

Comparison of SS and RP scoring using 41 PROMIS short forms

Miya Segawa

Introduction

PROMIS offers two scoring methods for fixed-length tests (Short Forms, SFs): summed score scoring (SS scoring) and response pattern scoring (RP scoring). SS scoring allows for simpler paper-and-pencil administration but is less accurate, resulting in increased uncertainty in SS scores (INCREASE). INCREASE, which varies across SFs, is critical information for deciding between SS and RP scoring. If INCREASE is negligible—for example, less than 2% of the standard error (SE)—it can be ignored. However, if it is substantial—for example, more than 10%—users should be strongly cautioned against using SS scoring. PROMIS does not provide INCREASE values for individual SFs; instead, it generally recommends RP scoring based on research from non-PROMIS SFs.

Objectives

We examine INCREASE for 41 PROMIS SFs to provide users with reference information when choosing between SS and RP scoring.

Methods

Among the 41 SFs, 21 are included in the PROMIS Adult Profiles (Cella et al. 2019), the most widely used PROMIS instruments. We refer to these 21 SFs, along with 5 two-item SFs, as standard SFs because their items were selected to maximize information (higher-slope items). The remaining 15 SFs are classified as non-standard SFs, where other substantive requirements may prevent the selection of highly informative items.

For each SF, we sampled one million latent values from a standard normal distribution, with the mean shifted by one standard deviation in the direction of severe symptoms (Segawa 2025). Using these latent values, we computed one million SS and RP scores along with their associated uncertainties (standard errors, SEs). INCREASE was then calculated using the average SEs of SS scores (SS SEs) and RP scores (RP SEs).

Results

Table 1: SS score SE, RP score SE and INCREASE in 41 PROMIS SFs

SF Name	SS score SE	RP score SE	INCREASE (%)
Anxiety			
2a	4.6	4.6	0.4

SF Name	SS score SE	RP score SE	INCREASE (%)
4a	3.3	3.3	0.7
6a	2.8	2.8	1.1
8a	2.4	2.4	1.2
7a	2.5	2.4	1.3
Sleep Disturbance			
4a	3.6	3.5	2.4
6a	3.1	3.0	3.4
8a	2.8	2.7	3.6
8b	2.8	2.7	4.2
6aOAK	3.2	3.0	4.9
Depression			
2a	3.9	3.9	0.2
4a	3.2	3.2	0.5
6a	2.6	2.6	1.0
8a	2.4	2.3	1.1
8b	2.3	2.3	1.3
4aOAK	3.3	3.2	4.5
Fatigue			
2a	3.5	3.5	0.1
4a	2.7	2.6	0.4
6a	2.3	2.3	0.7
8a	2.0	2.0	0.9
7a	3.1	2.8	10.9
7aMI	2.9	2.8	3.9
8aMS	2.2	2.2	2.5
8aOAK	2.3	2.2	3.8
10a	2.3	2.1	8.5
13a	2.1	1.9	8.0
Pain Interference			
2a	3.7	3.7	0.2
4a	3.1	3.0	0.6
6a	2.8	2.7	0.5
8a	2.6	2.5	0.5
6b	2.8	2.8	1.3
Physical Function			
4a	3.8	3.8	1.2
6b	3.0	3.0	1.3
8b	2.6	2.6	1.4
10a	2.6	2.5	3.6
10b	2.8	2.6	4.3
20a	2.1	2.0	4.6
Social Roles			
2a	3.8	3.8	1.4
4a	2.8	2.7	0.5
6a	2.4	2.4	0.6

SF Name	SS score SE	RP score SE	INCREASE (%)
8a	2.2	2.2	0.7
8aOAK	2.3	2.3	0.9

Table 1 presents the averages of SS SEs, RP SEs, and INCREASE for the 41 SFs. For all SFs included in the PROMIS Adult Profile, those with 4a, 6a, and 8a items had INCREASE values below 2% (inconsequential), except for the three Sleep Disturbance SFs, which ranged from 2.4% to 3.6%.

Table 2: INCREASE in 41 SFs by amount and standard

	Standard SF	Non-Standard SF	Total
< 2%	23	5	28
≥ 2%, < 10%	3	9	12
≥ 10%	0	1	1
Total	26	15	41

Table 2 summarizes the INCREASE data from Table 1. It shows that 28 SFs had inconsequential INCREASE (<2%), 12 had small INCREASE (≥2% and <10%), and 1 had large INCREASE (>10%). Among the 26 standard SFs, all were inconsequential except the three Sleep Disturbance SFs (small). Among the 15 non-standard SFs, INCREASE tended to be larger: 5 SFs were inconsequential, 9 were small, and 1 was large.

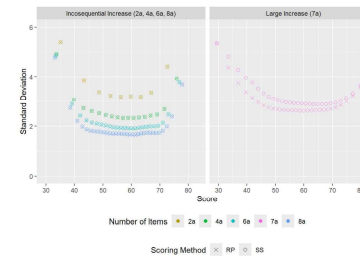


Figure 1: INCREASE in some Fatigue SFs

Figure (ref?)(fig:fig-fatigue1) illustrates the uncertainty (SE) in SS and RP scoring for selected Fatigue SFs. The left panel shows four standard SFs with inconsequential INCREASE, where SS SEs (circles) were only slightly larger than the corresponding RP SEs (crosses). The right panel shows an SF with large INCREASE (10.9%), where SS SEs were clearly larger than RP SEs.

Conclusion and Discussion

We provide Table 1, which includes the most widely used PROMIS SFs, as a reference for users to make

informed decisions when selecting either SS or RP scoring.

For the majority (65.9%) of the 41 SFs, INCREASE was inconsequential. For the SF with large INCREASE (10.9%), caution is warranted. These results indicate that INCREASE values vary widely, and recommendations should depend on these values. The current PROMIS guidance, which uniformly recommends RP scoring, may not be appropriate in all cases.

Within the PROMIS Adult Profiles, all SFs except Sleep Disturbance SFs had inconsequential INCREASE. For Sleep Disturbance, INCREASE was slightly above inconsequential, ranging from 2.4% (4a) to 3.6% (8a). Therefore, INCREASE can generally be considered inconsequential for PROMIS Adult Profiles, especially for the 21-item version.

Although our analysis covers only about one fifth of the more than 200 PROMIS SFs, inclusion of the most popular SFs ensures that our results are relevant to a large proportion of PROMIS users. Table 1 could be expanded to all PROMIS SFs, allowing all users to make informed choices between SS and RP scoring.

References

- Cella, David, Seung W Choi, David M Condon, Ben Schalet, Ron D Hays, Nan E Rothrock, Susan Yount, et al. 2019. "PROMIS Adult Health Profiles: Efficient Short-Form Measures of Seven Health Domains." *Value in Health* 22 (5): 537–44.
- Segawa, Eisuke. 2025. "Increase of Uncertainty in Summed-Score-Based Scoring in Non-Rasch IRT." *Applied Psychological Measurement*, 01466216251350342.