

K-CAT[®] INTERPRETIVE GUIDE

GUIDELINES FOR A GENERAL POPULATION

These recommended guidelines are based on client data and observations. To obtain the benefit from using the K-CAT[®], your organization should determine whether these interpretation guidelines require modifications to meet the needs and goals of your organization's efforts and user populations.

Copyright © 2026 Adaptive Testing Technologies. All rights reserved. Portions of this work were previously published.
Distributed by NCS Pearson, Inc.

AWS is a trademark of Amazon Technologies, Inc. Imodium is a registered trademark of Johnson & Johnson.

CONTENTS

CONTENTS	2
THE VALUE OF THE K-CAT®	3
UNDERSTANDING K-CAT® METRICS	4
UNDERSTANDING AREA UNDER THE CURVE (AUC)	7
MODULES OF THE K-CAT®	8
ANXIETY	8
MANIA/ HYPOMANIA	9
OPPOSITIONAL DEFIANT DISORDER (ODD).....	10
ATTENTION-DEFICIT / HYPERACTIVITY DISORDER (ADHD)	11
DEPRESSION	12
CONDUCT DISORDER (CD)	13
SUBSTANCE USE DISORDER (SUD).....	14
POST-TRAUMATIC STRESS DISORDER (PTSD)	15
SUICIDE.....	16
REFERENCES	17

THE VALUE OF THE K-CAT®

The K-CAT® is a uniquely effective and accurate approach to measuring, tracking, and screening for a variety of conditions and other mental health indicators among children 7–17 years old. The K-CAT® currently has modules to screen and measure:

- Anxiety
- Mania/Hypomania
- Oppositional Defiant Disorder (ODD)
- Attention-Deficit/Hyperactivity Disorder (ADHD)
- Depression
- Conduct Disorder (CD)
- Substance Use Disorder (SUD)
- Post-Traumatic Stress Disorder (PTSD)
- Suicide

Rapid and accurate assessment of child mental health disorders can be used to facilitate identification and referral, and potentially reduce the occurrence of functional disability that stems from early-onset mental disorders. The K-CAT® can also be used for measurement-based care in that repeated assessments at any interval in time can be used to measure change. K-CAT® modules use a multirater interview format where both the child and a parent or caregiver complete the assessment to maximize accuracy. It is possible to use the K-CAT® in situations where only the child or only the parent is available for the assessment, but the diagnostic classification (probability of diagnosis) will not be available. In order to obtain the diagnostic classification, users must administer the test to both the child and the parent and then select the multirater report option in Q-global. This report is available at no extra charge when both the child and parent reports have been generated.

For children, the full battery of modules requires, on average, 8 minutes, and parents require approximately 5 minutes. Additionally, K-CAT® modules can be completed on any web-enabled device via Q-global integrated within a HIPAA-compliant, secure Amazon Web Services (AWS™) platform.

Because the K-CAT® has such a low client burden, it is ideal for tracking symptoms remotely between visits, over time, and ultimately, screening children and adolescents to determine their level of need and type of care. However, the real power of the K-CAT® is the accuracy and clarity that it brings to assessment.

The K-CAT[®] has been validated against structured clinical interviews with precision matching or exceeding the agreement between two trained clinician interviewers. This means that the results of the K-CAT[®] are aligned with what a trained clinician would find when completing a full structured clinical diagnostic interview, taking an hour or more for both the child and parent.

UNDERSTANDING K-CAT[®] METRICS

Unlike most traditional measures, the K-CAT[®] results include four important metrics:

- **Severity:** a numeric score that indicates how severe a client's presentation for a condition is on a 100-point scale where 0 represents the lowest level of severity and 100 represents the highest level of severity.
- **Category:** a categorization of a client's severity score into the categories: normal, mild, moderate, or severe or low risk, medium risk, high risk, depending on the module.
- **Precision:** a number that represents the degree of uncertainty of the severity score. Adaptive testing allows the precision to be fixed to 5 points on a 100-point scale, where 0 would represent absolute certainty and 100 would represent absolute uncertainty. This allows the clinician to assess the clinical and statistical significance of change at the level of the individual patient at any point in time.
- **Probability of Diagnosis:** a number that represents the probability that a client would receive a given diagnosis if a full diagnostic interview was administered where 0 represents the lowest probability and 1 represents the highest probability. This is unique to the K-CAT[®].

The value of these metrics lies in their ability to help a clinician understand the type and degree of psychopathology a client is experiencing at a particular point in time and the degree of uncertainty in that knowledge. This precision has important implications for practice because it allows you to measure change over time in a statistically meaningful way, even at the level of the individual patient.

The **severity score** and accompanying **category** tell the clinician how a given client is presenting *in relation* to the general population that a given module has been designed to measure. Clinicians can use this information in consultation with the client to triage clients into higher or lower levels of care and determine, when multiple conditions might be present, which conditions should be prioritized.

Child and parent assessments may yield different severity scores and categories; generally, the higher metric provides the better estimate of the child's severity level. This is due to the child and parent scales using different severity thresholds, and the parent thresholds being generally higher. As a result, one administration can show a higher severity score for one rater but a more severe category for the other. Because scores and category labels are not directly comparable across raters, interpretation should follow the assessment's purpose. For screening, use the most conservative reading across raters: the highest severity score and the most severe category, even if they come from different raters. For monitoring change over time, track each rater's severity score separately rather than combining them. In both cases, the probability estimate is a useful additional indicator, since it draws on both the child and parent responses.

The **precision** estimate is useful in that it gives clinicians the ability to understand how much certainty they should hold about a given severity score. This is valuable for a few reasons. First, a clinician who sees a severity score of 75 with a precision score of 5.0, can know that scores in the range of 70–80 are within the limits of precision for this measured score of 75. If we were to repeat this assessment 100 times, the true severity would be contained within the interval of 65 to 85, 95% of the time (i.e. 95% confidence interval = severity score plus/minus $1.965 \times \text{precision}$). Second, precision can be used to determine if an individual's change over time is clinically and / or statistically significant.

K-CAT® categorizes change in three ways: uncertain, clinically significant, and statistically significant. Table 1 provides a useful example with definitions of each type of change.

TABLE 1: HOW TO USE PRECISION MEASURES TO INTERPRET CHANGE

Case	Precision	Test 1	Test 2	Absolute Distance	Type of Change	Definition of Change Type
A	5.0	75	70	$ 75 - 70 = 5$	<i>Uncertain</i>	A change is not clinically or statistically significant when the second score is still within the bounds of uncertainty (5 points). In this case, any second score between 80 and 70 (inclusive) would be considered an insignificant change.
B	5.0	75	69	$ 75 - 69 = 6$	<i>Clinically Significant</i>	A change is clinically significant when the absolute distance between the first and second scores is larger than the precision score. Change at this level may be clinically meaningful, but it has not met the higher bar of statistical significance. In this case, any second score below 70 or above 80 would be <i>at least</i> clinically significant.
C	5.0	75	86	$ 75 - 86 = 11$	<i>Statistically Significant</i>	A change is statistically significant when the absolute difference between the first and second scores is more than 2x the level of precision. Where the level of precision is 5, a statistically significant change would need to be more than 10 points away from the first score. Change at this level is likely to present real and meaningful change. In this case, any second score below 65 or above 85 would be statistically significant.

With this information, clinicians can truly understand how a client is doing on a monthly, weekly, or even daily basis, and tailor their care accordingly. Longitudinal analysis of K-CAT[®] test scores, when used in treatment planning, can help a clinician understand the efficacy of modalities and strategies being used. Traditional mental health measurement scales do not provide uncertainty estimates and therefore cannot be used to precisely estimate change at the level of the individual patient.

Lastly, the **probability of diagnosis** is a metric that will be produced for each of the modules if both the child and a parent have completed the module. (Child-only modules have module-specific probability metrics, as described in their sections.) This metric is not available if only the child or only the parent completes the assessment. The specific name of this metric will vary by module. For example, the result for the anxiety module would be shown as the probability of Generalized Anxiety Disorder (GAD). However, in each module, this metric can be understood as the probability that a child would receive a given diagnosis if they were given a full diagnostic interview. The benefit of this metric is that clinicians can see the likelihood of a given diagnosis without completing an interview that requires, on average, over an hour to administer. However, this metric is intended to support screening and clinical decision-making and should not be used as the sole basis for diagnostic decisions.

UNDERSTANDING AREA UNDER THE CURVE (AUC)

AUC is a concept that will be used regularly in the sections below where each module in the K-CAT® is detailed. AUC is a measure of diagnostic classification accuracy across the continuous range of the K-CAT® scale score. In this context, the AUC indicates how well, for example, the anxiety module can predict whether a client would be given a diagnosis of Generalized Anxiety Disorder if a Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS) interview were conducted. General guidelines around how to interpret different ranges of AUC values are listed below:

- $0.5 < \text{AUC} < 0.6$ indicates little to no discrimination
- $0.6 \leq \text{AUC} < 0.7$ indicates poor discrimination
- $0.7 \leq \text{AUC} < 0.8$ indicates acceptable discrimination
- $0.8 \leq \text{AUC} < 0.9$ indicates excellent discrimination
- $\text{AUC} \geq 0.9$ indicates outstanding discrimination

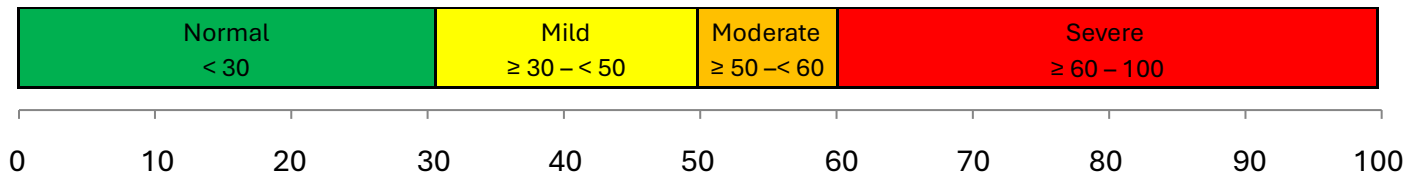
All K-CAT® modules have AUCs that fall within the excellent or outstanding range.

MODULES OF THE K-CAT®

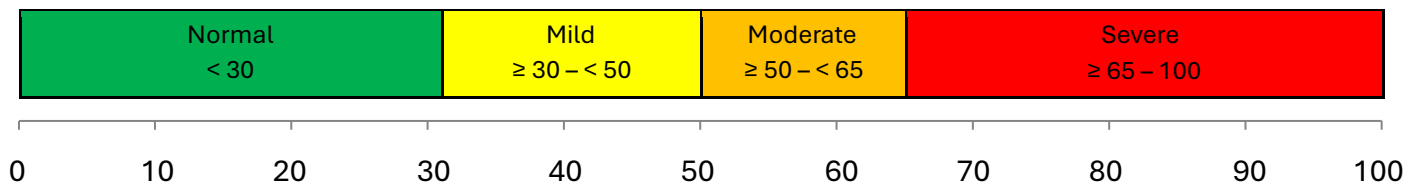
ANXIETY

The anxiety module effectively measures anxiety severity based on the child's experience of symptomatology as well as the parent's rating of the child's experience. The module has an item bank of 221 items for children or 228 items for parents but requires an average of 10 questions administered in an average of two minutes to accurately predict likelihood of a diagnosis of Generalized Anxiety Disorder (GAD) with an AUC = 0.83. The anxiety module is best used for screening and measuring anxiety and is ideal for measurement and evaluation over time. Note that the items are not limited to GAD.

Child Severity Thresholds



Parent Severity Thresholds



Metrics: Severity, Category, Precision, Probability of GAD

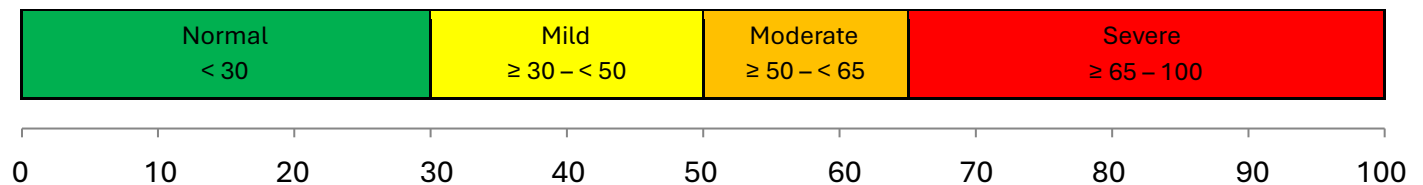
Notes:

1. When interpreting assessment results for multirater interviews, the maximum severity score is the best estimate of the child's current severity level.
2. Like the majority of modules, this module will not produce a Probability of GAD result unless both the child and the parent complete the assessment and the multirater report is selected. This is because the combined scores produce the highest diagnostic classification accuracy for all modules of the K-CAT®.

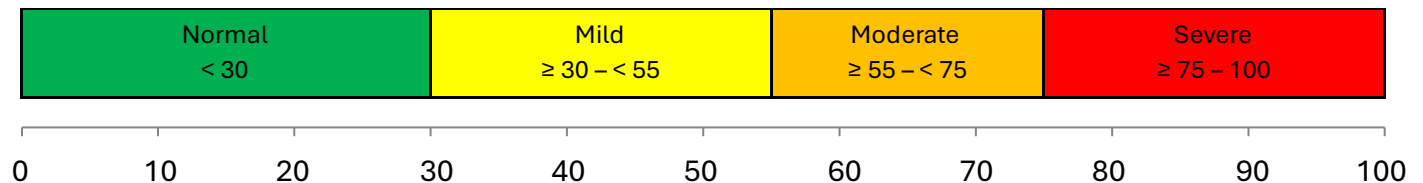
MANIA/ HYPOMANIA

The mania/hypomania module effectively measures mania/hypomania severity based on the child's experience of symptomatology as well as the parent's rating of the child's experience. The module has an item bank of 161 items for children and 152 items for parents and requires an average of 9 questions over 78 seconds to accurately predict the likelihood of a diagnosis of Bipolar I and II disorder with AUC = 0.85. The mania/hypomania module is best used for screening and is ideal for measurement and evaluation over time.

Child Severity Thresholds



Parent Severity Thresholds

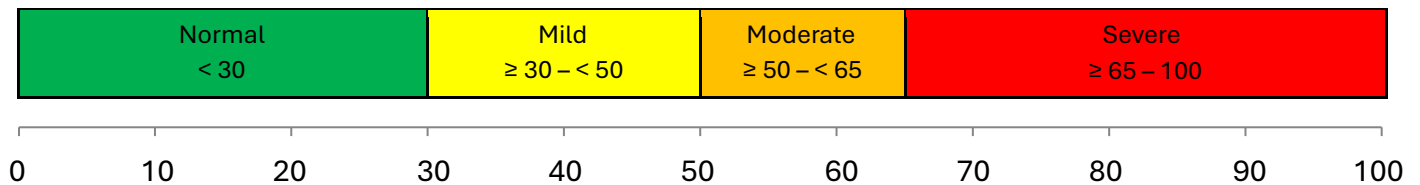


Metrics: Severity, Category, Precision, Probability of Bipolar I/II (BP)

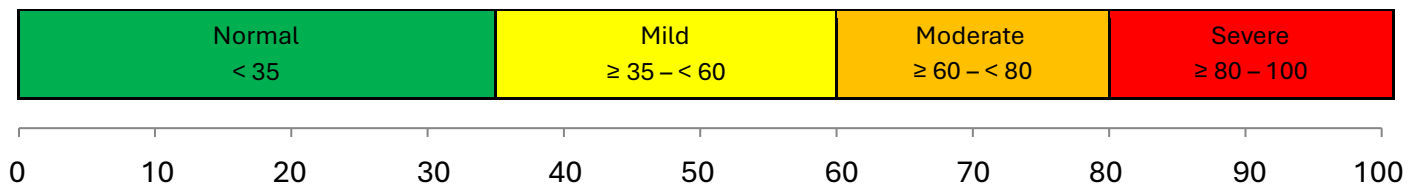
OPPOSITIONAL DEFIANT DISORDER (ODD)

The ODD module effectively measures oppositional defiant disorder symptomatology in the child from the perspective of both the child and parent. The module has an item bank of 140 items for children and 157 items for parents and requires an average of 8 questions administered in an average of 64 seconds to accurately predict the likelihood of a diagnosis of ODD with AUC = 0.88. The ODD module is best used for screening and is ideal for measurement and evaluation over time.

Child Severity Thresholds



Parent Severity Thresholds

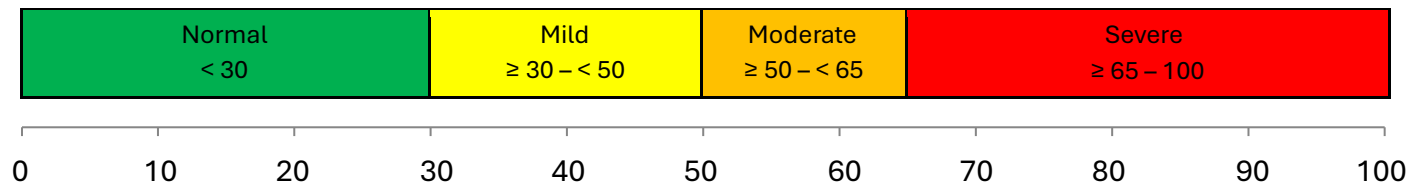


Metrics: Severity, Category, Precision, Probability of ODD

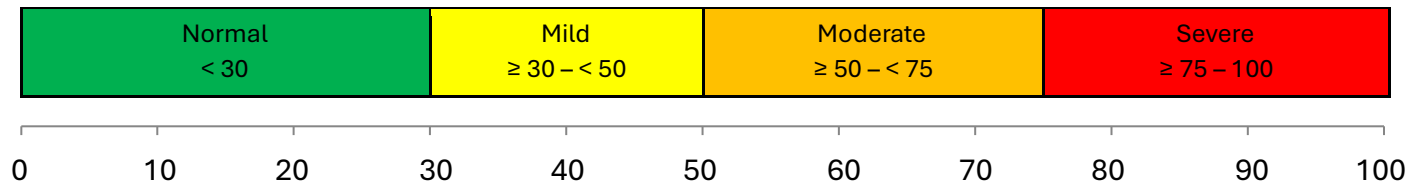
ATTENTION-DEFICIT / HYPERACTIVITY DISORDER (ADHD)

The ADHD module effectively measures attention deficit/hyperactivity disorder symptomatology in the child from the perspective of both the child and parent. The module has an item bank of 110 items for children and 150 items for parents and requires an average of 11 questions administered in an average of 109 seconds to accurately predict the likelihood of a diagnosis of ADHD with AUC = 0.86. The ADHD module is best used for screening and is ideal for measurement and evaluation over time.

Child Severity Thresholds



Parent Severity Thresholds

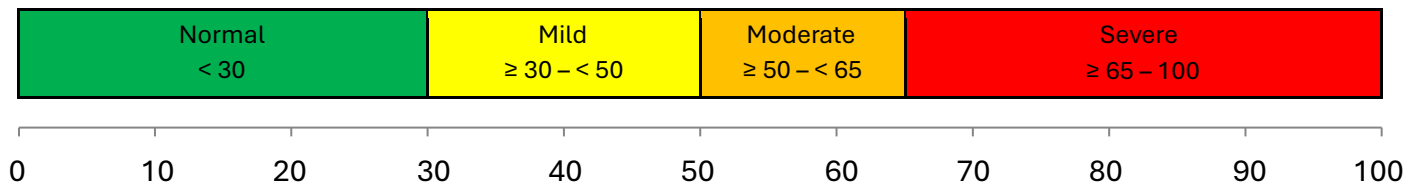


Metrics: Severity, Category, Precision, Probability of ADHD

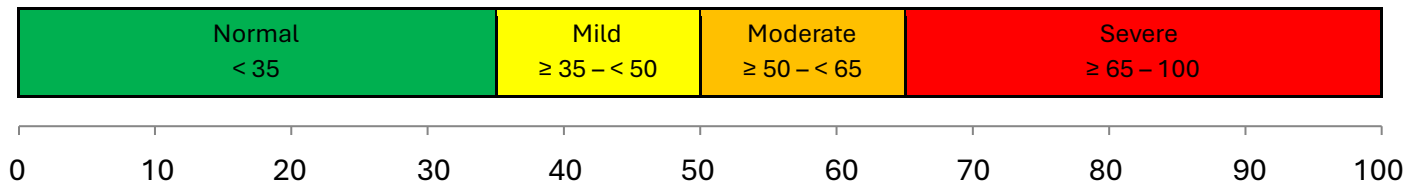
DEPRESSION

The depression module effectively measures depression symptomatology in the child from the perspective of both the child and parent. The module has an item bank of 150 items for both the child and parent, and requires an average of 10 questions administered in an average of 82 seconds to accurately predict the likelihood of a diagnosis of depression with AUC = 0.92. The depression module is best used for screening and is ideal for measurement and evaluation over time.

Child Severity Thresholds



Parent Severity Thresholds

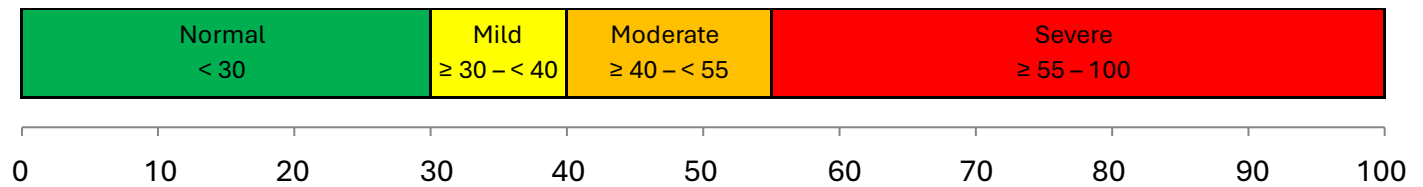


Metrics: Severity, Category, Precision, Probability of MDD

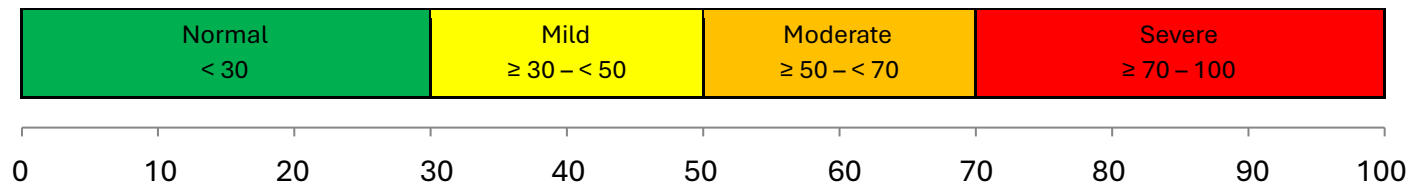
CONDUCT DISORDER (CD)

The conduct disorder module effectively measures conduct disorder symptomatology in the child from the perspective of both the child and parent. The module has an item bank of 123 items for children or 119 items for parents and requires an average of 10 questions administered in an average of 133 seconds to accurately predict the likelihood of a diagnosis of conduct disorder with AUC = 0.89. The conduct disorder module is best used for screening and is ideal for measurement and evaluation over time.

Child Severity Thresholds



Parent Severity Thresholds



Metrics: Severity, Category, Precision, Probability of CD

SUBSTANCE USE DISORDER (SUD)

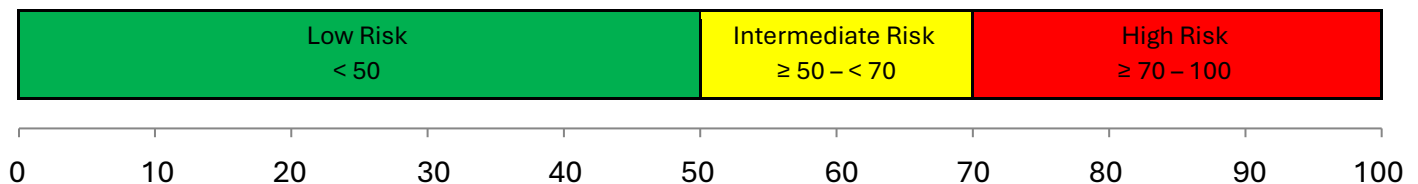
The substance use disorder (SUD) module uses a set of standard diagnostic questions as well as adaptive questions to measure the risk and severity of substance use among children.

The diagnostic element of the module includes questions related to use of 10 specific substances in the last 30 days including: alcohol, marijuana, nicotine, opioids, cocaine, amphetamines, hallucinogens, inhalants, sedatives, and over the counter medications such as cough syrup or Imodium®.

The adaptive SUD severity component provides a crosswalk between substance use and mental health symptoms (depression, anxiety, PTSD), social support and risky behaviors. It can detect SUD risk in children who have not yet started to use substances or refuse to admit that they are. The severity measure describes the impact of the specific SUD diagnoses on the person's current life.

This combination of diagnostic and adaptive questions enables the SUD module to be effective as a screening tool and for longitudinal measurement.

Child Severity Thresholds



Metrics: Severity, Category, Precision, Diagnosis of specific SUDs (either current and/or lifetime)

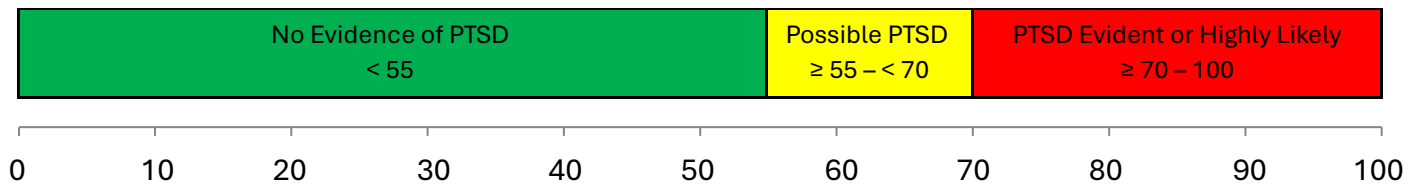
Notes:

1. This module will always measure behavior and risk over the last 30 days; this will occur even when other modules reference different timeframes.
2. This module does not have a parent component.

POST-TRAUMATIC STRESS DISORDER (PTSD)

The Child PTSD module is a dimensional measure of childhood trauma severity based on a modification of the Child PTSD Symptom Scale for *DSM-5* (CPSS-5), a measure of Post-Traumatic Stress Disorder (PTSD) for traumatized children based on *Diagnostic and Statistical Manual of Mental Disorders* (5th ed., American Psychiatric Association, 2013; [*DSM-5*]) criteria. The test can be adaptively administered using an average of 6 items and maintains a correlation of $r = 0.98$ with the total 20 item test score.

Child Severity Thresholds



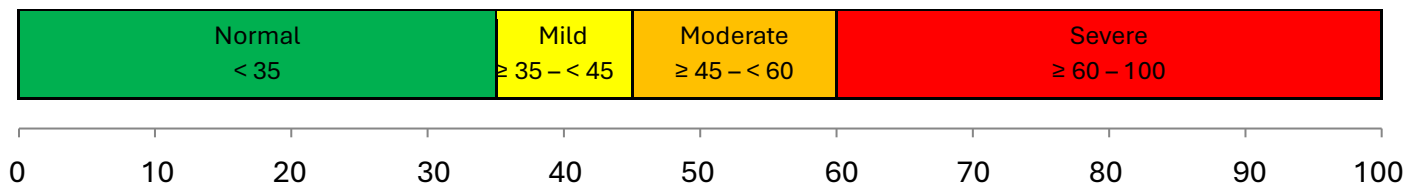
Metrics: Severity, Category, Precision, Evidence of PTSD

Notes: This module does not have a parent component.

SUICIDE

The suicide scale module effectively measures suicidal ideation and behavior in the child from the child's perspective. The module has an item bank of 64 items and requires an average of 10 questions administered in an average of 96 seconds to accurately predict the likelihood of suicidal ideation or behavior based on a structured clinical interview, with AUC = 0.996. This tool is best used to measure suicidality, stratify risk of a future suicide attempt over the next 3 to 6 months, and assess change in the severity of suicidality over time.

Child Severity Thresholds



Metrics: Severity, Category, Precision, Probability of Suicide Attempt (Next 3 Months), Alert

Notes:

1. This module does not have a parent component.

Probability of Attempt (Next 3 Months)

The K-CAT® Child/Suicide Scale estimates the probability of a suicide attempt within the next 3 months, with sensitivity and specificity of 83%. Because suicide attempts are relatively rare events, a probability threshold of 0.10 is used to identify individuals at elevated risk. A suicide risk indicator is prominently displayed on the first page of the report when any of the following conditions are met:

- The severity score falls within the Severe range (≥ 60)
- The estimated probability of attempt exceeds 0.10

The most accurate estimate of the probability of a suicide attempt within the next 3 months is obtained from multirater K-CAT® interviews that include the Child/Suicide Scale (AUC = 0.86). The probability can also be calculated using only the K-CAT® psychopathology measures (AUC = 0.85) or the Child/Suicide Scale alone (AUC = 0.83).

REFERENCES

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). [https://doi.org/10.1176.appi.books.9780890425596](https://doi.org/10.1176/appi.books.9780890425596)