: Well being varies by neuroticism (weighted r = -.59)



State level: Liberalism varies by Conscientiousness (r= -.42)



Summary and Conclusions

1. Ability, temperament and interests all provide useful information about human personality.
2. Intellectual and Personality development is the process of experiencing and choosing niches.
3. When we describe the intellectual requirements of a profession or a college major, we should not ignore that appropriate interests and temperaments guide occupational choice.
4. We need to consider appetites along with aptitudes.
5. The statistics, materials, methods, and data from all of these studies are done using Open Source Science.
6. Join us in this journey.
7. For more information and for these slides go to `http://personality-project.org/sapa.html`

Anderson, C. J., Bahník, Š., Barnett-Cowan, M., Bosco, F. A., Chandler, J., Chartier, C. R., & Cheung, F. (2016). Response to comment on “estimating the reproducibility of psychological science”. *Science*, *351*(6277), 1037–1037.

Anderson, J., Lin, H., Treagust, D., Ross, S., & Yore, L. (2007). Using large-scale assessment datasets for research in science and mathematics education: Programme for International Student Assessment (PISA). *International Journal of Science and Mathematics Education*, *5*(4), 591–614.

Ashton, M. C., Lee, K., & Goldberg, L. R. (2007). The IPIP-HEXACO scales: An alternative, public-domain measure of the personality constructs in the HEXACO model. *Personality and Individual Differences*, *42*(8), 1515–1526.

Bashaw, W. & Anderson Jr, H. E. (1967). A correction for replicated error in correlation coefficients. *Psychometrika*, *32*(4), 435–441.

Bouchard, T. (1997). Experience producing drive theory: how

genes drive experience and shape personality. *Acta Paediatrica*, *86*(S422), 60–64.

Brown, A. D. (2014). Simulating the MMCAR method: An examination of precision and bias in synthetic correlations when data are 'massively missing completely at random'. Master's thesis, Northwestern University, Evanston, Illinois.

Cattell, R. B. (1957). *Personality and motivation structure and measurement*. Oxford, England: World Book Co.

Cattell, R. B. & Stice, G. (1957). *Handbook for the Sixteen Personality Factor Questionnaire.* Champaign, Ill.: Institute for Ability and Personality Testing.

Cloninger, C. R., Przybeck, T. R., & Svrakic, D. M. (1994). *The Temperament and Character Inventory (TCI): A guide to its development and use*. center for psychobiology of personality, Washington University St. Louis, MO.

Comrey, A. L. (2008). The Comrey Personality Scales. In G. J. Boyle, G. Matthews, & D. H. Saklowfske (Eds.), *Sage handbook*

of personality theory and testing: Personality measurement and assessment, volume II (pp. 113–134). London: Sage.

Condon, D. M. (2014). *An organizational framework for the psychological individual differences: Integrating the affective, cognitive, and conative domains*. PhD thesis, Northwestern University.

Condon, D. M. (2015). The many little items of “big five” measures: Hierarchy, heterarchy, and predictive utility in personality structure. Long Beach, California. Society of Personality and Social Psychology.

Condon, D. M. (2017). The SAPA Personality Inventory: An empirically-derived, hierarchically-organized self-report personality assessment model. Technical report, Northwestern University.

Condon, D. M. & Revelle, W. (2014). The International Cognitive Ability Resource: Development and initial validation of a public-domain measure. *Intelligence*, *43*, 52–64.

Condon, D. M. & Revelle, W. (2015a). Selected ICAR data from the SAPA-Project: Development and initial validation of a public-domain measure. *Harvard Dataverse*.

Condon, D. M. & Revelle, W. (2015b). Selected personality data from the SAPA-Project: 08dec2013 to 26jul2014. *Harvard Dataverse*.

Condon, D. M. & Revelle, W. (2015c). Selected personality data from the SAPA-Project: On the structure of phrased self-report items. *Journal of Open Psychology Data*, *3*(1).

Condon, D. M. & Revelle, W. (2016). Selected ICAR data from the SAPA-Project: Development and initial validation of a public-domain measure. *Journal of Open Psychology Data*.

Costa, P. T. & McCrae, R. R. (1992). *NEO PI-R professional manual*. Odessa, FL: Psychological Assessment Resources, Inc.

Cureton, E. (1966). Corrected item-test correlations. *Psychometrika*, *31*(1), 93–96.

Deary, I. (2008). Why do intelligent people live longer? *Nature*, *456*(7219), 175–176.

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2002). Higher-order factors of the big five predict conformity: Are there neuroses of health? *Personality and Individual Differences*, *33*(4), 533–552.

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the big five. *Journal of Personality and Social Psychology*, *93*(5), 880–896.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, *41*, 417–440.

Digman, J. M. (1997). Higher-order factors of the big five. *Journal of Personality and Social Psychology*, *73*, 1246–1256.

Elleman, L. G. (2019). *Behavioral measures of personality*. PhD thesis, Northwestern University.

Eysenck, H. J. (1994). The big five or the giant three: Criteria for a paradigm. In C. F. Halverson, G. A. Kohnstamm, & R. P. Martin

(Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 37–51). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Eysenck, S. B. G., Eysenck, H. J., & Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and Individual Differences*, *6*(1), 21 – 29.

French, J. A. & Condon, D. M. (2015). SAPA Tools: Tools to analyze the SAPA Project. R package version 0.1.

Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D. (2016). Comment on “estimating the reproducibility of psychological science”. *Science*, *351*(6277), 1037.

Goldberg, L. R. (1990). An alternative “description of personality”: The big-five factor structure. *Journal of Personality and Social Psychology*, *59*(6), 1216–1229.

Goldberg, L. R. (1992). The development of markers for the big-five factor structure. *Psychological Assessment*, *4*(1), 26–42.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe*, volume 7 (pp. 7–28). Tilburg, The Netherlands: Tilburg University Press.

Goldberg, L. R. (2010). Personality, demographics and self reported acts: the development of avocational interest scales from estimates of the mount time spent in interest related activities. In C. Agnew, D. Carlston, W. Graziano, & J. Kelly (Eds.), *Then a miracle occurs: Focusing on the behavior in social psychological theory and research* (pp. 205–226). New York, NY: Oxford University Press.

Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, *24*(1), 79 – 132.

Gough, H. G. & Bradley, P. (1996). Cpi manual . palo alto.

Hayes, K. (1962). Genes, drives, and intellect. *Psychological Reports (Monograph Supplement 2)*, *10*, 299–342.

Hendriks, A. A. J. (1997). *The construction of the five-factor personality inventory (FFPI)*. PhD thesis, Rijksunivsiteit Groningen, Groningen, The Netherlands.

Hendriks, A. A. J., Hofstee, W. K., & De Raad, B. (1999). The five-factor personality inventory (FFPI). *Personality and Individual Differences*, *27*(2), 307 – 325.

Hofstee, W. K., de Raad, B., & Goldberg, L. R. (1992). Integration of the big five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology*, *63*(1), 146–163.

Hogan, R. & Hogan, J. (1995). *The Hogan personality inventory manual (2nd. ed.)*. Tulsa, OK: Hogan Assessment Systems.

Jackson, D. N. (1983). *JPI-R: Jackson Personality Inventory, Revised*. Sigma Assesment Systems, Incorporated.

Jackson, D. N., Paunonen, S. V., & Tremblay, P. F. (2000). *Six Factor Personality Questionnaire.* Port Huron, MI: Sigma Assesment Systems.

Johnson, W. (2010). Extending and testing tom bouchard's experience producing drive theory. *Personality and Individual Differences*, *49*(4), 296 – 301. Collected works from the Festschrift for Tom Bouchard, June 2009: A tribute to a vibrant scientific career.

Kelly, E. L. & Fiske, D. W. (1950). The prediction of success in the VA training program in clinical psychology. *American Psychologist*, *5*(8), 395 – 406.

Kelly, E. L. & Fiske, D. W. (1951). *The prediction of performance in clinical psychology*. Ann Arbor, Michigan: University of Michigan Press.

Kievit, R. A., Frankenhuis, W. E., Waldorp, L. J., & Borsboom, D. (2013). Simpson's paradox in psychological science: a practical guide. *Frontiers in Psychology*, *4*(513), 1–14.

Lee, K. & Ashton, M. C. (2004). Psychometric properties of the HEXACO personality inventory. *Multivariate Behavioral Research*, *39*(2), 329–358.

Liebert, M. (2006). A public-domain assessment of music preferences as a function of personality and general intelligence. Honors Thesis. Department of Psychology, Northwestern University.

Loe, B. S. & Rust, J. (2017). The perceptual maze test revisited: Evaluating the difficulty of automatically generated mazes. *Assessment*, *0*(0).

Lord, F. M. (1955). Estimating test reliability. *Educational and Psychological Measurement*, *15*, 325–336.

Möttus, R., Sinick, J., A.Terracciano, Hřebíckova, M., Kandler, C., & Jang, J. A. . . . K. L. (2018). Personality characteristics below facets: A replication and meta-analysis of cross-rater agreement, rank-order stability, heritability, and utility of personality nuances. *Journal of Personality and Social Psychology*.

Pozzebon, J. A., Visser, B. A., Ashton, M. C., Lee, K., & Goldberg, L. R. (2010). Psychometric characteristics of a public-domain self-report measure of vocational interests: The Oregon

Vocational Interest Scales. *Journal of Personality Assessment*, *92*(2), 168–174.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Revelle, W. (2018). *psych: Procedures for Personality and Psychological Research*. https://CRAN.r-project.org/package=psych: Northwestern University, Evanston. R package version 1.8.12.

Revelle, W. & Condon, D. M. (2012). Multilevel analysis of personality: Personality of college majors. Presented at the annual meeting of the Society of Multivariate Experimental Psychology.

Revelle, W. & Condon, D. M. (2015a). Ability, temperament, and interests: their joint predictive power for job choice. Albuquerque, New Mexico. International Society for the Study of Intelligence (September).

Revelle, W. & Condon, D. M. (2015b). A model for personality at three levels. *Journal of Research in Personality*, *56*, 70–81.

Revelle, W., Condon, D. M., Wilt, J., French, J. A., Brown, A., & Elleman, L. G. (2016). Web and phone based data collection using planned missing designs. In N. G. Fielding, R. M. Lee, & G. Blank (Eds.), *SAGE Handbook of Online Research Methods* (2nd ed.). chapter 37, (pp. 578–595). Sage Publications, Inc.

Revelle, W. & Wilt, J. (2016). The data box and within subject analyses: A comment on Nesselroade and Molenaar. *Multivariate Behavioral Research*, *51*(2-3), 419–421.

Revelle, W., Wilt, J., & Condon, D. (2011). Individual differences and differential psychology: A brief history and prospect. In T. Chamorro-Premuzic, A. Furnham, & S. von Stumm (Eds.), *Handbook of Individual Differences* chapter 1, (pp. 3–38). Oxford: Wiley-Blackwell.

Revelle, W., Wilt, J., & Rosenthal, A. (2010). Individual differences in cognition: New methods for examining the

personality-cognition link. In A. Gruszka, G. Matthews, & B. Szymura (Eds.), *Handbook of Individual Differences in Cognition: Attention, Memory and Executive Control* chapter 2, (pp. 27–49). New York, N.Y.: Springer.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, *2*(4), 313–345.

Rounds, J., Su, R., Lewis, P., & Rivkin, D. (2010). O* NET® interest profiler short form psychometric characteristics: Summary.

Shipley, W. C. (2009). *Shipley-2: manual*. Western Psychological Services.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *13*(2), 238–241.

Tellegen, A. & Waller, N. G. (2008). Exploring personality through test construction: Development of the multidimensional personality questionnaire. *The Sage handbook of personality theory and assessment*, *2*, 261–292.

Wilt, J. (2014). *A new form and function for personality.* PhD thesis, Northwestern University.

Wilt, J., Funkhouser, K., & Revelle, W. (2011). The dynamic relationships of affective synchrony to perceptions of situations. *Journal of Research in Personality*, *45*, 309–321.

Wilt, J. & Revelle, W. (2015). Affect, behaviour, cognition and desire in the big five: An analysis of item content and structure. *European Journal of Personality*, *29*(4), 478–497.

Yule, G. U. (1903). Notes on the theory of association of attributes in statistics. *Biometrika*, *2*(2), 121–134.