

Psychology 360: Personality Research

Methods of Scale Construction

William Revelle

Department of Psychology
Northwestern University
Evanston, Illinois USA



NORTHWESTERN
UNIVERSITY

November, 2022

Outline

Methods of scale construction

Empirical

Rational Keying

Scale Construction: the Pragmatics

Dimensional versus empirical, another example

Validity using peer ratings

References

Methods of scale construction

1. Empirical
 - MMPI
 - Strong Vocational Interest Blank
2. Rational
 - California Psychological Inventory
3. Theoretical
 - Measures of Need Achievement (e.g., Jackson PI)
4. Homeogeneous keying
 - Eysenck Personality Inventory
 - NEO
 - BFI
 - TIPI

Empirical

1. Ask items that discriminate known groups
 - People in general versus specific group
 - Choose items that are maximally independent and that have highest validities
2. Example:
 - MMPI
 - Strong-Campbell
 - sex and ethnic differences in personality and music
3. Problem:
 - What is the meaning of the scale?
 - Need to develop new scale for every new group

Early examples of SAPA analysis

1. Development of SAPA reported by [Revelle, Wilt & Rosenthal \(2010\)](#)
2. Work on music preference by Mellissa Liebert as an honors thesis
3. Other work on Honesty, (Trust [Evans & Revelle, 2008](#)), RWA

Sex differences at item level

Item	effect size
Get overwhelmed by emotions.	0.59
Sympathize with others' feelings.	0.45
Worry about things.	0.43
Feel others' emotions.	0.39
Get stressed out easily.	0.51
Have a soft heart.	0.38
Panic easily	0.50
Inquire about others' well-being.	0.41
Get upset by unpleasant thoughts that come into my mind.	0.38
Get upset easily.	0.37
Am indifferent to the feelings of others.	-0.33
Am not interested in other people's problems.	-0.33
Feel little concern for others.	-0.35
Am not easily bothered by things	-0.35
Love to help others.	0.34
Am not really interested in others.	-0.32
Think of others first.	0.30
Take offense easily.	0.29
Take time out for others.	0.33

Sex differences and music preference

effect size	Item
0.9	Broadway Musicals (e.g. Rent, Cats, Phantom of the Opera)
0.68	Top 40/Pop Vocal Music (e.g. Kelly Clarkson, Madonna, The Black Eyed Peas)
0.65	Broadway, Movie and TV Soundtrack Music in General
0.59	Contemporary Rhythm and Blues (e. g. Whitney Houston, Usher, Alicia Keys)
0.59	Modern Country Music (e.g. Garth Brooks, Dixie Chicks, Tim McGraw)
0.37	Country Music in General
0.37	Movie Soundtracks (e.g. Starwars, Good Will Hunting, Garden State)
0.36	Top 40 Music/Pop in General
0.32	Pop Rock (e.g. Maroon 5, Counting Crows, John Mayer)
0.31	Modern Religious Music (e.g. 4Him, Casting Crowns)
0.3	Soul Rock (e.g. Stevie Wonder, Earth Wind and Fire)
-0.3	Acid Rock (e.g. Pink Floyd, The Doors, Jefferson Airplane)
-0.4	Heavy Metal (e.g. Metallica, Marilyn Manson, System of a Down)

Ethnic differences and music preference

effect size	Item
1.26	Acid Rock (e.g. Pink Floyd, The Doors, Jefferson Airplane)
1	Alternative (e.g. Pearl Jam, Incubus, Radiohead)
0.97	Electronic Music in General
0.91	Rock Music In General
0.87	Jam Bands (e.g. The Grateful Dead, Phish, String Cheese Incident)
0.87	Classic Rock (e.g. The Beatles, The Rolling Stones, Led Zeppelin)
0.85	Country Rock (e.g. The Allman Brothers, Lynyrd Skynyrd)
0.61	Electronic Dance Music (e.g. DJ Tiesto, Paul Van Dyk, Keoki)
0.59	Folk Music in General (e.g. Bob Dylan, Iron and Wine, Simon and Garfunkel)
0.57	Pop Rock (e.g. Maroon 5, Counting Crows, John Mayer)
0.56	Country Music in General
0.51	Bluegrass (e.g. Alison Krauss, Lester Flatt, Nickel Creek)
-0.56	Contemporary Rhythm and Blues (e. g. Whitney Houston, Usher, Alicia Keys)
-0.6	Blues in General (e.g. Ray Charles, Stevie Ray Vaughn, B.B. King)
-0.63	Instrumental Hip-Hop (e.g. DJ Hi-Tek, RJD2, Prefuse 73)
-0.64	Gospel Soul (e.g. Aretha Franklin, Solomon Burke)
-0.67	Soul in General (e.g. Otis Redding, Marvin Gaye)
-0.84	Religious Music in General
-1.04	Soul Rock (e.g. Stevie Wonder, Earth Wind and Fire)
-1.11	Rhythm and Blues in General
-1.43	Religious Gospel (e.g. Andre Crouch, Gospel Quartet)

Rational Keying

1. Ask items with direct content relevance
2. Example: California Psychological Inventory
3. Problems
 - Not all items predict in obvious way
 - Need evidence for validity
 - Easy to fake

Theoretical Keying

1. Ask items with theoretical relevance
2. Example: Jackson Personality Research Form
3. Problems:
 - Theoretical circularity
 - Need evidence for validity

Homogeneous Keying

1. Select items to represent single domain
2. Exclude items based upon internal consistency
3. Examples:
 - 16PF
 - EPI/EPQ,
 - NEO/NEO-PIR
4. Problems
 - Garbage In, Garbage Out
 - Need evidence for validity

Methods of Homogeneous keying

1. Cluster analysis (e.g. iclust)
2. Principal Components analysis (e.g., pca)
3. Factor analysis (e.g., fa)

The Hase and Goldberg and Goldberg studies

1. Hase and Goldberg: a direct comparison of different techniques
 - Differential validity of scale construction
 - Factor analytic
 - Empirical Group discrimination
 - Intuitive theoretical
 - Intuitive rational
 - Stylistic-psychometric
 - Random
2. 200 University Freshman women
3. CPI items and 13 criteria

Hase and Goldberg: 13 criteria

1. Sorority Membership
2. An experimental measure of conformity
3. Peer ratings of
 - Dominance
 - Sociability
 - Responsibility
 - Psychological Mindedness
 - Femininity
4. Peer ratings of how well known the person is
5. Average number of dates per month
6. College Grade Point Average
7. College Achievement relative to ability
8. College Major
9. College Dropout

Does it make a difference?

1. Hase and Goldberg ([Hase & Goldberg, 1967](#)) No
2. [Goldberg \(1972\)](#) YES

Hase and Goldberg; mean values)

Original Hase and Goldberg showed no difference between methods, except that stylistic and random were much worse.

Factor	var	n	mean	sd	median	trimmed	mad	min	max	range	se
Theoretical	1	13	0.25	0.18	0.27	0.25	0.13	-0.05	0.57	0.62	0.05
Rational	2	13	0.25	0.16	0.26	0.25	0.18	0.01	0.52	0.51	0.04
Empirical	3	13	0.26	0.16	0.32	0.27	0.09	-0.08	0.49	0.57	0.04
Stylistic	4	13	0.26	0.11	0.30	0.26	0.06	0.04	0.44	0.40	0.03
Random	5	13	0.13	0.12	0.11	0.13	0.12	-0.07	0.35	0.42	0.03
	6	13	0.10	0.12	0.11	0.10	0.13	-0.08	0.30	0.38	0.03

○○○○○○
○○○○

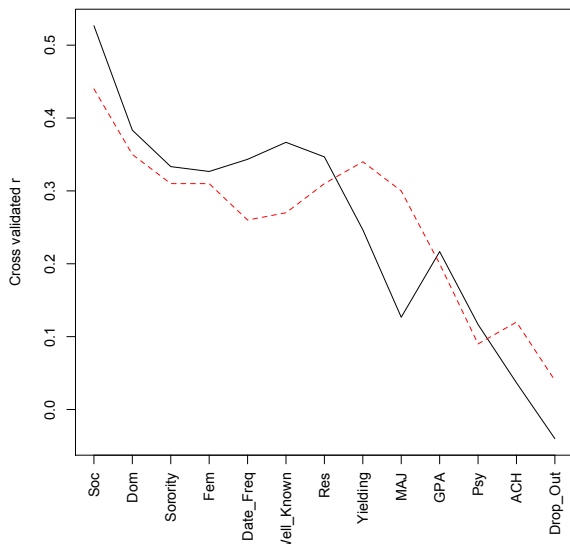
○○○○●

○○○○○○

○○○

Prediction depends upon criteria: Goldberg: 72

Hase and Goldberg



Another factorial versus empirical example

1. Sapa Personality Inventory best 135 item ([Condon \(2018\)](#))
 - From 1800 IPIP items, found that 696 were most common
 - Factor structure of these 696 showed 135 very clear items
 - 5/27 factors, but not hierarchically organized
2. 4,000 subjects on spi 135 in the *psych* package
3. 135 items plus 10 criteria variables

ooooo
oooo

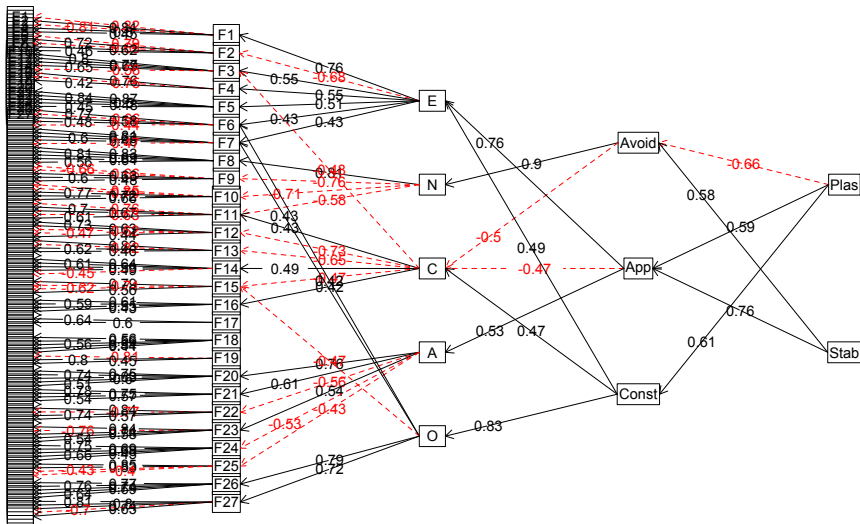
ooooo

●●oooo

ooo

Applying the 'Bass Ackward' function

BassAckward



Score the Big 5 and predict the criteria

R code

```
spi.scales <- scoreItems(spi.keys[1:5],spi)

cor2(spi[1:10],spi.scales$scores)
```

	Agree	Consc	Neuro	Extra	Open
age	0.18	0.19	-0.17	-0.02	0.13
sex	0.17	0.09	0.24	0.06	-0.15
health	0.11	0.23	-0.34	0.21	0.07
p1edu	0.02	-0.02	-0.05	0.06	0.07
p2edu	0.02	-0.04	-0.04	0.08	0.07
education	0.13	0.12	-0.17	-0.01	0.15
wellness	0.11	0.12	-0.02	0.11	0.01
exer	0.07	0.19	-0.18	0.13	0.10
smoke	-0.09	-0.11	0.06	0.06	0.09
ER	-0.03	-0.01	0.12	0.02	-0.02

ooooo
oooo

ooooo

ooo●oo

ooo

What about multiple R

R code

```
summary(setCor(1:10,11:15,data=spi.scores.df,plot=FALSE))
```

```
summary(setCor(1:10,11:15,data=spi.scores.df,plot=FALSE))
```

Multiple Regression from raw data

```
setCor(y = 1:10, x = 11:15, data = spi.scores.df, plot = FALSE)
```

Multiple Regression from matrix input

Beta weights

	age	sex	health	p1edu	p2edu	education	wellness	exer	smoke	ER
Agree	0.16	0.162	0.0063	0.015	0.014	0.116	0.0631	-0.0053	-0.083	-0.025
Consc	0.13	0.103	0.1715	-0.034	-0.049	0.065	0.1053	0.1613	-0.082	0.016
Neuro	-0.14	0.286	-0.2721	-0.036	-0.033	-0.147	0.0302	-0.1247	0.058	0.131
Extra	-0.11	0.086	0.1436	0.047	0.061	-0.086	0.0918	0.0876	0.084	0.050
Open	0.12	-0.122	0.0126	0.058	0.057	0.142	0.0031	0.0675	0.090	-0.012

Multiple R

	age	sex	health	p1edu	p2edu	education	wellness	exer	smoke	ER
	0.306	0.360	0.405	0.098	0.109	0.264	0.170	0.267	0.181	0.133

Multiple R2

	age	sex	health	p1edu	p2edu	education	wellness	exer	smoke	ER
	0.0939	0.1296	0.1642	0.0096	0.0118	0.0699	0.0288	0.0711	0.0329	0.0176

Cohen's set correlation R2

```
[1] 0.4
```

Squared Canonical Correlations

```
[1] 0.2394 0.1332 0.0620 0.0298 0.0079
```

Compare with best scales

R code

```
bs <- bestScales(spi[11:145], spi[1:10], dictionary=spi.dictionary, n.iter=20)
```

```
Call = bestScales(x = spi[11:145], criteria = spi[1:10], dictionary = spi.dictionary,
  n.iter = 20)
```

	derivation.mean	derivation.sd	validation.m	validation.sd	final.valid
age	0.37	0.014	0.360	0.021	0.35
sex	0.36	0.014	0.354	0.021	0.35
health	0.44	0.016	0.432	0.017	0.43
p1edu	0.15	0.030	0.124	0.026	NA
p2edu	0.17	0.027	0.098	0.024	NA
education	0.32	0.022	0.285	0.026	0.18
wellness	0.25	0.014	0.213	0.026	0.22
exer	0.32	0.018	0.283	0.023	0.30
smoke	0.28	0.016	0.255	0.024	0.27
ER	0.17	0.025	0.127	0.025	0.12

Repeat from setCor:

Multiple R

age	sex	health	p1edu	p2edu	education	wellness	exer	smoke	ER
0.306	0.360	0.405	0.098	0.109	0.264	0.170	0.267	0.181	0.133

○○○○○
○○○○

○○○○○

○○○○○●

○○○

What are the items?

Criterion = age

	Freq	mean.r	sd.r	item_id	item	item_scale	resp_type
q_4296	20	-0.23	0.01	q_4296	Tell a lot of lies.	EPQ:P	reg
q_4249	20	-0.21	0.02	q_4249	Would call myself a nervous person.	EPQ:N	reg
q_501	20	-0.21	0.01	q_501	Cheat to get ahead.	IPIP	reg
q_1024	18	-0.21	0.01	q_1024	Hang around doing nothing.	IPIP	reg
q_803	19	0.20	0.02	q_803	Express myself easily.	IPIP	reg
q_1081	18	-0.20	0.01	q_1081	Have difficulty expressing my feelings.	IPIP	reg

Criterion = sex

	Freq	mean.r	sd.r	item_id	item	item_scale	resp_type
q_1505	20	0.29	0.01	q_1505	Panic easily.	IPIP	re
q_979	20	0.29	0.01	q_979	Get overwhelmed by emotions.	IPIP	re
q_793	20	0.25	0.01	q_793	Experience my emotions intensely.	IPIP	re
q_174	20	-0.24	0.01	q_174	Am not easily affected by my emotions.	IPIP	re
q_1989	18	0.21	0.01	q_1989	Worry about things.	IPIP	re
q_851	19	0.21	0.01	q_851	Feel sympathy for those who are worse off than myself.	IPIP	re
q_1763	18	0.21	0.02	q_1763	Sympathize with others feelings.	IPIP	re
q_4252	18	0.20	0.01	q_4252	Am a worrier.	EPQ:N	re

Criterion = health

	Freq	mean.r	sd.r	item_id	item	item_scale	resp_type
q_820	20	0.35	0.02	q_820	Feel comfortable with myself.	IPIP	reg
q_2765	20	0.35	0.01	q_2765	Am happy with my life.	IPIP	reg
q_811	20	-0.34	0.01	q_811	Feel a sense of worthlessness or hopelessness.	IPIP	reg
q_578	20	-0.34	0.02	q_578	Dislike myself.	IPIP	reg
q_1371	20	0.32	0.02	q_1371	Love life.	IPIP	reg
q_56	20	0.28	0.01	q_56	Am able to control my cravings.	IPIP	reg
q_1505	20	-0.27	0.01	q_1505	Panic easily.	IPIP	reg
q_808	18	-0.26	0.02	q_808	Fear for the worst.	IPIP	reg

Validating SAPA using peer ratings (Zola, Condon & Revelle, 2021)

1. From 158,000 SAPA subjects
2. 1,554 peer reports on 921 targets
3. Asked a short form for peer ratings
4. Item by rating correlations in zola

○○○○○
○○○○

○○○○○

○○○○○○

●●○

The peer rating items from Zola et al. (2021)

\$Agreeableness

	item_id	item
q_3837-	q_3837	Is indifferent to others feelings.
q_3838	q_3838	Likes to help others.
q_3853-	q_3853	Tells people when they are frustrated.
q_3854	q_3854	Is patient and polite.

\$Conscientiousness

	item_id	item
q_3845-	q_3845	Neglects their work/duties.
q_3846	q_3846	Works hard.
q_3851-	q_3851	Is disorganized.
q_3852	q_3852	Likes things to be just right.

\$Stability

	item_id	item
q_3833-	q_3833	Is moody, easily upset.
q_3834	q_3834	Is composed, not easily annoyed.
q_3835-	q_3835	Is fearful, panics easily.
q_3836	q_3836	Faces danger confidently.

\$Extraversion

	item_id	item
q_3831-	q_3831	Prefers to let others lead.
q_3832	q_3832	Is assertive, takes charge.
q_3839-	q_3839	Keeps others at a distance.
q_3840	q_3840	Enjoys being with people.

\$IntellectOpenness

	item_id	item
q_3847-	q_3847	Cant handle a lot of information.
q_3848	q_3848	Understands things quickly.
q_3849-	q_3849	Is disinterested in abstract ideas.
q_3850	q_3850	Believes in the importance of art.

\$HonestyHumility

	item_id	item
q_3841-	q_3841	Disregards rules to get ahead.

Multi-Trait, Multi-Method correlations

R code

```
scores <- psych::scoreOverlap(zola.keys[c(1:5,33:37)],zola)
lowerMat(scores$cor)
```

Table: The Zola et al. (2021) MTMM correlations

Variable	Agribl	Cnscn	Nrtcs	Extrv	Opnnn	Agribl	Cnscn	Stblt	Extrv	IntlO
Agreeableness	1.00									
Conscientiousness	0.28	1.00								
Neuroticism	-0.12	-0.18	1.00							
Extraversion	0.25	0.12	-0.25	1.00						
Openness	0.08	0.05	-0.09	0.13	1.00					
Agreeableness	0.47	0.10	-0.01	0.00	-0.09	1.00				
Conscientiousness	0.15	0.55	-0.12	-0.01	-0.04	0.18	1.00			
Stability	0.13	0.16	-0.58	0.05	0.07	0.25	0.25	1.00		
Extraversion	0.23	0.28	-0.27	0.49	0.11	0.07	0.23	0.22	1.00	
IntellectOpenness	0.14	0.08	-0.15	0.09	0.30	0.19	0.24	0.27	0.15	1.00

- Condon, D. M. (2018). The SAPA Personality Inventory: An empirically-derived, hierarchically-organized self-report personality assessment model. *PsyArXiv*.
- Evans, A. M. & Revelle, W. (2008). Survey and behavioral measurements of interpersonal trust. *Journal of Research in Personality*, 42(6), 1585–1593.
- Goldberg, L. R. (1972). Parameters of personality inventory construction and utilization: A comparison of prediction strategies and tactics. *Multivariate Behavioral Research Monographs*. No 72-2, 7.
- Hase, H. D. & Goldberg, L. R. (1967). Comparative validity of different strategies of constructing personality inventory scales. *Psychological Bulletin*, 67(4), 231–248.
- 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.

Zola, A., Condon, D. M., & Revelle, W. (2021). [The Convergence of Self and Informant Reports in a Large Online Sample.](#) *Collabra: Psychology*, 7(1). 25983.