GENDER ROLE ATTITUDES IN THE PHILIPPINES AND ELSEWHERE

Harry BG Ganzeboom
(with Linda Luz Guerrero, Iremae Labucay, Gerardo Sandoval)

Department of Sociology, VU University Amsterdam

WAPOR
Seoul KR
July 29, 2023

Earlier presentations

- ISA-RC33, Nicosia (CY), Sep 9, 2021 (online)

Gender Role Attitudes GRA

- GRA is a 6-8 item likert-style attitude instrument on desirable gender roles / distribution of household tasks.
- It has been part of ISSP (module: Family and Gender Roles) since 1988, and is also included in EVS, WVS, US-GSS, ESS and probably many other social attitudes surveys.
- GRA is a key instrument in the sociologies of families, gender, and more broadly: value climate. We would like this instrument to have good measurement properties:
  - Validity – does it measure what we intend in to measure?
  - Reliability – is the measurement stable – does it show the same results wheb repeated?
- GRA is also an excellent instrument to introduce students to the art and craft of comparative measurement and comparative analysis.

Blasius & Thiessen (2006)

- Blasius, Jörg & Victor Thiessen. 2006. “Assessing Data Quality and Construct Comparability in Cross-National Surveys.” *European Sociological Review* 22 (3): 229–42. https://doi.org/10.1093/esr/jci054.
- Blasius & Thiessen examine the measurement quality of the GRA data in ISSP 1994 (11 items, 24 countrIes) using Multiple Correspondence Analysis, with far-reaching and dramatic conclusions for the PH data: ***Russia, and especially the Philippines (together they constitute the fifth cluster), raise even stronger doubts about the quality of the data. In the case of the Philippines we find clear evidence of response styles for the moderate and neutral categories (…) as well as a response style that can be described as a mixture of strong agreement and strong disagreement irrespective of the content of the items.(…) it probably makes no sense to compare support for single-/dual-earner family structures in the data from the Philippines with that of any from the other countries. (…) Including [data from] Russia and the Philippines seems to make no sense at all, since the quality of data is too low.***
- This statement has created considerable concern with the data collection agency – the Social Weather Stations.
- However, SWS researchers were not able to replicate the methods that Blasius & Thiessen used to arrive at their dramatic conclusion. I want to look at the issues with classical measurement theory.
- A methodologist who recommends you not to use the data is not very helpful.

Aim of the paper and research questions

- Re-examine the measurement quality of the PH-GRA data in the context of repeated ISSP Family and Gender Role surveys (4 waves are now available) and cast this in more interpretable language about measurement quality.
  - Can we reproduce the result of Blasius & Thiessen using more traditional instrument of validity assessment (latent factor analysis) and reliability assessment (CA internal consistency estimate)?
  - Is the finding of low measurement quality in the PH a unique incident, or does it arise more systematically?
  - To what extent does low measurement quality impact substantive conclusions?

Conclusions

- The PH GRA data are characterized by a low level of reliability (internal consistency estimate CA).
- This is NOT a unique feature of the ISSP-PH 1994:
  - It also arises in subsequent waves of the ISSP Gender Role module in PH.
  - It is not confined to PH, but also arises in the ISSP in other (developing) nations: BR, CN, IN, JP, MX, TH, TW and VE
- There is no evidence that strong multidimensionality (=systematic error) is responsible for the low measument quality of the GRA index in these countries. Rather, it arises because some of the items “do not work” in these countries – they are ambiguous. This leads to unreliability (random noise).
- However, low reliability is less harmful than one would think. It does not affect aggregate (mean) scale scores.

Table 1: Gender Role Attitude items in ISSP 1988-2012: single vd dual earnership

	1988	1994	2002	2012
a A working mother can establish just as warm and secure a relationship with her children as a mother who does not work (dual)	v4	v4	v4	v5
b A pre-school child is likely to suffer if his or her mother works (single)	v5	v5	v5	v6
c All in all, family life suffers when the woman has a full-time job (single)	v6	v6	v6	v7
d A job is all right, but what most women really want is a home and children (single)	v8	v7	v7	v8
e Being a housewife is just as fulfilling as working for pay (single)	v9	v8	v8	v9
f Having a job is the best way for a woman to be an independent person (dual)	v10	v9	v9	
h Both the man and woman should contribute to the household income (dual)	v11	v11	v10	v10
i A man's job is to earn money; a woman's job is to look after the home and family (single)		v12	v11	v11

GRA ITEM MEANS BY ISSP WAVE

WAVE		ax	b	c	d	e	fx	hx	i
1988	Mean	3.49	2.57	2.72	2.88	2.66	3.44	3.40	
	N	15343	15303	15344	14977	15023	15210	15214	
1994	Mean	3.57	2.74	2.81	2.91	2.81	3.53	3.79	3.19
	N	31024	30791	30898	30062	29998	30412	31109	31247
2002	Mean	3.67	2.79	2.88	2.87	2.85	3.73	3.98	3.35
	N	42724	42286	42399	41440	41090	42095	42779	42698
2012	Mean	3.74	2.98	3.07	2.93	2.84		4.02	3.34
	N	55914	55602	55669	54445	53682		56147	56062
Total	Mean	3.65	2.83	2.92	2.90	2.82	3.61	3.90	3.31
	N	145005	143982	144310	140924	139793	87717	145249	130007

Methodology

- Part A: traditional items analysis
  - CA Reliability and Factor analysis
  - Reject items with low reliability / partially invariant measurement equivalence model (SEM).
  - Simplex model for aggregate means to separate true change from measurement unreliability
- Part B: individual level analysis of GRA covariates with correction for attenuation

PART A: Aggregate level analyses

Scale analysis – steps taken

- Step 1: reverse the coding of items a f g
- Step 2: Within-country standardization (**`split file by cntry year`**)
- Step 3: Dimensional analysis: PAF / SEM on available data
- Step 4a: Calculate CA
- Step 4b: Calculate CA for each country
- Step 4c: Calculate CA for country*year
- Step 5: Omit items with low (<.10) item-rest correlations, and repeir step 3-4

➔ Partially invariant multiple groups measurement model.

Correlation Matrix eight GRA items, standardized within country and wave, available data

	C	B	I	D	AX	E	FX	HX
C	1.000	.565	.355	.350	.356	.186	.152	.072
B	.565	1.000	.327	.296	.353	.169	.123	.055
I	.355	.327	1.000	.423	.240	.283	.195	.082
D	.350	.296	.423	1.000	.167	.326	.069	.051
AX	.356	.353	.240	.167	1.000	.052	.222	.146
E	.186	.169	.283	.326	.052	1.000	.119	.151
FX	.152	.123	.195	.069	.222	.119	1.000	.300
HX	.072	.055	.082	.051	.146	.151	.300	1.000

Results of reliability analysis

- Low reliabilities emerge for a number of countries, not only PH: BR, CN, IN, JP, MX, PH, TH, TW and VE.
- The low reliabilities are consistently low when a country has repeated data – it is a systematic phenomenon

Table 2: Reliability Statistics

		Cronbach's Alpha
Rest	1994	.750
Rest	2002	.735
Rest	2012	.752
BR	2002	.495
CN	2012	.487
IN	2012	.192
JP	1994	.380
	2002	.448
	2012	.546
MX	2002	.519
	2012	.529
PH	1994	.300
	2002	.395
	2012	.434
TW	2002	.342
	2012	.467
VE	2012	.567

Correlation Matrix eight GRA items, standardized within country and wave, available data from countries with low reliabilities

	C	B	I	D	AX	E	FX	HX
C	1.000	.395	.187	.217	.139	.034	.014	-.024
B	.395	1.000	.161	.164	.108	.037	-.018	-.051
I	.187	.161	1.000	.299	.059	.133	.043	-.032
D	.217	.164	.299	1.000	.006	.202	-.042	-.064
AX	.139	.108	.059	.006	1.000	-.132	.131	.097
E	.034	.037	.133	.202	-.132	1.000	-.064	-.116
FX	.014	-.018	.043	-.042	.131	-.064	1.000	.253
HX	-.024	-.051	-.032	-.064	.097	-.116	.253	1.000

Factor analysis: PAF with one and two dimensions on available (pairwise missing) data

	Countries with HIGH reliabilities		Countries with LOW reliabilities		Countries with HIGH reliabilities	Countries with LOW reliabilities	
	Factor		Factor		Factor	Factor	Factor
	1	2	1	2	1	1	1
C	.793	-.109	.611	.132	.716	.573	.578
B	.749	-.124	.520	.077	.667	.510	.509
I	.565	.106	.404	-.064	.623	.421	.428
D	.548	-.001	.442	-.194	.551	.460	.459
AX	.447	.110	.183	.353	.501	.124	.183
E	.295	.200	.155	-.330	.394	.183	.220
FX	.087	.516	.024	.382	.316	-.021	.026
HX	-.056	.576	-.057	.400	.208	-.095	-.053
Correlation: 0.377			Correlation: -0.068				

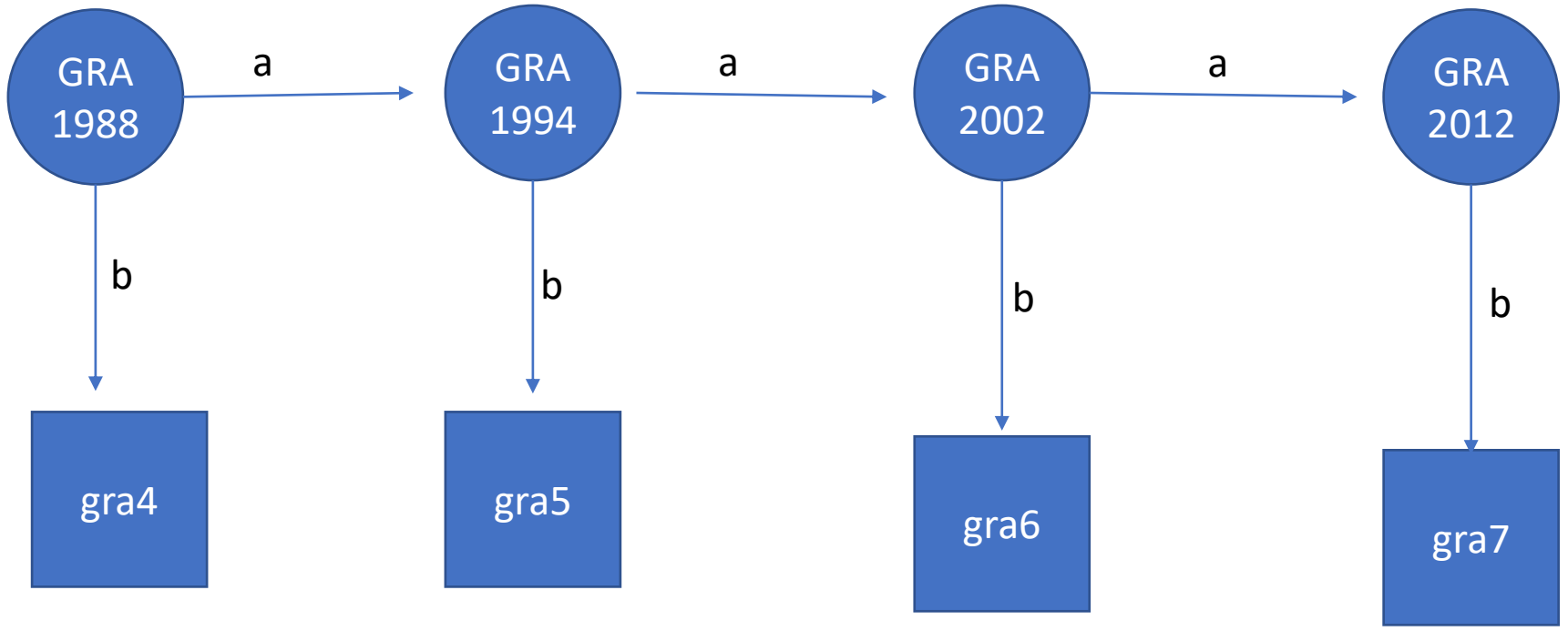
Country means by wave (extreme parts)

CNTRY	GRA	Gra1988	Gra1994	Gra2002	Gra2012
DK	3.97			3.87	4.07
DE2	3.96		3.89	3.94	4.13
IS	3.80				3.80
SE	3.69		3.51	3.68	3.94
FR	3.66			3.59	3.72
CA	3.55		3.55		3.54
FI	3.53			3.34	3.73
BE	3.52				3.52
NO	3.52		3.29	3.54	3.82
CY	3.45			3.45	
HR	3.45				3.45
IL	3.39		3.44	3.39	3.33
ES	3.36		3.20	3.39	3.50
NI	3.34		3.33	3.34	
UK	3.33	3.21	3.34	3.35	3.43
SI	3.30		3.10	3.28	3.56
US	3.29	3.20	3.31	3.34	3.32
VE	3.03				3.03
LV	3.03			3.10	2.96
PL	3.03		2.85	3.13	3.17
CN	3.01				3.01
BG	3.01		2.91	3.04	3.08
BR	3.00			3.00	
AR	2.98				2.98
PH	2.92		2.94	2.99	2.81
CL	2.91			2.89	2.94
MX	2.91			2.97	2.85
RU	2.89		2.72	3.04	2.93
KR	2.88				2.88
IN	2.82				2.82
HU	2.82	2.75	2.67	2.96	3.00
TR	2.81				2.81
Total	3.23	3.02	3.17	3.27	3.28

Table 5a: Correlations of country means by GRA

	GRA1988	GRA1994	GRA2002	GRA2012
GRA1988	1 N=9			
GRA1994	0.884 N=9	1 N=24		
GRA2002	0.694 N=7	0.921 N=22	1 N=35	
GRA2012	0.543 N=8	0.829 N=21	0.937 N=30	1 N=42

Simplex model of change in GRA



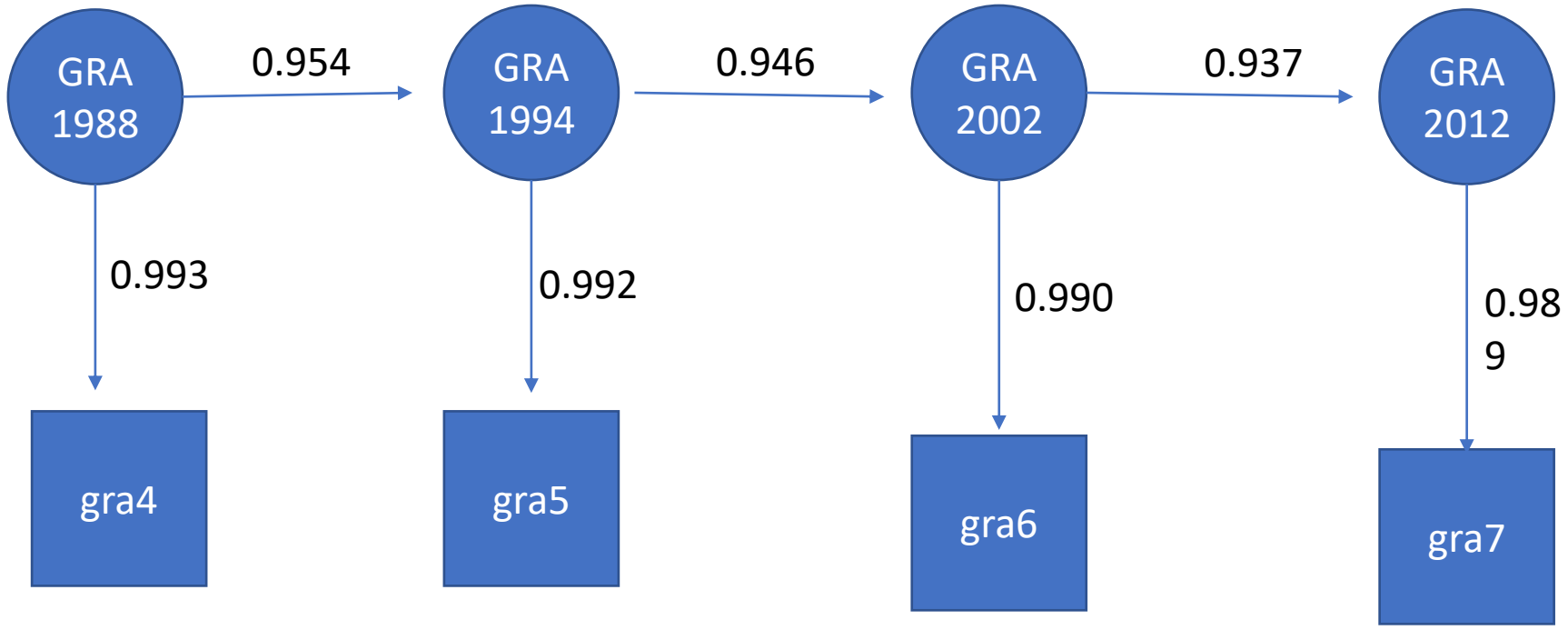
Correlations when there is NO real change, only measurement error

1.00			
0.81	1.00		
0.81	0.81	1.00	
0.81	0.81	0.81	1.00

Correlations when there is real change, NO measurement error

1.00			
0.90	1.00		
0.81	0.90	1.00	
0.73	0.81	0.90	1.00

Simplex model of change in GRA



ISSP, four GR waves (1988-2012), N=48 countries, L2=6.23 , df=7

Table 5b: Correlations of country means by GRA
after wiping out unreliable items per wave (item-rest < .10)

	GRA1988	GRA1994	GRA2002	GRA2012
GRA1988	1 N=9			
GRA1994	0.889 N=9	1 N=24		
GRA2002	0.686 N=7	0.921 N=22	1 N=35	
GRA2012	0.542 N=8	0.829 N=21	0.937 N=30	1 N=42

Conclusions

- The PH GRA data are characterized by a low level of reliability (internal consistency estimate CA).
- This is NOT a unique feature of the ISSP-PH 1994:
  - It also arises in subsequent waves of the ISSP Gender Role module in PH.
  - It is not confined to PH, but also arises in the ISSP in other (developing?) nations: BR, CN, IN, JP, MX, TH, TW and VE
- There is no evidence that strong multidimensionality (=systematic error) is responsible for the low measument quality of the GRA index in these countries. Rather, it arises because some of the items “do not work” in these countries – they are ambiguous. This leads to unreliability (random noise).
- However, low reliability is less harmful than one would think. It does not affect aggregate (mean) scale scores.

Interpretation: How does unreliability affect structural relations?

- Unreliability lowers ('attenuates') correlations.
- The effect on structural relationships depends upon the causal role of the variable:
  - Unreliability in a dependent variable lowers the correlation, explained variance, and statistical power. However, it does not affect (unstandardized) group differences, or regression coefficients.
  - Unreliability in an independent variable also lowers structural effects, such as group difference and regression coefficients [this is why you learned among your 'regression assumptions' that the X should **perfectly** measured]
  - Unreliability in mediating and confounding variables is even more problematic than in a cause or outcome variable. The attenuating effect of unreliability kicks in twice.

Discussion: How does unreliability affect aggregate means?

- Very little.
- This is so, because random measurement error does not affect estimated means: the means become a bit more uncertain but are not biased.
- We observe that the estimated means for countries, including the unreliable ones are indeed very stable (measurement coefficient > .99). The PH scores consistently low in the league table (= in favour on single earnership), which is true for most other unreliable countries.
- Explanation needs to be found in how the unreliable items are perceived in countries that are overwhelmingly in favour of single earnership.

PART B: Individual level analyses

Table 4: Structural relationships with GRA (7 and 4 items)

	GRA 7 items				GRA 4 items			
	CA > 0.60		CA < .60		CA > .60		CA < 0.60	
[YEAR=1994]	-.425	.006	-.181	.015	-.539	.008	-.267	.021
[YEAR=2002]	-.099	.006	-.002	.011	-.129	.007	-.017	.015
[YEAR=2012]	0		0		0		0	
FEMALE	.167	.004	.060	.008	.176	.005	.051	.011
AGE/10	-.082	.001	-.018	.003	-.110	.002	-.038	.004
DEGREE	.093	.002	.061	.003	.125	.002	.095	.004
NONRELIGIOUS	.185	.006	.071	.011	.231	.007	.117	.016

Correcting for attenuation

- In individual level model, unreliability will attenuate standardized effects.
- However, attenuation can be corrected, in multiple ways:
  - Full factor analytic measurement model per country & wave.
  - Randomly parceled factor analytic measurement model per country & wave.
  - Correction for attenuation with fixed value of unreliability.
- To be continued ...