Measurement Invariance of the Dirty Dozen: Student and Working Adult Samples

Yang Yang[1] & John Kulas[2]

[1]

[2] Montclair State University

Author Note

Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line.

Enter author note here.

Correspondence concerning this article should be addressed to Yang Yang, Shanghai, China. E-mail: yangyangsh@outlook.com

Abstract

Now we are evaluating the psychometric properties of the dirty dozen simplified Chinese version by using samples in real settings: job applicants and incumbents (in addition to students). We replicate a previous study using the student sample (Geng, Sun, Huang, Zhu, & Han, 2015), then continue to evaluate with organizational data. We find that the scales are non-invariant.

Keywords: keywords

Word count: X

Measurement Invariance of the Dirty Dozen: Student and Working Adult Samples

Initially we were interested in looking at reliance on student samples. Now we are evaluating the psychometric properties of the dirty dozen (DD) simplified Chinese version by using samples in real settings: job applicants and incumbents (in addition to students). We replicate a previous study using the student sample (Yang gonna send some articles), then continue to evaluate with organizational data. We find that the scales are non-invariant.

SDSME another version (27 items).

All studies investigating psychometric properties of these scales use University students.

Some groups may be expected to exhibit different item-construct associations due to shifting motivational forces.

ITC guidlines for translating and adapting tests recommends looking at possible differences across motives (Commission, 2017). For example,

Yang's references: Church et al. (2011), Schoot, Lugtig, and Hox (2012), Schmitt and Kuljanin (2008), Geng et al. (2015), Grigoras, Butucescu, Miulescu, Opariuc-Dan, and Iliescu (2020), Jonason and Webster (2010)

Methods

We applied three different nested multiple group confirmatory factor analysis models progressing through levels of restriction. These invariance tests were evaluations of configural-, weak-, and finally strict-invariance. The weak invariance models constrained factor loadings to be equal across groups and the strong invariance models also constrained intercepts to be equal across groups. We also look at intercorrelations among items within the samplings. We use the methodology of Grigoras et al. (2020)

Participants

In total 1106 individuals responded to the Dirty Dozen (as well as additional scales not the focus of the current presentation). This total was comprised of 208 working adults low-stakes, 527 working adults high-stakes, and 371 students low-stakes individuals. After screening for undifferentiated responses via the `R` package `careless` (Yentes & Wilhelm, 2021), we retained 1054 respondents who had no more than 6 sequentially identical responses across the 12 total items.

Materials

Dirty dozen version XX. Coefficiant alphas for the scales were 0.81 (Machiavellianism), 0.62 (Psychopathy), and 0.74 (Narcissism), with corresponding corrected item-total correlations ranging from 0.52 to 0.74 (Machiavellianism), 0.31 to 0.51 (Psychopathy), and 0.53 to 0.56 (Narcissism).

Procedure

Decrease in $\Delta\chi^2$ across models indicates a lack of invariance (typically not considered a "good thing"). Multiple indices can be consulted across models, including $\Delta\chi^2$, RMSEA, CFI, TLI, BIC, and AIC. Our determination of level of invariance achieved was informed by a likelihood ration test

Also want to look at correlations of the simplified Chinese version of the DD with the Honesty-Humility subscales (Sincerity, Fairness, Greed Avoidance, and Modesty).

Data analysis

We used R (Version 4.0.3; R Core Team, 2021) and the R-packages *careless* (Version 1.1.3; Yentes & Wilhelm, 2021), *corx* (Version 1.0.6.1; Conigrave, 2020), *foreign* (Version

0.8.80; R Core Team, 2020), *lavaan* (Version 0.6.8; Rosseel, 2012), *papaja* (Version 0.1.0.9997; Aust & Barth, 2020), and *semTools* (Version 0.5.4; Jorgensen, Pornprasertmanit, Schoemann, & Rosseel, 2021) for all analyses.

Results

Mean differences and scale correlations

Inter-scale correlations are presented in Tables X, Y, and Z.

Mean differences were noted across groups for all three dark triad scales: Machiavellianism ($F(2, 1, 101) = 56.59$, $MSE = 0.89$, $p < .001$, $\hat{\eta}^2_G = .093$), Psychopathy ($F(2, 1, 101) = 57.18$, $MSE = 0.62$, $p < .001$, $\hat{\eta}^2_G = .094$), and Narcissism ($F(2, 1, 101) = 31.46$, $MSE = 1.19$, $p < .001$, $\hat{\eta}^2_G = .054$). The Machiavellianism effect was driven by students exhibiting lower scores than both low- (t = -6.18, $p < .0001$) and high-stakes (t = -10.53, $p < .0001$) working adults. The Psychopathy effect occured across all contrasts: students exhibiting lower scores than both low- (t = -3.88, $p < .001$) and high-stakes (t = -10.64, $p < .0001$) working adults, as well as low-stakes working adults having higher Psychopathy scores than high-stakes working adults (t = 4.70, $p < .0001$). For Narcissism students once again exhibited lower scores than both low- (t = -6.35, $p <$ .0001) and high-stakes (t = -7.14, $p < .0001$) working adults.

Measurement Invariance

We looked at structural invariance as well as latent means (Meredith, 1993; Steinmetz, Schmidt, Tina-Booh, Wieczorek, & Schwartz, 2009). The models failed to exhibit metric invariance (Model 2 - Model 1 exhibited a significant Δ on both χ^2 as well as RMSEA)

Not sure how to pull table or identify object elements - `model1` object is too large to navigate easily.

Discussion

There is a lack of measurement invariance

References

Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown*. Retrieved from https://github.com/crsh/papaja

Church, A. T., Alvarez, J. M., Mai, N. T., French, B. F., Katigbak, M. S., & Ortiz, F. A. (2011). Are cross-cultural comparisons of personality profiles meaningful? Differential item and facet functioning in the revised neo personality inventory. *Journal of Personality and Social Psychology*, *101*(5), 1068–1089.

Commission, I. T. (2017). The itc guidelines for translating and adapting tests (second edition).

Conigrave, J. (2020). *Corx: Create and format correlation matrices*. Retrieved from https://CRAN.R-project.org/package=corx

Geng, Y., Sun, Q., Huang, J., Zhu, Y., & Han, X. (2015). Dirty dozen and short dark triad: A chinese validation of two brief measures of the dark triad. *Chinese Journal of Clinical Psychology*, *23*(2), 246–250.

Grigoras, M., Butucescu, A., Miulescu, A., Opariuc-Dan, C., & Iliescu, D. (2020). The measurement invariance of the short dark triad: Implications for high-and low-stakes contexts. *Journal of Individual Differences*, 1–12.

Jonason, P. K., & Webster, G. D. (2010). The dirty dozen: A concise measure of the dark triad. *Psychological Assessment*, *22*(2), 420–432.

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2021). *`semTools`: Useful tools for structural equation modeling*. Retrieved from https://CRAN.R-project.org/package=semTools

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543.

R Core Team. (2020). *Foreign: Read data stored by 'minitab', 's', 'sas', 'spss', 'stata', 'systat', 'weka', 'dBase', ...* Retrieved from https://CRAN.R-project.org/package=foreign

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. Retrieved from https://www.jstatsoft.org/v48/i02/

Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review*, *18*(4), 210–222.

Schoot, R. van de, Lugtig, P., & Hox, J. (2012). Developmetrics: A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*(4), 486–492.

Steinmetz, H., Schmidt, P., Tina-Booh, A., Wieczorek, S., & Schwartz, S. H. (2009). Testing measurement invariance using multigroup cfa: Differences between educational groups in human values measurement. *Quality & Quantity*, *43*(4), 599–616.

Yentes, R. D., & Wilhelm, F. (2021). *Careless: Procedures for computing indices of careless responding*.

Table 1

Scale intercorrelations (all participants).

	1	2	3	4	5	6	7	*M*	*SD*
1. Machiavelliansm	-							1.62	0.78
2. Narcissism	.29***	-						3.69	1.07
3. Psychopathy	.57***	.19***	-					1.51	0.62
4. Fairness	-.34***	-.02	-.45***	-				5.40	0.84
5. GreedAvoidance	-.26***	-.45***	-.24***	.27***	-			3.52	1.14
6. Modesty	-.23***	-.43***	-.17***	.15**	.43***	-		3.72	0.85
7. Sincerity	-.14**	.04	-.04	.23***	.11*	.18***	-	3.85	0.74
8. HonestyHumility	-.38***	-.37***	-.35***	.61***	.77***	.68***	.51***	4.12	0.59

Note. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 2

Scale intercorrelations (working adults low-stakes).

	1	2	3	4	5	6	7	*M*	*SD*
1. Machiavelliansm	-							1.74	0.86
2. Narcissism	.31***	-						3.64	1.10
3. Psychopathy	.61***	.20**	-					1.73	0.74
4. Fairness	-.35***	-.02	-.45***	-				5.27	0.88
5. GreedAvoidance	-.27***	-.45***	-.21**	.30***	-			3.53	1.08
6. Modesty	-.18**	-.42***	-.16*	.24***	.43***	-		3.72	0.82
7. Sincerity	-.15*	.10	-.08	.26***	.13	.14*	-	3.79	0.73
8. HonestyHumility	-.36***	-.33***	-.35***	.68***	.76***	.68***	.52***	4.08	0.59

Note. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 3

Scale intercorrelations (working adults high-stakes).

	1	2	3	4	5	6	7	*M*	*SD*
1. Machiavelliansm	-							1.57	0.74
2. Narcissism	.29***	-						3.71	1.06
3. Psychopathy	.53***	.21***	-					1.42	0.55
4. Fairness	-.33***	-.03	-.40***	-				5.54	0.78
5. GreedAvoidance	-.26***	-.46***	-.29***	.24***	-			3.52	1.19
6. Modesty	-.27***	-.43***	-.18**	.06	.42***	-		3.73	0.88
7. Sincerity	-.14*	-.02	.02	.17*	.09	.21**	-	3.90	0.74
8. HonestyHumility	-.39***	-.41***	-.34***	.54***	.78***	.68***	.50***	4.17	0.58

Note. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 4

Scale intercorrelations (students low-stakes).

	1	2	*M*	*SD*
1. Machiavelliansm	-		2.25	1.21
2. Narcissism	.39***	-	4.24	1.13
3. Psychopathy	.51***	.27***	1.99	1.06

Note. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 5

Machiavellianism

Effect	F	df_1	df_2	MSE	p	$\hat{\eta}^2_G$
Selection	56.59	2	1,101	0.89	$< .001$	.093

Note. Need to figure out how to get all three into one table.

Table 6

Psychopathy

Effect	F	df_1	df_2	MSE	p	$\hat{\eta}^2_G$
Selection	57.18	2	1,101	0.62	$< .001$	.094

Note. Need to figure out how to get all three into one table.

Table 7

Narcissism

Effect	F	df_1	df_2	MSE	p	$\hat{\eta}^2_G$
Selection	31.46	2	1,101	1.19	$< .001$	.054

Note. Need to figure out how to get all three into one table.

Table 8

Measurement invariance summary statistics.

	Df	AIC	BIC	Chisq	Chisq diff	Df diff	Pr(>Chisq)
configural	153	37,059.45	37,639.59	1,407.67	NA	NA	NA
weak	171	37,134.71	37,625.60	1,518.93	111.25	18	0.00
strong	189	37,230.25	37,631.89	1,650.47	131.55	18	0.00
strict	213	37,670.36	37,952.99	2,138.58	488.11	24	0.00

Note. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

	Constrained parameters	Free parameters	comparison model
configural	FMean (=0)	fl+inter+res+var	
Weak/loading invariance	fl+Fmean (=0)	inter+res+var	configural
Strong/scalar invariance	fl+inter	res+var+Fmean*	Weak/loading invariance
strict invariance	fl+inter+res	Fmean*+var	Strong/scalar invariance

Note. fl= factor loadings, inter = item intercepts, res = item residual variances, Fmean = mean of latent variable, var = variance of latent variable
*Fmean is fixed to 0 in group 1 and estimated in the other group(s)

Figure 1. Steps for measurement invariance (taken from Xu, 2012).