Synthetic Aperture Personality Assessment: An old technique applied with modern technology

part of a symposium:
Studying individual differences using the web:
A report from the SAPA project
International Society for the Study of Individual Differences
London, Ontario

William Revelle
Northwestern University
Evanston, Illinois USA
Partially supported by a grant from the National Science Foundation:
SMA-1419324

Outline

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-?
   - Structure of Ability (g - g_f, g_c, V-P-R ?
   - Hierarchical structure of interests people-things RAISEC
   - 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 items, we also want many people.
2. Resolution (fidelity) goes up with sample size, N (standard errors are a function of $\sqrt{N}$)

$$\sigma_r = \frac{1 - r^2}{\sqrt{N(1 + r^2)}}$$

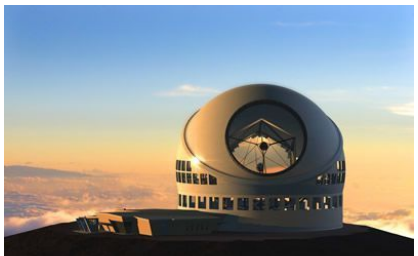
3. Also increases as number of items, n, measuring each construct (signal/noise ratio varies as number of items and average correlation of the items)

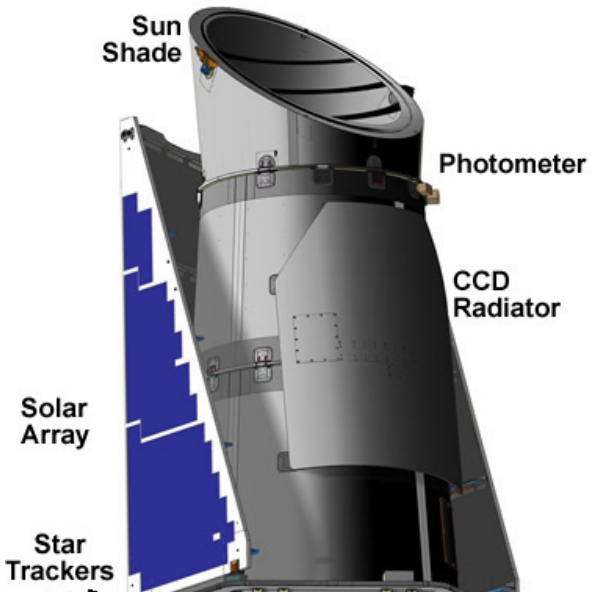
$$s/n = n\bar{r}/(1 - n\bar{r})$$

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

A short diversion: the history of optical telescopes

Resolution varies by aperture diameter (bigger is better)

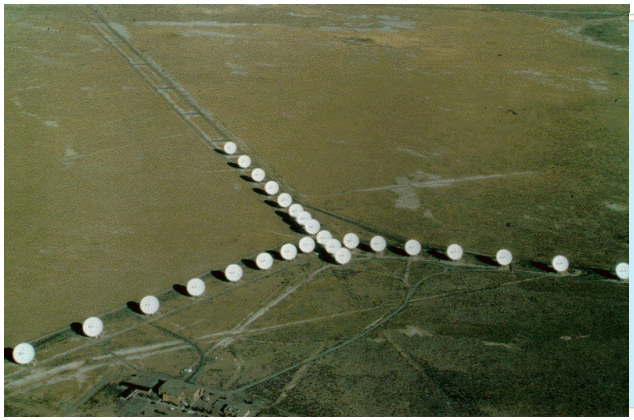


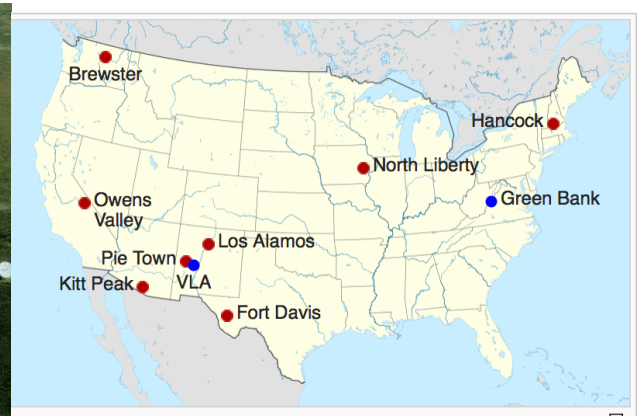


A short diversion: history of radio telescopes

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) and the concept of sampling people and 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) (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 Gallop or 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. Syntehetically combine data across subjects and across items. This will imply a missing data structure which is
   - Missing Completely At Random (MCAR), or even more descriptive:
   - Massively Missing Completely at Random (MMCAR)
4. 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
```

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.......5
......63.1.......6...5..4..2....
..2.4..5..........52.4.....44...
2.55.....2.....6.....6.....55...
```

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 $\mathbf{X}$ with N observations converted to deviation scores.
   - Then the item variance covariance matrix is $\mathbf{C} = \mathbf{X}\mathbf{X}'N^{-1}$
   - and scale scores, $\mathbf{S}$ are found by $\mathbf{S} = \mathbf{K}'\mathbf{X}$.
   - $\mathbf{K}$ is a keying matrix, with $\mathbf{K}_{ij} = 1$ 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, $\mathbf{C}_s$, is

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

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

NORTHWESTERN UNIVERSITY

The telescope analogy

1. Telescopes are shared resources between many researchers.
2. Similarly, SAPA is a shared resource, different users have different "time on the machine".
3. This is implemented by changing the probability any item is presented.
4. Low interest items are given perhaps p =.05, high interest to every subject (p=1)
5. Item statistics such as means, sds, vary by p_iN.
6. item covariances vary by $p_i p_j$ N

History

1. Originally developed with various honors undergraduates
   - Gregory Laun (Temperament and RWA)
   - Melissa Liebert (musical preferences, ability, and temperament)
   - First reported in Revelle, Wilt & Rosenthal (2010), but since elaborated in Condon & Revelle (2014); Revelle, Condon, Wilt, French, Brown & Elleman (2015)
2. Have now produced two dissertations, master theses and more honors theses using SAPA
   - Joshua Wilt (2014) and the structure of ABCDs of personality measurement
   - David Condon (2014) and the factor structure of personality inventories
   - Ashley Brown (2014) and the sampling theory behind SAPA
   - Lorien Elleman and demographic correlates of personality
   - Anthony Evans & Revelle (2008) and correlates of trust worthyness
   - Zara Wright and the measurement of psychopathology
   - Other studies include computer addiction, different methods

How does it work?: part I

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, 2015) using R (R Core Team, 2015)
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, Visser, Ashton, Lee & Goldberg, 2010)

How does it work?: part II

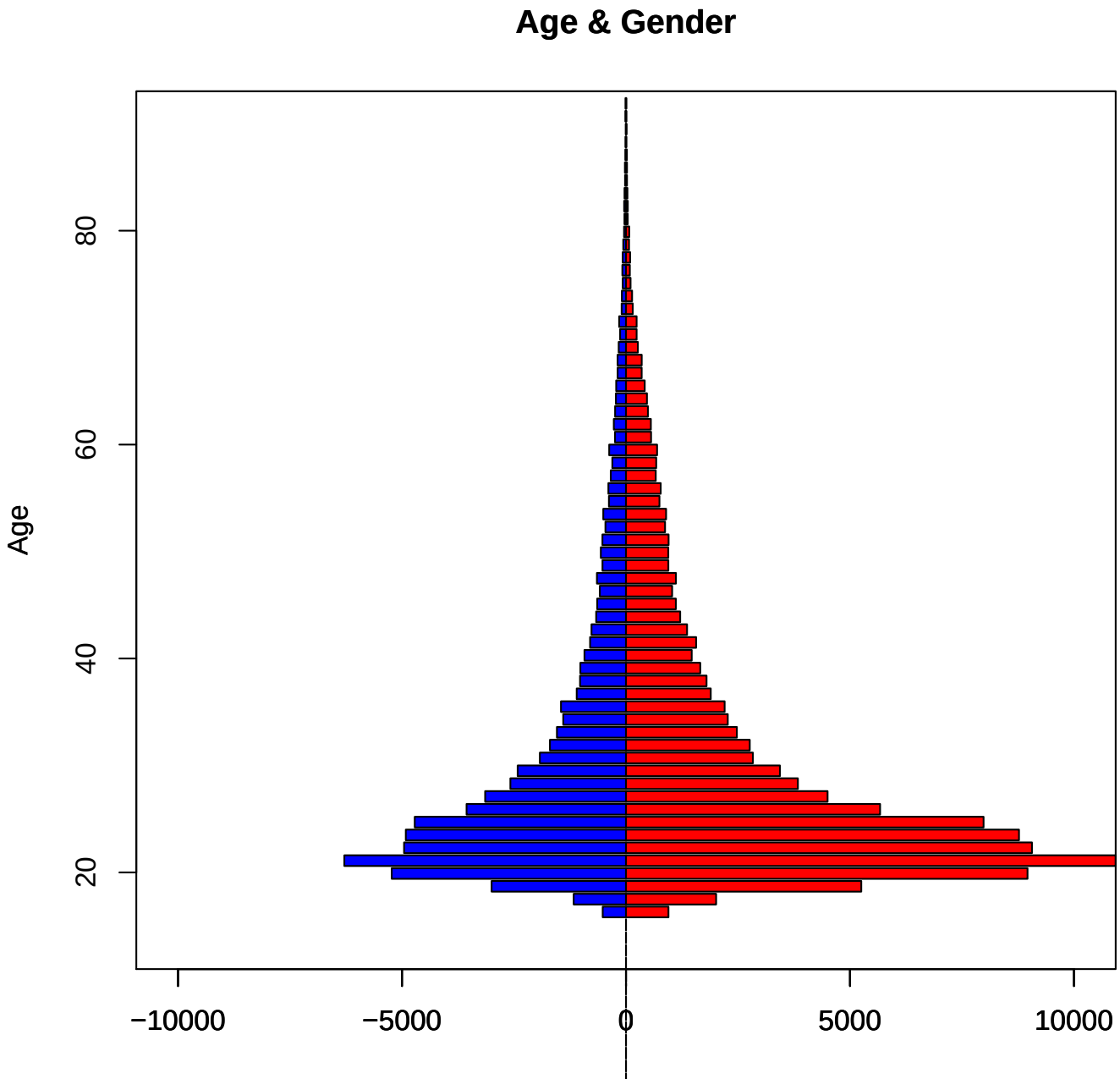
1. Participants find us by searching web for "personality tests", etc. and find personality-project.org or sapa-project.org
2. Each participant is given a number of web pages
   - Consent Form Basic description of project and question whether they have taken test before.
   - Demographics Age, sex, height, 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, 2015) for R.

Who are the participants and where do they come trom?

1. Public data set of $\approx 20,000$ participants is available at DataVerse and JOPD (e.g., Condon & Revelle, 2015a,b).
2. Current data set is from August 2010 and April 2015.
3. 177,048 participants,
4. Participants from this sample are 63% female.
5. Participants grew up in 214 countries, with the United States accounting for 74% of the sample.
6. Sixteen countries besides the U.S. have 500 or more participants, with the top three being Canada (7,703), the United Kingdom (4,561), and Australia (3,338)
7. Participants from the U.S. identify as 67% white, 10% African American, 9% Hispanic, 6% multiracial, and 4% Asian American.


Age & Gender
Age
20
40
60
80
−10000
−5000
0
5000
10000

Rest of symposium

Brown Ashley Brown: Standard Errors of SAPA correlations: a Monte Carlo analysis.

Wilt Joshua Wilt: Affect, Behavior, Cognition, and Desire in the Big Five: An Analysis of Item Content and Structure

Condon David Condon: Temperament and ability: predicting health outcomes

Elleman Lorien Elleman: Demographic Correlates of Temperament and Ability

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.

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.

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

NORTHWESTERN UNIVERSITY

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

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

Evans, A. M. & Revelle, W. (2008). Survey and behavioral measurements of interpersonal trust. *Journal of Research in Personality*, *42*(6), 1585–1593.

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

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.

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.

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

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 (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Revelle, W. (2015). *psych: Procedures for Personality and Psychological Research*.

http://cran.r-project.org/web/packages/psych/: Northwestern University, Evanston. R package version 1.5.6.

Revelle, W., Condon, D., Wilt, J., French, J. A., Brown, A. D., & Elleman, L. G. (2015). Web and phone based data collection using planned missing designs. In G. B. Nigel G. Fielding, Raymond M. Lee (Ed.), *he Sage Handbook of Online Research Methods* (2nd ed.). SAGE Publications.

Revelle, W. & Laun, G. (2004). Synthetic aperture personality assessment: A progress report and a proposal. Presented at the annual meeting of the Society of Multivariate Experimental Psychology.

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.

NORTHWESTERN UNIVERSITY

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

Wright, Z. E. (2014). Creating a self-report measure of psychopathy using items from the Personality Inventory for the DSM-5. Honors Thesis. Department of Psychology, Northwestern University.