Construct Validity

Dr. Ian Hussey

Digitalisation of Psychology

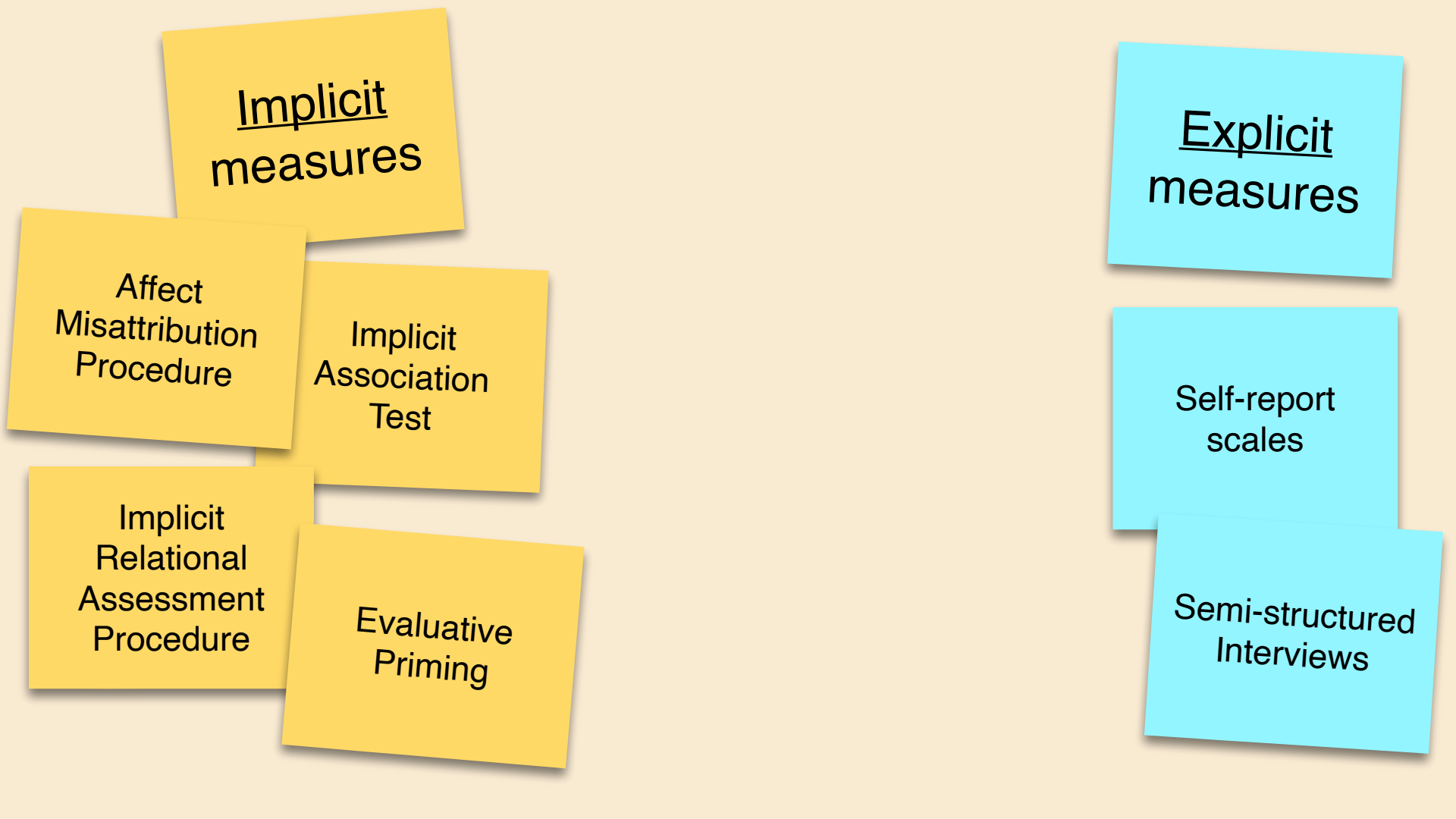
Implicit measures

Implicit Association Test

Explicit measures

Self-report scales

Semi-structured Interviews


Implicit measures
Affect Misattribution Procedure
Implicit Association Test
Implicit Relational Assessment Procedure
Evaluative Priming
Explicit measures
Self-report scales
Semi-structured Interviews


Implicit measures
Behavioral / cognitive tasks ???
Explicit measures
Affect Misattribution Procedure
Implicit Association Test
Dot probe task
Iowa gambling task
Implicit Relational Assessment Procedure
Evaluative Priming
Shooter bias task
Self-report scales
Semi-structured Interviews

…Why?

…Why?

Rarely/never:

“They have been shown to measure something different”

…Why?

Rarely/never:

“They have been shown to measure something different”

Often:

“It’s in the name!”
“Because I say so!”

Schmeasurement

Systematic disinterest in validity

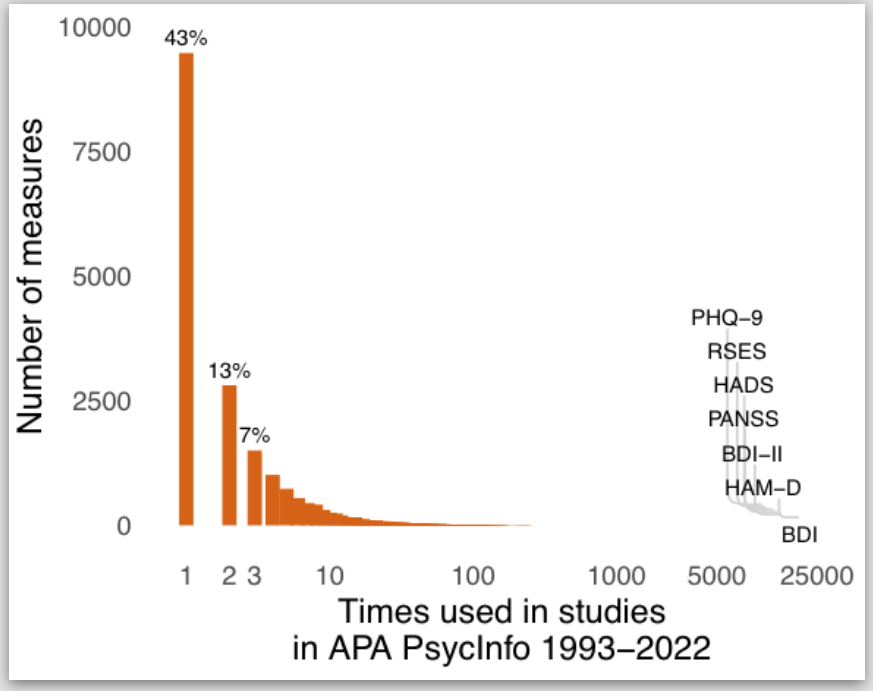
Through inaction

Through negligence & hacking

Measurement TO DO list (essential steps)

- Create whatever items or task you want
- Name it

Disposable measures

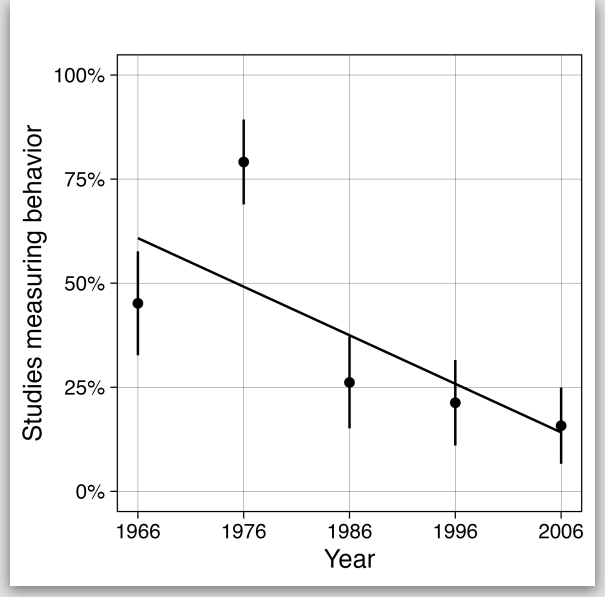


43% of measures used just 1 time

80% used ≤ 10 times

60,000 measures since 1896
(Elson, Hussey, Alsalti, & Arslan, 2023)

Psychology:
The science of mind & behavior?



"Psychology as the Science of Self-Reports and Finger Movements:
Whatever Happened to Actual Behavior?"

Baumeister et al. (2007)

Common measurement fallacies

Common measurement fallacies

A measure's name tells you what it measures

Nominal Fallacy

Common measurement fallacies

A measure's name tells you what it measures

Nominal Fallacy

Measures with the same name measure the same thing

Jingle Fallacy

Common measurement fallacies

A measure's name tells you what it measures

Nominal Fallacy

Measures with the same name measure the same thing

Jingle Fallacy

Measures with different names measure different things

Jangle Fallacy

Have you wished you were dead or wished you could go to sleep and not wake up?

Do you ever feel as if you do not want to live anymore?

Jangle Fallacy

Measures with different names measure different things

Conscientiousness
‘Orderliness, diligence, and dependability’

Grit
‘Perseverance and passion for long-term goals’
Duckworth et al. 2007

They measure the same thing: $r = .84$
Meta-analysis by Crede et al. 2017

> Policy implications

Jangle Fallacy

Measures with the same name measure the same thing

Depression

Poor content overlap between popular depression scales

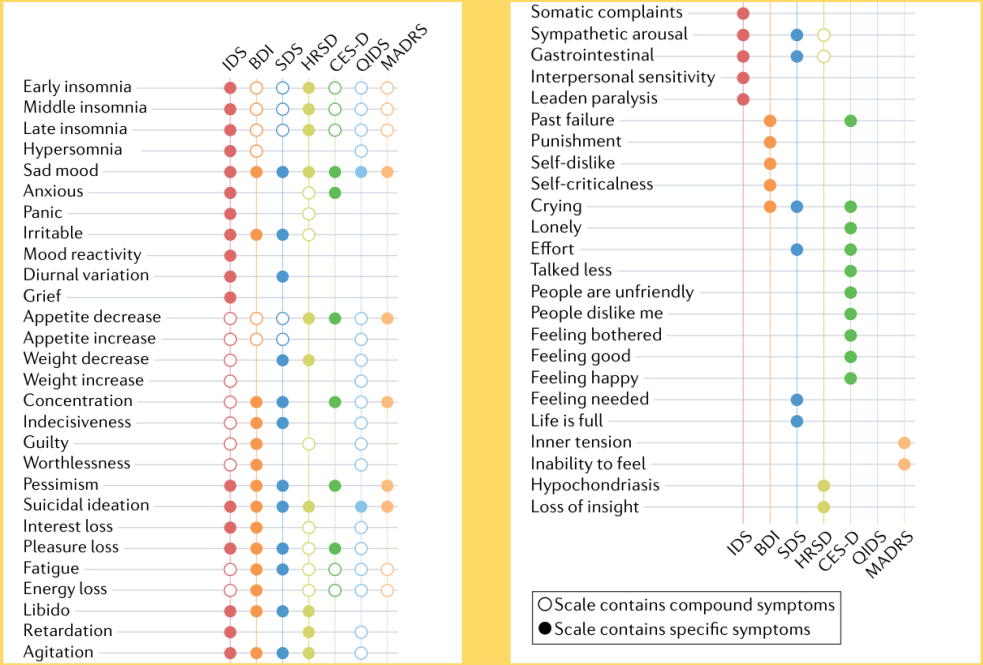
If they ask about different symptoms, do they measure the same disorder?

Fried et al. 2016, Fried, 2017, Fried et al. 2022

> Diagnosis & treatment implications

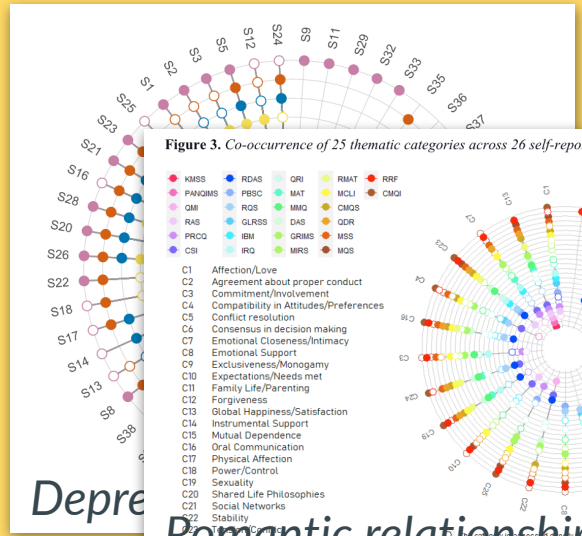
Jingle Fallacy

Measures with the same name measure the same thing



Jingle Fallacy

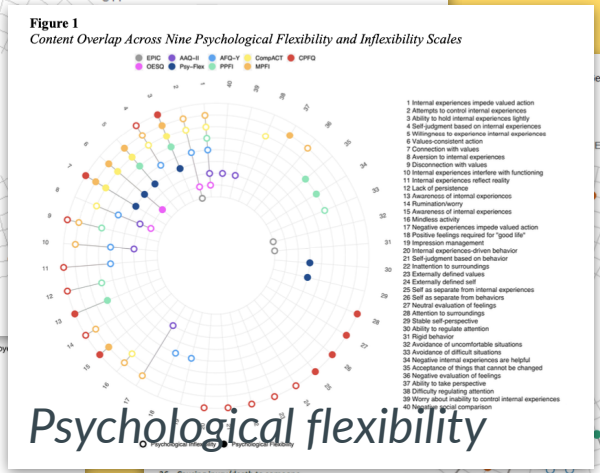
Measures with the same name measure the same thing



Depre

Romantic relationships

Anxiety



Psychological flexibility

Trauma

Jingle Fallacy

A measure's name tells you what it measures

Both Jingle and Jangle fallacies are sub-types of Nominal Fallacy

Measures with the same name measure the same thing

Measures with different names measure different things

Nominal Fallacy

Common measurement fallacies

A measure's name tells you what it measures

Nominal Fallacy

Measures with the same name measure the same thing

Jingle Fallacy

Measures with different names measure different things

Jangle Fallacy

What measurement fallacies may be at work here?

Measures of Implicit Attitudes / Bias (vs. explicit attitudes / self-report)

- Different implicit measures correlate poorly with one another (r = .30)
- Implicit and explicit measures correlate well with one-another (r = .60)

What measurement fallacies may be at work here?

Correlations between different Theory of Mind tasks are near zero

Warnell & Redcay (2019)

Table 6
Relations among theory of mind tasks in adulthood.

	Spontaneous ToM	Belief Reasoning Speed	Pragmatics	Adult Eyes	Higher-Order ToM
Spontaneous ToM	–	−0.023	0.015	−0.115	0.125
Belief Reasoning Speed		–	0.056	0.115	0.048
Pragmatics			–	0.068	−0.051
Adult Eyes				–	−0.069
High-Order ToM					–

Note. Correlation values are Pearson's *r* controlling for age and gender.

Theory of Mind

What measurement fallacies may be at work here?

The Dot-Probe Task is used as a measure of attentional bias in anxiety

- Core foundation of experimental psychopathology

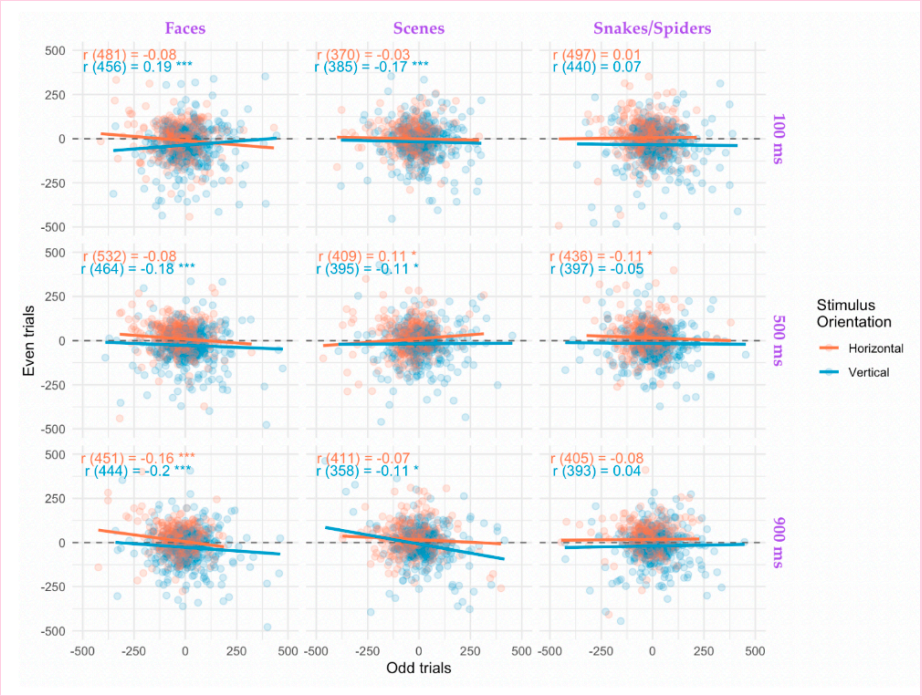
There is no one ‘Dot-Probe’ task, but many variants

- Stimuli domains
- Stimulus orientations
- Scoring comparisons
- Stimulus onset asynchrony

Dot-Probe Task

What measurement fallacies may be at work here?

Comparing 36 versions
in 9000 participants,
reliability was consistently near zero.
Xu et al. 2022



Dot-Probe Task

What measurement fallacies may be at work here?

Social Media is a Major Cause of the Mental Illness Epidemic in Teen Girls. Here's the Evidence.

Journalists should stop saying that the evidence is just correlational

JON HAIDT

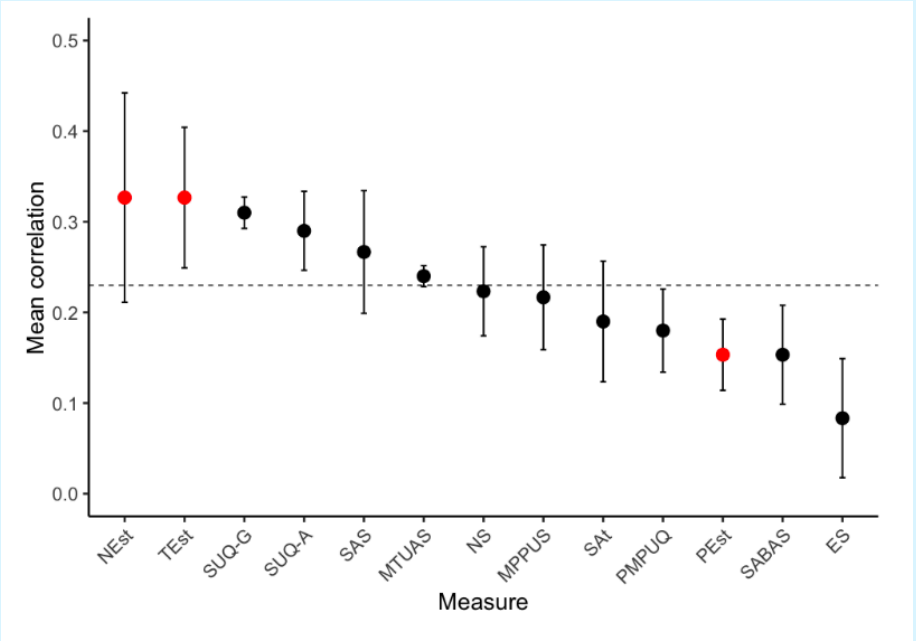
22 FEB 2023

Smartphone Use

What measurement fallacies may be at work here?

Self-reported smartphone usage
is weakly correlated
with actual usage

Ellis et al. 2019



Smartphone Use

What measurement fallacies may be at work here?

How well do scores on individual items of the BDI-II and other depression scales measure actual changes in behaviour?

10. Crying

0	I don't cry anymore than I used to.
1	I cry more than I used to.
2	I cry over every little thing.
3	I feel like crying, but I can't.

13. Indecisiveness

0	I make decisions about as well as ever.
1	I find it more difficult to make decisions than usual.
2	I have much greater difficulty in making decisions than I used to.
3	I have trouble making any decisions.

16. Changes in Sleeping Pattern

0	I have not experienced any change in my sleeping pattern.
1a	I sleep somewhat more than usual.
1b	I sleep somewhat less than usual.
2a	I sleep a lot more than usual.
2b	I sleep a lot less than usual.
3a	I sleep most of the day.
3b	I wake up 1–2 hours early and can't get back to sleep.

18. Changes in Appetite

0	I have not experienced any change in my appetite.
1a	My appetite is somewhat less than usual.
1b	My appetite is somewhat greater than usual.
2a	My appetite is much less than before.
2b	My appetite is much greater than usual.
3a	I have no appetite at all.
3b	I crave food all the time.

Depression

What measurement fallacies may be at work here?

In the field of Sex Research, many researchers wish to understand the context, causes, consequences, etc. of pornography use.

What could go wrong here, in terms of measurement validity?

[hint: think stupider]

Pornography use

What measurement fallacies may be at work here?

“57 of the 100 studies sampled
relied on idiosyncratic measures of pornography use
that were not found in other published research.”
Kohut et al. 2020

Issues with measuring pornography use

- What is “pornography”?
- What is “use”?
- What timeframe? (day, week, month, year, lifetime)
- What dimension? (frequency, first usage, specific content)

Pornography use

What measurement fallacies may be at work here?

“We controlled for socio-economic status”

SES

What measurement fallacies may be at work here?

“We controlled for socio-economic status”

There is no agreement how to measure SES
Antonoplis (2022)

Indicators and Modeling Procedures of Socioeconomic Status

Operationalizations	Percentage	Number (n=495)	Example
Indicators			
None given	0.20	1	–
Assets/Housing	8.10	40	Home value, Own home/car, Neighborhood wealth/cohesion
Composite	5.25	26	Hollingshead Index, Brazilian ABEP
Demographic	0.40	2	Race, Gender
Education	27.47	136	Personal education (highest degree attained)
Family structure	4.24	21	Teen mom, Father present, Number of children
Income/Poverty	22.83	113	Family income, Neighborhood poverty rate
Mannerisms	2.83	14	Extracurricular activities, Verb use, Name, Clothing
Occupation	16.36	81	Parental Duncan’s SEI
Subjective	10.30	51	MacArthur ladder
Uncategorized	2.02	10	Health insurance, Welfare/aid, Food insecurity

SES

Credibility checks

- Ad hoc measure?
    - ie created just for this study
- Consider Nominal/Jingle/Jangle fallacies

- Construct validity evidence
    - Reported in the article
    - Previous research
    - *Assume this evidence is also hacked and biased to some degree!*
    - *Evidence of predictive validity is not evidence of construct validity!*