

Assessment of differential item functioning in the Perceived Stress Scale-10

Stephen R Cole

Differential item functioning is the degree to which items that comprise a measurement scale are systematically related to various exogenous variables (for example, sex, race, etc) after conditioning on the latent variable of interest, which is often represented by the total scale score. When an item functions differentially with respect to an exogenous variable and the source of this differential functioning is irrelevant to the construct measured the item is termed "biased." Description of item functioning, a relatively modern psychometric development,¹ is tantamount to providing evidence for, or against, construct validity, and is therefore of imminent importance to the use of measurement scales in epidemiological research.²

The 10-item Perceived Stress Scale (PSS-10) is a popular Likert-type rating scale developed to quantify the psychological construct "perceived stress."³ Herein, we present the first results of differential item functioning in the PSS-10.

Methods

A nationally representative sample (n = 2387) was collected by random digit dialling for a previously reported⁴ US Department of Health and Human Services study that focused on psychological stress, health, and well being. The present study limits the sample to 2264 (95%) with complete data for variables of interest. The average age was 42 (SD 17) years.

Differential item functioning was assessed for each item on the exogenous variables sex (59% female), race (85% white, 8% African-

American, 4% Hispanic, 3% other) and education (48% greater than a high school education). The statistical formulation of the problem is: $I_i \perp E_i | \theta$. Paraphrased, each studied item (I_i) should be unrelated to each exogenous variable (E_i), after conditioning on the latent variable (θ). A test of the statistical formulation above where the I_i are ordered variables is by use of Kendall's Tau-b (τ_b).⁵ Tau-b measures concordant and discordant pairs while correcting for ties, and has the range $-1 \leq \tau_b \leq 1$. An independent association between any studied item and exogenous variable ($\tau_b \neq 0$) provides evidence that item is functioning differently by level of the exogenous variable. While methods developed from item response theory or structural equations may be used to assess differential item functioning, the method outlined above is intuitively appealing, defensible⁶ and has (in a similar form) previously been applied to other health measurement scales.^{2,7}

Confidence intervals (95% CI) were generated using a bootstrap method, resampling with replacement 1000 times for each τ_b . In addition, exogenous variable level specific means, adjusted for the overall PSS-10, are provided by use of a general linear model. All analyses were conducted using SAS version 6.12.

Results

Table 1 shows the sample estimates of τ_b , 95% CIs for τ_b , and exogenous variable level specific adjusted means. Briefly, items 3, 6, 7, 8, and 10 functioned differently by sex, with women

Division of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, 900 Commonwealth Avenue, Boston, MA 02215, USA

Accepted for publication 13 November 1998

Table 1 Association between PSS-10 items and exogenous variables: conditioning on total PSS-10 Scale (n = 2264)

Studied Items*	Exogenous Variables†					
	Sex		Race‡		Education§	
	Partial τ_b (95% CI)¶	Adjusted mean female/male	Partial τ_b (95% CI)	Adjusted mean white/non-white	Partial τ_b (95% CI)	Adjusted mean >HS/≤HS
1 ...upset because of something happened unexpectedly	0.01 (-0.03, 0.04)	2.4/2.5	-0.03 (-0.07, 0.01)	2.4/2.5	0.02 (-0.02, 0.05)	2.5/2.4
2 ...unable to control the important things ...	0.01 (-0.02, 0.05)	2.3/2.3	-0.03 (-0.07, 0.01)	2.3/2.3	0.01 (-0.03, 0.04)	2.3/2.3
3 ...felt nervous and "stressed"	0.07 (0.03, 0.10)	2.8/2.7	0.09 (0.05, 0.13)	2.8/2.5	0.09 (0.05, 0.12)	2.9/2.7
4 ...dealt successfully with irritating life hassles	0.02 (-0.02, 0.06)	1.8/1.8	0.05 (0.01, 0.08)	1.9/1.7	-0.04 (-0.07, 0.00)	1.7/1.8
5 ...ineffectively coping with important changes ...	-0.03 (-0.06, 0.01)	2.2/2.3	0.03 (-0.01, 0.07)	2.2/2.2	-0.02 (-0.05, 0.02)	2.2/2.3
6 ...felt confident about your ability to handle personal problems	0.04 (0.01, 0.08)	2.3/2.3	-0.03 (-0.07, 0.01)	2.3/2.3	-0.02 (-0.06, 0.01)	2.3/2.3
7 ...felt things were going your way	0.06 (0.02, 0.10)	2.2/2.0	-0.01 (-0.04, 0.03)	2.1/2.2	-0.01 (-0.03, 0.04)	2.1/2.1
8 ...felt you could not cope with all the things you had to do?	0.08 (0.05, 0.12)	2.2/2.0	-0.01 (-0.04, 0.04)	2.1/2.1	-0.05 (-0.09, -0.02)	2.0/2.2
9 ...been able to control irritations in your life?	-0.03 (-0.06, 0.01)	2.8/2.8	0.01 (-0.03, 0.05)	2.8/2.8	0.06 (0.02, 0.09)	2.9/2.7
10 ...felt you were on top of things?	0.04 (0.01, 0.08)	2.2/2.1	-0.02 (-0.07, 0.01)	2.2/2.2	-0.02 (-0.05, 0.02)	2.1/2.1

* All items answered on a 5 point Likert type scale (never, almost never, sometimes, fairly often, very often).
 † All results adjusted for overall PSS-10 scale score.
 ‡ Race coded as white versus non-white.
 § Education coded as > high school education versus ≤ high school education.
 ¶ 95% confidence intervals generated with 1000 bootstrap resamples.

reporting higher perceived stress. Items 3 and 4 functioned differently by ethnicity, with white populations reporting increased perceived stress. Items 3 and 9 functioned differently by education, with higher educated participants reporting increased perceived stress, while items 4 and 8 functioned such that lower educated participants reported increased perceived stress. The adjusted mean item scores were never more than one third of a point different (on 5 point scale).

Discussion

While statistically significant differential item functioning was observed for several items, all of the items appeared practically invariant to the exogenous variables, demonstrated by extremely small τ_i s with narrow confidence intervals and the remarkably similar adjusted item means. These findings may be compared to a reported 2 mm reduction in blood pressure, with the 95% CI ranging from 1 mm to 3 mm; while an association is observed, it is near meaningless.

Strong evidence for unidimensionality of the PSS-10 (a prerequisite to performing such analyses) was previously provided.⁴ Using a purified sub-scale (total PSS-10 minus statistically significantly biased items) as the proxy for the latent variable did not appreciably change the results. The present analysis strategy may be insensitive to low levels of differential item functioning when the measurement scale consists of less than < 19 items,⁸ therefore analysis using methods derived from item response theory would confirm the present

findings. Combination of the potentially slightly biased items may explain the apparent test level bias towards women reporting higher levels of perceived stress in the PSS-10.⁴

In summary, while several items exhibited statistically significant (p value < 0.05) differential item functioning by sex, race and/or education, none exhibited practically meaningful differential item functioning for the PSS-10 in a large ($n = 2264$) random sample of US adults. The finding of relative item invariance to sex, race, and education suggests continued widespread use of the PSS-10. However, the possibility of test level bias resulting from combined item level differences in the PSS-10 should be explored.

Funding: none.

Conflicts of interest: none.

- 1 Nunally J, Bernstein I. *Psychometric theory*. 3rd ed. New York: McGraw-Hill, 1994.
- 2 Dean K, Holst E, Kreiner S, et al. Measurement issues in research on social support and health. *J Epidemiol Community Health* 1994;48:201-6.
- 3 Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav* 1983;24:386-96.
- 4 Cohen S, Williamson G. Perceived stress in a probability sample if the United States. In: Spacapan S, Oskamp S, eds. *The social psychology of health*. Newbury Park: Sage, 1988.
- 5 Goodman L, Kruskal W. Measures of association for cross-classification IV. *Journal of American Statistical Association* 1972;67:415-21.
- 6 Stricker L. Identifying test items that perform differently in population subgroups: a partial correlation index. *Applied Psychological Measurement* 1982;6:261-73.
- 7 Groenvold M, Bjorner J, Klee M, et al. Test for item bias in a quality of life questionnaire. *J Clin Epidemiol* 1995;48:805-16.
- 8 Holland P, Wainer H. *Differential item functioning: theory and practice*. Hillsdale: Lawrence Erlbaum Associates, 1993.