



# DEPRESSION SCORING MANUAL

A brief guide to scoring the PROMIS® Depression instruments:

ADULT	PEDIATRIC	EARLY CHILDHOOD PARENT-REPORT	PARENT PROXY
PROMIS Item Bank v1.0 – Depression PROMIS Item Bank v1.0 – Depression (recommended) PROMIS Item Bank v1.0 – Depression (screen-to-CAT) PROMIS Short Form v1.0 – Depression 4a PROMIS Short Form v1.0 – Depression 6a PROMIS Short Form v1.0 – Depression 8a PROMIS Short Form v1.0 – Depression 8b PROMIS-Ca Bank v1.0 – Depression	PROMIS Pediatric Item Bank v2.0 – Depressive Symptoms PROMIS Pediatric Item Bank v2.0 – Depressive Symptoms (recommended) PROMIS Pediatric Item Bank v2.0 – Depressive Symptoms (screen-to-CAT) PROMIS Pediatric Short Form v2.0 – Depressive Symptoms 8a PROMIS Pediatric Item Bank v1.0 – Depressive Symptoms* PROMIS Pediatric Short Form v1.0 – Depressive Symptoms 8a* PROMIS Pediatric Item Bank v1.1 – Depressive Symptoms* PROMIS Pediatric Short Form v1.1 – Depressive Symptoms 8b*	PROMIS Early Childhood Parent-Report Bank v1.0 – Depressive Symptoms PROMIS Early Childhood Parent-Report Short Form v1.0 – Depressive Symptoms 4a PROMIS Early Childhood Parent-Report Short Form v1.0 – Depressive Symptoms 8a	PROMIS Parent Proxy Item Bank v2.0 – Depressive Symptoms PROMIS Parent Proxy Item Bank v2.0 – Depressive Symptoms (recommended) PROMIS Parent Proxy Item Bank v2.0 – Depressive Symptoms (screen-to-CAT) PROMIS Parent Proxy Short Form v2.0 – Depressive Symptoms 6a PROMIS Parent Proxy Item Bank v1.0 – Depressive Symptoms* PROMIS Parent Proxy Short Form v1.0 – Depressive Symptoms 6a* PROMIS Parent Proxy Item Bank v1.1 – Depressive Symptoms* PROMIS Parent Proxy Short Form v1.1 – Depressive Symptoms 6b*

\*Retired measure

## COMPARING SCORES ACROSS VERSIONS

Some PROMIS domains have multiple versions of instruments (i.e. v1.0, v1.1, v2.0). Generally, **it is recommended that you use the most recent version available which can be identified as the instrument with the highest version number.** In most cases, an instrument that has a decimal increase (v1.0 to v1.1) retains the same item-level parameters as well as instrument reliability and validity. In cases where a version number increases by a whole number (e.g., v1.0 to v2.0), the changes to the instrument are more substantial.

For depression, v2.0 pediatric and parent proxy measures replaced v1.0/v1.1. The v2.0 measures 1) changed from using response scores of 0-4 to use 1-5 (item IDs amended with an “r”) and 2) added new items (item IDs start with 7000). The calibrations between v1.0, v1.1, and v2.0 are identical. Therefore, T-scores from v1.0, v1.1, and v2.0 are comparable. The pediatric and parent proxy v1.1 item bank existed briefly – they eliminated one item from the original v1.0 banks. This affected the 8-item short forms as well; the v1.0 and v1.1 short forms have slightly different items included. The v1.1 short form 8b and v2.0 short form 8a include the same items but have different response scores.

## WHICH CALIBRATION SAMPLE SHOULD I USE?

Some PROMIS Parent Proxy instruments (Anxiety, Depressive Symptoms, Fatigue, Mobility, Pain Interference, Peer Relationships) have two calibration samples – “Parent Proxy” and “Parent Proxy Without Local Dependence.” The former (Parent Proxy) includes calibrations for all items. This is the default calibration sample. If you aren’t sure which calibration sample to use, utilize this one. The Parent Proxy Without Local



Dependence does not include calibrations for some items. The items without calibrations are enemy items. That is, a dyad or triad of items was identified in which there are psychometric reasons to only administer one of those items to a given respondent. For example, item Pf2depr11 and Pf2depr3 are enemy items. A participant should only see one of these items in a CAT.

## 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\\_scoring-service](https://www.assessmentcenter.net/ac_scoring-service)) or a data collection tool that automatically calculates scores (e.g., 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.

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 adult 8-item form, the lowest possible raw score is 8; the highest possible raw score is 40 (see all short form scoring tables in Appendix 1). **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](https://www.assessmentcenter.net/ac_scoring-service)) to generate a final score.

Locate the applicable score conversion table in Appendix 1 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 Adult PROMIS Depression 8b short form, a raw score of 10 converts to a T-score of 46.2 with a standard error (SE) of 2.8 (see scoring table for the 8b short form in appendix). Thus, the 95% confidence interval around the observed score ranges from 40.7 to 51.7 (T-score  $\pm$  (1.96\*SE) or  $46.2 \pm (1.96*2.8)$ ).

**CAT:** A minimum number of items (e.g., 4) must be answered in order to receive a score for the Depression 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. The CAT will continue until either the standard error drops below a specified level (e.g., on the T-score metric 3.0), or the participant has answered the maximum number of questions (e.g., 12), whichever occurs first. For some CATs, specifically “recommended” and “screen-to-CAT” there are additional stopping rules. These include stopping when the standard error isn’t improving much or if a respondent is asymptomatic. For details on the exact stopping rules for Depression CATs, view the Measure Differences summary.

## SCORES

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 depression, a T-score of 60 is one SD worse than average. By comparison, a depression T-score of 40 is one SD better than average.*

**Standard Error (SE):** A PROMIS score includes a T-score and a standard error (SE). The standard error is a measure of the variability for a given T-score across hypothetical repeated measurements. The standard error can be used to construct confidence intervals around a T-score. A 95% confidence interval is common. A 95% confidence interval means there is a 95% probability that the true T-score is within this range. The formula for a 95% confidence interval is  $(T\text{-score} \pm (1.96 * SE))$ . For example, if  $T=52$  and  $SE=2$ , the lower boundary of the confidence interval is  $(52 - (1.96 * 2)) = 48$  and the upper boundary is  $(52 + (1.96 * 2)) = 56$ .

## FREQUENTLY ASKED QUESTIONS (FAQs)

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

Review the HealthMeasures website at [www.healthmeasures.net](http://www.healthmeasures.net).

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 1 — SCORING TABLES

<b>Depression 4a - Adult v1.0</b>		
<i>Short Form Conversion Table</i>		
Raw Score	T-Score	SE*
4	41.0	6.2
5	49.0	3.2
6	51.8	2.7
7	53.9	2.4
8	55.7	2.3
9	57.3	2.3
10	58.9	2.3
11	60.5	2.3
12	62.2	2.3
13	63.9	2.3
14	65.7	2.3
15	67.5	2.3
16	69.4	2.3
17	71.2	2.4
18	73.3	2.4
19	75.7	2.6
20	79.4	2.6

SE\* = Standard Error on T-Score

<b>Depression 6a - Adult v1.0</b>		
<i>Short Form Conversion Table</i>		
Raw Score	T-Score	SE*
6	38.4	5.8
7	45.2	3.4
8	48.3	2.8
9	50.4	2.4
10	52.0	2.2
11	53.4	2.1
12	54.7	2.0
13	55.9	2.0
14	57.0	1.9
15	58.2	1.9
16	59.3	2.0
17	60.5	2.0
18	61.7	2.0
19	62.9	2.0
20	64.2	2.0
21	65.5	2.0
22	66.7	2.0
23	68.0	2.0
24	69.3	2.0
25	70.6	2.0
26	72.0	2.0
27	73.4	2.0
28	75.0	2.1
29	76.9	2.4
30	80.3	3.5

SE\* = Standard Error on T-Score

<b>Depression 8a - Adult v1.0</b>		
<i>Short Form Conversion Table</i>		
Raw Score	T-Score	SE*
8	38.2	5.7
9	44.7	3.3
10	47.5	2.7
11	49.4	2.3
12	50.9	2.0
13	52.1	1.9
14	53.2	1.8
15	54.1	1.8
16	55.1	1.7
17	55.9	1.7
18	56.8	1.7
19	57.7	1.7
20	58.5	1.7
21	59.4	1.7
22	60.3	1.7
23	61.2	1.7
24	62.1	1.8
25	63.0	1.8
26	63.9	1.8
27	64.9	1.8
28	65.8	1.8
29	66.8	1.8
30	67.7	1.8
31	68.7	1.8
32	69.7	1.8
33	70.7	1.8
34	71.7	1.8
35	72.8	1.8
36	73.9	1.8
37	75.0	1.9
38	76.4	2.0
39	78.2	2.4
40	81.3	3.4

SE\* = Standard Error on T-Score

<b>Depression 8b - Adult v1.0</b>		
<i>Short Form Conversion Table</i>		
Raw Score	T-score	SE*
8	37.1	5.5
9	43.3	3.4
10	46.2	2.8
11	48.2	2.4
12	49.8	2.2
13	51.2	2.0
14	52.3	1.9
15	53.4	1.8
16	54.3	1.8
17	55.3	1.7
18	56.2	1.7
19	57.1	1.7
20	57.9	1.7
21	58.8	1.7
22	59.7	1.8
23	60.7	1.8
24	61.6	1.8
25	62.5	1.8
26	63.5	1.8
27	64.4	1.8
28	65.4	1.8
29	66.4	1.8
30	67.4	1.8
31	68.3	1.8
32	69.3	1.8
33	70.4	1.8
34	71.4	1.8
35	72.5	1.8
36	73.6	1.8
37	74.8	1.9
38	76.2	2.0
39	77.9	2.4
40	81.1	3.4

SE\* = Standard Error on T-Score

<b>Pediatric v2.0 - Depressive Symptoms 8a</b>		
<i>Short Form Conversion Table</i>		
Raw Score	T-Score	SE*
8	35.2	5.8
9	40.4	4.6
10	43.2	4.2
11	45.5	3.9
12	47.4	3.7
13	49.1	3.5
14	50.6	3.3
15	52.0	3.2
16	53.3	3.2
17	54.5	3.1
18	55.7	3.1
19	56.8	3.0
20	57.9	3.0
21	59.0	3.0
22	60.0	3.0
23	61.1	3.0
24	62.1	3.0
25	63.1	3.0
26	64.1	3.0
27	65.1	3.0
28	66.1	3.0
29	67.2	2.9
30	68.2	2.9
31	69.3	3.0
32	70.3	3.0
33	71.4	3.0
34	72.6	3.0
35	73.8	3.1
36	75.1	3.2
37	76.5	3.3
38	78.1	3.5
39	79.9	3.6
40	82.4	3.7

SE\* = Standard Error on T-Score



<b>Depressive Symptoms 4a – Early Childhood Parent-Report v1.0</b>		
<i>Short Form Conversion Table</i>		
Raw Score	T-score	SE *
4	42.4	6.8
5	50.5	4.4
6	54.5	3.8
7	57.6	3.5
8	60.2	3.4
9	62.5	3.4
10	64.8	3.4
11	67.0	3.4
12	69.1	3.4
13	70.9	3.4
14	72.7	3.4
15	74.6	3.3
16	76.4	3.3
17	78.4	3.3
18	80.4	3.3
19	82.4	3.2
20	84.9	3.1

\*SE = T-score standard error

<b>Depressive Symptoms 8a – Early Childhood Parent-Report v1.0</b>		
<i>Short Form Conversion Table</i>		
Raw Score	T-score	SE *
8	41.0	6.5
9	47.8	4.4
10	51.0	3.7
11	53.5	3.1
12	55.4	2.8
13	57.1	2.6
14	58.5	2.5
15	59.9	2.5
16	61.2	2.4
17	62.4	2.4
18	63.6	2.4
19	64.8	2.4
20	65.9	2.4
21	67.1	2.4
22	68.2	2.4
23	69.3	2.4
24	70.3	2.4
25	71.3	2.4
26	72.3	2.4
27	73.3	2.4
28	74.3	2.3
29	75.3	2.3
30	76.2	2.3
31	77.2	2.3
32	78.2	2.3
33	79.2	2.3
34	80.2	2.3
35	81.2	2.3
36	82.3	2.4
37	83.5	2.4
38	84.8	2.4
39	86.1	2.3
40	87.4	2.0

SE\* = Standard Error on T-Score

<b>Depressive Symptoms 6a - Parent Proxy v2.0</b>		
<i>Short Form Conversion Table</i>		
Raw Score	T-Score	SE*
6	36.2	5.9
7	42.1	4.4
8	45.4	4.0
9	48.0	3.6
10	50.2	3.4
11	52.1	3.3
12	53.9	3.2
13	55.6	3.1
14	57.2	3.1
15	58.9	3.1
16	60.5	3.1
17	62.1	3.1
18	63.7	3.1
19	65.3	3.1
20	66.8	3.1
21	68.4	3.1
22	70.0	3.1
23	71.5	3.0
24	73.1	3.0
25	74.7	3.0
26	76.3	3.1
27	78.1	3.2
28	80.2	3.3
29	82.5	3.4
30	84.7	3.2

SE\*=Standard Error on T-Score

All scoring tables are based on default Parent Proxy calibrations.

## APPENDIX 2 - SCORING TABLES FOR RETIRED MEASURES

<b>Depressive Symptoms 8a - Pediatric v1.0</b>		
<i>Short Form Conversion Table</i>		
Raw Score	T-Score	SE*
0	35.2	5.8
1	40.4	4.6
2	43.2	4.2
3	45.5	3.8
4	47.4	3.6
5	49.1	3.4
6	50.5	3.3
7	51.9	3.2
8	53.2	3.1
9	54.3	3.1
10	55.5	3.0
11	56.6	3.0
12	57.7	3.0
13	58.7	3.0
14	59.7	2.9
15	60.7	2.9
16	61.8	2.9
17	62.8	2.9
18	63.7	2.9
19	64.7	2.9
20	65.7	2.9
21	66.8	2.9
22	67.8	2.9
23	68.8	2.9
24	69.9	2.9
25	70.9	2.9
26	72.1	3.0
27	73.2	3.0
28	74.5	3.1
29	75.9	3.2
30	77.5	3.4
31	79.3	3.5
32	81.9	3.7

SE\*=Standard Error on T-Score

<b>Depressive Symptoms 8b - Pediatric v1.1</b>		
<i>Short Form Conversion Table</i>		
Raw Score	T-score	SE*
0	35.2	5.8
1	40.4	4.6
2	43.2	4.2
3	45.5	3.9
4	47.4	3.7
5	49.1	3.5
6	50.6	3.3
7	52.0	3.2
8	53.3	3.2
9	54.5	3.1
10	55.7	3.1
11	56.8	3.0
12	57.9	3.0
13	59.0	3.0
14	60.0	3.0
15	61.1	3.0
16	62.1	3.0
17	63.1	3.0
18	64.1	3.0
19	65.1	3.0
20	66.1	3.0
21	67.2	2.9
22	66.2	2.9
23	69.3	3.0
24	70.3	3.0
25	71.4	3.0
26	72.6	3.0
27	73.8	3.1
28	75.1	3.2
29	76.5	3.3
30	78.1	3.5
31	79.9	3.6
32	82.4	3.7

SE\*=Standard Error on T-Score

<b>Depressive Symptoms 6a - Parent Proxy v1.0</b>		
<i>Short Form Conversion Table</i>		
Raw Score	T-Score	SE*
0	36.0	6.0
1	42.0	4.0
2	45.0	4.0
3	48.0	4.0
4	50.0	3.0
5	52.0	3.0
6	54.0	3.0
7	55.0	3.0
8	57.0	3.0
9	59.0	3.0
10	60.0	3.0
11	62.0	3.0
12	64.0	3.0
13	65.0	3.0
14	67.0	3.0
15	68.0	3.0
16	70.0	3.0
17	72.0	3.0
18	73.0	3.0
19	75.0	3.0
20	77.0	3.0
21	78.0	3.0
22	80.0	3.0
23	83.0	4.0
24	86.0	4.0

SE\*=Standard Error on T-Score

<b>Depressive Symptoms 6b - Parent Proxy v1.1</b>		
<i>Short Form Conversion Table</i>		
Raw Score	T-Score	SE*
0	36.0	5.9
1	42.1	4.4
2	45.4	4.0
3	48.0	3.6
4	50.2	3.4
5	52.1	3.3
6	53.9	3.2
7	55.6	3.1
8	57.2	3.1
9	58.9	3.1
10	60.5	3.1
11	62.1	3.1
12	63.7	3.1
13	65.3	3.1
14	66.8	3.1
15	68.4	3.1
16	70.0	3.1
17	71.5	3.0
18	73.1	3.0
19	74.7	3.0
20	76.3	3.1
21	78.1	3.2
22	80.2	3.3
23	82.5	3.4
24	84.7	3.2

SE\*=Standard Error on T-Score