Patient-Reported Outcomes Measurement Information System
Dynamic Tools to Measure Health Outcomes From the Patient Perspective

PHYSICAL STRESS EXPERIENCES

A brief guide to the PROMIS® Physical Stress Experiences instruments:

PEDIATRIC	PARENT PROXY
PROMIS Pediatric Bank v1.0 - Physical Stress Experiences	PROMIS Parent Proxy Bank v1.0 - Physical Stress Experiences
PROMIS Pediatric Short Form v1.0 - Physical Stress Experiences 4a	PROMIS Parent Proxy Short Form v1.0 Physical Stress Experiences 4a
PROMIS Pediatric Short Form v1.0 - Physical Stress Experiences 8a	PROMIS Parent Proxy Short Form v1.0 Physical Stress Experiences 8a

ABOUT PHYSICAL STRESS EXPERIENCES

The PROMIS Pediatric and PROMIS Parent Proxy Physical Stress Experiences item banks assess the physically experienced sensations associated with responses to internal or external challenges including arousal, agitation, pain, and gastrointestinal distress. The Pediatric and Parent Proxy Physical Stress Experiences item banks use a 7 day recall period and each contains 26 items.

Physical Stress Experiences instruments are available for pediatric self-report (ages 8-17) and for parents serving as proxy reporters for their child (youth ages 5-17).

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing physical stress experiences: short forms and computerized adaptive tests (CATs). When administering a short form, instruct participants to answer all of the items (i.e., questions or statements) presented. With a CAT, participant responses guide the system's choice of subsequent items from the full item bank (26 items in total for both pediatric and parent proxy). Although items differ across respondents taking a CAT, scores are comparable across participants. Some administrators may prefer to ask the same question of all respondents or of the same respondent over time, to enable a more direct comparability across people or time. In these cases, or when paper administration is preferred, a short form would be more desirable than a CAT. This guide provides information on all physical stress experiences short forms and CAT instruments.

Whether one uses a short form or a CAT, the score metric is Item Response Theory (IRT), a family of statistical models that link individual questions to a presumed underlying trait or concept of physical stress experiences represented by all items in the item bank. When choosing between a CAT and a short form, it is useful to consider the demands of computer-based assessment, and the psychological, physical, and cognitive burden placed on respondents as a result of the number of questions asked.

SHORT FORM DIFFERENCES

Pediatric and Parent Proxy Short Forms

There are 2 Pediatric and 2 Parent Proxy short forms. Items were selected based on content and psychometric characteristics.

Selecting a Short Form

In selecting between short forms, the difference is instrument length. The reliability and precision of the short forms within a domain is highly similar. If you are working with a sample in which you want the most precise measure, select the longest short form. If you have little room for additional measures but really wanted to capture something as a secondary outcome, select one of the shorter instruments (e.g., 4-item short form).

SELECTING A PEDIATRIC OR PARENT PROXY INSTRUMENT

In selecting whether to use the pediatric or parent proxy instrument for this domain, it is important to consider both the population and the domain which you are studying. Pediatric self-report should be considered the standard for measuring patient-reported outcomes among children. However, circumstances exist when the child is too young, cognitively impaired, or too ill to complete a patient-reported outcome instrument. While information derived from self-report and proxy-report is not equivalent, it is optimal to assess both the child and the parent since their perspectives may be independently related to healthcare utilization, risk factors, and quality of care.

SCORING THE INSTRUMENT

Short Forms: PROMIS instruments are scored using item-level calibrations. This means that the most accurate way to score a PROMIS instrument is to use the HealthMeasures Scoring Service (https://www.assessmentcenter.net/ac_scoringservice) or a data collection tool that automatically calculates scores (e.g., Assessment Center, REDCap auto-score). This method of scoring uses responses to each item for each participant. We refer to this as "response pattern scoring." Because response pattern scoring is more accurate than the use of raw score/scale score look up tables included in this manual, it is preferred. Response pattern scoring is especially useful when there is missing data (i.e., a respondent skipped an item), different groups of participants responded to different items, or you have created a new questionnaire using a subset of questions from a PROMIS item bank.

To use the scoring tables in this manual, calculate a summed score. Each question usually has five response options ranging in value from one to five. To find the total raw score for a short form with all questions answered, sum the values of the response to each question. For example, for the v1.0 pediatric 4-item form, the lowest possible raw score is 44.5; the highest possible raw score is 84.7 (see all short form scoring tables in the Appendix). **All questions must be answered in order to produce a valid score using the scoring tables**. If a participant has skipped a question, use the HealthMeasures Scoring Service (https://www.assessmentcenter.net/ac_scoringservice) to generate a final score.

With the total raw score for a measure, locate the applicable score conversion table in the Appendix and use this table to translate the total raw score into a T-score for each participant. The T-score rescales the raw score into a standardized score with a mean of 50 and a standard deviation (SD) of 10. Therefore a person with a T-score of 40 is one SD below the mean.

For the pediatric PROMIS Physical Stress Experiences 4a short form v1.0, a raw score of 10 converts to a T-score of 65.9 with a standard error (SE) of 4.2 (see scoring table for the 4a v1.0 short form in the Appendix). Thus, the 95% confidence interval around the observed score ranges from 57.7 to 74.1 (T-score $\pm$ (1.96*SE) or 65.9$\pm$ (1.96*4.2).

<u>CAT:</u> A minimum number of items (5 for peds and parent proxy CATs) must be answered in order to receive a score for the Physical Stress Experiences CAT. The response to the first item will guide the system's choice of the next item for the participant. The participant's response to the second item will dictate the selection of the following question, and so on. As additional items are administered, the potential for error is reduced and confidence in the respondent's score increases. A CAT will continue until either the standard error drops below a specified level (on the T-score metric 4.0 for peds and parent proxy CATs), or the participant has answered the maximum number of questions (12), whichever occurs first.
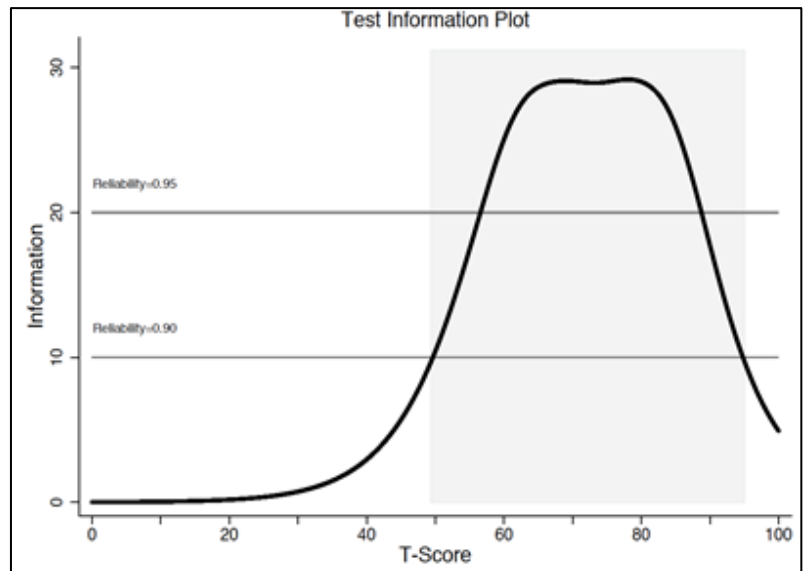
For most PROMIS instruments, a score of 50 is the average for the United States general population with a standard deviation of 10 because calibration testing was performed on a large sample of the general population. You can read more about the calibration and centering samples on HealthMeasures.net (http://www.healthmeasures.net/score-and-interpret/interpret-scores/promis). The T-score is provided with an error term (Standard Error or SE). The Standard Error is a statistical measure of variance and represents the "margin of error" for the T-score.

Important: *A higher PROMIS T-score represents more of the concept being measured.* For negatively-worded concepts like Physical Stress Experiences, a T-score of 40 is one SD better than average. By comparison, a Physical Stress Experiences T-score of 60 is one SD worse than average.

STATISTICAL CHARACTERISTICS

There are four key features of the score for physical stress experiences:
- **Reliability**: The degree to which a measure is free of error. It can be estimated by the internal consistency of the responses to the measure, or by correlating total scores on the measure from two time points when there has been no true change in what is being measured (for z-scores, reliability = $1 - SE^2$).
- **Precision**: The consistency of the estimated score (reciprocal of error variance).
- **Information:** The precision of an item or multiple items at different levels of the underlying continuum (for z-scores, information = $1/SE^2$).
- **Standard Error (SE):** The possible range of the actual final score based upon the scaled T-score. For example, with a T-score of 52 and a SE of 2, the 95% confidence interval around the actual final score ranges from 48.1 to 55.9 (T-score $\pm$ (1.96*SE) = 52 $\pm$ 3.9 = 48.1 to 55.9).

Figure 1

Patient-Reported Outcomes Measurement Information System
Dynamic Tools to Measure Health Outcomes From the Patient Perspective

PROMIS

The final score is represented by the T-score, a standardized score with a mean of 50 and a standard deviation (SD) of 10.

In Figure 1(Pediatric 8a short form) and Figure 2 (Parent Proxy 8a short form), the four horizontal lines each represent a degree of internal consistency reliability (i.e., .90 or .95) typically regarded as sufficient for an accurate individual score. The shaded regions mark the ranges of the scales where measurement precision is comparable to the reliability of .90 for the eight-item forms. Figure 2 also tells us where on the scales the forms are most informative based upon the T-scores. These forms would typically be more informative than physical stress experiences forms with fewer items.

Figure 2

Figure 3 (Pediatric 4a and 8a short forms) and Figure 4 (Parent Proxy 4a and 8a short forms) also tell us where on the scales the forms are most informative based upon the T-scores: the 8-item forms are more informative than the the 4-item forms.
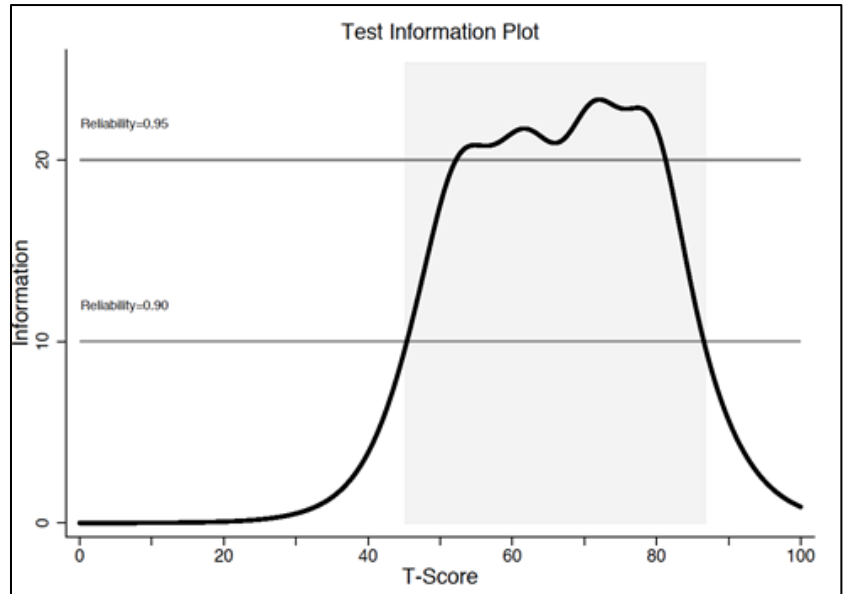
More information is available at HealthMeasures.net.
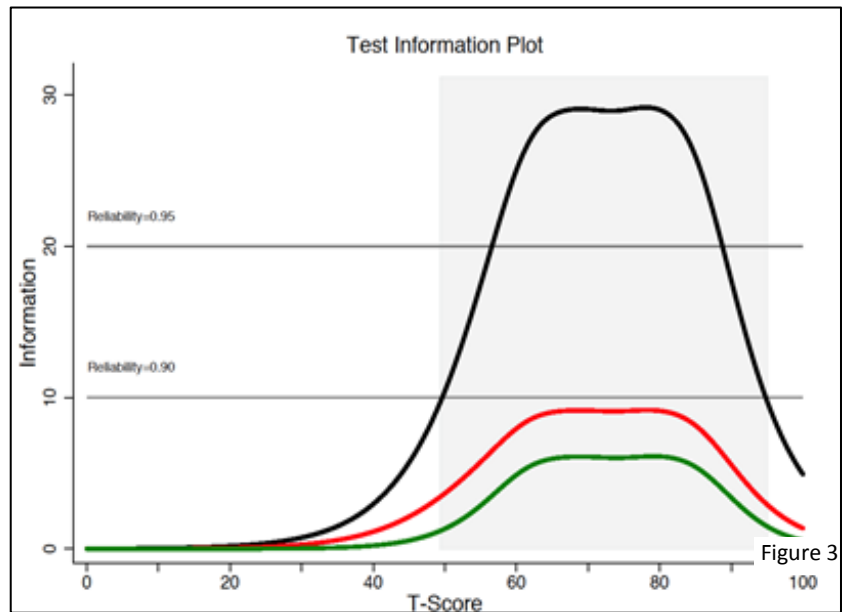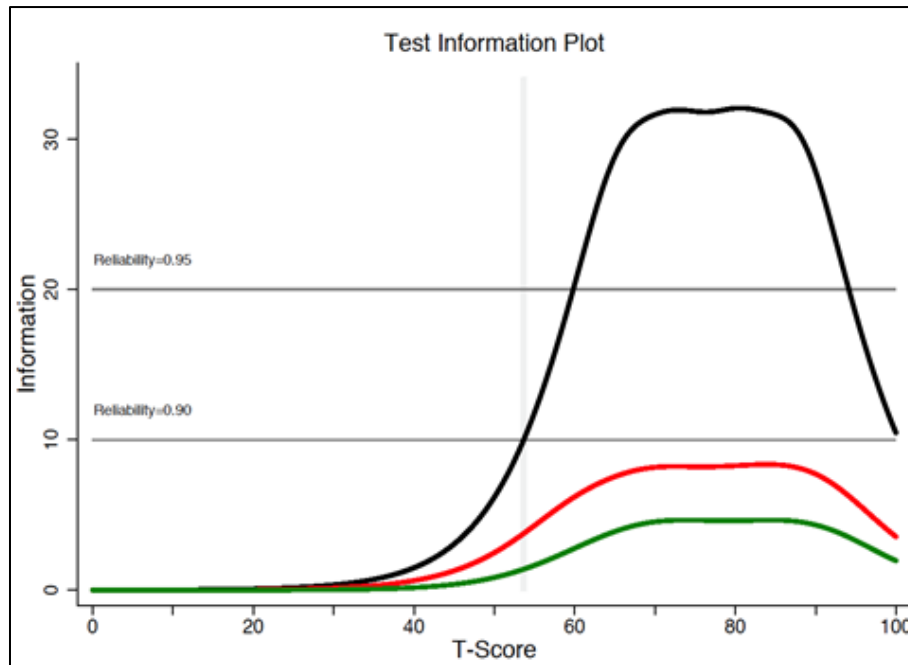

Figure 3

Figure 4

PREVIEWS OF SAMPLE ITEMS

Figure 5 and Figure 6 are excerpts from the paper version of the pediatric and parent proxy eight-item short forms, respectively. These are the paper version formats used for all physical stress experiences instruments. It is important to note, a CAT is not available for paper administration.

In the past 7 days...		Never	Rarely	Sometimes	Often	Always
EoS_S_032R1	My heart beat faster than usual, even when I was not exercising or playing hard	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
EoS_S_033R1	I had trouble breathing, even when I was not exercising or playing hard	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Figure 5

In the past 7 days...		Never	Rarely	Sometimes	Often	Always
EoS_S_032_PXR1	My child's heart beat faster than usual, even when he/she was not exercising or playing hard	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
EoS_S_033_PXR1	My child had trouble breathing, even when he/she was not exercising or playing hard	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Figure 6

FREQUENTLY ASKED QUESTIONS (FAQs)

Q: I am interested in learning more. Where can I do that?

Review the HealthMeasures website at www.healthmeasures.net.

Q: Do I need to register with PROMIS to use these instruments?

No.

Q: Are these instruments available in other languages?

Yes! Look at the HealthMeasures website (www.healthmeasures.net) for current information on PROMIS translations.

Q: Can I make my own short form?

Yes, custom short forms can be made by selecting any items from an item bank. This can be scored using the Scoring Service (https://www.assessmentcenter.net/ac_scoringservice).

Q: How do I handle multiple responses when administering a short form on paper?

Guidelines on how to deal with multiple responses have been established. Resolution depends on the responses noted by the research participant.

- If two or more responses are marked by the respondent, and they are next to one another, then a data entry specialist will be responsible for randomly selecting one of them to be entered and will write down on the form which answer was selected. Note: To randomly select one of two responses, the data entry specialist will flip a coin (heads - higher number will be entered; tails – lower number will be entered).To randomly select one of three (or more) responses, a table of random numbers should be used with a statistician's assistance.
- If two or more responses are marked, and they are NOT all next to one another, the response will be considered missing.

Q: What is the minimum change on a PROMIS instrument that represents a clinically meaningful difference?

To learn more about research on the meaning of a change in scores, we suggest conducting a literature review to identify the most current information. The HealthMeasures website (http://www.healthmeasures.net/score-and-interpret/interpret-scores/promis) has additional information on interpreting scores.

APPENDIX-SCORING TABLES

Pediatric v1.0 – Physical Stress Experiences 4a		
Raw Summed Score	T-Score	SE*
4	44.5	7.6
5	52.8	5.6
6	55.9	5.5
7	59.0	5.1
8	61.4	4.9
9	63.9	4.3
10	65.9	4.2
11	67.8	4.1
12	69.6	4.0
13	71.3	4.0
14	73.1	4.0
15	74.9	4.0
16	76.7	4.0
17	78.6	4.0
18	80.6	3.9
19	82.6	3.7
20	84.7	3.4
*SE = Standard Error on T-score metric		

Pediatric v1.0 – Physical Stress Experiences 8a		
Raw Summed Score	T-Score	SE*
8	39.4	6.8
9	45.0	5.6
10	47.9	5.4
11	50.6	4.9
12	52.7	4.7
13	54.7	4.4
14	56.4	4.1
15	58.0	4.0
16	59.5	3.8
17	60.9	3.7
18	62.2	3.6
19	63.5	3.5
20	64.7	3.5
21	65.8	3.4
22	67.0	3.4
23	68.1	3.3
24	69.2	3.3
25	70.3	3.3
26	71.3	3.3
27	72.4	3.3
28	73.5	3.3
29	74.6	3.3
30	75.7	3.3
31	76.8	3.3
32	78.0	3.3
33	79.2	3.3
34	80.4	3.3
35	81.6	3.2
36	82.9	3.2
37	84.1	3.0
38	85.2	2.8
39	86.2	2.5
40	87.1	2.2
*SE = Standard Error on T-score metric		

Parent Proxy v1.0 – Physical Stress Experiences 4a		
Raw Summed Score	T-Score	SE*
4	45.3	8.0
5	54.0	6.3
6	57.2	6.2
7	60.7	5.8
8	63.4	5.6
9	66.2	5.1
10	68.5	4.9
11	70.7	4.7
12	72.7	4.6
13	74.7	4.6
14	76.6	4.5
15	78.5	4.5
16	80.4	4.4
17	82.3	4.2
18	83.9	3.8
19	85.3	3.4
20	86.3	3.0
*SE = Standard Error on T-score metric		

Parent Proxy v1.0 – Physical Stress Experiences 8a		
Raw Summed Score	T-Score	SE*
8	41.6	7.1
9	48.0	5.8
10	51.1	5.5
11	54.0	5.0
12	56.4	4.7
13	58.5	4.3
14	60.3	4.1
15	61.9	3.9
16	63.5	3.8
17	64.9	3.7
18	66.3	3.6
19	67.6	3.6
20	68.8	3.5
21	70.1	3.5
22	71.3	3.5
23	72.5	3.5
24	73.7	3.5
25	74.8	3.4
26	76.0	3.4
27	77.2	3.4
28	78.3	3.4
29	79.5	3.4
30	80.7	3.4
31	81.8	3.3
32	83.0	3.2
33	84.0	3.1
34	85.0	2.9
35	85.9	2.6
36	86.6	2.4
37	87.2	2.1
38	87.7	1.9
39	88.1	1.7
40	88.4	1.5
*SE = Standard Error on T-score metric		