MEANING AND PURPOSE

A brief guide to the PROMIS® Meaning and Purpose instruments:

ADULT	PEDIATRIC	PARENT PROXY
PROMIS Item Bank v1.0 Meaning and Purpose PROMIS Short Form v1.0 Meaning and Purpose 4a PROMIS Short Form v1.0 Meaning and Purpose 6a PROMIS v1.0 Meaning and Purpose 8a	PROMIS Pediatric Item Bank v1.0 Meaning and Purpose PROMIS Pediatric Short Form v1.0 Meaning and Purpose 4a PROMIS Pediatric Short Form v1.0 Meaning and Purpose 8a	PROMIS Proxy Item Bank v1.0 Meaning and Purpose PROMIS Proxy Short Form v1.0 Meaning and Purpose 4a PROMIS Proxy Short Form v1.0 Meaning and Purpose 8a

ABOUT MEANING AND PURPOSE

The PROMIS Meaning and Purpose item banks assess one's sense of life having purpose and that there are good reasons for living. Higher scores indicate hopefulness, optimism, goal-directedness, and feelings that one's life is worthy. All item banks do not use a recall period.

The adult PROMIS Meaning and Purpose item bank expanded the Toolbox Meaning and Purpose item bank from 18 to 37 items to increase breadth of content and improve measurement precision.

Meaning and Purpose instruments are available for adults (ages 18+), pediatric self-report (ages 8-17) and for parents serving as proxy reporters for their child (youth ages 5-17).

INTRODUCTION TO ASSESSMENT OPTIONS

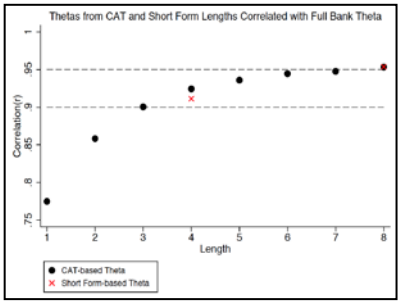

Figure 1

There are two administration options for assessing Meaning and Purpose: 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 (37 items in total for adult and 44 items for 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 Meaning and Purpose short forms and CAT instruments.

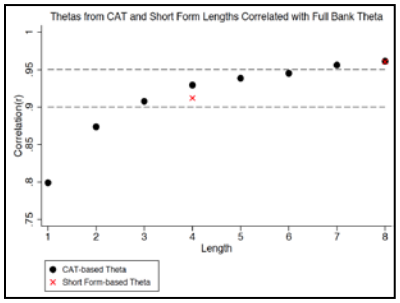

Figure 2

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 meaning and purpose represented by all items in the item bank. 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 and Figure 2 illustrate the correlations (strength of relationship) of the pediatric and parent proxy full banks, respectively, with CATs and with short forms of varying length. The correlations of CAT scores with the full bank scores are 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.

SHORT FORM DIFFERENCES

Adult Short Forms
There are 2 adult short forms. Items were selected based on content and psychometric characteristics. Short form items are nested or overlap (e.g., an 8 item short form is the 4-item short form plus two additional items).

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 that 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_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 adult 4-item form, the lowest possible raw score is 4; the highest possible raw score is 20 (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 Meaning and Purpose 4a short form v1.0, a raw score of 10 converts to a T-score of 31.8 with a standard error (SE) of 3.0 (see scoring table for the 4a v1.0 short form in the Appendix). Thus, the 95% confidence interval around the observed score ranges from 25.9 to 37.7 (T-score ± (1.96*SE) or 31.8 ± (1.96*3.0).

CAT: A minimum number of items (4 for adult and 5 for peds and parent proxy CATs) must be answered in order to receive a score for the Pediatric and Parent Proxy Meaning and Purpose 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. 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 positively-worded concepts like Meaning and Purpose, a T-score of 60 is one SD better than average. By comparison, a Meaning and Purpose T-score of 40 is one SD worse than average.

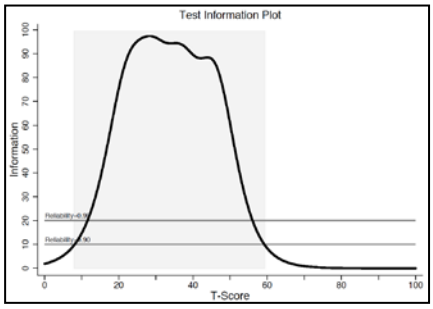

Figure 3

STATISTICAL CHARACTERISTICS

There are four key features of the score for Meaning and Purpose:

- **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).

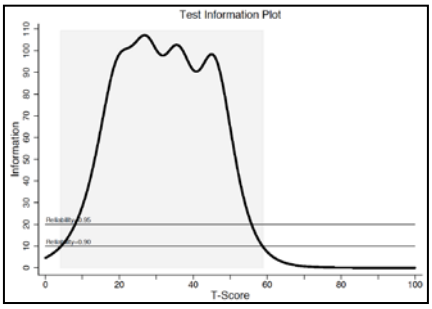

Figure 4

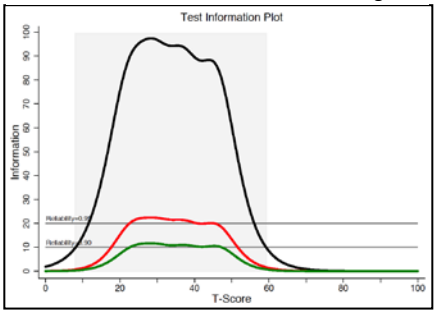

Figure 5

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 3 (Pediatric 8-item short form) and Figure 4 (Parent Proxy 8-item short form), the two dotted horizontal lines in each figure each 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. Figure 3 and Figure 4 also tell us where on the scales are the forms most informative based upon the T-scores. These forms would typically be more informative than the 4-item Meaning and Purpose short forms.

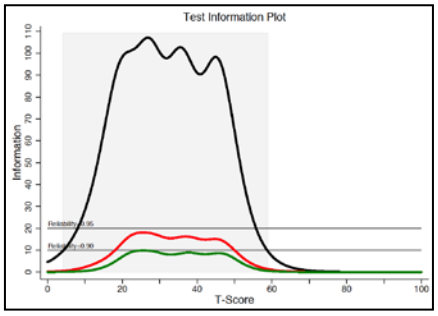

Figure 6

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

Figure 7 and Figure 8 are samples of the statistical information available in Assessment Center for the Meaning and Purpose CAT.

More information is available HealthMeasures.net.

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

PREVIEW OF SAMPLE ITEM

Figure 9 is an excerpt from the paper version of the adult eight-item short form. This is the paper version formats used for all Meaning and Purpose 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 8

		Not at all	A little bit	Somewhat	Quite a bit	Very much
PA215_C	I have a reason for living.....	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
PA216_C	My life has been productive.....	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Figure 9

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 – ADULT

V1.0 SF4a Sum Score to T-Score Conversion Table

Raw	Scale	SE*
4	21.2	4.1
5	25.2	3.8
6	28.2	3.6
7	30.7	3.6
8	33.0	3.6
9	35.2	3.6
10	37.4	3.6
11	39.5	3.6
12	41.7	3.6
13	43.9	3.6
14	46.1	3.6
15	48.4	3.6
16	50.8	3.7
17	53.5	3.8
18	56.5	4.0
19	60.0	4.4
20	65.5	5.7

*Standard Error on T-score metric

V1.0 SF6a Sum Score to T-Score Conversion Table

Raw	Scale	SE*
6	17.9	3.8
7	20.5	3.7
8	22.9	3.6
9	25.1	3.5
10	27.1	3.4
11	28.9	3.3
12	30.6	3.3
13	32.2	3.3
14	33.8	3.3
15	35.4	3.3
16	37.0	3.3
17	38.6	3.3
18	40.2	3.3
19	41.8	3.3
20	43.5	3.4
21	45.2	3.4
22	46.9	3.4
23	48.7	3.4
24	50.6	3.4
25	52.6	3.4
26	54.7	3.5
27	57.0	3.7
28	59.6	4.0
29	62.9	4.5
30	67.5	5.5

*Standard Error on T-score metric

V1.0 SF8a Sum Score to T-Score Conversion Table

Raw	Scale	SE*
8	16.3	3.3
9	18.6	3.3
10	20.6	3.2
11	22.5	3.0
12	24.1	2.9
13	25.7	2.9
14	27.1	2.8
15	28.5	2.8
16	29.8	2.8
17	31.0	2.7
18	32.2	2.7
19	33.5	2.7
20	34.7	2.7
21	35.9	2.7
22	37.1	2.8
23	38.3	2.8
24	39.5	2.8
25	40.7	2.8
26	42.0	2.8
27	43.3	2.8
28	44.6	2.9
29	46.0	2.9
30	47.4	2.9
31	48.8	2.9
32	50.3	2.9
33	51.8	2.9
34	53.4	2.9
35	55.0	3.0
36	56.8	3.1
37	58.8	3.3
38	61.1	3.6
39	64.1	4.2
40	68.4	5.2

*Standard Error on T-score metric

APPENDIX-SCORING TABLES – PEDIATRIC

V1.0 SF8a Sum Score to T-Score Conversion Table

Raw	Scale	SE*
8	16.5	3.0
9	19.0	2.7
10	20.6	2.5
11	22.0	2.4
12	23.3	2.3
13	24.3	2.2
14	25.4	2.2
15	26.3	2.2
16	27.2	2.1
17	28.2	2.1
18	29.1	2.1
19	29.9	2.2
20	30.8	2.2
21	31.7	2.2
22	32.6	2.2
23	33.5	2.2
24	34.5	2.2
25	35.4	2.2
26	36.4	2.2
27	37.3	2.2
28	38.3	2.3
29	39.4	2.3
30	40.4	2.3
31	41.5	2.3
32	42.6	2.3
33	43.7	2.3
34	44.9	2.4
35	46.2	2.5
36	47.7	2.6
37	49.3	2.9
38	51.5	3.3
39	54.3	3.9
40	60.6	6.1

*Standard Error on T-score metric

V1.0 SF4a Sum Score to T-Score Conversion Table

Raw	Scale	SE*
4	19.2	3.6
5	22.5	3.2
6	24.7	3.1
7	26.6	3.0
8	28.4	2.9
9	30.1	3.0
10	31.8	3.0
11	33.5	3.0
12	35.3	3.0
13	37.1	3.0
14	39.0	3.0
15	41.0	3.0
16	43.0	3.1
17	45.3	3.2
18	47.9	3.4
19	51.3	4.0
20	58.8	6.4

*Standard Error on T-score metric

APPENDIX-SCORING TABLES – PARENT PROXY

V1.0 SF8a Sum Score to T-Score Conversion Table		
Raw	Scale	SE*
8	15.3	2.8
9	17.3	2.8
10	19.1	2.7
11	20.6	2.5
12	21.9	2.4
13	23.1	2.3
14	24.1	2.2
15	25.1	2.2
16	26.1	2.2
17	27.1	2.2
18	28.1	2.2
19	29.1	2.2
20	30.0	2.2
21	31.1	2.2
22	32.1	2.3
23	33.1	2.3
24	34.2	2.3
25	35.3	2.3
26	36.4	2.3
27	37.6	2.3
28	38.8	2.3
29	40.0	2.3
30	41.2	2.3
31	42.4	2.3
32	43.7	2.4
33	45.1	2.4
34	46.7	2.6
35	48.4	2.8
36	50.7	3.3
37	53.7	4.0
38	60.2	6.1

*Standard Error on T-score metric

V1.0 SF4a Sum Score to T-Score Conversion Table		
Raw	Scale	SE*
4	18.4	3.5
5	21.4	3.3
6	23.7	3.1
7	25.6	3.1
8	27.5	3.1
9	29.4	3.1
10	31.3	3.2
11	33.2	3.2
12	35.2	3.2
13	37.3	3.2
14	39.5	3.2
15	41.8	3.2
16	44.2	3.3
17	46.9	3.5
18	50.5	4.0
19	58.3	6.5

*Standard Error on T-score metric