PAIN BEHAVIOR

A brief guide to the PROMIS® Pain Behavior instruments:

ADULT	PEDIATRIC	PARENT PROXY
PROMIS Item Bank v1.0 – Pain Behavior* PROMIS Short Form v1.0 – Pain Behavior 7a* PROMIS Item Bank v1.1 – Pain Behavior* PROMIS Short Form v1.1 – Pain Behavior 7a* PROMIS Item Bank v2.0 – Pain Behavior PROMIS Scale v2.0 – Pain Behavior 20a	PROMIS Pediatric Item Bank v1.0 – Pain Behavior PROMIS Pediatric Short Form v1.0 – Pain Behavior 8a	PROMIS Parent Proxy Item Bank v1.0 – Pain Behavior PROMIS Parent Proxy Short Form v1.0 – Pain Behavior 8a

*retired measure

ABOUT PAIN BEHAVIOR

The PROMIS Pain Behavior item banks measure self-reported external manifestations of pain: behaviors that typically indicate to others that an individual is experiencing pain. These actions or reactions can be verbal or nonverbal, and involuntary or deliberate. They include observable displays (sighing, crying), pain severity behaviors (resting, guarding, facial expressions, and asking for help), and verbal reports of pain. The Pain Behavior short form is universal rather than disease-specific. It assesses pain behavior over the past seven days.

Pain Behavior instruments are available for adults (ages 18+).

INTRODUCTION TO ASSESSMENT OPTIONS

There are three administration options for assessing Pain Behavior: short forms, a scale, and a computerized adaptive test (CAT). A short form is a subset of items from a larger set (i.e., item bank). A scale is the complete set of available items. When administering a short form or scale, 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 (20 items in total). 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 or scale would be more desirable than a CAT. This guide provides information on all Pain Behavior short forms, scales and CAT instruments.

Whether one uses a short form, scale 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 pain behavior 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.

VERSION DIFFERENCES

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 (though not in adult Pain Behavior), 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.

There are multiple versions of the adult PROMIS Pain Behavior item bank. The v1.1 item bank includes the same items as v1.0 items but uses revised item-level calibrations. The v2.0 item bank is the most current bank and the preferred version. There are 20 items in the v2.0 item bank. Six of the v2.0 items are from the v1.0/v1.1 item bank but use a revised item ID (e.g., PAINBE28 → PAINBE28r) and response scores were slightly revised (i.e., "Had no pain" does not contribute to the final measure score). The v2.0 calibrations are on the same metric as v1.1. This means that scores from v1.1 and v2.0 can be compared. Scores from v1.0 should not be compared with v1.1 or v2.0 measures.

To learn more about PROMIS Pain Behavior v2.0, see:
Cook KF, Keefe F, Jensen MP, Roddey TS, Callahan LF, Revicki D, Bamer AM, Kim J, Chung H, Salem R, & Amtmann D. (2013) Development and validation of a new self-report measure of pain behaviors. *Pain, 154*, 2867-2876.

SHORT FORM DIFFERENCES

Adult Short Forms

The v1.0 and v1.1 measures included one adult short form. Items were selected based on content and psychometric characteristics.

The v2.0 measure is available as an item bank or a scale. An item bank is intended to be administered as a computer adaptive test or a subset of items can be selected to create a custom fixed length short form. The v2.0 scale is identical to the item bank. That is, it contains all 20 items. The scale is intended to be administered in its entirety.

Pediatric and Parent Proxy Short Forms

There is one Pediatric and one Parent Proxy short form. Items were selected based on content and psychometric characteristics.

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/Scale: 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.

If you are scoring either a v1.0 or v1.1 adult pain behavior measure, use the v1.1 scoring table included in this manual. The HealthMeasures Scoring Service does not include the v1.0 or v1.1 measures. Scores calculated from the v1.1 table in this manual can be compared with v2.0 pain behavior measures. If you would like to transform v1.0 scores to the v1.1/v2.0 metric using statistical software, please contact us.

If you are scoring a scale or custom short form, the HealthMeasures Scoring Service is the preferred approach. If you are using a scale, use the v2.0 scoring table included in this manual.

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 v2.0 adult 20-item form, the lowest possible raw score is 204; the highest possible raw score is 100 (see short form scoring table 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_scoringservice) to generate a final score.

With the total raw score for a measure, 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 Pain Behavior 20a scale version v2.0, a raw score of 20 converts to a T-score of 32.9 with a standard error (SE) of 0.53 (see scoring table for the 7a v2.0 short form in Appendix 1). Thus, the 95% confidence interval around the observed score ranges from 31.9 to 33.9 (T-score $\pm$ (1.96*SE) or 32.9 $\pm$ (1.96*0.53).

CAT: A minimum number of items (4 for adult CAT) must be answered in order to receive a score for the Pain Behavior 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. CAT will continue until either the standard error drops below a specified level (on the T-score metric 3.0 for adult 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 Pain Behavior, a T-score of 60 is one SD worse than average. By comparison, a Pain Behavior T-score of 40 is one SD better than average.

STATISTICAL CHARACTERISTICS

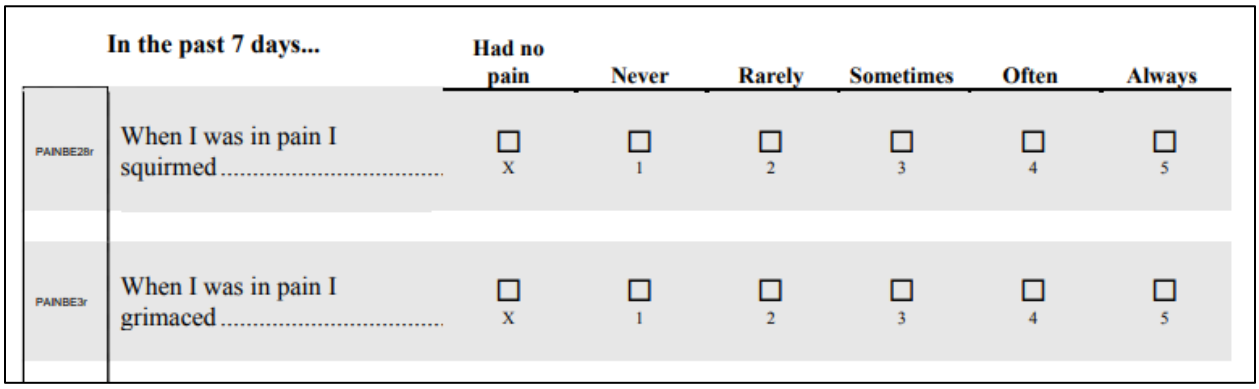
There are four key features of the score for Pain Behavior:

- **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.

More information is available at HealthMeasures.net.

PREVIEW OF SAMPLE ITEM



In the past 7 days...	Had no pain	Never	Rarely	Sometimes	Often	Always
PAINBE28r When I was in pain I squirmed	☐ X	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
PAINBE3r When I was in pain I grimaced	☐ X	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Figure 1

Figure 1 is an excerpt from the paper version of the adult v2.0 scale. This is the paper version format used for all Pain Behavior instruments. It is important to note, CAT is not available for paper administration.

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?

These instruments are currently not available in other languages. The HealthMeasures website 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 1-SCORING TABLE MEASURE

Pain Behavior v2.0 Scale 20a								
Conversion Table								
Summary Score	SE	Tscore	Summary Score	SE	Tscore	Summary Score	SE	Tscore
20	0.53	32.9	51	0.14	57.8	82	0.14	68.7
21	0.38	38.0	52	0.14	58.1	83	0.14	69.1
22	0.33	40.6	53	0.14	58.5	84	0.15	69.6
23	0.30	42.4	54	0.13	58.8	85	0.15	70.0
24	0.27	43.8	55	0.13	59.1	86	0.15	70.5
25	0.25	45.0	56	0.14	59.5	87	0.15	70.9
26	0.23	46.0	57	0.14	59.8	88	0.16	71.4
27	0.21	46.9	58	0.14	60.1	89	0.16	71.9
28	0.20	47.7	59	0.14	60.5	90	0.17	72.5
29	0.19	48.4	60	0.14	60.8	91	0.17	73.0
30	0.18	49.0	61	0.14	61.1	92	0.18	73.7
31	0.18	49.6	62	0.14	61.5	93	0.19	74.3
32	0.17	50.2	63	0.14	61.8	94	0.20	75.1
33	0.16	50.7	64	0.14	62.2	95	0.22	76.0
34	0.16	51.2	65	0.14	62.5	96	0.24	76.9
35	0.16	51.7	66	0.14	62.9	97	0.26	78.1
36	0.15	52.1	67	0.14	63.2	98	0.29	79.5
37	0.15	52.6	68	0.14	63.5	99	0.33	81.2
38	0.15	53.0	69	0.14	63.9	100	0.39	83.7
39	0.15	53.4	70	0.14	64.3			
40	0.14	53.8	71	0.14	64.6			
41	0.14	54.2	72	0.14	65.0			
42	0.14	54.6	73	0.14	65.3			
43	0.14	55.0	74	0.14	65.7			
44	0.14	55.3	75	0.14	66.0			
45	0.14	55.7	76	0.14	66.4			
46	0.14	56.0	77	0.14	66.8			
47	0.14	56.4	78	0.14	67.2			
48	0.14	56.7	79	0.14	67.5			
49	0.14	57.1	80	0.14	67.9			
50	0.14	57.4	81	0.14	68.3			

Pediatric v1.0 Pain Behavior 8a					
Short Form Conversion Table					
Raw Summed Score	T-Score	SE*	Raw Summed Score	T-Score	SE*
8	20.0	N/A	29	51.0	2.2
9	28.4	3.3	30	51.8	2.2
10	30.7	2.8	31	52.5	2.2
11	32.6	2.6	32	53.2	2.2
12	34.2	2.5	33	53.9	2.2
13	35.7	2.5	34	54.7	2.2
14	37.1	2.5	35	55.4	2.3
15	38.5	2.5	36	56.2	2.3
16	39.8	2.5	37	57.0	2.3
17	40.9	2.5	38	57.7	2.3
18	41.9	2.5	39	58.5	2.3
19	42.9	2.5	40	59.4	2.3
20	43.9	2.4	41	60.3	2.4
21	44.8	2.4	42	61.2	2.4
22	45.7	2.4	43	62.2	2.5
23	46.5	2.3	44	63.3	2.7
24	47.3	2.3	45	64.6	2.8
25	48.1	2.3	46	66.2	3.2
26	48.8	2.3	47	67.9	3.4
27	49.6	2.2	48	80.0	N/A
28	50.3	2.2			
*SE = Standard Error on T-score metric					
Marginal reliability of the scaled scores for summed scores = 0.93604					

Parent Proxy v1.0 Pain Behavior 8a					
Short Form Conversion Table					
Raw Summed Score	T-Score	SE*	Raw Summed Score	T-Score	SE*
8	10.0	N/A	29	47.6	2.7
9	19.0	3.1	30	48.7	2.6
10	21.2	2.9	31	49.6	2.4
11	23.3	2.8	32	50.5	2.4
12	25.2	2.6	33	51.4	2.6
13	27.0	2.9	34	52.5	2.7
14	29.1	3.0	35	53.7	2.6
15	31.0	3.0	36	54.6	2.4
16	32.9	3.0	37	55.5	2.4
17	34.5	2.9	38	56.4	2.6
18	35.9	2.9	39	57.6	2.7
19	37.3	2.9	40	58.7	2.6
20	38.6	2.7	41	59.7	2.3
21	39.6	2.5	42	60.6	2.4
22	40.6	2.5	43	61.7	2.7
23	41.6	2.7	44	63.1	2.9
24	42.8	2.7	45	64.6	2.9
25	43.8	2.5	46	66.2	3.1
26	44.7	2.3	47	68.1	3.5
27	45.5	2.4	48	80.0	N/A
28	46.5	2.6			
*SE = Standard Error on T-score metric					
Marginal reliability of the scaled scores for summed scores = 0.92771					

APPENDIX 2-SCORING TABLE FOR RETIRED MEASURE

PROMIS SF V1.1 Pain Behavior 7a *Short Form Conversion Table*		
Raw Score	**T-score**	**SE***
7	34.1	4.9
8	39.3	2.8
9	41.5	2.2
10	43.1	2.0
11	44.5	2.0
12	45.9	2.0
13	47.3	2.0
14	48.6	2.1
15	49.9	2.1
16	51.1	2.1
17	52.3	2.1
18	53.4	2.1
19	54.4	2.0
20	55.4	2.0
21	56.4	1.9
22	57.3	1.9
23	58.1	1.9
24	59.0	1.9
25	59.8	1.8
26	60.6	1.8
27	61.4	1.8
28	62.2	1.8
29	63.0	1.8
30	63.7	1.8
31	64.5	1.8
32	65.3	1.8
33	66.1	1.8
34	66.9	1.8
35	67.8	1.9
36	68.7	1.9
37	69.7	2.0
38	70.8	2.1
39	72.0	2.2
40	73.5	2.4
41	75.3	2.7
42	78.9	3.6
*** SE = Standard error**		