

PSYCHOLOGICAL STRESS EXPERIENCES

A brief guide to the PROMIS® Pediatric and Parent Proxy Report Psychological Stress Experiences instruments:

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

ABOUT PSYCHOLOGICAL STRESS EXPERIENCES

The PROMIS Pediatric and Parent Proxy Psychological Stress Experiences item banks assess the thoughts or feelings about self and the world in the context of environmental or internal challenges. Items represent 3 facets of psychological stress reactions: feeling overwhelmed, perceived lack of control of capacity to manage one’s life, and cognitive-perceptual disruption. The Pediatric and Parent Proxy Psychological Stress Experiences item bank uses a 7-day recall period and consists of 19 items.

Psychological 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 psychological stress experiences: short forms and computerized adaptive test (CAT). When administering a short form, instruct participants to answer all of the items (i.e., questions or statements) presented. With CAT, participant responses guide the system’s choice of subsequent items from the full item bank (19 items in total). Although items differ across respondents taking 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 CAT. This guide provides information on all Psychological Stress Experiences short forms and CAT instruments.

Whether one uses a short form or 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 Psychological Stress Experiences represented by all items in the item banks. When choosing between 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.

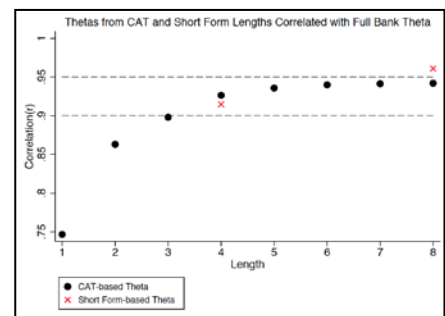


Figure 1

Figure 1 (pediatric) and Figure 2 (parent proxy) illustrate the correlations (strength of relationship) of the full banks with CATs and with short forms of varying length. The correlation of CAT scores with the full bank score is greater than a short form of any length. The correlation of CAT scores with the full bank score is greater than a short form of any length. A longer CAT or longer short form offers greater correlation, as well as greater precision. When evaluating precision, not all questions are equally informative. The flexibility of CAT to choose more informative questions offers more precision.

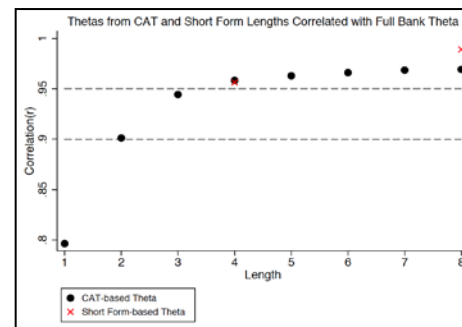


Figure 2

SHORT FORM DIFFERENCES

There are 2 Psychological Stress Experiences short forms for child and 2 Psychological Stress Experiences short forms for proxy report: 8-item forms (SF-8) and 4-item forms (SF-4). Short form items were selected based on analyses conducted using data from the child-report and parent-report calibration samples. Items in the 8 and 4 short forms were selected based on rankings using two psychometric criteria: (1) maximum interval information; and 2) CAT simulations. Item rankings were similar for both criteria. For the maximum interval criterion, each item information function was integrated (without weighting) for the interval from the mean to 2 SDs worse than the mean. For the CAT simulations, items were rank ordered based on their average administration rank over the simulees. The domain team constructed the short forms with a focus on representing the range of the trait and also representing the content of the item bank. Domain experts reviewed short forms to give input on the relevance of each item. Domain experts reviewed the items and rankings to select items for the SF-4 and SF-8. For each domain, 4-item and 8-items have been selected so that the items are nested/overlap (e.g., the 8-item form is the 4-item form plus four additional items). The longer version of the short form will provide a more precise score with less error than the shorter short form. However, this does increase the respondent burden.

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. Since 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_scoringervice) 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 pediatric 4-item form, the lowest possible raw score is 4; the highest possible raw score is 20 (see all short form scoring tables in 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_scoring-service) to generate a final score.

With the total raw score for a measure, locate the applicable score conversion table in 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 PROMIS Pediatric Psychological Stress Experiences 4a short form, a raw score of 10 converts to a T-score of 58.0 with a standard error (SE) of 3.1 (see scoring table for the 4a short form in Appendix). Thus, the 95% confidence interval around the observed score ranges from 51.9 to 64.1 (T-score \pm (1.96*SE) or 58.0 \pm (1.96*3.1)).

CAT: A minimum number of items (5 for peds and parent proxy CATs) must be answered in order to receive a score for the Pediatric and Parent Proxy Psychological Stress Experiences CATs. 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. CATs 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 Psychological Stress Experiences, a T-score of 60 is one SD worse than average. By comparison, a Psychological Stress Experiences T-score of 40 is one SD better than average.

STATISTICAL CHARACTERISTICS

There are four key features of the score for Psychological 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).

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

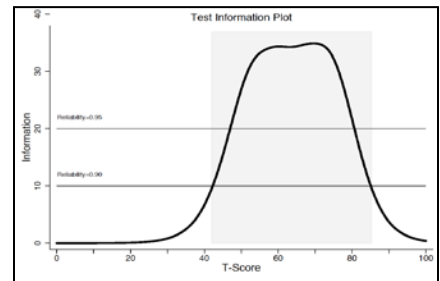


Figure 3

In Figure 3 (Pediatric 8-item short form) and Figure 4 (Parent Proxy 8-item short form), the two dotted horizontal lines in each graph represent a degree of internal consistency reliability (i.e., .90 or .95) typically regarded as sufficient for an accurate individual score. The shaded blue regions mark the ranges of the scales where measurement precision is comparable to the reliability of .90 for the eight-item forms. Both of these figures also tell us where on the scales the forms are most informative based upon the T-scores. These forms would typically be more informative than the 4-item Psychological Stress Experiences short forms.

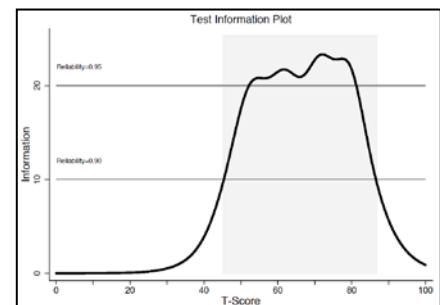


Figure 4

Figure 5 (Pediatric 4- and 8-item short forms) and Figure 6 (Parent Proxy 4- and 8-item short forms) also tell us where on the scales the forms are most informative based upon the T-scores: the 8-item forms (red lines) are more informative than the 4-item forms (green lines).

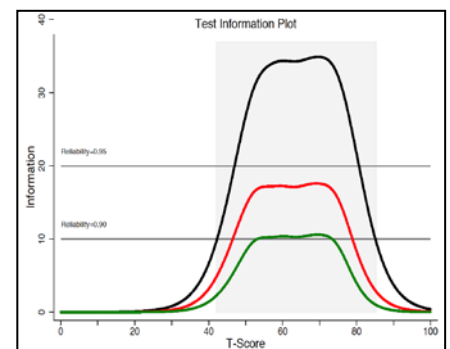


Figure 5

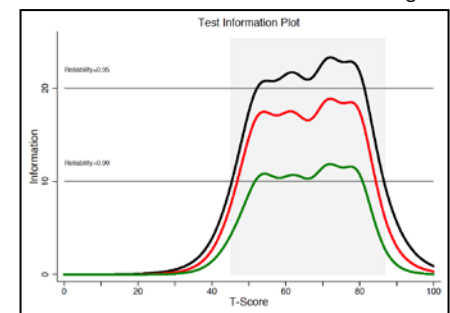


Figure 6



Figure 7 and Figure 8 are samples of the statistical information available in Assessment Center for a CAT. More information is available online at HealthMeasures.net.

PREVIEW OF SAMPLE ITEM

Figure 9 (pediatric) and Figure 10 (parent proxy) are excerpts from the paper versions of the eight-item short forms. These are the paper version formats used for all Psychological Stress Experiences instruments. It is important to note, CAT is not available for paper administration.

| | |
|------------------------------------|-----------------------------|
| Scaling Model Used For Calibration | Graded Response Model (GRM) |
| Total Number of Items | 28 |

| | | |
|-------------------------|-----|-------------------|
| Sample | N | Alpha Reliability |
| PROMIS Wave 1 Full Bank | 782 | 0.98 |

| Score Distributions | | | | | | | | | |
|---------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | Mean | SD | P5 | P10 | P25 | P50 | P75 | P90 | P95 |
| Raw | 49.16 | 22.58 | 28.00 | 29.00 | 31.00 | 41.00 | 60.00 | 83.00 | 96.95 |
| Scale | 49.31 | 9.63 | 33.49 | 37.66 | 42.42 | 48.62 | 55.25 | 62.01 | 66.00 |

| | | | | | | | | | | Min | Max |
|-------------|-------|------|------|------|------|------|------|------|------|-----|-----|
| Scale Score | 10.0 | 20.0 | 30.0 | 40.0 | 50.0 | 60.0 | 70.0 | 80.0 | 90.0 | 0.0 | 0.0 |
| SE | 13.90 | 4.00 | 1.10 | .30 | .10 | .10 | .10 | .20 | .50 | | |
| Reliability | .00 | .00 | .00 | .92 | .98 | .99 | .99 | .97 | .75 | | |

Figure 7

| | |
|------------------------------------|-----------------------------|
| Scaling Model Used For Calibration | Graded Response Model (GRM) |
| Total Number of Items | 28 |

| | | |
|-------------------------|-----|-------------------|
| Sample | N | Alpha Reliability |
| PROMIS Wave 1 Full Bank | 782 | 0.98 |

| Score Distributions | | | | | | | | | |
|---------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | Mean | SD | P5 | P10 | P25 | P50 | P75 | P90 | P95 |
| Raw | 49.16 | 22.58 | 28.00 | 29.00 | 31.00 | 41.00 | 60.00 | 83.00 | 96.95 |
| Scale | 49.31 | 9.63 | 33.49 | 37.66 | 42.42 | 48.62 | 55.25 | 62.01 | 66.00 |

| | | | | | | | | | | Min | Max |
|-------------|-------|------|------|------|------|------|------|------|------|-----|-----|
| Scale Score | 10.0 | 20.0 | 30.0 | 40.0 | 50.0 | 60.0 | 70.0 | 80.0 | 90.0 | 0.0 | 0.0 |
| SE | 13.90 | 4.00 | 1.10 | .30 | .10 | .10 | .10 | .20 | .50 | | |
| Reliability | .00 | .00 | .00 | .92 | .98 | .99 | .99 | .97 | .75 | | |

Figure 8

In the past 7 days...

| | | Never | Rarely | Sometimes | Often | Always |
|-------------|---|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| EOS_P_011R1 | I felt stressed. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| | | 1 | 2 | 3 | 4 | 5 |
| EOS_P_064R1 | I felt that my problems kept piling up. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| | | 1 | 2 | 3 | 4 | 5 |

Figure 9

In the past 7 days...

| | | Never | Rarely | Sometimes | Often | Always |
|----------------|---|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| EOS_P_011_PXR1 | My child felt stressed. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| | | 1 | 2 | 3 | 4 | 5 |
| EOS_P_064_PXR1 | My child felt that his/her problems kept piling up. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| | | 1 | 2 | 3 | 4 | 5 |

Figure 10



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_scoring-service).

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 SF8a Sum Score to T-Score Conversion Table | | |
|--|-------|-----|
| Raw | Scale | SE |
| 8 | 37.0 | 5.7 |
| 9 | 42.6 | 4.0 |
| 10 | 45.2 | 3.6 |
| 11 | 47.2 | 3.2 |
| 12 | 48.9 | 2.9 |
| 13 | 50.3 | 2.8 |
| 14 | 51.6 | 2.6 |
| 15 | 52.8 | 2.6 |
| 16 | 54.0 | 2.5 |
| 17 | 55.0 | 2.5 |
| 18 | 56.1 | 2.5 |
| 19 | 57.1 | 2.4 |
| 20 | 58.1 | 2.4 |
| 21 | 59.1 | 2.4 |
| 22 | 60.1 | 2.4 |
| 23 | 61.0 | 2.4 |
| 24 | 62.0 | 2.4 |
| 25 | 63.0 | 2.4 |
| 26 | 63.9 | 2.4 |
| 27 | 64.9 | 2.4 |
| 28 | 65.9 | 2.4 |
| 29 | 66.8 | 2.4 |
| 30 | 67.8 | 2.4 |
| 31 | 68.8 | 2.4 |
| 32 | 69.8 | 2.4 |
| 33 | 70.8 | 2.4 |
| 34 | 71.9 | 2.4 |
| 35 | 73.0 | 2.5 |
| 36 | 74.3 | 2.6 |
| 37 | 75.6 | 2.7 |
| 38 | 77.2 | 2.9 |
| 39 | 79.1 | 3.1 |
| 40 | 81.8 | 3.5 |

| V1.0 SF4a Sum Score to T-Score Conversion Table | | |
|--|-------|-----|
| Raw | Scale | SE |
| 4 | 39.5 | 6.2 |
| 5 | 46.2 | 4.2 |
| 6 | 49.4 | 3.8 |
| 7 | 52.0 | 3.4 |
| 8 | 54.2 | 3.2 |
| 9 | 56.2 | 3.1 |
| 10 | 58.0 | 3.1 |
| 11 | 59.8 | 3.1 |
| 12 | 61.6 | 3.1 |
| 13 | 63.3 | 3.1 |
| 14 | 65.0 | 3.1 |
| 15 | 66.8 | 3.1 |
| 16 | 68.5 | 3.1 |
| 17 | 70.4 | 3.1 |
| 18 | 72.4 | 3.2 |
| 19 | 74.8 | 3.3 |
| 20 | 78.4 | 4.0 |



APPENDIX SCORING TABLES – PARENT PROXY

| V1.0 SF8a Sum Score to T-Score Conversion Table | | |
|--|-------|-----|
| Raw | Scale | SE |
| 8 | 37.6 | 5.8 |
| 9 | 43.5 | 4.0 |
| 10 | 46.3 | 3.5 |
| 11 | 48.5 | 3.1 |
| 12 | 50.3 | 2.8 |
| 13 | 51.9 | 2.7 |
| 14 | 53.3 | 2.6 |
| 15 | 54.7 | 2.5 |
| 16 | 56.0 | 2.5 |
| 17 | 57.2 | 2.5 |
| 18 | 58.4 | 2.5 |
| 19 | 59.6 | 2.5 |
| 20 | 60.8 | 2.5 |
| 21 | 62.0 | 2.5 |
| 22 | 63.2 | 2.5 |
| 23 | 64.4 | 2.5 |
| 24 | 65.6 | 2.5 |
| 25 | 66.7 | 2.5 |
| 26 | 67.9 | 2.5 |
| 27 | 69.0 | 2.5 |
| 28 | 70.1 | 2.5 |
| 29 | 71.2 | 2.4 |
| 30 | 72.3 | 2.4 |
| 31 | 73.4 | 2.4 |
| 32 | 74.5 | 2.4 |
| 33 | 75.6 | 2.4 |
| 34 | 76.7 | 2.4 |
| 35 | 77.9 | 2.5 |
| 36 | 79.2 | 2.5 |
| 37 | 80.5 | 2.6 |
| 38 | 82.1 | 2.7 |
| 39 | 83.7 | 2.7 |
| 40 | 85.4 | 2.6 |

| V1.0 SF4a Sum Score to T-Score Conversion Table | | |
|--|-------|-----|
| Raw | Scale | SE |
| 4 | 39.6 | 6.1 |
| 5 | 46.8 | 4.0 |
| 6 | 50.5 | 3.4 |
| 7 | 53.4 | 3.1 |
| 8 | 55.9 | 3.0 |
| 9 | 58.3 | 3.0 |
| 10 | 60.6 | 3.0 |
| 11 | 62.9 | 3.0 |
| 12 | 65.1 | 3.0 |
| 13 | 67.1 | 3.1 |
| 14 | 69.2 | 3.0 |
| 15 | 71.1 | 3.0 |
| 16 | 73.1 | 2.9 |
| 17 | 75.0 | 2.9 |
| 18 | 77.1 | 3.0 |
| 19 | 79.5 | 3.1 |
| 20 | 82.7 | 3.3 |