SADNESS/APATHY/DEPRESSION MEASURE DIFFERENCES

A brief guide to differences between the NIH Toolbox® Sadness/Apathy/Depression instruments:

ADULT	PEDIATRIC	PARENT REPORT
NIH Toolbox Apathy FF Ages 18+ v3.0 NIH Toolbox Sadness/Depression CAT Ages 18+ v3.0 NIH Toolbox Sadness FF Ages 18+ v2.0 NIH Toolbox Apathy FF Ages 18+ v2.0* NIH Toolbox Sadness CAT Ages 18+ v2.0*	NIH Toolbox Sadness/Depression CAT Ages 8-17 v3.0 NIH Toolbox Sadness FF Ages 8-17 v2.0	NIH Toolbox Sadness/Depression Parent Report CAT Ages 3-7 v3.0 NIH Toolbox Sadness/Depression Parent Report CAT Ages 8-12 v3.0 NIH Toolbox Sadness Parent Report FF Ages 3-7 v2.0 NIH Toolbox Sadness Parent Report FF Ages 8-12 v2.0 NIH Toolbox Sadness Parent Report CAT Ages 8-12 v2.0*

*Retired Measure

CAT = Computer Adaptive Test, FF = Fixed Form

ABOUT SADNESS/APATHY/DEPRESSION

NIH Toolbox Sadness/Apathy/Depression instruments are part of the Negative Affect domain. Sadness/Depression measures assess poor mood and negative perceptions of the self, the world, and the future. Apathy is defined by deficits in goal-oriented behavior and decrements in goal-related thought content.

Sadness measures are available for adults (ages 18+), pediatric self-report (ages 8-17) and for parents serving as proxy reporters for their child (youth ages 3-12).

Apathy measures are available for adults (ages 18+).

Measures are appropriate for the general population and individuals with chronic health conditions. The measures are universal rather than disease-specific.

INTRODUCTION TO ASSESSMENT OPTIONS

There are two administration options for assessing Sadness/Depression: fixed forms and computer adaptive tests (CATs).

There is one administrative option for assessing Apathy: fixed form.

Fixed Form

In a fixed form, respondents answer all of the items (i.e., questions or statements) presented.

CAT

In a CAT, items are tailored to the respondent. A minimum number of items (e.g., 4) must be answered in order to receive a score for a Sadness/Depression CAT. The response to the first item will guide the system's choice of the next item for the respondent. The response to the second item will dictate the selection of the following item, and so on. As additional items are administered, the potential for error is reduced and confidence in the respondent's score increases. The CAT will continue until either the standard error drops below a specified level (e.g., on the T-score metric 3.0), or the participant has answered the maximum number of questions (e.g., 12), whichever occurs first. Although items differ across respondents completing a CAT, scores are comparable across respondents.

Fixed Form versus CAT

For both fixed forms and CATs, scores are based in Item Response Theory (IRT), a family of statistical models that link individual items to a presumed underlying trait or concept of Sadness/Depression represented by all items in the item bank. Some administrators may prefer to ask the same questions 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 fixed form would be more desirable than a CAT. In cases where brevity and measurement precision are priorities, CATs are likely a better fit.

VERSION DIFFERENCES BETWEEN V2.0 AND V3.0

Some NIH Toolbox domains have multiple versions of measures (e.g., v2.0, v3.0). Generally, **it is recommended that you use the most recent version available which can be identified as the measure with the highest version number.** The following information describes differences between versions of measures.

In 2023, NIH Toolbox v3.0 replaced NIH Toolbox v2.0. In most (but not all) cases, NIH Toolbox v2.0 and v3.0 measures include the same items. However, there are three main differences between v2.0 and v3.0. These are:

1. Different item administration engines. V3.0 includes new features to allow more flexibility in CAT stopping rules and uses streamlined code. Measure scoring is the same.
2. Different CAT stopping rules (see CAT Stopping Rules below).
3. V2.0 measures previously administered as fixed forms are now administered as CATs in v3.0.

Scores from NIH Toolbox v2.0 measures are comparable to scores from the equivalent v3.0 measure when v2.0 and v3.0 use the same items or item banks (as they do in most cases). In the case of Sadness/Depression, please see specific version differences outlined below.

Adult

- The NIH Toolbox Apathy FF Age 18+ v2.0 and NIH Toolbox Apathy FF Ages 18+ v3.0 measures contain the same items. Thus, scores from these v2.0 and v3.0 NIH Toolbox measures can be compared.
- The NIH Toolbox Sadness/Depression CAT Ages 18+ v3.0 measure contains the same item content as the NIH Toolbox Sadness CAT Age 18+ v2.0, but uses slightly different stopping rules. See CAT Stopping Rules below for details.

NIH Toolbox v2.0 Measure	Corresponding NIH Toolbox v3.0 Measure	
	Name in NIH Toolbox v3.0 App	**Name on Paper Form**
NIH Toolbox Apathy FF Age 18+ v2.0	Apathy FF Ages 18+	NIH Toolbox Apathy FF Ages 18+ v3.0
NIH Toolbox Sadness CAT Age 18+ v2.0	Sadness/Depression CAT Ages 18+	NIH Toolbox Sadness/Depression CAT Ages 18+ v3.0

- The NIH Toolbox Sadness FF Age 18+ v2.0 was removed in NIH Toolbox v3.0:

NIH Toolbox v2.0 Measure	Corresponding NIH Toolbox v3.0 Measure	
	Name in NIH Toolbox v3.0 App	**Name on Paper Form**
NIH Toolbox Sadness FF Age 18+ v2.0	N/A (v3.0 unavailable)	N/A (v3.0 unavailable)

Pediatric and Parent Report

- Items from the NIH Toolbox Sadness/Depression CAT Ages 8-17 v3.0 measure were taken from the PROMIS Pediatric Item Bank GenPop v3.0 – Depressive Symptoms. Thus, scores from the v2.0 and v3.0 NIH Toolbox Pediatric Sadness/Depression CAT cannot be compared.
- The NIH Toolbox Sadness Parent Report FF Ages 3-7 v2.0 included 7 items. It was replaced by the NIH Toolbox Sadness/Depression Parent Report CAT Ages 3-7 v3.0, which uses the same 7 items, but administers them as a CAT. Consequently, respondents will answer a variable number of items. See CAT Stopping Rules below for details.
- The NIH Toolbox Sadness Parent Report CAT Ages 8-12 v2.0 and NIH Toolbox Sadness/Depression Parent Report CAT Ages 8-12 v3.0 measures contain the same items. Thus, scores from these v2.0 and v3.0 NIH Toolbox measures can be compared.

NIH Toolbox v2.0 Measure	Corresponding NIH Toolbox v3.0 Measure	
	Name in NIH Toolbox v3.0 App	Name on Paper Form
NIH Toolbox Sadness FF Ages 8-17 v2.0	Sadness/Depression CAT Ages 8-17	NIH Toolbox Sadness/Depression CAT Ages 8-17 v3.0
NIH Toolbox Sadness Parent Report FF Ages 3-7 v2.0	Sadness/Depression Parent Report CAT Ages 3-7	NIH Toolbox Sadness/Depression Parent Report CAT Ages 3-7 v3.0
NIH Toolbox Sadness Parent Report CAT Ages 8-12 v2.0	Sadness/Depression Parent Report CAT Ages 8-12	NIH Toolbox Sadness/Depression Parent Report CAT Ages 8-12 v3.0

- The NIH Toolbox Sadness Parent Report FF Ages 8-12 v2.0 was removed in NIH Toolbox v3.0:

NIH Toolbox v2.0 Measure	Corresponding NIH Toolbox v3.0 Measure	
	Name in NIH Toolbox v3.0 App	Name on Paper Form
NIH Toolbox Sadness Parent Report FF Ages 8-12 v2.0	N/A (v3.0 unavailable)	N/A (v3.0 unavailable)

CAT Stopping Rules

Measure	Stopping Rules

NIH Toolbox Sadness/Depression CAT Ages 18+ v3.0
 	
NIH Toolbox Sadness/Depression Parent Report CAT Ages 8-12 v3.0
	• Minimum = 4 items
	• Stop when one of these occurs:
	  ○ 8 items are administered OR
	  ○ The standard error is below 0.224 (theta metric)

NIH Toolbox Sadness/Depression Parent Report CAT Ages 3-7 v3.0	• Minimum = 4 items
	• Stop when one of these occurs:
	  ○ 7 items are administered OR
	  ○ The standard error is below 0.224 (theta metric)

NIH Toolbox Sadness CAT Age 18+ v2.0	• Minimum = 4 items
	• Stop when one of these occurs:
	  ○ 8 items are administered OR
	  ○ The standard error is below 0.3 (theta metric) OR
	  ○ The standard error changes by less than 0.01 on the theta metric (0.1 on the T-score metric)

NIH Toolbox Sadness Parent Report CAT Ages 8-12 v2.0	• Minimum = 4 items
	• Stop when one of these occurs:
	  ○ 11 items are administered OR
	  ○ The standard error is below 0.3 (theta metric) OR
	  ○ The standard error changes by less than 0.01 on the theta metric (0.1 on the T-score metric)