# PROMIS SCORING GUIDE: TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview</td>
<td>2</td>
</tr>
<tr>
<td>Adult Version 1.0 Short Forms: Introduction</td>
<td>3</td>
</tr>
<tr>
<td>Anger</td>
<td>4</td>
</tr>
<tr>
<td>Anxiety</td>
<td>7</td>
</tr>
<tr>
<td>Depression</td>
<td>10</td>
</tr>
<tr>
<td>Fatigue</td>
<td>13</td>
</tr>
<tr>
<td>Pain Behavior</td>
<td>16</td>
</tr>
<tr>
<td>Pain Interference</td>
<td>19</td>
</tr>
<tr>
<td>Physical Function</td>
<td>22</td>
</tr>
<tr>
<td>Satisfaction with Participation in Discretionary Social Activities</td>
<td>25</td>
</tr>
<tr>
<td>Satisfaction with Participation in Social Roles</td>
<td>28</td>
</tr>
<tr>
<td>Sleep Disturbance</td>
<td>31</td>
</tr>
<tr>
<td>Sleep-Related Impairment</td>
<td>34</td>
</tr>
<tr>
<td>Adult Profile Short Forms: Introduction</td>
<td>37</td>
</tr>
<tr>
<td>Anxiety</td>
<td>38</td>
</tr>
<tr>
<td>Depression</td>
<td>41</td>
</tr>
<tr>
<td>Fatigue</td>
<td>44</td>
</tr>
<tr>
<td>Pain Interference</td>
<td>47</td>
</tr>
<tr>
<td>Physical Function</td>
<td>50</td>
</tr>
<tr>
<td>Satisfaction with Participation in Social Roles</td>
<td>53</td>
</tr>
<tr>
<td>Sleep Disturbance</td>
<td>56</td>
</tr>
<tr>
<td>Adult Computerized Adaptive Testing: Introduction</td>
<td>59</td>
</tr>
<tr>
<td>Anxiety</td>
<td>60</td>
</tr>
<tr>
<td>Depression</td>
<td>63</td>
</tr>
<tr>
<td>Fatigue</td>
<td>66</td>
</tr>
<tr>
<td>Pain Interference</td>
<td>69</td>
</tr>
<tr>
<td>Physical Function</td>
<td>72</td>
</tr>
<tr>
<td>Satisfaction with Participation in Social Roles</td>
<td>75</td>
</tr>
<tr>
<td>Sleep Disturbance</td>
<td>78</td>
</tr>
<tr>
<td>Glossary</td>
<td>81</td>
</tr>
<tr>
<td>Citations</td>
<td>82</td>
</tr>
</tbody>
</table>
PROMIS SCORING GUIDE: OVERVIEW

This document includes brief reports on three sets of PROMIS instruments:

- Version 1.0 short forms
- Profile instruments
- Computerized adaptive testing (CAT)

Version 1.0 short forms were the first instruments developed from PROMIS item banks. Careful attention was paid to item content and the ability of the short form to measure the full range of the concept or ‘domain’ measured by the item bank. Domain experts independently reviewed short forms to give input on the relevance of each item. Psychometric properties and clinical input were considered together. With the exception of one physical function short form, which is 20 items in length, the version 1.0 short forms range in length from six to 10 items.

The profile instruments are a collection of profile short forms, where each form contains items from one of seven primary PROMIS domains (depression, anxiety, physical function, pain interference, fatigue, sleep disturbance, and satisfaction with participation in social roles). Profile short forms, which are four, six, and eight items in length, are meant to be administered together with other profile short forms from each of the other primary domains, thus generating a ‘profile’ of seven scores. There are three profile instruments: the PROMIS-29 profile instrument, so named because it contains 29 items, includes four items from each primary domain plus a single pain intensity rating; the PROMIS-43 profile instrument includes six items from each primary domain plus the pain intensity item; and the PROMIS-57 profile instrument includes eight items from each primary domain plus the pain intensity item. Each profile short form can also be administered independently.

Computerized adaptive testing (CAT) is a flexible, computer-driven assessment that can use any items in the item bank to measure each domain. CAT selects only those items that sharpen the estimate of a respondent’s score on the domain being measured. CAT length varies but usually includes four to seven items.

All PROMIS instruments are generated from item banks. The development of items for each domain’s item bank consisted of six phases: identification of items, item classification and selection, item review and revision, focus group input on domain coverage, cognitive interviews with individual items, and final revision before field testing. Identification of items refers to the systematic search for existing items in currently available scales. Item review and revision was conducted by domain experts who reviewed the wording of each item and revised as appropriate for conventions adopted by PROMIS. Focus groups were used to confirm domain definitions and to identify new areas of item development for future PROMIS item banks. Cognitive interviews were used to examine individual items. Items successfully screened through this process were sent to field testing and subjected to innovative scale construction procedures. Final revisions were then made based on the compiled information.
There are two types of PROMIS short forms: the version 1.0 short forms that were initially developed in 2007 and 2008, and the profile short forms that were created in 2008 for PROMIS profile assessments. Version 1.0 short forms were the first instruments developed from PROMIS item banks. Careful attention was paid to item content and the ability of the short form to measure the full range of the concept or ‘domain’ measured by the item bank. Domain experts independently reviewed short forms to give input on the relevance of each item. Psychometric properties and clinical input were considered together. With the exception of one physical function short form, which is 20 items in length, the version 1.0 short forms range in length from six to 10 items.

By contrast, the profile short forms are ‘high information’ short forms focusing on seven primary PROMIS domains (anxiety, depression, sleep disturbance, fatigue, pain interference, physical function, and satisfaction with participation in social roles). For each domain, there are three profile short forms, which are four, six, and eight items in length. The items for each of the forms are overlapping (e.g., the eight-item form is the six-item form plus two additional items). These short forms are so named because when a short form of the same length from each of the seven primary domains is administered together, this offers a ‘profile’ of seven scores. There are three PROMIS profile instruments: the PROMIS-29 profile instrument contains 29 items, including four items from each primary domain plus a single pain intensity rating; the PROMIS-43 profile instrument includes six items from each primary domain plus the pain intensity item; and the PROMIS-57 profile instrument includes eight items from each primary domain plus the pain intensity item.

When selecting a short form, the main difference is instrument length. Reliability and precision of short forms within a domain are highly similar. If you are working with a sample in which you expect large variability in a domain and you want to include the full range of item content from that domain, you would probably prefer a version 1.0 short form. If you are working with a sample in which you want the most precise measure, the eight-item profile short form would be your best bet. On the other hand, if you are hoping to capture secondary outcomes data, but have little room for additional measures, you would probably prefer the four-item profile short form.
ADULT ANGER VERSION 1.0 SHORT FORM
A brief guide to the 8-item PROMIS Short Form v1.0 – Anger 8a

ABOUT ANGER*

The anger item bank measures self-reported angry mood (irritability, frustration), negative social cognitions (interpersonal sensitivity, envy, disagreeableness), and efforts to control anger. Often associated with episodes of frustration that impede goal-directed behavior, anger is marked by attitudes of hostility and cynicism. Specific components relate to verbal and non-verbal evidence of anger. Physical aggression items are not included. The anger short form is generic rather than disease-specific. It assesses anger over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing anger: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (29 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 the eight-item anger version 1.0 short form. The strength of the eight-item instrument lies in its focus on item content and its ability to assess the full range of anger measured by the anger item bank. When selecting a short form, the main difference is instrument length. Reliability and precision of short forms within a domain are highly similar. Longer short forms generally offer greater correlation (strength of relationship) with the full item bank, as well as greater precision. If you are working with a sample in which you expect large variability in a domain and you want to include the full range of item content from that domain, you would probably prefer this eight-item version 1.0 short form. On the other hand, if you are hoping to capture secondary outcomes data, but have little room for additional measures, you would probably prefer a very brief (four-item) profile short form.

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. Longer CAT offers greater correlation with the full item bank, 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.

Whether one uses a short form or CAT, the score metric is based upon Item Response Theory (IRT), a family of statistical models that link individual questions to a presumed underlying trait or concept of anger represented by all items in the item bank.

Graphical reports, which visually illustrate results, are available in Assessment Center if you choose to administer a PROMIS Profile, which includes a profile short form from seven PROMIS domains (physical function, depression, anxiety, pain interference, fatigue, sleep disturbance, and satisfaction with participation in social roles). To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the eight-item form, the lowest possible raw score is 8; the highest possible raw score is 40 (Table 1).

You can use Table 1 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 8-item form, a raw score of 24 converts to a T-score of 60.7 with a standard error (SE) of 2.6. Thus, the 95% confidence interval around the actual observed score ranges from 55.6 to 65.8 (T-score ± (1.96*SE) = 60.7 ± 5.1 = 55.6 to 65.8).

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

You can upload data to a free computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at nihpromis.org.

STATISTICAL CHARACTERISTICS

There are three key features of the score for anger:

- **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 – SE²).

- **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/SE²).

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, the dotted horizontal line represents a degree of internal consistency reliability (i.e., .90) typically regarded as sufficient for an accurate individual score. The shaded blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the eight-item form. Figure 1 also tells us where on the scale the form is most informative based upon the T-score. This form would typically be more informative than an anger form with fewer items.

<table>
<thead>
<tr>
<th>Raw Score</th>
<th>T-score</th>
<th>SE*</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>92.4</td>
<td>2.3</td>
</tr>
<tr>
<td>30</td>
<td>90.6</td>
<td>2.3</td>
</tr>
<tr>
<td>28</td>
<td>88.9</td>
<td>2.3</td>
</tr>
<tr>
<td>26</td>
<td>87.2</td>
<td>2.3</td>
</tr>
<tr>
<td>24</td>
<td>85.5</td>
<td>2.3</td>
</tr>
<tr>
<td>22</td>
<td>83.7</td>
<td>2.3</td>
</tr>
<tr>
<td>20</td>
<td>82.0</td>
<td>2.3</td>
</tr>
<tr>
<td>18</td>
<td>80.3</td>
<td>2.3</td>
</tr>
</tbody>
</table>

*SE = Standard Error on T-score

Table 1

**TEST INFORMATION PLOT**

Figure 1
PREVIEW OF SAMPLE ITEMS

Figure 2 is an excerpt from the paper version of the eight-item short form. This instrument is also available for online administration. There are a variety of formatting options for the online version.

<table>
<thead>
<tr>
<th>In the past 7 days...</th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>I was irritated more than people knew ...</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>I made myself angry about something I was irritating about ...</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Figure 2

FREQUENTLY ASKED QUESTIONS

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

This short form is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 2 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

**Q: Do I need to register with PROMIS to use this short form?**

Yes, to get a copy of this short form, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

**Q: Is this short form available in other languages?**

Yes, this short form is currently available in Spanish. The PROMIS group is also working to translate this form into other languages. Information on available translations is updated periodically at nihpromis.org.

**Q: Can I make my own short form?**

Yes, custom anger short forms can be made by selecting any items from the item bank. Instructions can be found in the Assessment Center user manual. The full item bank is available at assessmentcenter.net.
ABOUT ANXIETY*

The anxiety item bank measures self-reported fear (fearfulness, panic), anxious misery (worry, dread), hyperarousal (tension, nervousness, restlessness), and somatic symptoms related to arousal (racing heart, dizziness). Anxiety is best differentiated by symptoms that reflect autonomic arousal and experience of threat. Only one behavioral avoidance item is included in the item bank; therefore, behavioral fear avoidance is not fully evaluated. The anxiety short form is generic rather than disease-specific. It assesses anxiety over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing anxiety: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (124 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 the seven-item anxiety version 1.0 short form. The strength of the seven-item instrument lies in its focus on item content and its ability to assess the full range of anxiety measured by the anxiety item bank. When selecting a short form, the main difference is instrument length. Reliability and precision of short forms within a domain are highly similar. Longer short forms generally offer greater correlation (strength of relationship) with the full item bank, as well as greater precision. If you are working with a sample in which you expect large variability in a domain and you want to include the full range of item content from that domain, you would probably prefer this seven-item version 1.0 short form. On the other hand, if you are hoping to capture secondary outcomes data, but have little room for additional measures, you would probably prefer a very brief (four-item) profile short form.

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. Longer CAT offers greater correlation with the full item bank, 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.

Whether one uses a short form or CAT, the score metric is based upon Item Response Theory (IRT), a family of statistical models that link individual questions to a presumed underlying trait or concept of anxiety represented by all items in the item bank.

Graphical reports, which visually illustrate results, are available in Assessment Center if you choose to administer a PROMIS Profile, which includes a profile short form from seven PROMIS domains (physical function, depression, anxiety, pain interference, fatigue, sleep disturbance, and satisfaction with participation in social roles). To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the seven-item form, the lowest possible raw score is 7; the highest possible raw score is 35 (Table 1).

You can use Table 1 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 7-item form, a raw score of 21 converts to a T-score of 61.3 with a standard error (SE) of 2.2. Thus, the 95% confidence interval around the actual observed score ranges from 57.0 to 65.6 (T-score ± (1.96*SE) = 61.3 ± 4.3 = 57.0 to 65.6).

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

You can upload data to a free computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at nihpromis.org.

STATISTICAL CHARACTERISTICS

There are three key features of the score for anxiety:

- **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 – SE²).
- **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/SE²).

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, the two dotted 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 blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the seven-item form. Figure 1 also tells us where on the scale the form is most informative based upon the T-score. This form would typically be more informative than an anxiety form with fewer items.
PREVIEW OF SAMPLE ITEMS

Figure 1 is an excerpt from the paper version of the seven-item short form. This instrument is also available for online administration. There are a variety of formatting options for the online version.

<table>
<thead>
<tr>
<th>In the past 7 days...</th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>I felt fearful.........</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>I felt anxious.........</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

FREQUENTLY ASKED QUESTIONS

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

This short form is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 1 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

**Q: Do I need to register with PROMIS to use this short form?**

Yes, to get a copy of this short form, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

**Q: Is this short form available in other languages?**

Yes, this short form is currently available in Spanish. The PROMIS group is also working to translate this form into other languages. Information on available translations is updated periodically at nihpromis.org.

**Q: Can I make my own short form?**

Yes, custom anxiety short forms can be made by selecting any items from the item bank. Instructions can be found in the Assessment Center user manual. The full item bank is available at assessmentcenter.net.
ADULT DEPRESSION VERSION 1.0 SHORT FORM
A brief guide to the 8-item PROMIS Short Form v1.0 – Depression 8b

ABOUT DEPRESSION*

The depression item bank assesses self-reported negative mood (sadness, guilt), views of self (self-criticism, worthlessness), and social cognition (loneliness, interpersonal alienation), as well as decreased positive affect and engagement (loss of interest, meaning, and purpose). Somatic symptoms (changes in appetite, sleeping patterns) are not included, which eliminates consideration of these items’ confounding effects when assessing patients with comorbid physical conditions. The depression short form is generic rather than disease-specific. It assesses depression over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing depression: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (28 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 the eight-item depression version 1.0 short form. The strength of the eight-item version 1.0 instrument lies in its focus on item content and its ability to assess the full range of depression measured by the depression item bank. When selecting a short form, the main difference is instrument length. Reliability and precision of short forms within a domain are highly similar. Longer short forms generally offer greater correlation (strength of relationship) with the full item bank, as well as greater precision. If you are working with a sample in which you expect large variability in a domain and you want to include the full range of item content from that domain, you would probably prefer this eight-item version 1.0 short form. On the other hand, if you are hoping to capture secondary outcomes data, but have little room for additional measures, you would probably prefer a very brief (four-item) profile short form.

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. Longer CAT offers greater correlation with the full item bank, 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.

Whether one uses a short form or CAT, the score metric is based upon Item Response Theory (IRT), a family of statistical models that link individual questions to a presumed underlying trait or concept of depression represented by all items in the item bank.

Graphical reports, which visually illustrate results, are available in Assessment Center if you choose to administer a PROMIS Profile, which includes a profile short form from seven PROMIS domains (physical function, depression, anxiety, pain interference, fatigue, sleep disturbance, and satisfaction with participation in social roles). To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the eight-item form, the lowest possible raw score is 8; the highest possible raw score is 40 (Table 1).

You can use Table 1 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 8-item form, a raw score of 24 converts to a T-score of 61.6 with a standard error (SE) of 1.8. Thus, the 95% confidence interval around the actual observed score ranges from 58.1 to 65.1 (T-score ± (1.96*SE) = 61.6 ± 3.5 = 58.1 to 65.1).

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.

You can upload data to a free computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at nihpromis.org.

STATISTICAL CHARACTERISTICS

There are three key features of the score for depression:

- **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 – SE²).

- **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/SE²).

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, the two dotted 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 blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the eight-item form. Figure 1 also tells us where on the scale the form is most informative based upon the T-score. This form would typically be more informative than a depression form with fewer items.
PREVIEW OF SAMPLE ITEMS

Figure 2 is an excerpt from the paper version of the eight-item short form. This instrument is also available for online administration. There are a variety of formatting options for the online version.

FREQUENTLY ASKED QUESTIONS

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

This short form is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 2 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

Q: Do I need to register with PROMIS to use this short form?

Yes, to get a copy of this short form, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

Q: Is this short form available in other languages?

Yes, this short form is currently available in Spanish. The PROMIS group is also working to translate this form into other languages. Information on available translations is updated periodically at nihpromis.org.

Q: Can I make my own short form?

Yes, custom depression short forms can be made by selecting any items from the item bank. Instructions can be found in the Assessment Center user manual. The full item bank is available at assessmentcenter.net.
ABOUT FATIGUE*

The fatigue item bank evaluates a range of self-reported symptoms, from mild subjective feelings of tiredness to an overwhelming, debilitating, and sustained sense of exhaustion that likely decreases one’s ability to execute daily activities and function normally in family or social roles. Fatigue is divided into the experience of fatigue (frequency, duration, and intensity) and the impact of fatigue on physical, mental, and social activities. The fatigue short form is generic rather than disease-specific. It assesses fatigue over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing fatigue: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (95 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 the seven-item fatigue version 1.0 short form. The strength of the seven-item instrument lies in its focus on item content and its ability to assess the full range of fatigue measured by the fatigue item bank. When selecting a short form, the main difference is instrument length. Reliability and precision of short forms within a domain are highly similar. Longer short forms generally offer greater correlation (strength of relationship) with the full item bank, as well as greater precision. If you are working with a sample in which you expect large variability in a domain and you want to include the full range of item content from that domain, you would probably prefer this seven-item version 1.0 short form. On the other hand, if you are hoping to capture secondary outcomes data, but have little room for additional measures, you would probably prefer a very brief (four-item) profile short form.

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. Longer CAT offers greater correlation with the full item bank, 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.

Whether one uses a short form or CAT, the score metric is based upon Item Response Theory (IRT), a family of statistical models that link individual questions to a presumed underlying trait or concept of fatigue represented by all items in the item bank.

Graphical reports, which visually illustrate results, are available in Assessment Center if you choose to administer a PROMIS Profile, which includes a profile short form from seven PROMIS domains (physical function, depression, anxiety, pain interference, fatigue, sleep disturbance, and satisfaction with participation in social roles). To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the seven-item form, the lowest possible raw score is 7; the highest possible raw score is 35 (Table 1).

You can use Table 1 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 7-item form, a raw score of 21 converts to a T-score of 57.8 with a standard error (SE) of 2.9. Thus, the 95% confidence interval around the actual observed score ranges from 52.1 to 63.5 (T-score ± (1.96*SE) = 57.8 ± 5.7 = 52.1 to 63.5). Important: A higher PROMIS T-score represents more of the concept being measured. For negatively-worded concepts like fatigue, a T-score of 60 is one SD worse than average. By comparison, a fatigue T-score of 40 is one SD better than average.

You can upload data to a free computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at nihpromis.org.

STATISTICAL CHARACTERISTICS

There are three key features of the score for fatigue:

- **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 – SE²).
- **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/SE²).

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, the dotted horizontal line represents a degree of internal consistency reliability (i.e., .90) typically regarded as sufficient for an accurate individual score. The shaded blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the seven-item form. Figure 1 also tells us where on the scale the form is most informative based upon the T-score. This form would typically be more informative than a fatigue form with fewer items.
PREVIEW OF SAMPLE ITEMS

Figure 2 is an excerpt from the paper version of the seven-item short form. This instrument is also available for online administration. There are a variety of formatting options for the online version.

<table>
<thead>
<tr>
<th>In the past 7 days...</th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>How often did you feel tired?..............</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>How often did you experience extreme exhaustion?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

FREQUENTLY ASKED QUESTIONS

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

This short form is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 2 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

Q: Do I need to register with PROMIS to use this short form?

Yes, to get a copy of this short form, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

Q: Is this short form available in other languages?

Yes, this short form is currently available in Spanish. The PROMIS group is also working to translate this form into other languages. Information on available translations is updated periodically at nihpromis.org.

Q: Can I make my own short form?

Yes, custom fatigue short forms can be made by selecting any items from the item bank. Instructions can be found in the Assessment Center user manual. The full item bank is available at assessmentcenter.net.
ADULT PAIN BEHAVIOR VERSION 1.0 SHORT FORM
A brief guide to the 7-item PROMIS Short Form v1.0 – Pain Behavior 7a

ABOUT PAIN BEHAVIOR*

The pain behavior item bank measures 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 non-verbal, 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 generic rather than disease-specific. It assesses pain behavior over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing pain behavior: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (39 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 the seven-item pain behavior version 1.0 short form. The strength of the seven-item instrument lies in its focus on item content and its ability to assess the full range of pain behavior measured by the pain behavior item bank. When selecting a short form, the main difference is instrument length. Reliability and precision of short forms within a domain are highly similar. Longer short forms generally offer greater correlation (strength of relationship) with the full item bank, as well as greater precision. If you are working with a sample in which you expect large variability in a domain and you want to include the full range of item content from that domain, you would probably prefer this seven-item version 1.0 short form. On the other hand, if you are hoping to capture secondary outcomes data, but have little room for additional measures, you would probably prefer a very brief (four-item) profile short form.

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. Longer CAT offers greater correlation with the full item bank, 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.

Whether one uses a short form or CAT, the score metric is based upon 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.

Graphical reports, which visually illustrate results, are available in Assessment Center if you choose to administer a PROMIS Profile, which includes a profile short form from seven PROMIS domains (physical function, depression, anxiety, pain interference, fatigue, sleep disturbance, and satisfaction with participation in social roles). To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has six response options ranging in value from one to six. To find the total raw score, sum the values of the response to each question. For example, for the seven-item form, the lowest possible raw score is 7; the highest possible raw score is 42 (Table 1).

You can use Table 1 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 7-item form, a raw score of 24 converts to a T-score of 58.6 with a standard error (SE) of 1.6. Thus, the 95% confidence interval around the actual observed score ranges from 55.7 to 61.5 (T-score ± (1.96*SE) = 58.6 ± 2.9 = 55.7 to 61.5).

**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.

You can upload data to a free computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at nihpromis.org.

### STATISTICAL CHARACTERISTICS

There are three 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 – SE²).
- **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/SE²).

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, the two dotted 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 blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the seven-item form. Figure 1 also tells us where on the scale the form is most informative based upon the T-score. This form would typically be more informative than a pain behavior form with fewer items.

<table>
<thead>
<tr>
<th>Pain Behavior 7a Short-Form Conversion Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw Score</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>36.7</td>
</tr>
<tr>
<td>42.5</td>
</tr>
<tr>
<td>44.4</td>
</tr>
<tr>
<td>45.7</td>
</tr>
<tr>
<td>46.9</td>
</tr>
<tr>
<td>48.9</td>
</tr>
<tr>
<td>50.1</td>
</tr>
<tr>
<td>51.1</td>
</tr>
<tr>
<td>52.4</td>
</tr>
<tr>
<td>53.0</td>
</tr>
<tr>
<td>53.9</td>
</tr>
<tr>
<td>54.8</td>
</tr>
<tr>
<td>55.8</td>
</tr>
<tr>
<td>56.4</td>
</tr>
<tr>
<td>57.1</td>
</tr>
<tr>
<td>57.2</td>
</tr>
<tr>
<td>58.6</td>
</tr>
<tr>
<td>59.2</td>
</tr>
<tr>
<td>59.9</td>
</tr>
<tr>
<td>60.6</td>
</tr>
<tr>
<td>61.2</td>
</tr>
<tr>
<td>61.9</td>
</tr>
<tr>
<td>62.5</td>
</tr>
<tr>
<td>63.1</td>
</tr>
<tr>
<td>63.8</td>
</tr>
<tr>
<td>64.4</td>
</tr>
<tr>
<td>65.1</td>
</tr>
<tr>
<td>65.8</td>
</tr>
<tr>
<td>66.6</td>
</tr>
<tr>
<td>67.4</td>
</tr>
<tr>
<td>68.3</td>
</tr>
<tr>
<td>69.3</td>
</tr>
<tr>
<td>70.5</td>
</tr>
<tr>
<td>72.1</td>
</tr>
<tr>
<td>75.9</td>
</tr>
</tbody>
</table>

Table 1
PREVIEW OF SAMPLE ITEMS

Figure 2 is an excerpt from the paper version of the seven-item short form. This instrument is also available for online administration. There are a variety of formatting options for the online version.

<table>
<thead>
<tr>
<th>In the past 7 days....</th>
<th>Had no Pain</th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>When I was in pain I became irritable.................................</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>When I was in pain I grimaced ...........</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FREQUENTLY ASKED QUESTIONS

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

This short form is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 2 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

Q: Do I need to register with PROMIS to use this short form?

Yes, to get a copy of this short form, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

Q: Is this short form available in other languages?

Yes, this short form is currently available in Spanish. The PROMIS group is also working to translate this form into other languages. Information on available translations is updated periodically at nihpromis.org.

Q: Can I make my own short form?

Yes, custom pain behavior short forms can be made by selecting any items from the item bank. Instructions can be found in the Assessment Center user manual. The full item bank is available at assessmentcenter.net.
ADULT PAIN INTERFERENCE VERSION 1.0 SHORT FORM
A brief guide to the 6-item PROMIS Short Form v1.0 – Pain Interference 6b

ABOUT PAIN INTERFERENCE*
The pain interference item bank measures the self-reported consequences of pain on relevant aspects of one’s life. This includes the extent to which pain hinders engagement with social, cognitive, emotional, physical, and recreational activities. Pain interference also incorporates items probing sleep and enjoyment in life, though the item bank only contains one sleep item. The pain interference short form is generic rather than disease-specific. It assesses pain interference over the past seven days. (*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS
There are two administration options for assessing pain interference: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (41 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 the six-item pain interference version 1.0 short form. The strength of the six-item version 1.0 instrument lies in its focus on item content and its ability to assess the full range of pain interference measured by the pain interference item bank. When selecting a short form, the main difference is instrument length. Reliability and precision of short forms within a domain are highly similar. Longer short forms generally offer greater correlation (strength of relationship) with the full item bank, as well as greater precision. If you are working with a sample in which you expect large variability in a domain and you want to include the full range of item content from that domain, you would probably prefer this six-item version 1.0 short form. On the other hand, if you are hoping to capture secondary outcomes data, but have little room for additional measures, you would probably prefer a very brief (four-item) profile short form.

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. Longer CAT offers greater correlation with the full item bank, 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.

Whether one uses a short form or CAT, the score metric is based upon Item Response Theory (IRT), a family of statistical models that link individual questions to a presumed underlying trait or concept of pain interference represented by all items in the item bank.

Graphical reports, which visually illustrate results, are available in Assessment Center if you choose to administer a PROMIS Profile, which includes a profile short form from seven PROMIS domains (physical function, depression, anxiety, pain interference, fatigue, sleep disturbance, and satisfaction with participation in social roles). To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the six-item form, the lowest possible raw score is 6; the highest possible raw score is 30 (Table 1).

You can use Table 1 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 6-item form, a raw score of 18 converts to a T-score of 61.8 with a standard error (SE) of 1.6. Thus, the 95% confidence interval around the actual observed score ranges from 58.7 to 64.9 (T-score ± (1.96*SE) = 61.8 ± 3.1 = 58.7 to 64.9).

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

You can upload data to a free computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at nihpromis.org.

STATISTICAL CHARACTERISTICS

There are three key features of the score for pain interference:

- **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 – SE²).

- **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/SE²).

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, the two dotted 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 blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the six-item form. Figure 1 also tells us where on the scale the form is most informative based upon the T-score. This form would typically be more informative than a pain interference form with fewer items.
PREVIEW OF SAMPLE ITEMS

Figure 2 is an excerpt from the paper version of the six-item short form. This instrument is also available for online administration. There are a variety of formatting options for the online version.

<table>
<thead>
<tr>
<th>In the past 7 days…</th>
<th>Not at all</th>
<th>A little bit</th>
<th>Somewhat</th>
<th>Quite a bit</th>
<th>Very much</th>
</tr>
</thead>
<tbody>
<tr>
<td>PANINF1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How much did pain interfere with your day-to-day activities?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PANINF2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How much did pain interfere with work around the home?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PANINF3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How much did pain interfere with your ability to participate in social activities?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FREQUENTLY ASKED QUESTIONS

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

This short form is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 2 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

**Q: Do I need to register with PROMIS to use this short form?**

Yes, to get a copy of this short form, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

**Q: Is this short form available in other languages?**

Yes, this short form is currently available in Spanish. The PROMIS group is also working to translate this form into other languages. Information on available translations is updated periodically at nihpromis.org.

**Q: Can I make my own short form?**

Yes, custom pain interference short forms can be made by selecting any items from the item bank. Instructions can be found in the Assessment Center user manual. The full item bank is available at assessmentcenter.net.
ADULT PHYSICAL FUNCTION VERSION 1.0 SHORT FORMS

A brief guide to the 10-item PROMIS Short Form v1.0 – Physical Function 10a and the 20-item PROMIS Short Form v1.0 – Physical Function 20a

ABOUT PHYSICAL FUNCTION*

The physical function item bank measures self-reported capability rather than actual performance of physical activities. This includes the functioning of one’s upper extremities (dexterity), lower extremities (walking or mobility), and central regions (neck, back), as well as instrumental activities of daily living, such as running errands. A single physical function capability score is obtained from a short form. Each physical function short form is appropriate for the adult general population and adults with chronic health conditions. The forms are generic rather than disease-specific. Each form assesses current function rather than function over a specified time period.

(*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing physical function: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (124 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 the 10- and 20-item physical function version 1.0 short forms. The strength of the 10- and 20-item instruments lies in their focus on item content and their ability to assess the full range of physical function measured by the physical function item bank. When selecting a short form, the main difference is instrument length. Reliability and precision of short forms within a domain are highly similar. Longer short forms generally offer greater correlation (strength of relationship) with the full item bank, as well as greater precision. If you are working with a sample in which you expect large variability in a domain and you want to include the full range of item content from that domain, you would probably prefer a version 1.0 form. On the other hand, if you are hoping to capture secondary outcomes data, but have little room for additional measures, you would probably prefer a very brief (four-item) profile short form.

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. Longer CAT offers greater correlation with the full item bank, 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.

Whether one uses a short form or CAT, the score metric is based upon Item Response Theory (IRT), a family of statistical models that link individual questions to a presumed underlying trait or concept of physical function represented by all items in the item bank.

Graphical reports, which visually illustrate results, are available in Assessment Center if you choose to administer a PROMIS Profile, which includes a profile short form from seven PROMIS domains (physical function, depression, anxiety, pain interference, fatigue, sleep disturbance, and satisfaction with participation in social roles). To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the 10-item form, the lowest possible raw score is 10; the highest possible raw score is 50 (Table 1).

You can use Tables 1 and 2 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 10-item form, a raw score of 30 converts to a T-score of 35.0 with a standard error (SE) of 1.7. Thus, the 95% confidence interval around the actual observed score ranges from 31.7 to 38.3 (T-score ± (1.96*SE) = 35.0 ± 3.3 = 31.7 to 38.3).

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

You can upload data to a free computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual can be downloaded at nihpromis.org.

STATISTICAL CHARACTERISTICS

There are three key features of the score for physical function:

- **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 – SE²).

- **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/SE²).

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, the two dotted 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 blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the 20-item form. Figure 1 also tells us where on the scale the forms are most informative based upon the T-score: the 20-item form is more informative than the 10-item form, and the 20-item form offers sufficient reliability over a wider range of T-scores than the 10-item form.
PREVIEW OF SAMPLE ITEMS

Figure 2 is an excerpt from the paper version of the 10- and 20-item short forms. These instruments are also available for online administration. There are a variety of formatting options for the online version.

<table>
<thead>
<tr>
<th>PROMIS</th>
<th>Does your health now limit you in doing vigorous activities, such as running, lifting heavy objects, participating in strenuous sports?</th>
<th>Not at all</th>
<th>Very little</th>
<th>Somewhat</th>
<th>Quite a lot</th>
<th>Cannot do</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF0A</td>
<td></td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>PF20A</td>
<td></td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PROMIS</th>
<th>Does your health now limit you in walking more than a mile?</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF0A</td>
<td></td>
</tr>
<tr>
<td>PF20A</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2

FREQUENTLY ASKED QUESTIONS

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

Each of the short forms is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 2 is an excerpt from a paper short form.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

Q: *Do I need to register with PROMIS to use these short forms?*

Yes, to get a copy of these short forms, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

Q: *Are these short forms available in other languages?*

Yes, these short forms are currently available in Spanish. The PROMIS group is also working to translate these forms into other languages. Information on available translations is updated periodically at nihpromis.org.

Q: *Can I make my own short form?*

Yes, custom physical function short forms can be made by selecting any items from the item bank. Instructions can be found in the Assessment Center user manual. The full item bank is available at assessmentcenter.net.
ADULT SATISFACTION WITH PARTICIPATION IN DISCRETIONARY SOCIAL ACTIVITIES (DSA)
VERSION 1.0 SHORT FORM

A brief guide to the 7-item PROMIS Short Form v1.0 – Satisfaction with Participation in DSA 7a

ABOUT SATISFACTION WITH PARTICIPATION IN DSA*

The satisfaction with participation in discretionary social activities item bank analyzes self-reported contentment with leisure interests and relationships with friends. This is separate from social roles, which include work and family responsibilities, and does not include an evaluation of one’s ability to participate in social contexts. The satisfaction with participation in discretionary social activities short form is generic rather than disease-specific. It assesses satisfaction over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing satisfaction with DSA: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (12 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 if paper administration is preferred, a short form would be more desirable than CAT.

This guide provides information on the seven-item satisfaction with DSA version 1.0 short form. The strength of the seven-item instrument lies in its focus on item content and its ability to assess the full range of satisfaction with DSA measured by the satisfaction with DSA item bank. When selecting a short form, the main difference is instrument length. Reliability and precision of short forms within a domain are highly similar. Longer short forms generally offer greater correlation (strength of relationship) with the full item bank, as well as greater precision. If you are working with a sample in which you expect large variability in a domain and you want to include the full range of item content from that domain, you would probably prefer a version 1.0 short form. On the other hand, if you are hoping to capture secondary outcomes data, but have little room for additional measures, you would probably prefer a very brief (four-item) profile short form.

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. Longer CAT offers greater correlation with the full item bank, 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.

Whether one uses a short form or CAT, the score metric is based upon Item Response Theory (IRT), a family of statistical models that link individual questions to a presumed underlying trait or concept of satisfaction with DSA represented by all items in the item bank.

Graphical reports, which visually illustrate results, are available in Assessment Center if you choose to administer a PROMIS Profile, which includes a profile short form from seven PROMIS domains (physical function, depression, anxiety, pain interference, fatigue, sleep disturbance, and satisfaction with participation in social roles). To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the seven-item form, the lowest possible raw score is 7; the highest possible raw score is 35 (Table 1).

You can use Table 1 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 7-item form, a raw score of 21 converts to a T-score of 47.2 with a standard error (SE) of 1.9. Thus, the 95% confidence interval around the actual observed score ranges from 43.5 to 50.9 (T-score ± (1.96*SE) = 47.2 ± 3.7 = 43.5 to 50.9).

Important: A higher PROMIS T-score represents more of the concept being measured. For positively-worded concepts like satisfaction with DSA, a T-score of 60 is one SD better than average. By comparison, a satisfaction with DSA T-score of 40 is one SD worse than average.

You can upload data to a free computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at nihpromis.org.

STATISTICAL CHARACTERISTICS

There are three key features of the score for satisfaction with DSA:

- **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).

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, the two dotted 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 blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the seven-item form. Figure 1 also tells us where on the scale the form is most informative based upon the T-score. This form would typically be more informative than a satisfaction with DSA form with fewer items.
PREVIEW OF SAMPLE ITEMS

Figure 2 is an excerpt from the paper version of the seven-item short form. This instrument is also available for online administration. There are a variety of formatting options for the online version.

<table>
<thead>
<tr>
<th>In the past 7 days…</th>
<th>Not at all</th>
<th>A little bit</th>
<th>Somewhat</th>
<th>Quite a bit</th>
<th>Very much</th>
</tr>
</thead>
<tbody>
<tr>
<td>I am satisfied with my ability to do things for fun at home (like reading, listening to music, etc.)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>I am satisfied with my ability to do things for my friends</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

FREQUENTLY ASKED QUESTIONS

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

This short form is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 2 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

Q: Do I need to register with PROMIS to use this short form?

Yes, to get a copy of this short form, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

Q: Is this short form available in other languages?

Yes, this short form is currently available in Spanish. The PROMIS group is also working to translate this form into other languages. Information on available translations is updated periodically at nihpromis.org.

Q: Can I make my own short form?

Yes, custom satisfaction with participation in discretionary social activities short forms can be made by selecting any items from the item bank. Instructions can be found in the Assessment Center user manual. The full item bank is available at assessmentcenter.net.
ADULT SATISFACTION WITH PARTICIPATION IN SOCIAL ROLES VERSION 1.0 SHORT FORM
A brief guide to the 7-item PROMIS Short Form v1.0 – Satisfaction with Participation in Social Roles 7a

ABOUT SATISFACTION WITH PARTICIPATION IN SOCIAL ROLES (PSR)*

The satisfaction with participation in social roles item bank examines self-reported contentment with social roles, such as work and family responsibilities. This is separate from discretionary social activities, which include leisure interests and relationships with friends, and does not include an evaluation of one’s ability to participate in social contexts. The satisfaction with participation in social roles short form is generic rather than disease-specific. It assesses satisfaction over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing satisfaction with PSR: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (14 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 if paper administration is preferred, a short form would be more desirable than CAT.

This guide provides information on the seven-item satisfaction with PSR version 1.0 short form. The strength of the seven-item version 1.0 instrument lies in its focus on item content and its ability to assess the full range of satisfaction with PSR measured by the satisfaction with PSR item bank. When selecting a short form, the main difference is instrument length. Reliability and precision of short forms within a domain are highly similar. Longer short forms generally offer greater correlation (strength of relationship) with the full item bank, as well as greater precision. If you are working with a sample in which you expect large variability in a domain and you want to include the full range of item content from that domain, you would probably prefer a version 1.0 short form. On the other hand, if you are hoping to capture secondary outcomes data, but have little room for additional measures, you would probably prefer a very brief (four-item) profile short form.

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. Longer CAT offers greater correlation with the full item bank, 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.

Whether one uses a short form or CAT, the score metric is based upon Item Response Theory (IRT), a family of statistical models that link individual questions to a presumed underlying trait or concept of satisfaction with PSR represented by all items in the item bank.

Graphical reports, which visually illustrate results, are available in Assessment Center if you choose to administer a PROMIS Profile, which includes a profile short form from seven PROMIS domains (physical function, depression, anxiety, pain interference, fatigue, sleep disturbance, and satisfaction with participation in social roles). To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the seven-item form, the lowest possible raw score is 7; the highest possible raw score is 35 (Table 1).

You can use Table 1 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 7-item form, a raw score of 21 converts to a T-score of 44.4 with a standard error (SE) of 1.8. Thus, the 95% confidence interval around the actual observed score ranges from 40.9 to 47.9 (T-score = 44.4 ± 1.96*1.8 = 40.9 to 47.9).

Important: A higher PROMIS T-score represents more of the concept being measured. For positively-worded concepts like satisfaction with PSR, a T-score of 60 is one SD better than average. By comparison, a satisfaction with PSR T-score of 40 is one SD worse than average.

You can upload data to a free computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at nihpromis.org.

STATISTICAL CHARACTERISTICS

There are three key features of the score for satisfaction with PSR:

- **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).

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, the two dotted 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 blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the seven-item form. Figure 1 also tells us where on the scale the form is most informative based upon the T-score. This form would typically be more informative than a satisfaction with PSR form with fewer items.
PREVIEW OF SAMPLE ITEMS

Figure 2 is an excerpt from the paper version of the seven-item short form. This instrument is also available for online administration. There are a variety of formatting options for the online version.

<table>
<thead>
<tr>
<th>In the past 7 days…</th>
<th>Not at all</th>
<th>A little bit</th>
<th>Somewhat</th>
<th>Quite a bit</th>
<th>Very much</th>
</tr>
</thead>
<tbody>
<tr>
<td>I am satisfied with how much work I can do (include work at home)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am satisfied with my ability to work (include work at home)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am satisfied with my ability to do regular personal and household</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FREQUENTLY ASKED QUESTIONS

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

This short form is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 2 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

Q: Do I need to register with PROMIS to use this short form?

Yes, to get a copy of this short form, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

Q: Is this short form available in other languages?

Yes, this short form is currently available in Spanish. The PROMIS group is also working to translate this form into other languages. Information on available translations is updated periodically at nihpromis.org.

Q: Can I make my own short form?

Yes, custom satisfaction with participation in social roles short forms can be made by selecting any items from the item bank. Instructions can be found in the Assessment Center user manual. The full item bank is available at assessmentcenter.net.
ABOUT SLEEP DISTURBANCE*

The sleep disturbance item bank assesses self-reported perceptions of sleep quality, sleep depth, and restoration associated with sleep. This includes perceived difficulties and concerns with getting to sleep or staying asleep, as well as perceptions of the adequacy of and satisfaction with sleep. Sleep disturbance does not focus on symptoms of specific sleep disorders, nor does it provide subjective estimates of sleep quantities (total amount of sleep, time to fall asleep, amount of wakefulness during sleep). The sleep disturbance short form is generic rather than disease-specific. It assesses sleep disturbance over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing sleep disturbance: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (27 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 the eight-item sleep disturbance version 1.0 short form. The strength of the eight-item version 1.0 instrument lies in its focus on item content and its ability to assess the full range of sleep disturbance measured by the sleep disturbance item bank. When selecting a short form, the main difference is instrument length. Reliability and precision of short forms within a domain are highly similar. Longer short forms generally offer greater correlation (strength of relationship) with the full item bank, as well as greater precision. If you are working with a sample in which you expect large variability in a domain and you want to include the full range of item content from that domain, you would probably prefer this version 1.0 short form. On the other hand, if you are hoping to capture secondary outcomes data, but have little room for additional measures, you would probably prefer a very brief (four-item) profile short form.

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. Longer CAT offers greater correlation with the full item bank, 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.

Whether one uses a short form or CAT, the score metric is based upon Item Response Theory (IRT), a family of statistical models that link individual questions to a presumed underlying trait or concept of sleep disturbance represented by all items in the item bank.

Graphical reports, which visually illustrate results, are available in Assessment Center if you choose to administer a PROMIS Profile, which includes a profile short form from seven PROMIS domains (physical function, depression, anxiety, pain interference, fatigue, sleep disturbance, and satisfaction with participation in social roles). To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the eight-item form, the lowest possible raw score is 8; the highest possible raw score is 40 (Table 1).

You can use Table 1 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 8-item form, a raw score of 24 converts to a T-score of 54.3 with a standard error (SE) of 2.5. Thus, the 95% confidence interval around the actual observed score ranges from 49.4 to 59.2 (T-score ± (1.96*SE) = 54.3 ± 4.9 = 49.4 to 59.2).

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

You can upload data to a free computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at nihpromis.org.

STATISTICAL CHARACTERISTICS

There are three key features of the score for sleep disturbance:

- **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 – SE²).

- **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/SE²).

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, the dotted horizontal line represents a degree of internal consistency reliability (i.e., .90) typically regarded as sufficient for an accurate individual score. The shaded blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the eight-item form. Figure 1 also tells us where on the scale the form is most informative based upon the T-score. This form would typically be more informative than a sleep disturbance form with fewer items.
PREVIEW OF SAMPLE ITEMS

Figure 2 is an excerpt from the paper version of the eight-item short form. This instrument is also available for online administration. There are a variety of formatting options for the online version.

<table>
<thead>
<tr>
<th>Sleep 11</th>
<th>In the past 7 days...</th>
<th>Not at all</th>
<th>A little bit</th>
<th>Somewhat</th>
<th>Quite a bit</th>
<th>Very much</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sleep 2</td>
<td>My sleep was refreshing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleep 3</td>
<td>I had a problem with my sleep</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleep 4</td>
<td>I had difficulty falling asleep</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FREQUENTLY ASKED QUESTIONS

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

This short form is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 2 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

**Q: Do I need to register with PROMIS to use this short form?**

Yes, to get a copy of this short form, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

**Q: Is this short form available in other languages?**

Yes, this short form is currently available in Spanish. The PROMIS group is also working to translate this form into other languages. Information on available translations is updated periodically at nihpromis.org.

**Q: Can I make my own short form?**

Yes, custom sleep disturbance short forms can be made by selecting any items from the item bank. Instructions can be found in the Assessment Center user manual. The full item bank is available at assessmentcenter.net.
ABOUT SLEEP-RELATED IMPAIRMENT*

The sleep-related impairment item bank focuses on self-reported perceptions of alertness, sleepiness, and tiredness during usual waking hours, and the perceived functional impairments during wakefulness associated with sleep problems or impaired alertness. Though sleep-related impairment does not directly assess cognitive, affective, or performance impairment, it does measure waking alertness, sleepiness, and function within the context of overall sleep-wake function. The sleep-related impairment short form is generic rather than disease-specific. It assesses sleep-related impairment over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing sleep-related impairment: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (16 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 if paper administration is preferred, a short form would be more desirable than CAT.

This guide provides information on the eight-item sleep-related impairment version 1.0 short form. The strength of the eight-item instrument lies in its focus on item content and its ability to assess the full range of sleep-related impairment measured by the sleep-related impairment item bank. When selecting a short form, the main difference is instrument length. Reliability and precision of short forms within a domain are highly similar. Longer short forms generally offer greater correlation (strength of relationship) with the full item bank, as well as greater precision. If you are working with a sample in which you expect large variability in a domain and you want to include the full range of item content from that domain, you would probably prefer a version 1.0 short form. On the other hand, if you are hoping to capture secondary outcomes data, but have little room for additional measures, you would probably prefer a very brief (four-item) profile short form.

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. Longer CAT offers greater correlation with the full item bank, 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.

Whether one uses a short form or CAT, the score metric is based upon Item Response Theory (IRT), a family of statistical models that link individual questions to a presumed underlying trait or concept of sleep-related impairment represented by all items in the item bank.

Graphical reports, which visually illustrate results, are available in Assessment Center if you choose to administer a PROMIS Profile, which includes a profile short form from seven PROMIS domains (physical function, depression, anxiety, pain interference, fatigue, sleep disturbance, and satisfaction with participation in social roles). To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the eight-item form, the lowest possible raw score is 8; the highest possible raw score is 40 (Table 1).

You can use Table 1 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 8-item form, a raw score of 24 converts to a T-score of 59.3 with a standard error (SE) of 4.7. Thus, the 95% confidence interval around the actual observed score ranges from 54.6 to 64.0 (T-score ± (1.96*SE) = 59.3 ± 4.7 = 54.6 to 64.0).

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

You can upload data to a free computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at nihpromis.org.

STATISTICAL CHARACTERISTICS

There are three key features of the score for sleep-related impairment:

- **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 – SE²).
- **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/SE²).

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, the two dotted 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 blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the eight-item form. Figure 1 also tells us where on the scale the form is most informative based upon the T-score. This form would typically be more informative than a sleep-related impairment form with fewer items.
PREVIEW OF SAMPLE ITEMS

Figure 2 is an excerpt from the paper version of the eight-item short form. This instrument is also available for online administration. There are a variety of formatting options for the online version.

FREQUENTLY ASKED QUESTIONS

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

This short form is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 2 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

Q: Do I need to register with PROMIS to use this short form?

Yes, to get a copy of this short form, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

Q: Is this short form available in other languages?

Yes, this short form is currently available in Spanish. The PROMIS group is also working to translate this form into other languages. Information on available translations is updated periodically at nihpromis.org.

Q: Can I make my own short form?

Yes, custom sleep-related impairment short forms can be made by selecting any items from the item bank. Instructions can be found in the Assessment Center user manual. The full item bank is available at assessmentcenter.net.
PROMIS ADULT PROFILE SHORT FORMS: INTRODUCTION

The profile instruments are a collection of profile short forms, where each form contains items from one of seven primary PROMIS domains (depression, anxiety, physical function, pain interference, fatigue, sleep disturbance, and satisfaction with participation in social roles). Profile short forms, which are four, six, and eight items in length, are meant to be administered together with other profile short forms from each of the other primary domains, thus generating a ‘profile’ of seven scores. There are three profile instruments: the PROMIS-29 profile instrument, so named because it contains 29 items, which includes four items from each primary domain plus a single pain intensity rating; the PROMIS-43 profile instrument, which includes six items from each primary domain plus the pain intensity item; and the PROMIS-57 profile instrument, which includes eight items from each primary domain plus the pain intensity item. Each profile short form can also be administered independently.
ADULT ANXIETY PROFILE SHORT FORMS

A brief guide to the 4-item PROMIS Short Form v1.0 – Anxiety 4a, the 6-item PROMIS Short Form v1.0 – Anxiety 6a, and the 8-item PROMIS Short Form v1.0 – Anxiety 8a

ABOUT ANXIETY*

The anxiety item bank measures self-reported fear (fearfulness, panic), anxious misery (worry, dread), hyperarousal (tension, nervousness, restlessness), and somatic symptoms related to arousal (racing heart, dizziness). Anxiety is best differentiated by symptoms that reflect autonomic arousal and experience of threat. Only one behavioral avoidance item is included in the item bank; therefore, behavioral fear avoidance is not fully evaluated. The anxiety short forms are generic rather than disease-specific. Each assesses anxiety over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing anxiety: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (29 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 the 4-, 6-, and 8-item anxiety short forms designed for use with short forms of similar length from other domains (physical function, depression, pain interference, fatigue, sleep disturbance, and satisfaction with participation in social roles). When a short form from each of the seven domains is administered together, this offers a ‘profile’ of the respondent. While these short forms are designed for use as part of a PROMIS Profile, they can also be administered individually.

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 anxiety 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 illustrates the correlations (strength of relationship) of the full bank with CAT 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. 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.

Graphical reports, which visually illustrate results, are available if you choose to administer a PROMIS Profile. To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the 4-item form, the lowest possible raw score is four; the highest possible raw score is 20 (Table 1).

You can use Tables 1, 2, and 3 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 4-item form, a raw score of 10 converts to a T-score of 59.5 with a standard error (SE) of 2.6. Thus, the 95% confidence interval around the observed score ranges from 54.4 to 64.6 (T-score ± 1.96*SE = 59.5 ± 5.1 = 54.4 to 64.6).

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

You can upload data to a computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at nihpromis.org.

STATISTICAL CHARACTERISTICS

There are three key features of the score for anxiety:

- **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 multiple time points when there has been no true change in what is being measured (for z-scores, reliability = 1 – SE²).

- **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/SE²).

In Figure 2, the two dotted horizontal lines each represent a degree of reliability (i.e., .90 or .95) typically regarded as sufficient for an accurate individual score. The shaded blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the 8-item form. Figure 2 also tells us where on the scale the form is most informative based upon the T-score: the 8-item form is more informative than the 6-item form, which is more informative than the 4-item form.
PREVIEW OF SAMPLE ITEMS

Figure 3 is an excerpt from the paper version of the 4-, 6-, and 8-item short forms. These instruments are also available for online administration. There are a variety of formatting options for the online version. **Note:** This excerpt comes from the version for investigators; there is also a version for participants which does NOT include the labels for response scores (1 through 5) or the identification tag for each item (e.g., EDANX01).

![Image of sample items](image)

FREQUENTLY ASKED QUESTIONS

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

Each of these short forms is available on the PROMIS website through Assessment Center, an online research management tool which houses all PROMIS instruments. Assessment Center enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 3 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at [nihpromis.org](http://nihpromis.org) and [assessmentcenter.net](http://assessmentcenter.net). To learn more, contact help@assessmentcenter.net.

**Q: Do I need to register with PROMIS to use these short forms?**

Yes, to get a copy of one or more of these short forms, we ask that you register with Assessment Center so that we are better able to track who has accessed instruments for research. Assessment Center is available at [assessmentcenter.net](http://assessmentcenter.net), along with the PROMIS terms and conditions of use.

**Q: Are these short forms available in other languages?**

Yes, these short forms are available in Spanish. The PROMIS group is working to translate these forms into other languages. Information on available translations is updated periodically at [nihpromis.org](http://nihpromis.org).

**Q: Can I make my own short form?**

Yes, custom anxiety short forms of any length can be made by selecting items from the full item bank. Instructions can be found in the Assessment Center User Manual at [assessmentcenter.net](http://assessmentcenter.net).
ADULT DEPRESSION PROFILE SHORT FORMS

A brief guide to the 4-item PROMIS Short Form v1.0 – Depression 4a, the 6-item PROMIS Short Form v1.0 – Depression 6a, and the 8-item PROMIS Short Form v1.0 – Depression 8a

ABOUT DEPRESSION*

The depression item bank assesses self-reported negative mood (sadness, guilt), views of self (self-criticism, worthlessness), and social cognition (loneliness, interpersonal alienation), as well as decreased positive affect and engagement (loss of interest, meaning, and purpose). Somatic symptoms (changes in appetite, sleeping patterns) are not included, which eliminates consideration of these items' confounding effects when assessing patients with comorbid physical conditions. The depression short forms are generic rather than disease-specific. Each assesses depression over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing depression: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (28 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 the 4-, 6-, and 8-item depression short forms designed for use with short forms of similar length from other domains (physical function, anxiety, pain interference, fatigue, sleep disturbance, and satisfaction with participation in social roles). When a short form from each of the seven domains is administered together, this offers a ‘profile’ of the respondent. While these short forms are designed for use as part of a PROMIS Profile, they can also be administered individually.

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 depression 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 illustrates the correlations (strength of relationship) of the full bank with CAT 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. 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.

Graphical reports, which visually illustrate results, are available if you choose to administer a PROMIS Profile. To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the 4-item form, the lowest possible raw score is four; the highest possible raw score is 20 (Table 1).

You can use Tables 1, 2, and 3 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 4-item form, a raw score of 10 converts to a T-score of 58.9 with a standard error (SE) of 2.3. Thus, the 95% confidence interval around the observed score ranges from 54.4 to 63.4 (T-score \( 58.9 \pm 4.5 = 54.4 \) to 63.4).

**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.

You can upload data to a computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at [nihpromis.org](https://nihpromis.org).

STATISTICAL CHARACTERISTICS

There are three key features of the score for depression:

- **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 multiple 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).

In Figure 2, the two dotted horizontal lines each represent a degree of reliability (i.e., .90 or .95) typically regarded as sufficient for an accurate individual score. The shaded blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the 8-item form. Figure 2 also tells us where on the scale the form is most informative based upon the T-score: the 8-item form is more informative than the 6-item form, which is more informative than the 4-item form.
PREVIEW OF SAMPLE ITEMS

Figure 3 is an excerpt from the paper version of the 4-, 6-, and 8-item short forms. These instruments are also available for online administration. There are a variety of formatting options for the online version. Note: This excerpt comes from the version for investigators; there is also a version for participants which does NOT include the labels for response scores (1 through 5) or the identification tag for each item (e.g., EDDEP04).

<table>
<thead>
<tr>
<th>In the past 7 days....</th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>I felt worthless........</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I felt that I had nothing to look forward</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FREQUENTLY ASKED QUESTIONS

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

Each of these short forms is available on the PROMIS website through Assessment Center, an online research management tool which houses all PROMIS instruments. Assessment Center enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 3 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org and assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

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

Yes, to get a copy of one or more of these short forms, we ask that you register with Assessment Center so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the PROMIS terms and conditions of use.

Q: Are these short forms available in other languages?

Yes, these short forms are available in Spanish. The PROMIS group is working to translate these forms into other languages. Information on available translations is updated periodically at nihpromis.org.

Q: Can I make my own short form?

Yes, custom depression short forms of any length can be made by selecting items from the full item bank. Instructions can be found in the Assessment Center User Manual at assessmentcenter.net.
ABOUT FATIGUE*

The fatigue item bank evaluates a range of self-reported symptoms, from mild subjective feelings of tiredness to an overwhelming, debilitating, and sustained sense of exhaustion that likely decreases one’s ability to execute daily activities and function normally in family or social roles. Fatigue is divided into the experience of fatigue (frequency, duration, and intensity) and the impact of fatigue on physical, mental, and social activities. The fatigue short forms are generic rather than disease-specific. Each assesses fatigue over the past seven days.

(*)abbreviated definition: see nihpromis.org for the full version

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing fatigue: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (95 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 the 4-, 6-, and 8-item fatigue short forms designed for use with short forms of similar length from other domains (physical function, depression, pain interference, anxiety, sleep disturbance, and satisfaction with participation in social roles). When a short form from each of the seven domains is administered together, this offers a ‘profile’ of the respondent. While these short forms are designed for use as part of a PROMIS Profile, they can also be administered individually.

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 fatigue 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 illustrates the correlations (strength of relationship) of the full bank with CAT 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. 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.

Graphical reports, which visually illustrate results, are available if you choose to administer a PROMIS Profile. To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the 4-item form, the lowest possible raw score is four; the highest possible raw score is 20 (Table 1).

You can use Tables 1, 2, and 3 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 4-item form, a raw score of 10 converts to a T-score of 53.1 with a standard error (SE) of 2.4. Thus, the 95% confidence interval around the observed score ranges from 48.4 to 57.8 (T-score ± 1.96*SE = 53.1 ± 4.7 = 48.4 to 57.8).

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

You can upload data to a computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at nihpromis.org.

STATISTICAL CHARACTERISTICS

There are three key features of the score for fatigue:

- **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 multiple time points when there has been no true change in what is being measured (for z-scores, reliability = 1 – SE²).
- **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/SE²).

In Figure 2, the two dotted horizontal lines each represent a degree of reliability (i.e., .90 or .95) typically regarded as sufficient for an accurate individual score. The shaded blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the 8-item form. Figure 2 also tells us where on the scale the form is most informative based upon the T-score: the 8-item form is more informative than the 6-item form, which is more informative than the 4-item form.
PREVIEW OF SAMPLE ITEMS

Figure 3 is an excerpt from the paper version of the 4-, 6-, and 8-item short forms. These instruments are also available for online administration. There are a variety of formatting options for the online version. Note: This excerpt comes from the version for investigators; there is also a version for participants which does NOT include the labels for response scores (1 through 5) or the identification tag for each item (e.g., FATEXP20).

<table>
<thead>
<tr>
<th>In the past 7 days...</th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>How often did you feel tired?...............</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How often did you experience extreme exhaustion?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FREQUENTLY ASKED QUESTIONS

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

Each of these short forms is available on the PROMIS website through Assessment Center, an online research management tool which houses all PROMIS instruments. Assessment Center enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 3 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org and assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

**Q: Do I need to register with PROMIS to use these short forms?**

Yes, to get a copy of one or more of these short forms, we ask that you register with Assessment Center so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the PROMIS terms and conditions of use.

**Q: Are these short forms available in other languages?**

Yes, these short forms are available in Spanish. The PROMIS group is working to translate these forms into other languages. Information on available translations is updated periodically at nihpromis.org.

**Q: Can I make my own short form?**

Yes, custom fatigue short forms of any length can be made by selecting items from the full item bank. Instructions can be found in the Assessment Center User Manual at assessmentcenter.net.
ADULT PAIN INTERFERENCE PROFILE SHORT FORMS

A brief guide to the 4-item PROMIS Short Form v1.0 – Pain Interference 4a, 
the 6-item PROMIS Short Form v1.0 – Pain Interference 6a, 
and the 8-item PROMIS Short Form v1.0 – Pain Interference 8a

ABOUT PAIN INTERFERENCE*

The pain interference item bank measures the self-reported consequences of pain on relevant aspects of one’s life. This includes the extent to which pain hinders engagement with social, cognitive, emotional, physical, and recreational activities. Pain interference also incorporates items probing sleep and enjoyment in life, though the item bank only contains one sleep item. The pain interference short forms are generic rather than disease-specific. Each assesses pain interference over the past seven days. (*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing pain interference: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (41 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 the 4-, 6-, and 8-item pain interference short forms designed for use with short forms of similar length from other domains (physical function, depression, anxiety, fatigue, sleep disturbance, and satisfaction with participation in social roles). When a short form from each of the seven domains is administered together, this offers a ‘profile’ of the respondent. While these short forms are designed for use as part of a PROMIS Profile, they can also be administered individually.

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 pain interference 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 illustrates the correlations (strength of relationship) of the full bank with CAT 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. 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.

Graphical reports, which visually illustrate results, are available if you choose to administer a PROMIS Profile. To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the 4-item form, the lowest possible raw score is four; the highest possible raw score is 20 (Table 1).

You can use Tables 1, 2, and 3 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 4-item form, a raw score of 10 converts to a T-score of 58.5 with a standard error (SE) of 1.8. Thus, the 95% confidence interval around the observed score ranges from 55.0 to 62.0 (T-score ± 1.96*SE) = 58.5 ± 3.5 = 55.0 to 62.0).

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

You can upload data to a computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at nihpromis.org.

STATISTICAL CHARACTERISTICS

There are three key features of the score for pain interference:

- **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 multiple time points when there has been no true change in what is being measured (for z-scores, reliability = 1 – SE²).
- **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/SE²).

In Figure 2, the two dotted horizontal lines each represent a degree of reliability (i.e., .90 or .95) typically regarded as sufficient for an accurate individual score. The shaded blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the 8-item form. Figure 2 also tells us where on the scale the form is most informative based upon the T-score: the 8-item form is more informative than the 6-item form, which is more informative than the 4-item form.
PREVIEW OF SAMPLE ITEMS

Figure 3 is an excerpt from the paper version of the 4-, 6-, and 8-item short forms. These instruments are also available for online administration. There are a variety of formatting options for the online version. Note: This excerpt comes from the version for investigators; there is also a version for participants which does NOT include the labels for response scores (1 through 5) or the identification tag for each item (e.g., PAININ9).

<table>
<thead>
<tr>
<th>In the past 7 days…</th>
<th>Not at all</th>
<th>A little bit</th>
<th>Somewhat</th>
<th>Quite a bit</th>
<th>Very much</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAININ9 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>How much did pain interfere with your day to day activities?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>PAININ9 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>How much did pain interfere with work around the home?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

FREQUENTLY ASKED QUESTIONS

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

Each of these short forms is available on the PROMIS website through Assessment Center, an online research management tool which houses all PROMIS instruments. Assessment Center enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 3 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org and assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

**Q: Do I need to register with PROMIS to use these short forms?**

Yes, to get a copy of one or more of these short forms, we ask that you register with Assessment Center so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the PROMIS terms and conditions of use.

**Q: Are these short forms available in other languages?**

Yes, these short forms are available in Spanish. The PROMIS group is working to translate these forms into other languages. Information on available translations is updated periodically at nihpromis.org.

**Q: Can I make my own short form?**

Yes, custom pain interference short forms of any length can be made by selecting items from the full item bank. Instructions can be found in the Assessment Center User Manual at assessmentcenter.net.
ADULT PHYSICAL FUNCTION PROFILE SHORT FORMS
A brief guide to the 4-item PROMIS Short Form v1.0 – Physical Function 4a, the 6-item PROMIS Short Form v1.0 – Physical Function 6a, and the 8-item PROMIS Short Form v1.0 – Physical Function 8a

ABOUT PHYSICAL FUNCTION*
The physical function item bank measures self-reported capability rather than actual performance of physical activities. This includes the functioning of one’s upper extremities (dexterity), lower extremities (walking or mobility), and central regions (neck, back), as well as instrumental activities of daily living, such as running errands. A single physical function capability score is obtained from a short form. Each physical function short form is appropriate for the adult general population and adults with chronic health conditions. The forms are generic rather than disease-specific. Each form assesses current function rather than function over a specified time period.

(*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS
There are two administration options for assessing physical function: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (124 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 the 4-, 6-, and 8-item physical function short forms designed for use with short forms of similar length from other domains (depression, anxiety, pain interference, fatigue, sleep disturbance, and satisfaction with participation in social roles). When a short form from each of the seven domains is administered together, this offers a ‘profile’ of the respondent. While these short forms are designed for use as part of a PROMIS Profile, they can also be administered individually.

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 physical function 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 illustrates the correlations (strength of relationship) of the full bank with CAT 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. 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.

Graphical reports, which visually illustrate results, are available if you choose to administer a PROMIS Profile. To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the 4-item form, the lowest possible raw score is four; the highest possible raw score is 20 (Table 1).

You can use Tables 1, 2, and 3 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 4-item form, a raw score of 10 converts to a T-score of 34.4 with a standard error (SE) of 2.1. Thus, the 95% confidence interval around the observed score ranges from 30.3 to 38.5 (T-score $-1.96 \times SE = 34.4 - 4.1 = 30.3$ to 38.5).

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

You can upload data to a computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at nihpromis.org.

STATISTICAL CHARACTERISTICS

There are three key features of the score for physical function:

- **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 multiple 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 = $1/SE^2$).

In Figure 2, the two dotted horizontal lines each represent a degree of reliability (i.e., .90 or .95) typically regarded as sufficient for an accurate individual score. The shaded blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the 8-item form. Figure 2 also tells us where on the scale the form is most informative based upon the T-score: the 8-item form is more informative than the 6-item form, which is more informative than the 4-item form.
PREVIEW OF SAMPLE ITEMS

Figure 3 is an excerpt from the paper version of the 4-, 6-, and 8-item short forms. These instruments are also available for online administration. There are a variety of formatting options for the online version. Note: This excerpt comes from the version for investigators; there is also a version for participants which does NOT include the labels for response scores (1 through 5) or the identification tag for each item (e.g., PFA11).

<table>
<thead>
<tr>
<th></th>
<th>Without any difficulty</th>
<th>With a little difficulty</th>
<th>With some difficulty</th>
<th>With much difficulty</th>
<th>Unable to do</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PFA11</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are you able to do chores such as vacuuming or yard work?</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td><strong>PFA16</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are you able to dress yourself, including tying shoelaces and doing buttons?</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

FREQUENTLY ASKED QUESTIONS

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

Each of these short forms is available on the PROMIS website through Assessment Center, an online research management tool which houses all PROMIS instruments. Assessment Center enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 3 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org and assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

**Q: Do I need to register with PROMIS to use these short forms?**

Yes, to get a copy of one or more of these short forms, we ask that you register with Assessment Center so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the PROMIS terms and conditions of use.

**Q: Are these short forms available in other languages?**

Yes, these short forms are available in Spanish. The PROMIS group is working to translate these forms into other languages. Information on available translations is updated periodically at nihpromis.org.

**Q: Can I make my own short form?**

Yes, custom physical function short forms of any length can be made by selecting items from the full item bank. Instructions can be found in the Assessment Center User Manual at assessmentcenter.net.
ADULT SATISFACTION WITH PARTICIPATION IN SOCIAL ROLES PROFILE SHORT FORMS

A brief guide to the 4-item PROMIS Short Form v1.0 – Satisfaction with Participation in Social Roles 4a, the 6-item PROMIS Short Form v1.0 – Satisfaction with Participation in Social Roles 6a, and the 8-item PROMIS Short Form v1.0 – Satisfaction with Participation in Social Roles 8a

ABOUT SATISFACTION WITH PARTICIPATION IN SOCIAL ROLES (PSR)*

The satisfaction with participation in social roles item bank examines self-reported contentment with social roles, such as work and family responsibilities. This is separate from discretionary social activities, which include leisure interests and relationships with friends, and does not include an evaluation of one’s ability to participate in social contexts. The satisfaction with participation in social roles short forms are generic rather than disease-specific. Each assesses satisfaction over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing satisfaction with PSR: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (14 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 if paper administration is preferred, a short form would be more desirable than CAT.

This guide provides information on the 4-, 6-, and 8-item satisfaction with PSR short forms designed for use with short forms of similar length from other domains (physical function, depression, pain interference, anxiety, fatigue, and sleep disturbance). When a short form from each of the seven domains is administered together, this offers a ‘profile’ of the respondent. While these short forms are designed for use as part of a PROMIS Profile, they can also be administered individually.

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 satisfaction with PSR 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 illustrates the correlations (strength of relationship) of the full bank with CAT and with short forms of varying length. The correlation of CAT scores with the full bank score is similar to 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.

Graphical reports, which visually illustrate results, are available if you choose to administer a PROMIS Profile. To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the 4-item form, the lowest possible raw score is four; the highest possible raw score is 20 (Table 1).

You can use Tables 1, 2, and 3 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 4-item form, a raw score of 10 converts to a T-score of 41.7 with a standard error (SE) of 2.1. Thus, the 95% confidence interval around the observed score ranges from 37.6 to 45.8 (T-score $^2 (1.96 \times SE) = 41.7^2 4.1 = 37.6$ to 45.8).

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

You can upload data to a computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at nihpromis.org.

STATISTICAL CHARACTERISTICS

There are three key features of the score for satisfaction with PSR:

- **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 multiple 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 = $1/SE^2$).

In Figure 2, the two dotted horizontal lines each represent a degree of reliability (i.e., .90 or .95) typically regarded as sufficient for an accurate individual score. The shaded blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the 8-item form. Figure 2 also tells us where on the scale the form is most informative based upon the T-score: the 8-item form is more informative than the 6-item form, which is more informative than the 4-item form.
PREVIEW OF SAMPLE ITEMS

Figure 3 is an excerpt from the paper version of the 4-, 6-, and 8-item short forms. These instruments are also available for online administration. There are a variety of formatting options for the online version. Note: This excerpt comes from the version for investigators; there is also a version for participants which does NOT include the labels for response scores (1 through 5) or the identification tag for each item (e.g., SRPSAT07).

<table>
<thead>
<tr>
<th>Item</th>
<th>Question</th>
<th>Response Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I am satisfied with how much work I can do (include work at home)</td>
<td>1 (Not at all), 2 (A little bit), 3 (Somewhat), 4 (Quite a bit), 5 (Very much)</td>
</tr>
<tr>
<td>2</td>
<td>I am satisfied with my ability to work (include work at home)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>I am satisfied with my ability to do tasks related to personal and household tasks</td>
<td></td>
</tr>
</tbody>
</table>

FREQUENTLY ASKED QUESTIONS

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

Each of these short forms is available on the PROMIS website through Assessment Center, an online research management tool which houses all PROMIS instruments. Assessment Center enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 3 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org and assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

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

Yes, to get a copy of one or more of these short forms, we ask that you register with Assessment Center so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the PROMIS terms and conditions of use.

Q: Are these short forms available in other languages?

Yes, these short forms are available in Spanish. The PROMIS group is working to translate these forms into other languages. Information on available translations is updated periodically at nihpromis.org.

Q: Can I make my own short form?

Yes, custom satisfaction with PSR short forms of any length can be made by selecting items from the full item bank. Instructions can be found in the Assessment Center User Manual at assessmentcenter.net.
ADULT SLEEP DISTURBANCE PROFILE SHORT FORMS
A brief guide to the 4-item PROMIS Short Form v1.0 – Sleep Disturbance 4a, the 6-item PROMIS Short Form v1.0 – Sleep Disturbance 6a, and the 8-item PROMIS Short Form v1.0 – Sleep Disturbance 8a

ABOUT SLEEP DISTURBANCE*

The sleep disturbance item bank assesses self-reported perceptions of sleep quality, sleep depth, and restoration associated with sleep. This includes perceived difficulties and concerns with getting to sleep or staying asleep, as well as perceptions of the adequacy of and satisfaction with sleep. Sleep disturbance does not focus on symptoms of specific sleep disorders, nor does it provide subjective estimates of sleep quantities (total amount of sleep, time to fall asleep, amount of wakefulness during sleep). The sleep disturbance short forms are generic rather than disease-specific. Each assesses sleep disturbance over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing sleep disturbance: short forms and computerized adaptive testing (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 computer’s choice of subsequent items from the full item bank (27 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 the 4-, 6-, and 8-item sleep disturbance short forms designed for use with short forms of similar length from other domains (physical function, depression, anxiety, pain interference, fatigue, and satisfaction with participation in social roles). When a short form from each of the seven domains is administered together, this offers a ‘profile’ of the respondent. While these short forms are designed for use as part of a PROMIS Profile, they can also be administered individually.

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 sleep disturbance 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 illustrates the correlations (strength of relationship) of the full bank with CAT and with short forms of varying length. The correlation of CAT scores with the full bank score is similar to 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.

Graphical reports, which visually illustrate results, are available if you choose to administer a PROMIS Profile. To access a sample report, complete the CAT demo at nihpromis.org.
SCORING THE INSTRUMENT

Each question has five response options ranging in value from one to five. To find the total raw score, sum the values of the response to each question. For example, for the 4-item form, the lowest possible raw score is four; the highest possible raw score is 20 (Table 1).

You can use Tables 1, 2, and 3 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 who has a T-score of 40 is one SD below the U.S. general population mean. These conversions are accurate ONLY when all questions on the short form have been answered. For the 4-item form, a raw score of 10 converts to a T-score of 50.5 with a standard error (SE) of 3.4. Thus, the 95% confidence interval around the observed score ranges from 43.8 to 57.2 (T-score ± 1.96*SE = 50.5 ± 6.7 = 43.8 to 57.2).

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

You can upload data to a computer program called PROMIScore, which will score your data one person at a time or as a group. PROMIScore is particularly useful because it can calculate T-scores even when there are missing responses. The PROMIScore software and user manual are available for download at nihpromis.org.

STATISTICAL CHARACTERISTICS

There are three key features of the score for sleep disturbance:

- **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 multiple time points when there has been no true change in what is being measured (for z-scores, reliability = 1 – SE²).

- **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/SE²).

In Figure 2, the dotted horizontal line represents a degree of reliability (.90) typically regarded as sufficient for an accurate individual score. The shaded blue region marks the range of the scale where measurement precision is comparable to the reliability of .90 for the 8-item form. Figure 2 also tells us where on the scale the form is most informative based upon the T-score: the 8-item form is more informative than the 6-item form, which is more informative than the 4-item form.
PREVIEW OF SAMPLE ITEMS

Figure 3 is an excerpt from the paper version of the 4-, 6-, and 8-item short forms. These instruments are also available for online administration. There are a variety of formatting options for the online version. Note: This excerpt comes from the version for investigators; there is also a version for participants which does NOT include the labels for response scores (1 through 5) or the identification tag for each item (e.g., Sleep116).

<table>
<thead>
<tr>
<th></th>
<th>Not at all</th>
<th>A little bit</th>
<th>Somewhat</th>
<th>Quite a bit</th>
<th>Very much</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sleep116</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Sleep20</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Sleep44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In the past 7 days...</td>
<td>My sleep was refreshing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I had a problem with my sleep</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>I had difficulty falling asleep</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FREQUENTLY ASKED QUESTIONS

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

Each of these short forms is available on the PROMIS website through Assessment Center, an online research management tool which houses all PROMIS instruments. Assessment Center enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for administration on paper. Figure 3 is an excerpt from a paper short form for investigators.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org and assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

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

Yes, to get a copy of one or more of these short forms, we ask that you register with Assessment Center so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the PROMIS terms and conditions of use.

Q: Are these short forms available in other languages?

Yes, these short forms are available in Spanish. The PROMIS group is working to translate these forms into other languages. Information on available translations is updated periodically at nihpromis.org.

Q: Can I make my own short form?

Yes, custom sleep disturbance short forms of any length can be made by selecting items from the full item bank. Instructions can be found in the Assessment Center User Manual at assessmentcenter.net.
PROMIS ADULT COMPUTERIZED ADAPTIVE TESTING

Computerized adaptive testing (CAT) is a flexible, computer-driven assessment that can use any items in the item bank to measure each domain. CAT selects only those items that sharpen the estimate of a respondent’s score on the domain being measured. CAT length varies but usually includes four to seven items.
ABOUT ANXIETY*

The anxiety item bank measures self-reported fear (fearfulness, panic), anxious misery (worry, dread), hyperarousal (tension, nervousness, restlessness), and somatic symptoms related to arousal (racing heart, dizziness). Anxiety is best differentiated by symptoms that reflect autonomic arousal and experience of threat. Only one behavioral avoidance item is included in the item bank; therefore, behavioral fear avoidance is not fully evaluated. Anxiety computerized adaptive testing (CAT) is generic rather than disease-specific. It assesses anxiety over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

PREVIEW OF SAMPLE ITEMS

Figure 1 shows three anxiety items (i.e., questions or statements) from the full item bank that can be included in CAT. Several formats for presenting the items are available for computer-based administration through Assessment Center (see FAQ on page 3). CAT is not available for paper administration.

SCORING THE INSTRUMENT

A minimum number of items (4) must be answered in order to receive a score for anxiety CAT. The first item is selected because it provides the most information about the U.S. general population. The response to this item will guide the computer’s choice of the next item for the participant. The participant’s response to this 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, or the participant has answered the maximum number of questions (12), whichever occurs first.

After finishing CAT, the participant’s pattern of responses is converted into a standardized T-score, with a mean of 50 based on the U.S. general population, and a standard deviation (SD) of 10. Thus, a person who has a T-score of 40 is one SD below the U.S. mean. The standardized T-score is reported as the final score for each participant.

Important: A higher T-score always represents more of the concept being measured. For negatively-worded concepts like anxiety, a T-score of 60 is one SD worse than average. By comparison, an anxiety T-score of 40 is one SD better than average.
STATISTICAL CHARACTERISTICS

Figure 2 is a sample of the statistical information available in Assessment Center. Two key features are:

1) **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 - SE²).

2) **Standard Error (SE)**: The possible range of the actual final score based upon the scaled T-score. 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 ± (1.96*SE) = 52 ± 3.9 = 48.1 to 55.9).

More information is available online via Assessment Center (assessmentcenter.net).

DATA REPORTS

Upon completion of CAT, a data report is available in Assessment Center. Figure 3 is a sample report:

Your score on the Anxiety CAT is __. The average score is 50. Your score indicates that your level of anxiety is higher (worse) than:

- ___ percent of people in the general population
- ___ percent of (age group)
- ___ percent of (gender)

Your score is shown below on the graph. The diamond (♦) is placed where we think your score lies. It represents your T-score, which is a standardized measure that is based on an average score of 50, based on responses to the same questions by the general population in the United States.

Your Score  SE
Anxiety  52  2

The horizontal, solid black line on either side of the diamond in your profile report shows the Standard Error (SE), a statistical measure of variance that represents the possible range of your actual score around this estimated score. It is very likely that your score is in the range between these lines.

To access a sample report for anxiety, complete the CAT demo at nihpromis.org. More than one CAT domain can be completed at a time; a separate report will be generated for each completed domain.

Data reports are also available if you choose to administer a PROMIS profile instrument, which includes a short form from seven PROMIS domains (physical function, depression, anxiety, fatigue, pain interference, satisfaction with participation in social roles, and sleep disturbance).
COMPARING COMPUTERIZED ADAPTIVE TESTING TO SHORT FORMS

There are two administration options for assessing anxiety: short forms and CAT. When administering a short form, instruct participants to answer all of the items presented. With CAT, participant responses guide the computer’s choice of subsequent items from the full item bank (29 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.

Figure 4 illustrates the correlations (strength of relationship) of the full bank with CAT 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. 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.

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 anxiety 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.

FREQUENTLY ASKED QUESTIONS

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

The full version of this item bank is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free, online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for non-adaptive administration on paper.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

Q: Do I need to register with PROMIS to use this CAT?

Yes, to gain access to this CAT, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

Q: Is this CAT available in other languages?

This CAT is not currently available in other languages. The PROMIS group is working to translate this CAT into Spanish and other languages. Information on available translations is updated periodically at nihpromis.org.
ABOUT DEPRESSION*

The depression item bank assesses self-reported negative mood (sadness, guilt), views of self (self-criticism, worthlessness), and social cognition (loneliness, interpersonal alienation), as well as decreased positive affect and engagement (loss of interest, meaning, and purpose). Somatic symptoms (changes in appetite, sleeping patterns) are not included, which eliminates consideration of these items’ confounding effects when assessing patients with comorbid physical conditions. Depression computerized adaptive testing (CAT) is generic rather than disease-specific. It assesses depression over the past seven days. (*abbreviated definition: see nihpromis.org for the full version)

PREVIEW OF SAMPLE ITEMS

Figure 1 shows three depression items (i.e., questions or statements) from the full item bank that can be included in CAT. Several formats for presenting the items are available for computer-based administration through Assessment Center (see FAQ on page 3). CAT is not available for paper administration.

SCORING THE INSTRUMENT

A minimum number of items (4) must be answered in order to receive a score for depression CAT. The first item is selected because it provides the most information about the U.S. general population. The response to this item will guide the computer’s choice of the next item for the participant. The participant’s response to this 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, or the participant has answered the maximum number of questions (12), whichever occurs first.

After finishing CAT, the participant’s pattern of responses is converted into a standardized T-score, with a mean of 50 based on the U.S. general population, and a standard deviation (SD) of 10. Thus, a person who has a T-score of 40 is one SD below the U.S. mean. The standardized T-score is reported as the final score for each participant.

Important: A higher T-score always 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.
STATISTICAL CHARACTERISTICS

Figure 2 is a sample of the statistical information available in Assessment Center. Two key features are:

1) **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 - SE²).

2) **Standard Error (SE)**: The possible range of the actual final score based upon the scaled T-score. 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 ± 1.96*SE = 52 ± 3.9 = 48.1 to 55.9).

More information is available online via Assessment Center ([assessmentcenter.net](http://assessmentcenter.net)).

DATA REPORTS

Upon completion of CAT, a data report is available in Assessment Center. Figure 3 is a sample report:

Your score on the Depression CAT is ___. The average score is 50. Your score indicates that your level of depression is higher (worse) than:

- __ percent of people in the general population
- __ percent of (age group)
- __ percent of (gender)

Your score is shown below on the graph. The diamond (♦) is placed where we think your score lies. It represents your T-score, which is a standardized measure that is based on an average score of 50, based on responses to the same questions by the general population in the United States.

Your Score  SE
Depression  52  2

The horizontal, solid black line on either side of the diamond in your profile report shows the Standard Error (SE), a statistical measure of variance that represents the possible range of your actual score around this estimated score. It is very likely that your score is in the range between these lines.

To access a sample report for depression, complete the CAT demo at [nihpromis.org](http://nihpromis.org). More than one CAT domain can be completed at a time; a separate report will be generated for each completed domain.

Data reports are also available if you choose to administer a PROMIS profile instrument, which includes a short form from seven PROMIS domains (physical function, depression, anxiety, fatigue, pain interference, satisfaction with participation in social roles, and sleep disturbance).
COMPARING COMPUTERIZED ADAPTIVE TESTING TO SHORT FORMS

There are two administration options for assessing depression: short forms and CAT. When administering a short form, instruct participants to answer all of the items presented. With CAT, participant responses guide the computer’s choice of subsequent items from the full item bank (28 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.

Figure 4 illustrates the correlations (strength of relationship) of the full bank with CAT 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. 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.

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 depression 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.

FREQUENTLY ASKED QUESTIONS

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

The full version of this item bank is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free, online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for non-adaptive administration on paper.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

Q: Do I need to register with PROMIS to use this CAT?

Yes, to gain access to this CAT, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

Q: Is this CAT available in other languages?

This CAT is not currently available in other languages. The PROMIS group is working to translate this CAT into Spanish and other languages. Information on available translations is updated periodically at nihpromis.org.
ADULT FATIGUE

COMPUTERIZED ADAPTIVE TESTING (CAT)

A brief guide to the PROMIS Item Bank v1.0 – Fatigue

ABOUT FATIGUE*

The fatigue item bank evaluates a range of self-reported symptoms, from mild subjective feelings of tiredness to an overwhelming, debilitating, and sustained sense of exhaustion that likely decreases one’s ability to execute daily activities and function normally in family or social roles. Fatigue is divided into the experience of fatigue (frequency, duration, and intensity) and the impact of fatigue on physical, mental, and social activities. Fatigue computerized adaptive testing (CAT) is generic rather than disease-specific. It assesses fatigue over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

PREVIEW OF SAMPLE ITEMS

Figure 1 shows three fatigue items (i.e., questions or statements) from the full item bank that can be included in CAT. Several formats for presenting the items are available for computer-based administration through Assessment Center (see FAQ on page 3). CAT is not available for paper administration.

SCORING THE INSTRUMENT

A minimum number of items (4) must be answered in order to receive a score for fatigue CAT. The first item is selected because it provides the most information about the U.S. general population. The response to this item will guide the computer’s choice of the next item for the participant. The participant’s response to this 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, or the participant has answered the maximum number of questions (12), whichever occurs first.

After finishing CAT, the participant’s pattern of responses is converted into a standardized T-score, with a mean of 50 based on the U.S. general population, and a standard deviation (SD) of 10. Thus, a person who has a T-score of 40 is one SD below the U.S. mean. The standardized T-score is reported as the final score for each participant.

Important: A higher T-score always represents more of the concept being measured. For negatively-worded concepts like fatigue, a T-score of 60 is one SD worse than average. By comparison, a fatigue T-score of 40 is one SD better than average.
STATISTICAL CHARACTERISTICS

Figure 2 is a sample of the statistical information available in Assessment Center. Two key features are:

1) **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 - SE²).

2) **Standard Error (SE)**: The possible range of the actual final score based upon the scaled T-score. With a T-score of 32 and a SE of 4, the 95% confidence interval around the actual final score ranges from 24.2 to 39.8 (T-score = 32 ± 7.8 = 24.2 to 39.8).

More information is available online via Assessment Center (assessmentcenter.net).

DATA REPORTS

Upon completion of CAT, a data report is available in Assessment Center. Figure 3 is a sample report:

Your score on the Fatigue CAT is __. The average score is 50. Your score indicates that your level of fatigue is higher (worse) than:

- __ percent of people in the general population
- __ percent of (age group)
- __ percent of (gender)

Your score is shown below on the graph. The diamond (♦) is placed where we think your score lies. It represents your T-score, which is a standardized measure that is based on an average score of 50, based on responses to the same questions by the general population in the United States.

Your Score  SE
Fatigue  32  4

The horizontal, solid black line on either side of the diamond in your profile report shows the Standard Error (SE), a statistical measure of variance that represents the possible range of your actual score around this estimated score. It is very likely that your score is in the range between these lines.

To access a sample report for fatigue, complete the CAT demo at nihpromis.org. More than one CAT domain can be completed at a time; a separate report will be generated for each completed domain.

Data reports are also available if you choose to administer a PROMIS profile instrument, which includes a short form from seven PROMIS domains (physical function, depression, anxiety, fatigue, pain interference, satisfaction with participation in social roles, and sleep disturbance).
COMPARING COMPUTERIZED ADAPTIVE TESTING TO SHORT FORMS

There are two administration options for assessing fatigue: short forms and CAT. When administering a short form, instruct participants to answer all of the items presented. With CAT, participant responses guide the computer’s choice of subsequent items from the full item bank (95 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.

Figure 4 illustrates the correlations (strength of relationship) of the full bank with CAT 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. 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.

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 fatigue 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.

FREQUENTLY ASKED QUESTIONS

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

The full version of this item bank is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free, online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for non-adaptive administration on paper.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

Q: Do I need to register with PROMIS to use this CAT?

Yes, to gain access to this CAT, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

Q: Is this CAT available in other languages?

This CAT is not currently available in other languages. The PROMIS group is working to translate this CAT into Spanish and other languages. Information on available translations is updated periodically at nihpromis.org.
ABOUT PAIN INTERFERENCE*

The pain interference item bank measures the self-reported consequences of pain on relevant aspects of one’s life. This includes the extent to which pain hinders engagement with social, cognitive, emotional, physical, and recreational activities. Pain interference also incorporates items probing sleep and enjoyment in life, though the item bank only contains one sleep item. Pain interference computerized adaptive testing (CAT) is generic rather than disease-specific. It assesses pain interference over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

PREVIEW OF SAMPLE ITEMS

Figure 1 shows three pain interference items (i.e., questions or statements) from the full item bank that can be included in CAT. Several formats for presenting the items are available for computer-based administration through Assessment Center (see FAQ on page 3). CAT is not available for paper administration.

SCORING THE INSTRUMENT

A minimum number of items (4) must be answered in order to receive a score for pain interference CAT. The first item is selected because it provides the most information about the U.S. general population. The response to this item will guide the computer’s choice of the next item for the participant. The participant’s response to this 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, or the participant has answered the maximum number of questions (12), whichever occurs first.

After finishing CAT, the participant’s pattern of responses is converted into a standardized T-score, with a mean of 50 based on the U.S. general population, and a standard deviation (SD) of 10. Thus, a person who has a T-score of 40 is one SD below the U.S. mean. The standardized T-score is reported as the final score for each participant.

Important: A higher T-score always represents more of the concept being measured. For negatively-worded concepts like pain interference, a T-score of 60 is one SD worse than average. By comparison, a pain interference T-score of 40 is one SD better than average.
STATISTICAL CHARACTERISTICS

Figure 2 is a sample of the statistical information available in Assessment Center. Two key features are:

1) **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 - SE²).

2) **Standard Error (SE)**: The possible range of the actual final score based upon the scaled T-score. With a T-score of 32 and a SE of 4, the 95% confidence interval around the actual final score ranges from 24.2 to 39.8 (T-score ± (1.96*SE) = 32 ± 7.8 = 24.2 to 39.8).

More information is available online via Assessment Center (assessmentcenter.net).

DATA REPORTS

Upon completion of CAT, a data report is available in Assessment Center. Figure 3 is a sample report:

Your score on the Pain Interference CAT is ___. The average score is 50. Your score indicates that your level of pain interference is higher (worse) than:

- __ percent of people in the general population
- __ percent of (age group)
- __ percent of (gender)

Your score is shown below on the graph. The diamond (♦) is placed where we think your score lies. It represents your T-score, which is a standardized measure that is based on an average score of 50, based on responses to the same questions by the general population in the United States.

Your Score SE

Pain Interference 32 4

The horizontal, solid black line on either side of the diamond in your profile report shows the Standard Error (SE), a statistical measure of variance that represents the possible range of your actual score around this estimated score. It is very likely that your score is in the range between these lines.

To access a sample report for pain interference, complete the CAT demo at nihpromis.org. More than one CAT domain can be completed at a time; a separate report will be generated for each completed domain.

Data reports are also available if you choose to administer a PROMIS profile instrument, which includes a short form from seven PROMIS domains (physical function, depression, anxiety, fatigue, pain interference, satisfaction with participation in social roles, and sleep disturbance).
COMPARING COMPUTERIZED ADAPTIVE TESTING TO SHORT FORMS

There are two administration options for assessing pain interference: short forms and CAT. When administering a short form, instruct participants to answer all of the items presented. With CAT, participant responses guide the computer’s choice of subsequent items from the full item bank (41 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.

Figure 4 illustrates the correlations (strength of relationship) of the full bank with CAT 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. 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.

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 pain interference 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.

FREQUENTLY ASKED QUESTIONS

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

The full version of this item bank is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free, online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for non-adaptive administration on paper.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

Q: Do I need to register with PROMIS to use this CAT?

Yes, to gain access to this CAT, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

Q: Is this CAT available in other languages?

This CAT is not currently available in other languages. The PROMIS group is working to translate this CAT into Spanish and other languages. Information on available translations is updated periodically at nihpromis.org.
ADULT PHYSICAL FUNCTION
COMPUTERIZED ADAPTIVE TESTING (CAT)
A brief guide to the PROMIS Item Bank v1.0 – Physical Function

ABOUT PHYSICAL FUNCTION*
The physical function item bank measures self-reported capability rather than actual performance of physical activities. This includes the functioning of one’s upper extremities (dexterity), lower extremities (walking or mobility), and central regions (neck, back), as well as instrumental activities of daily living, such as running errands. Physical function computerized adaptive testing (CAT) is appropriate for the adult general population and adults with chronic health conditions. CAT is generic rather than disease-specific. It assesses current function rather than function over a specified time period.

(*abbreviated definition: see nihpromis.org for the full version)

PREVIEW OF SAMPLE ITEMS
Figure 1 shows three physical function items (i.e., questions or statements) from the full item bank that can be included in CAT. Several formats for presenting the items are available for computer-based administration through Assessment Center (see FAQ on page 3). CAT is not available for paper administration.

SCORING THE INSTRUMENT
A minimum number of items (4) must be answered in order to receive a score for physical function CAT. The first item is selected because it provides the most information about the U.S. general population. The response to this item will guide the computer’s choice of the next item for the participant. The participant’s response to this 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, or the participant has answered the maximum number of questions (12), whichever occurs first.

After finishing CAT, the participant’s pattern of responses is converted into a standardized T-score, with a mean of 50 based on the U.S. general population, and a standard deviation (SD) of 10. Thus, a person who has a T-score of 40 is one SD below the U.S. mean. The standardized T-score is reported as the final score for each participant.

Important: A higher T-score always represents more of the concept being measured. For positively-worded concepts like physical function, a T-score of 60 is one SD better than average. By comparison, a physical function T-score of 40 is one SD worse than average.
STATISTICAL CHARACTERISTICS

Figure 2 is a sample of the statistical information available in Assessment Center. Two key features are:

1) **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 - SE²).

2) **Standard Error (SE)**: The possible range of the actual final score based upon the scaled T-score. With a T-score of 61 and a SE of 3, the 95% confidence interval around the actual final score ranges from 55.1 to 66.9 (T-score ± (1.96*SE) = 61 ± 5.9 = 55.1 to 66.9).

More information is available online via Assessment Center (assessmentcenter.net).

DATA REPORTS

Upon completion of CAT, a data report is available in Assessment Center. Figure 3 is a sample report:

To access a sample report for physical function, complete the CAT demo at nihpromis.org. More than one CAT domain can be completed at a time; a separate report will be generated for each completed domain.

Data reports are also available if you choose to administer a PROMIS profile instrument, which includes a short form from seven PROMIS domains (physical function, depression, anxiety, fatigue, pain interference, satisfaction with participation in social roles, and sleep disturbance).
COMPARING COMPUTERIZED ADAPTIVE TESTING TO SHORT FORMS

There are two administration options for assessing physical function: short forms and CAT. When administering a short form, instruct participants to answer all of the items presented. With CAT, participant responses guide the computer’s choice of subsequent items from the full item bank (124 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.

Figure 4 illustrates the correlations (strength of relationship) of the full bank with CAT 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. 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.

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 physical function 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.

FREQUENTLY ASKED QUESTIONS

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

The full version of this item bank is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain. Assessment Center is a free, online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for non-adaptive administration on paper.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

Q: Do I need to register with PROMIS to use this CAT?

Yes, to gain access to this CAT, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

Q: Is this CAT available in other languages?

This CAT is not currently available in other languages. The PROMIS group is working to translate this CAT into Spanish and other languages. Information on available translations is updated periodically at nihpromis.org.
ADULT SATISFACTION WITH PARTICIPATION IN SOCIAL ROLES
COMPUTERIZED ADAPTIVE TESTING (CAT)
A brief guide to the PROMIS Item Bank v1.0 – Satisfaction with Participation in Social Roles

ABOUT SATISFACTION WITH PARTICIPATION IN SOCIAL ROLES (PSR)*
The satisfaction with participation in social roles item bank examines self-reported contentment with social roles, such as work and family responsibilities. This is separate from discretionary social activities, which include leisure interests and relationships with friends, and does not include an evaluation of one’s ability to participate in social contexts. Satisfaction with participation in social roles computerized adaptive testing (CAT) is generic rather than disease-specific. It assesses satisfaction over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

PREVIEW OF SAMPLE ITEMS

Figure 1 shows three satisfaction with PSR items (i.e., questions or statements) from the full item bank that can be included in CAT. Several formats for presenting the items are available for computer-based administration through Assessment Center (see FAQ on page 3). CAT is not available for paper administration.

SCORING THE INSTRUMENT

A minimum number of items (4) must be answered in order to receive a score for satisfaction with participation in social roles CAT. The first item is selected because it provides the most information about the U.S. general population. The response to this item will guide the computer’s choice of the next item for the participant. The participant’s response to this 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, or the participant has answered the maximum number of questions (12), whichever occurs first.

After finishing CAT, the participant’s pattern of responses is converted into a standardized T-score, with a mean of 50 based on the U.S. general population, and a standard deviation (SD) of 10. Thus, a person who has a T-score of 40 is one SD below the U.S. mean. The standardized T-score is reported as the final score for each participant.

Important: A higher T-score always represents more of the concept being measured. For positively-worded concepts like satisfaction with participation in social roles, a T-score of 60 is one SD better than average. By comparison, a satisfaction with participation in social roles T-score of 40 is one SD worse than average.
STATISTICAL CHARACTERISTICS

Figure 2 is a sample of the statistical information available in Assessment Center. Two key features are:

1) **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 - SE²).

2) **Standard Error (SE)**: The possible range of the actual final score based upon the scaled T-score. With a T-score of 61 and a SE of 3, the 95% confidence interval around the actual final score ranges from 55.1 to 66.9 (T-score ± 1.96*SE) = 61 ± 5.9 = 55.1 to 66.9).

More information is available online via Assessment Center (assessmentcenter.net).

DATA REPORTS

Upon completion of CAT, a data report is available in Assessment Center. Figure 3 is a sample report:

Your score on the Satisfaction with Participation in Social Roles CAT is ___. The average score is 50. Your score indicates that your level of satisfaction with participation in social roles is higher (better) than:

- ___ percent of people in the general population
- ___ percent of (age group)
- ___ percent of (gender)

Your score is shown below on the graph. The diamond (♦) is placed where we think your score lies. It represents your T-score, which is a standardized measure that is based on an average score of 50, based on responses to the same questions by the general population in the United States.

The horizontal, solid black line on either side of the diamond in your profile report shows the Standard Error (SE), a statistical measure of variance that represents the possible range of your actual score around this estimated score. It is very likely that your score is in the range between these lines.

To access a sample report for satisfaction with PSR, complete the CAT demo at nihpromis.org. More than one CAT domain can be completed at a time; a separate report will be generated for each completed domain.

Data reports are also available if you choose to administer a PROMIS profile instrument, which includes a short form from seven PROMIS domains (physical function, depression, anxiety, fatigue, pain interference, satisfaction with participation in social roles, and sleep disturbance).
COMPARING COMPUTERIZED ADAPTIVE TESTING TO SHORT FORMS

There are two administration options for assessing satisfaction with PSR: short forms and CAT. When administering a short form, instruct participants to answer all of the items presented. With CAT, participant responses guide the computer’s choice of subsequent items from the full item bank (14 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.

Figure 4 illustrates the correlations (strength of relationship) of the full bank with CAT 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. 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.

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 satisfaction with PSR 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.

FREQUENTLY ASKED QUESTIONS

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

The full version of this item bank is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free, online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for non-adaptive administration on paper.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

Q: Do I need to register with PROMIS to use this CAT?

Yes, to gain access to this CAT, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

Q: Is this CAT available in other languages?

This CAT is not currently available in other languages. The PROMIS group is working to translate this CAT into Spanish and other languages. Information on available translations is updated periodically at nihpromis.org.
ABOUT SLEEP DISTURBANCE*

The sleep disturbance item bank assesses self-reported perceptions of sleep quality, sleep depth, and restoration associated with sleep. This includes perceived difficulties and concerns with getting to sleep or staying asleep, as well as perceptions of the adequacy of and satisfaction with sleep. Sleep disturbance does not focus on symptoms of specific sleep disorders, nor does it provide subjective estimates of sleep quantities (total amount of sleep, time to fall asleep, amount of wakefulness during sleep). Sleep disturbance CAT is generic rather than disease-specific. It assesses sleep disturbance over the past seven days.

(*abbreviated definition: see nihpromis.org for the full version)

PREVIEW OF SAMPLE ITEMS

Figure 1 shows three sleep disturbance items (i.e., questions or statements) from the full item bank that can be included in CAT. Several formats for presenting the items are available for computer-based administration through Assessment Center (see FAQ on page 3). CAT is not available for paper administration.

SCORING THE INSTRUMENT

A minimum number of items (4) must be answered in order to receive a score for sleep disturbance CAT. The first item is selected because it provides the most information about the U.S. general population. The response to this item will guide the computer’s choice of the next item for the participant. The participant’s response to this 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, or the participant has answered the maximum number of questions (12), whichever occurs first.

After finishing CAT, the participant’s pattern of responses is converted into a standardized T-score, with a mean of 50 based on the U.S. general population, and a standard deviation (SD) of 10. Thus, a person who has a T-score of 40 is one SD below the U.S. mean. The standardized T-score is reported as the final score for each participant.

Important: A higher T-score always represents more of the concept being measured. For negatively-worded concepts like sleep disturbance, a T-score of 60 is one SD worse than average. By comparison, a sleep disturbance T-score of 40 is one SD better than average.
STATISTICAL CHARACTERISTICS

Figure 2 is a sample of the statistical information available in Assessment Center. Two key features are:

1) **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 - \(SE^2\)).

2) **Standard Error (SE)**: The possible range of the actual final score based upon the scaled T-score. 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 ± (1.96*SE) = 52 ± 3.9 = 48.1 to 55.9).

More information is available online via Assessment Center ([assessmentcenter.net](http://assessmentcenter.net)).

DATA REPORTS

Upon completion of CAT, a data report is available in Assessment Center. Figure 3 is a sample report:

Your score on the Sleep Disturbance CAT is ____. The average score is 50. Your score indicates that your level of sleep disturbance is higher (worse) than:

- ____ percent of people in the general population
- ____ percent of (age group)
- ____ percent of (gender)

Your score is shown below on the graph. The diamond (♦) is placed where we think your score lies. It represents your T-score, which is a standardized measure that is based on an average score of 50, based on responses to the same questions by the general population in the United States.

The horizontal, solid black line on either side of the diamond in your profile report shows the Standard Error (SE), a statistical measure of variance that represents the possible range of your actual score around this estimated score. It is very likely that your score is in the range between these lines.

To access a sample report for sleep disturbance, complete the CAT demo at [nihpromis.org](http://nihpromis.org). More than one CAT domain can be completed at a time; a separate report will be generated for each completed domain.

Data reports are also available if you choose to administer a PROMIS profile instrument, which includes a short form from seven PROMIS domains (physical function, depression, anxiety, fatigue, pain interference, satisfaction with participation in social roles, and sleep disturbance).
COMPARING COMPUTERIZED ADAPTIVE TESTING TO SHORT FORMS

There are two administration options for assessing sleep disturbance: short forms and CAT. When administering a short form, instruct participants to answer all of the items presented. With CAT, participant responses guide the computer’s choice of subsequent items from the full item bank (277 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.

Figure 4 illustrates the correlations (strength of relationship) of the full bank with CAT and with short forms of varying length. The correlation of CAT scores with the full bank score is similar to 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.

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 sleep disturbance 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.

FREQUENTLY ASKED QUESTIONS

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

The full version of this item bank is available on the PROMIS website through Assessment Center, which houses all PROMIS instruments for each domain.

Assessment Center is a free, online research management tool. It enables researchers to create study-specific websites for capturing participant data securely. Studies can include measures within the Assessment Center library, as well as custom instruments created or entered by the researcher. PROMIS instruments (short forms, CAT, profiles) are a central feature of the instrument library within Assessment Center. Any PROMIS measure can be included in an online study or downloaded for non-adaptive administration on paper.

Detailed statistical information and development history about PROMIS items and instruments are available for review at nihpromis.org or assessmentcenter.net. To learn more, contact help@assessmentcenter.net.

Q: Do I need to register with PROMIS to use this CAT?

Yes, to gain access to this CAT, we ask that you register with Assessment Center, so that we are better able to track who has accessed instruments for research. Assessment Center is available at assessmentcenter.net, along with the terms and conditions of use.

Q: Is this CAT available in other languages?

This CAT is not currently available in other languages. The PROMIS group is working to translate this CAT into Spanish and other languages. Information on available translations is updated periodically at nihpromis.org.
PROMIS SCORING GUIDE: GLOSSARY

- **Computerized Adaptive Testing (CAT):** A flexible, computer-driven assessment that can use any items in the item bank to measure each domain. CAT selects only those items that sharpen the estimate of a respondent’s score on the domain being measured. CAT length varies but usually includes four to seven items.

- **Domain:** The trait or conceptual area that represents an instrument’s content. In item response theory, a domain is unidimensional, measuring a single trait or concept. It can be a symptom (such as pain) or a functional capability (physical function). A domain can also be divided into several related traits, concepts, or constructs labeled as ‘subdomains.’ For instance, the domain of pain is separated into the subdomains of pain interference and pain behavior, each of which possesses its own item bank, while fatigue is a domain that is not separated into any subdomains. This is decided empirically through the analysis of data collected under the domain and subdomain headings.

- **Instrument:** A measure, survey, or questionnaire. In PROMIS, this can be a short form, profile, or CAT.

- **Item:** A question or statement (and its response choices) in a survey.

- **Item Bank:** A collection of carefully selected items that provide an operational definition of a trait or construct. A good item bank covers the entire continuum of the latent trait being measured, capturing different severity levels along the continuum. A well-calibrated item bank makes it possible to compare the amount of a given trait for individuals who complete different sets of items in the bank. Not only does this allow for ‘adaptive’ testing, because all items are calibrated onto one common scale, it is possible to compare scores across diverse groups of patients and item sets. A well-organized item bank with wide-ranging item difficulties can also enable the creation of a wide variety of short forms, depending on the target populations and purpose of assessment. At a given difficulty level, any chosen item should provide the maximum amount of information to estimate an individual’s score on the domain of interest.

- **Item Response Theory (IRT):** A family of latent trait modeling in which a latent trait is an unobservable latent dimension that is thought to give rise to a set of observed item responses. Each item can be compared to other items measuring the same trait. Items that measure the same construct can be aggregated into longer assessments. IRT offers a framework for evaluating an assessment’s effectiveness, as well as evaluating the effectiveness of individual items within the assessment. IRT differs from classic test theory (CTT) in its assumptions about reliability and validity. The reliability for an IRT-based measure is calculated based on each respondent’s ‘ability’ and varies across the continuum; in CTT, instrument reliability remains the same regardless of respondent ‘ability.’ For IRT-based instruments, validity is assessed across the entire item bank, and any instrument created from the item bank inherits this validity; for CTT instruments, validity is based on the individual instrument and must be reassessed if an instrument is modified in any way.

- **Trait:** Synonymous with the terms ‘construct’ and ‘concept.’
PROMIS SCORING GUIDE: CITATIONS


