Title

The UCLA PTSD Reaction Index for DSM-5 Brief Form: A Screening Tool for Trauma-Exposed Youths

Journal

Journal of the American Academy of Child & Adolescent Psychiatry, 59(3)

ISSN

0890-8567

Authors

Rolon-Arroyo, Benjamin
Oosterhoff, Benjamin
Layne, Christopher M
et al.

Publication Date

2020-03-01

DOI

10.1016/j.jaac.2019.06.015

Peer reviewed

NEW RESEARCH

Check for updates

The UCLA PTSD Reaction Index for *DSM-5* Brief Form: A Screening Tool for Trauma-Exposed Youths

Benjamin Rolon-Arroyo, PhD, Benjamin Oosterhoff, PhD, Christopher M. Layne, PhD, Alan M. Steinberg, PhD, Robert S. Pynoos, MD, MPH, Julie B. Kaplow, PhD, ABPP

Objective: Children and adolescents who experience potentially traumatic events are at risk for developing posttraumatic stress disorder (PTSD). Although psychometrically sound measures are now available to assess these youths, brief tools are currently needed for screening purposes. Two studies were conducted to develop and validate the UCLA PTSD Reaction Index for *DSM-5*–Brief Form (RI-5-BF).

Method: Study 1 used item response theory models to derive the RI-5-BF from the UCLA PTSD Reaction Index for *DSM-5* and assess its internal consistency using a sample of 486 trauma-exposed youths (mean age = 13.32 years, SD = 2.90) recruited through a practice research network. Study 2 used receiver operating characteristic analyses and diagnostic efficiency statistics to assess the discriminant-groups validity and clinical utility of the RI-5-BF in identifying children at different levels of PTSD risk using a sample of 41 treatment-seeking youths (mean age = 12.44 years, SD = 2.99).

Results: In study 1, item response theory models identified the 11 most informative items across their respective subscales. The RI-5-BF exhibited excellent internal consistency in both studies ($\alpha > .93$). In study 2, receiver operating characteristic analyses indicated that an RI-5-BF score of 21 maximized sensitivity and specificity. Moreover, diagnostic likelihood ratios across multiple levels of scores provided support for the measure's clinical utility in identifying different levels of PTSD risk.

Conclusion: These findings provide support for both the psychometric properties of the RI-5-BF as a brief screening measure for PTSD in children and adolescents and its utility for identifying youths meriting further assessment and consideration for treatment.

Key words: assessment, PTSD, screening, trauma

J Am Acad Child Adolesc Psychiatry 2020;59(3):434–443.

Evidence-based assessment (EBA) focuses on supporting three major phases of clinical activity: prediction of diagnoses or other criteria, prescription of treatment or moderating factors, and assessing therapeutic process.[1] EBA is an organizing framework for supporting treatment through tailored assessment, as guided by such goals as improving accuracy, enhancing clinical outcomes, and increasing the efficiency of services. Risk screening falls under the prediction of diagnoses or other criteria arm of EBA. Screening has been defined as the identification of an unrecognized condition by the application of tests, examinations, or other procedures that can be applied rapidly.[2] Screening tests sort out persons who probably have a condition from persons who probably do not have the condition. A screening test is not intended to be diagnostic. Established recommendations for screening suggest that a screening test must identify a harmful condition; the condition must be prevalent; early detection through screening should lead to improved outcomes; and screening methods should be effective, efficient, and practical.[2,3] A fusion between the principles of EBA and simplified Bayesian methods has emerged in the past decade to create a hybrid method that is more client centered, and it defines a set of applied research topics that are highly clinically relevant.[4] Guided by these principles, the present study focused on developing a valid and clinically useful screening measure for posttraumatic stress disorder (PTSD) in children and adolescents.

Studies demonstrate that PTSD in children and adolescents meets the first three recommendations for screening (ie, PTSD is harmful, it is prevalent, and its early detection can lead to improved outcomes). However, few psychometrically sound and clinically useful screening methods for PTSD are available for youths.[5] Research studies estimate that more than half of children and adolescents in the United States experience at least one potentially traumatic event during their lifetime.[6] Approximately 15.9% of children and adolescents exposed to trauma go on to develop PTSD[7]—a condition that is likely to persist without intervention.[8] Consistent with principles of trauma-informed systems,[9] early identification is a key component of preventing or mitigating the effects of the

full disorder.[10] However, assessing the full *DSM-5*[11] criteria for PTSD can be time-consuming, especially during primary care or school-based health clinic visits, where trauma-exposed youths are likely to first seek help.[12] Consequently, providers need screening tools that can accurately and efficiently sort between trauma-exposed youths who likely need further assessment and trauma-focused intervention versus youths who likely do not further assessment and intervention.

To our knowledge, only two screening measures have been developed for PTSD as defined by *DSM-5* criteria. The Child Trauma Screen[13] is a 10-item screening tool derived from *DSM-IV* PTSD measures by selecting items that maximized item-level correlations with overall PTSD symptom severity and are broadly represented in *DSM-5* PSTD diagnostic criteria. The Child Posttraumatic Stress Scale for *DSM-5* Screen[14] was created by identifying the six most frequently endorsed items in the validation study of the full scale. Internal consistency, test-retest reliability, convergent and divergent validity, and criterion validity were reportedly adequate or better for both screening measures. Creators of both screening tools also conducted receiver operating characteristic (ROC) analyses to derive cutoff scores to determine PTSD risk; however, both screening tools utilized full scale scores from tests used to derive the screening items themselves as diagnostic benchmarks. This practice is not recommended, as shared method variance between predictor and criterion inflates the accuracy of the ROC analyses and biases the selected cutoff scores.[15] The field thus remains in need of screening tools for PTSD in youths.

Consistent with best-practice test construction methods,[16] we employed a two-study design to derive and validate a screening tool for *DSM-5* PTSD from the UCLA PTSD Reaction Index for *DSM-5* (RI-5), a psychometrically sound clinician-administered assessment tool for PTSD in children and adolescents.[17] Study 1 used a large sample of youths recruited through a practice research network of sites that provide supportive services for youths with trauma and bereavement exposure, with the aim of creating the UCLA PTSD Reaction Index for *DSM-5*–Brief Form (RI-5-BF) and evaluating its reliability. We conducted item response theory (IRT) models, as IRT is the preferred method for developing brief measures.[18] IRT relates item responses to a latent trait and allows the examination of each item's ability to discriminate between persons who vary in the degree to which they possess an underlying trait.[19] Such considerations are relevant when developing screening tools given that items can exhibit different levels of discriminative precision at different levels of the underlying trait. Our goal was to select three screening items from each of the RI-5 subscales that would provide the most information at the 50th percentile of the corresponding latent traits (subscale B [intrusion; 5 items], subscale D [negative alterations in cognitions and mood; 13 items], and subscale E [arousal and reactivity; 7 items]; subscale C contains only two items that were preselected by default). Three items per subscale is optimal for latent variable modeling techniques and recommended for measurement refinement[20] and would provide information regarding all PTSD criteria. We then computed item-total correlations to evaluate subscale reliability and calculated Cronbach's α to measure the reliability of the full scale. Study 1 hypotheses were as follows: First, IRT models would successfully identify candidate RI-5-BF items (three items for subscales B, D, and E). Second, RI-5-BF subscale and full scale scores would exhibit acceptable internal consistency reliability as indicated by item-total correlation coefficients $>.30$ and Cronbach's α values $>.70$.[21]

Study 2 used a separate sample of treatment-seeking youths with trauma and bereavement exposure to replicate the reliability of RI-5-BF (utilizing the same reliability coefficients used in study 1) and evaluate its discriminant-groups validity and clinical utility. Discriminant-groups validity is of great importance in the EBA approach to prediction because it reflects the ability of a tool to improve classification or discrimination between groups.[1] Clinical utility is also important, as discussions of EBA have focused on reliability and classical concepts of psychometric validity, but not application to individual decision making.[22] We used ROC analyses to evaluate the discriminant-groups validity of the RI-5-BF subscales and full scale scores, gauged by their capacity to discriminate between trauma-exposed youths who met PTSD diagnostic criteria and trauma-exposed youths who did not meet PTSD diagnostic criteria as benchmarked against a gold standard semistructured clinical interview. We then evaluated the clinical utility of the RI-5-BF by estimating multilevel score thresholds to classify different levels of PTSD risk. Study 2 hypotheses were as follows: First, the RI-5-BF subscale and full scale scores will exhibit acceptable internal consistency as gauged by item-total correlation coefficients and Cronbach's α coefficients, respectively.[21] Second, the RI-5-BF will show evidence of discriminant-groups validity, such that its subscale scores and the full scale score will discriminate between youths with PTSD and youths without PTSD. Third, the RI-5-BF full scale score will provide multilevel score thresholds that accurately classify youths at different levels of PTSD risk.

METHOD

Study 1: Participants

Study 1 participants included 486 children and adolescents from two urban cities across the United States. Sample characteristics included 54% girls; age range from 7 to 18 years identified as black, 37% self-identified as white/

Caucasian, 7% self-identified as mixed/biracial, 3% self-identified as other, and 10% not reported. Participants reported an average of 3.64 different types of potentially traumatic events (SD = 2.30). More information regarding the sample is published elsewhere.[17]

Procedure

Study 1 participants were recruited as part of a large practice research network.[17] The network comprises school-based health clinics, grief support centers, community clinics, and academic medical center settings that serve trauma-exposed and/or bereaved youths. The network uses a shared battery of common assessment tools to create a data repository with the aim of validating trauma-informed and bereavement-informed assessment tools across diverse populations. Participating youths assented to contribute their anonymous deidentified data to the repository. Children younger than age 7 or older than age 18 were excluded from the study. All eligible youths agreed to participate, resulting in a 100% response rate. Notably, participating youths were already seeking mental or behavioral health support services and were administered assessment tools as part of standard clinical practice. The study received institutional review board approval from the hub institution.

Measures

RI-5. RI-5 is a clinician-administered tool that includes six sections: Trauma History Profile, Trauma Details, PTSD Symptom Scale, Frequency Rating Sheet, Distress and Impairment in Functioning, and a scoring sheet.[17] The present study focused on the PTSD symptom scale and the frequency rating sheet. RI-5 subscales and individual items correspond directly with *DSM-5* symptom criteria.[17] Subscale B has 5 items, subscale C has 2 items, subscale D has 13 items, and subscale E has 7 items. Some symptoms are assessed by more than one item (ie, D2, D3, D4, and E2). In the present study, all items are treated independently from each other within their subscale.

The PTSD symptom scale and frequency rating sheet are used to rate the number of days during the past month in which the child or adolescent experienced each PTSD symptom. RI-5 uses a Likert scale to assess the frequency with which the symptom was experienced during the past month (0 = never, 1 = little, 2 = some, 3 = much, and 4 = most). Information regarding item development and scoring procedures is published elsewhere.[17] RI-5 subscales (with the exception of subscale C in some samples) and the full scale have exhibited acceptable or better internal consistency.[17,23,24] Moreover, recent evidence has provided support for factor structure,[23,24] convergent validity, and discriminant validity of RI-5.[17,24]

Data Analytic Plan

We used IRT models utilizing Mplus 8 to select RI-5 items for the screening tool.[25] More specifically, the graded response model (GRM)[26] framework was used given the Likert-scale response design (ie, polytomous items) of RI-5 items. The primary objective of GRMs is to specify associations between item responses and the latent trait the subscale is theorized to measure. We first tested the IRT model's assumption of unidimensionality by conducting a four-factor confirmatory factor analysis, in which each factor corresponded to one of the four RI-5 PTSD symptom subscale categories (subscales B, C, D, and E). Factors were allowed to intercorrelate. We used a mean-adjusted and variance-adjusted weighted least squares estimator, given its appropriateness as an estimator for categorical data, and evaluated goodness of fit using three indicators: root mean square error of approximation <0.08, comparative fit index >0.90, and Tucker Lewis Index >0.90.[27]

We then estimated separate GRMs for each of the four RI-5 subscales to maintain their independence in determining parameter estimates, which helps ensure that generated subscales can be modeled individually. Maximum likelihood and a logit link function were used to obtain item parameter estimates and item information functions. GRMs produce discrimination (α) and difficulty (β) parameter estimates.[28] A discrimination parameter refers to an item's ability to differentiate between individuals along the latent construct in question (higher parameters indicate better discrimination).[19] In contrast, item difficulty parameters represent a given construct (eg, arousal) level as a function of the maximum probability of choosing each response option. Three items were selected from each subscale (ie, except subscale C, as it comprises only two items). To identify the most informative items for screening purposes, we examined item information functions. We approached this goal by choosing items with the most information at the 50th percentile of the corresponding latent traits of each of the RI-5 subscales (except subscale C), as this is the area of the latent trait of greatest interest for screening purposes.[29] Items that contribute the most information at higher levels of the underlying symptoms would not be useful for screening purposes, as they are likely to miss cases that fall in the mild-to-moderate risk level (ie, increase specificity at the expense of sensitivity and thus less appropriate for a screening tool). We used item-total correlation coefficients to evaluate the reliability of each RI-5-BF subscale as recommended for scales with very few items and Cronbach's α to evaluate the internal consistency of the RI-5-BF full-scale score.[21]

Study 2: Participants

Study 2 participants included 41 treatment-seeking children and adolescents (59% girls) with age range

from 7 to 17 years (mean = 12.44 years, SD = 2.99). Of these, 39% identified as black; 27% identified as Caucasian; 20% identified as Latino; and 15% identified as mixed/biracial. On average, participants reported 4.02 different types of potentially traumatic events (SD = 2.30). More information regarding the sample is published elsewhere.[17]

Procedure

Participants were recruited through an outpatient clinic housed in a large academic medical center that provides EBA and intervention services to youths 7 to 17 years of age who have experienced potentially traumatic events. Clinic procedures involve administering a standardized assessment protocol, including measures of PTSD, to all children and adolescents presenting at the clinic. For purposes of study 2, participants were administered an additional semistructured clinical interview, the Clinician-Administered PTSD Scale for *DSM-5*–Child/Adolescent Version (CAPS-CA-5)[29] and were compensated for their time. Of 41 youths approached about the study (on a rolling basis as they presented to the clinic), all agreed to participate (100% response rate).

TABLE 1 Study 1: Graded Response Model Item Parameter Estimates for UCLA PTSD Reaction Index for *DSM-5* (RI-5) Items

RI-5 Subscale	Item Description	α (SE)	β_1	β_2	β_3	β_4
B	5. Dissociative reaction(s) as if trauma was recurring. (B3)	1.93 (0.19)	−.28	.91	1.45	2.05
	10. Recurrent distressing dreams related to trauma. (B2)	1.94 (0.20)	−.19	.42	1.11	1.68
	11. Psychological distress to external and internal trauma reminders. (B4)	**2.68 (0.25)**	**−.65**	**.11**	**.63**	**1.33**
	14. Physiological reactivity to external and internal trauma reminders. (B5)	**2.34 (0.24)**	**−.22**	**.28**	**.83**	**1.50**
	18. Recurrent, involuntary, and intrusive distressing memories of trauma. (B1)	**3.03 (0.33)**	**−.18**	**.31**	**.91**	**1.41**
C	**3. Avoidance of external reminders related to trauma. (C2)**	**2.11 (12.01)**	**−.23**	**.33**	**.85**	**1.40**
	13. Avoidance of thoughts, feelings, or conversations related to trauma. (C1)	**2.17 (20.03)**	**−.49**	**.05**	**.66**	**1.14**
D	2. Persistent and exaggerated negative beliefs. (D2)	1.48 (0.15)	−.14	.79	1.59	2.31
	6. Persistent negative emotional state. (D4)	1.34 (0.15)	.26	.77	1.33	1.90
	7. Markedly diminished interest in activities. (D5)	1.66 (0.18)	.39	.92	1.49	2.13
	9. Persistent and exaggerated negative beliefs. (D2)	1.34 (0.13)	−.53	.20	.94	1.89
	12. Persistent inability to experience positive emotions. (D7)	**2.51 (0.24)**	**.15**	**.66**	**1.14**	**1.65**
	15. Persistent and distorted cognitions. (D3)	1.38 (0.15)	.01	.61	1.17	1.83
	16. Persistent and exaggerated negative beliefs. (D2)	**2.38 (0.22)**	**.07**	**.59**	**1.11**	**1.72**
	17. Feelings of detachment or estrangement. (D6)	**2.68 (0.25)**	**−.01**	**.45**	**.97**	**1.47**
	19. Persistent and distorted cognitions. (D3)	2.03 (0.22)	.36	.96	1.52	2.10
	22. Persistent negative emotional state. (D4)	2.37 (0.24)	.45	.93	1.45	1.78
	23. Inability to recall aspects of the trauma. (D1)	1.05 (0.14)	.53	1.36	2.36	3.14
	25. Persistent negative emotional state. (D4)	1.98 (0.18)	−.13	.57	1.41	1.92
	27. Persistent negative emotional state. (D4)	1.24 (0.18)	.94	1.43	1.91	2.32
E	1. Hypervigilance. (E3)	.96 (0.12)	−.81	.05	.95	1.71
	4. Irritable behavior and angry outbursts. (E1)	**1.50 (0.17)**	**−.53**	**.14**	**.89**	**1.61**
	8. Problems with concentration. (E5)	**1.94 (0.22)**	**−.72**	**−.05**	**.71**	**1.31**
	20. Reckless or self-destructive behaviors. (E2)	1.84 (0.32)	1.23	1.62	2.15	2.58
	21. Sleep disturbance. (E6)	**1.78 (0.20)**	**−.56**	**.11**	**.66**	**1.17**
	24. Exaggerated startle response. (E4)	1.20 (0.15)	−.61	.20	.93	1.65
	26. Reckless or self-destructive behaviors. (E2)	1.61 (0.27)	.79	1.36	1.84	2.46

Note: *Boldface type indicates chosen items.* α *values.* α = *discrimination parameters (slope at each difficulty parameter);* β = *difficulty parameters (latent trait level at which the probability of next item response is .50).*

Measures

Demographics. Demographic information, including sex, age, and race/ethnicity, was obtained through in-person interviews with caregivers.

RI-5. RI-5 was identical to that used in study 1.

CAPS-CA-5. CAPS-CA-5, a developmentally modified version of its adult counterpart,[30] is a 30-item clinician-administered semistructured interview designed to assess *DSM-5* PTSD diagnostic criteria for youths 7 years of age and older. A detailed description of CAPS-CA-5 is available elsewhere.[17]

For purposes of study 2, the clinic director (J.B.K.) trained two clinicians (one master's-level and one doctoral-level) to administer and score CAPS-CA-5. All participant interviews were videotaped; 38% were double-coded and randomly selected at intermittent points throughout the study to evaluate interrater reliability. The team used intraclass correlation coefficients to assess interrater reliability between coders for continuous ratings and κ statistics to evaluate interrater reliability between coders for diagnostic and subtype status. Intraclass correlation coefficients for all symptom cluster scores were very good (0.80–1.0),[31] and κ coefficients for diagnostic and subtype status were excellent (all 1.0). Any between-rater discrepancies were discussed with the clinic director to reach consensus.

Data Analytic Plan

We used item-total correlation coefficients to evaluate the reliability of each RI-5-BF subscale as recommended for scales with very few items and Cronbach's α to evaluate the internal consistency of the RI-5-BF full scale score.[21] ROC analyses and related diagnostic efficiency statistics were conducted to evaluate discriminant-groups validity and clinical utility of RI-5-BF.[3,15] ROC analyses estimate the diagnostic accuracy of a measure by comparing it with a criterion test (eg, gold standard structured interview), producing thresholds that maximize a measure's sensitivity (ie, accuracy of RI-5-BF in including youths who meet PTSD diagnostic criteria) and/or specificity (ie, accuracy of RI-5-BF in excluding youths who do not meet PTSD diagnostic criteria). RI-5-BF subscales B, C, D, and E, and the full scale score (sum of subscales B, C, D, and E) served as input, and PTSD diagnosis according to CAPS-CA-5 served as the gold standard criterion. The probability of accurately classifying children with PTSD is estimated by the area under the curve coefficient, which is tested against the null hypothesis of chance performance (0.50). We used the following area under the curve benchmarks: ≥ 0.9 = excellent, ≥ 0.80 = good, ≥ 0.70 = fair, and < 0.70 = poor.[32]

Multilevel diagnostic likelihood ratios (DLRs) were conducted to assess different levels of PTSD risk based on RI-5 full scale scores.[15] DLRs are data-driven diagnostic

TABLE 2 Corrected Item Total Correlations (ITCs) for UCLA PTSD Reaction Index for *DSM-5*–Brief Form (RI-5-BF) Subscales and Cronbach's α for the Full Scale in Study 1 and Study 2

Subscale	Item	Study 1 ITC[a]	Study 2 ITC[b]
B	11. Psychological distress to external and internal trauma reminders. (B4)	.71	.72
	14. Physiological reactivity to external and internal trauma reminders. (B5)	.65	.63
	18. Recurrent, involuntary, and intrusive distressing memories of trauma. (B1)	.69	.79
C	3. Avoidance of external reminders related to trauma. (C2)	.51	.55
	13. Avoidance of thoughts, feelings, or conversations related to trauma. (C1)	.51	.55
D	12. Persistent inability to experience positive emotions. (D7)	.67	.76
	16. Persistent and exaggerated negative beliefs. (D2)	.68	.67
	17. Feelings of detachment or estrangement. (D6)	.74	.82
E	4. Irritable behavior and angry outbursts. (E1)	.46	.50
	8. Problems with concentration. (E5)	.58	.48
	21. Sleep disturbance. (E6)	.46	.44
	Total scale Cronbach's α	.90	.93

Note: Values represent correlations among a given item with the mean total score for the remaining items in the subscale and full scale, respectively.
[a]*n = 448–473.*
[b]*n = 38–41.*

efficiency statistics that provide information about the change in the odds of a diagnosis associated with a particular test score.[33] DLRs are derived from the sensitivity and specificity, independent of the base rate, and are more likely to generalize outside the sample where it was developed.[34] DLRs are calculated via the formula (proportion of cases with the diagnosis who score within a given score range)/(proportion of cases without the diagnosis who score within the same range). DLRs were first computed based on quintiles to provide estimates with which to calculate more informative thresholds.[15] Posterior probabilities were then estimated based on Bayes' theorem by synthesizing the prior probability of the diagnosis (base rate) with the information from the test result (DLRs).[35] Because the base rate of a diagnosis directly affects overall classification accuracy as well as the positive and negative predictive powers of a test,[34] posterior probabilities were first computed for the study 2 sample base rate (32%). Then, to provide more conservative estimates, we searched the literature for a base rate for outpatient clinic settings that do not necessarily specialize in treating trauma-exposed youths. Our search led us to calculate posterior probabilities based on a 15% base rate.[36] DLRs <1.0 lower the odds of a PTSD diagnosis, DLRs between 2 and 5 represent a moderate increase, DLRs between 5 and 10 represent a large increase, and DLRs >10 are often clinically decisive odds changes.[33]

RESULTS

Study 1: Preliminary Results

The confirmatory factor analysis conducted to test the RI-5 factor structure yielded good fit, $\chi^2_{318} = 819.68$, comparative fit index = 0.96, Tucker Lewis Index = 0.99, root mean square error of approximation = 0.06 (90% CI 0.05–0.06). All items loaded onto their respective factor (Table S1, available online), and all factors intercorrelated strongly ($r = .70–.89$).

IRT Models

Table 1 presents GRM item discrimination and difficulty parameter estimates for all RI-5 items by subscale. Figures S1–S4, available online, present items information functions for each RI-5 subscale. We selected three items from subscales B (items 11, 14, and 18), D (items 12, 16, and 17), and E (items 4, 8, and 21) and two from subscale C (items 3 and 13). Selected items in subscales B and D exhibited the highest degree of discrimination at the 50th percentile of their corresponding latent traits. In subscale E, two items (items 20 and 26) exhibited a high degree of discrimination, both reflecting reckless and self-destructive behaviors. However, these items provided the most information at high levels of the latent trait. Thus, consistent with our goal of selecting items providing the most information at the 50th percentile of the latent trait for subscale E to increase the sensitivity of the screening tool, we selected items 21 and 8 from subscale E. Furthermore, difficulty parameters of selected items indicated appropriate spread across their latent trait continuum (Table 1). Owing to the wide age range in the sample, exploratory analyses were conducted by age group (ie, 7–13 years old and 14–18 years old). Results were similar to the results obtained with the full sample (Tables S2 and S3, available online).

Reliability of RI-5-BF

Table 2 presents reliability estimates in the form of item-total correlation coefficients for the RI-5-BF subscales and Cronbach's α for the full scale. Item-total correlation coefficients for all items by subscale indicated acceptable or better internal consistency.[21] Cronbach's α for the full scale indicated excellent internal consistency ($\alpha = .90$).[21] Owing to the wide age range in the sample, we conducted exploratory analyses by age group (ie, 7–13 years old and 14–18 years old). Results were similar to the results obtained with the full sample (Table S4, available online).

Study 2: Reliability of RI-5-BF

Table 2 presents reliability estimates in the form of item-total correlation coefficients for the RI-5-BF subscales and Cronbach's α for the full scale. Item-total correlation coefficients for all items by subscale indicated acceptable or better internal consistency.[21] Cronbach's α for the full scale indicated excellent internal consistency ($\alpha = .93$).[21]

Discriminant-Groups Validity for RI-5-BF

ROC analyses indicated that the RI-5-BF scores were able to accurately discriminate between cases with PTSD and cases without PTSD. Table 3 presents area under the curve statistics. RI-5-BF subscales and full scale scores discriminated PTSD significantly better than chance ($p < .001$). According to used benchmarks, RI-5-BF subscale C demonstrated good performance (0.87, SE = 0.06), whereas subscales B, D, and E and the full scale score demonstrated excellent performance (0.90–0.96).[32] ROC analyses indicated that a cutoff score of 21 on RI-5-BF maximized sensitivity (1.00), specificity (0.86), and overall classification accuracy (0.92).

Table 4 presents multilevel DLRs based on quintiles and on more informative thresholds (the first three quintiles were collapsed into one group to reduce redundancy, as they

TABLE 3 Study 2: Area Under the Curve (AUC) From Receiver Operating Characteristic Analyses

Index Test	AUC	SE	p	95% CI Lower	95% CI Upper	Effect Size[a]
RI-5-BF						
Subscale B score	0.96	0.03	< .001	0.91	1.00	*d* = 2.48
Subscale C score	0.87	0.06	< .001	0.76	0.98	*d* = 1.59
Subscale D score	0.90	0.05	< .001	0.81	0.99	*d* = 1.81
Subscale E score	0.92	0.05	< .001	0.83	1.00	*d* = 1.99
Full scale score	0.96	0.03	< .001	0.90	1.00	*d* = 2.48

Note: RI-5-BF = UCLA PTSD Reaction Index for DSM-5–Brief Form.
[a]*d = Cohen's d.*

all exhibited DLRs <1). DLRs are interpreted for clinical purposes by transforming the likelihood of having a diagnosis of PTSD (ie, prior probability) to the odds of having PTSD given a specified full scale score on the RI-5-BF (ie, posterior probability).[3] Posterior probabilities were computed based on two different prior probabilities/PTSD prevalence rates (32% and 15%). Based on the more informative thresholds, participants with RI-5-BF full scale scores in the low range (ie, 0–20) had a very low PTSD risk (DLRs <1) using both base rates; participants who scored 21 to 35 had a moderate increase in PTSD risk (DLR = 4.32), with posterior probabilities at 67% (base rate = 32%) and 43% (base rate = 15%), respectively. Participants with full scale scores of 36 to 44 (DLR = 14.94) carried the highest risk for PTSD, with posterior probabilities of 88% (base rate = 32%) and 73% (base rate = 15%), respectively. Overall, these findings indicate that RI-5-BF full scale scores ≥21 denote a clinically meaningful increase in the probability of a PTSD diagnosis and indicate moderate-to-high risk for PTSD.

DISCUSSION

The aim of this research was to derive a brief and clinically useful screening tool capable of accurately identifying trauma-exposed youths at significant risk for PTSD—RI-5-BF. To this end, we conducted two studies. Study 1 identified 11 RI-5 items via IRT analyses from RI-5 and evaluated their reliability. These items cover all *DSM-5* PTSD symptom criteria (B–E) and were selected for screening applications. Reliability (gauged by internal consistency) indicated that all subscales and the full scale were acceptable or better. Study 2, conducted with a sample of treatment-seeking trauma-exposed youths, examined the reliability, discriminant-groups validity, and clinical utility of RI-5-BF. Similar to study 1, all RI-5-BF subscales and the full scale demonstrated acceptable or better internal consistency. RI-5-BF discriminated between groups of trauma-exposed youths who met full *DSM-5* PTSD diagnostic criteria and trauma-exposed youths who did not meet full *DSM-5* PTSD diagnostic criteria as benchmarked against a gold standard semistructured clinical interview. Given these favorable properties, it was possible to generate multilevel DLRs to estimate different levels of risk for PTSD using the

TABLE 4 Study 2: Diagnostic Likelihood Ratios (DLRs) for UCLA PTSD Reaction Index for *DSM-5*–Brief Form (RI-5-BF) Full Scale Scores

RI-5-BF Full Scale Score	Sensitivity	Specificity	Level[a]	DLR+[b]	Prevalence of 32% Prior Probability	Prevalence of 32% Posterior Probability	Prevalence of 15% Prior Probability	Prevalence of 15% Posterior Probability
Maximum κ: 21	1.00	0.86	0.37	6.99	.32	.77	.15	.56
Multilevel DLRs (based on quintiles)								
0–5	—	—	~0.20	0.00	.32	.00	.15	.00
6–11	—	—	~0.20	0.00	.32	.00	.15	.00
12–23	—	—	~0.20	0.27	.32	.11	.15	.05
24–35	—	—	~0.20	5.42	.32	.72	.15	.49
36–44	—	—	~0.20	14.94	.32	.88	.15	.73
Multilevel DLRs (informative thresholds)								
0–20	—	—	~0.60	0.00	.32	.00	.15	.00
21–35	—	—	~0.60	4.32	.32	.67	.15	.43
36–44	—	—	~0.60	14.94	.32	.88	.15	.73

Note: [a]*Test positive rate.*
[b]*Positive diagnostic likelihood ratio.*

RI-5-BF full scale score. As expected, multilevel DLRs reflected different levels of risk, with a significant increase in risk at scores of 21 and higher.[33] Overall, these findings provide supporting evidence for reliability, discriminant-groups validity, and clinical utility of RI-5-BF.

Building on prior efforts,[37] the present study fills a significant gap in the literature given the dearth of established *DSM-5* PTSD screening tools for youths.[13,14] RI-5-BF has notable strengths that extend beyond those of other *DSM-5* brief screening tools, including high construct coverage, robust psychometric properties, and clear utility for clinical and research applications that involve assessing the likelihood of a PTSD diagnosis. Moreover, RI-5-BF was derived from RI-5, a developmentally informed and psychometrically sound assessment tool for trauma-exposed youths.[17] Tools for PTSD such as the RI-5-BF are needed especially in settings where potentially traumatic events are widespread and time and other resources are limited. Based on personal communications with clinic directors from two centers where RI-5-BF is currently in use, administration time for RI-5-BF ranges from 2–3 minutes to 5–8 minutes depending on the complexity of trauma exposure and method of administration (self-report versus provider-administered).

A major clinical strength of RI-5-BF is the capacity it offers to clinicians and organizations to adjust for the prevalence of PTSD in their communities when determining risk. This allows users to customize their approach for different populations as a function of the local level of risk. The use of multilevel thresholds also allows users to focus on identifying youths at highest risk for PTSD—a pressing priority in many child-serving settings, given evidence that PTSD is associated with impaired academic and intellectual functioning.[38,39] Thus, risk detection, risk stratification, and appropriate referral and triage are of vital importance to ensuring that youths are able to work on key developmental tasks, including school achievement and interpersonal relationships.[40] Accurate risk detection is especially necessary for organizations with limited resources in that it reduces classification errors, misspent resources (for false positives), and lost opportunities to intervene in timely ways (for false negatives).

Regarding research implications, RI-5-BF offers multiple advantages over current assessment tools. First, the brevity of RI-5-BF offers the capacity for large-scale studies. Second, despite its brevity, RI-5-BF is sufficiently comprehensive to permit latent variable modeling, confirmatory factor analysis, path analysis, and latent growth modeling. Third, the brevity and psychometric soundness of RI-5-BF as a screening tool makes it feasible for non-trauma/non-PTSD researchers to incorporate trauma-informed assessments into their study protocols. This will allow them to parcel out the predictive effects of trauma exposure and associated PTSD symptoms in the form of covariates, thereby ruling out competing explanations. This practice will also permit the creation of data archives that invite secondary data analysis focusing on policy-related questions involving trauma exposure and PTSD. Questions that could be profitably addressed include burdens that trauma exposure places on child-servicing systems,[41] needs assessment for specific populations and settings,[42,43] sequelae of early trauma exposure,[44] and the emergence of risky behavior in later developmental periods.[45]

Regarding limitations, both studies 1 and 2 relied solely on child-reported symptoms. Future studies can benefit from greater use of multiple methods and/or informants, including observational, parent-report, and archival data (eg, school records) of mental health, wellness, and functioning. Such measures can serve as external criteria for evaluating other forms of test validity, including predictive validity and sensitivity to clinical change. Nevertheless, child-reported internalizing symptoms, including PTSD symptoms, have exhibited greater predictive validity than parent reports of children's symptoms.[46] Both studies also used samples of children seeking mental or behavioral health support services for trauma or loss, potentially limiting the generalizability of our findings to youths not seeking support. Furthermore, study 2 used a modest sample size, which potentially limited study generalizability and the power to test for potential group differences. Nevertheless, our study met Kraemer's[3] guidelines when reporting the sensitivity and specificity of a test (ie, at least 10 cases that meet the diagnosis, 10 cases that do not meet the diagnosis, 10 cases that test positive, and 10 cases that test negative). Future research can use larger samples to examine whether the discriminant-groups validity of RI-5-BF varies as a function of demographic characteristics, including age, gender, race, and culture. Moreover, the accuracy of RI-5-BF in identifying children with PTSD decreased as a function of a lower base rate of PTSD (32% > 15%), placing limits on its utility for settings with lower PTSD base rates. Last, despite covering a wide age range (7–18 years), RI-5-BF does not extend to the assessment of trauma-exposed children younger than 7 years of age, underscoring the need to create and validate screening measures for young children.

Despite these limitations, studies 1 and 2 developed a psychometrically sound brief screening tool (RI-5-BF) for PTSD as defined by *DSM-5* criteria in trauma-exposed youths. This promising tool can be used in clinical settings with varying PTSD base rates given its capacity to efficiently identify trauma-exposed youths at risk for developing PTSD. Its brevity and developmental

appropriateness make it ideally suited for routine use in time-limited settings and situations (eg, primary care offices, schools), where brief yet accurate assessment tools are sorely needed. Other uses include risk screening in postdisaster settings, where assessing trauma and loss exposure and assessing consequent mental health needs are essential steps in situation analysis and needs assessment, respectively.[43]

Accepted July 26, 2019.

Drs. Rolon-Arroyo, Layne, Steinberg, and Pynoos are with the University of California, Los Angeles. Dr. Oosterhoff is with Montana State University, Bozeman. Drs. Layne, Steinberg, and Pynoos are also with UCLA/Duke University National Center for Child Traumatic Stress, Los Angeles, CA. Dr. Kaplow is with Baylor College of Medicine, Houston, TX.

This research was supported in part by a grant from the New York Life Foundation (PIs: Kaplow and Layne) and by grants from the Substance Abuse and Mental Health Service Administration (SAMHSA; SM-16008 and SM-062111), both given to the senior author (Kaplow). Dr. Rolon-Arroyo was also supported by a postdoctoral Ruth L. Kirchstein Institutional National Research Service Award (T32) at UCLA (MH-073517).

The authors thank Karen Calhoun, MA, of Michigan Institute for Clinical and Health Research at University of Michigan, Daphne Marbury, LLPC, of St. John Providence Health System, Laura Johnson-Hughes, MSW, of Starfish Family Services, Molly Kriesel, LMSW, of Starfish Family Services, Mary Beth Staine, CEO, of Bo's Place, Marian Mankin, LCSW, of Bo's Place, and LaTanya Porter-Howard, LMSW, of Henry Ford Health System, for assisting with data collection. The authors also thank Hayley Goldenthal, MA, of Northwestern Medical School, and Timothy Brown, MPH, of Baylor College of Medicine and Texas Children's Hospital, for assisting with data management and oversight.

Disclosure: Dr. Steinberg is Chief Science Officer of Behavioral Health Innovations, LLC. This company, by agreement with the UCLA Office of Intellectual Property, licenses and receives payment for use of the UCLA PTSD Reaction Index for *DSM-5*. Dr. Pynoos is Chief Medical Officer of Behavioral Health Innovations, LLC. This company, by agreement with the UCLA Office of Intellectual Property, licenses and receives payment for use of the UCLA PTSD Reaction Index for *DSM-5*. Drs. Rolon-Arroyo, Oosterhoff, Layne, and Kaplow has reported no biomedical financial interests or potential conflicts of interest.

Correspondence to Julie B. Kaplow, PhD, ABPP, 6701 Fannin Street, CC1630, Houston, TX 77030; e-mail: Julie.Kaplow@bcm.edu

https://doi.org/10.1016/j.jaac.2019.06.015

REFERENCES

1. Youngstrom EA, Van Meter A, Frazier TW, Hunsley J, Prinstein MJ, Youngstrom JK. Evidence-based assessment as an integrative model for applying psychological science to guide the voyage of treatment. Clin Psychol. 2017;24:331-363.
2. Rampes H, Warner JP, Blizard R. How to appraise an article on diagnosis. Psychiatr Bull R Coll Psychiatr. 1998;22:506-509.
3. Kraemer HC. Evaluating Medical Tests: Objective and Quantitative Guidelines. Newbury Park, CA: Sage; 1992.
4. Youngstrom EA. Future directions in psychological assessment: combining evidence-based medicine innovations with psychology's historical strengths to enhance utility. J Clin Child Adolesc Psychol. 2013;42:139-159.
5. Eklund K, Rossen E, Koriakin T, Chafouleas SM, Resnick C. A systematic review of trauma screening measures for children and adolescents. Sch Psychol Q. 2018;33:30-43.
6. Finkelhor D, Turner H, Ormrod R, Hamby SL. Violence, abuse and crime exposure in a national sample of children and youth. Pediatrics. 2009;124:1411-1423.
7. Alisic E, Zalta AK, van Wesel F, *et al*. Rates of post-traumatic stress disorder in trauma-exposed children and adolescents: meta-analysis. Br J Psychiatry. 2014;204:335-340.
8. Hiller RM, Meiser-Stedman R, Fearon P, *et al*. Research review: changes in the prevalence and symptom severity of child post-traumatic stress disorder in the year following trauma—a meta-analytic study. J Child Psychol Psychiatry. 2016;57:884-898.
9. Ko SJ, Ford JD, Kassam-Adams N, *et al*. Creating trauma-informed systems: child welfare, education, first responders, health care, juvenile justice. Prof Psychol Res Pr. 2008;39:396-404.
10. Issakidis C, Sanderson K, Corry J, Andrews G, Lapsley H. Modeling the population cost-effectiveness of current and evidence-based optimal treatment for anxiety disorders. Psychol Med. 2004;34:19-35.
11. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th ed. Washington, D.C.: American Psychiatric Association; 2013.
12. Wallerstein N, Calhoun K, Eder M, Kaplow J, Hopkins Wilkins C. Voices from the field: building a cross-disciplinary culture. In: Hall K, Croyle R, Vogel A, eds. Advancing Social and Behavioral Health Research Through Cross-Disciplinary Team Science: Principles for Success. New York: Springer Publishing; 2018.
13. Lang JM, Connell CM. Development and validation of a brief trauma screening measure for children: the child trauma screen. Psychol Trauma. 2017;9:390-398.
14. Foa EB, Asnaani A, Zang Y, Capaldi S, Yeh R. Psychometrics of the child PTSD symptom scale for DSM-5 for trauma-exposed children and adolescents. J Clin Child Adolesc Psychol. 2018;47:38-46.
15. Youngstrom EA. A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: we are ready to ROC. J Pediatr Psychol. 2014;39:204-221.
16. DeVellis RF. Scale Development: Theory and Applications. 3rd ed., Vol. 26. Thousand Oaks, CA: Sage; 2012.
17. Kaplow JB, Rolon-Arroyo B, Layne CM, *et al*. Validation of the UCLA PTSD reaction index for DSM-5: a developmentally informed assessment tool for youth [Epub ahead of print Apr 3]. J Am Acad Child Adolesc Psychiatry. 2020;59:186-194.
18. Fraley RC, Waller NG, Brennan KA. An item response theory analysis of self-report measures of adult attachment. J Pers Soc Psychol. 2000;78:350-365.
19. Baker FB. The Basics of Item Response Theory. 2nd ed. College Park, MD: ERIC; 2001.
20. Little TD, Cunningham WA, Shahar G, Widaman KF. To parcel or not to parcel: exploring the question, weighing the merits. Struct Equ Modeling. 2002;9:151-173.
21. Nunnally JC, Bernstein IH. The Assessment of Reliability. Psychometric Theory. 3rd ed. New York: McGraw-Hill; 1994.
22. Mash EJ, Hunsley J. Evidence-based assessment of child and adolescent disorders: issues and challenges. J Clin Child Adolesc Psychol. 2005;34:362-379.
23. Takada S, Kameoka S, Okuyama M, *et al*. Feasibility and psychometric properties of the UCLA PTSD reaction index for DSM-5 in Japanese youth: a multi-site study. Asian J Psychiatr. 2018;33:93-98.
24. Doric A, Stevanovic D, Stupar D, *et al*. UCLA PTSD reaction index for DSM-5 (PTSD-RI-5): a psychometric study of adolescents sampled from communities in eleven countries. Eur J Psychotraumatol. 2019;10:1605282.
25. Muthén LK, Muthén BO. Mplus User's Guide. 8th ed. Los Angeles: Muthén and Muthén; 2017.
26. Samejima F. Graded response model. In: Van der Linden WJ, Hambleton RK, eds. Handbook of Modern Item Response Theory. New York: Springer-Verlag; 1997.
27. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling. 1999;6:1-55.
28. Reeve BB, Fayers P. Applying item response theory modeling for evaluating questionnaire item and scale properties. In: Fayers P, Hays RD, eds. Assessing Quality of Life in Clinical Trials: Methods and Practice. 2nd ed. New York: Oxford University Press; 2005.
29. Yang FM, Jones RN, Inouye S, *et al*. Selecting optimal screening items for delirium: an application of item response theory. BMC Med Res Methodol. 2013;13.
30. Pynoos RS, Weathers FW, Steinberg AM, *et al*. Clinician-administered PTSD scale for DSM-5–child/adolescent version. National Center for PTSD. Available at: https://www.ptsd.va.gov/professional/assessment/child/caps-ca.asp. Accessed October 9, 2019.
31. Gwet KL. Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Raters. 3rd ed. Gaithersburg, MD: Advanced Analytics; 2012.
32. Swets JA. Measuring the accuracy of diagnostic systems. Science. 1988;240:1285-1293.
33. Strauss SE, Tetroe JM, Graham ID. Knowledge translation is the use of knowledge in health care decision making. J Clin Epidemiol. 2011;64:6-10.
34. Pepe MS. The Statistical Evaluation of Medical Tests for Classification and Prediction. New York: Wiley; 2003.
35. Schwarts A. Diagnostic Test Calculator; http://araw.mede.uic.edu/cgi-bin/testcalc.pl. Accessed May 16, 2019.
36. Gabbay V, Oatis MD, Silva RR, Hirsch G. Epidemiological aspects of PTSD in children and adolescents. In: Silva RR, ed. Posttraumatic Stress Disorders in Children and Adolescents: Handbook. New York: WW Norton and Co; 2004.

37. Kenardy JA, Spence SH, Macleod AC. Screening for posttraumatic stress disorder in children after accidental injury. Pediatrics. 2006;118:1002-1009.
38. Oosterhoff B, Kaplow JB, Layne CM. Links between bereavement due to sudden death and academic functioning: results from a nationally representative sample of adolescents. Sch Psychol Q. 2018;33:372-380.
39. Saigh PA, Yasik AE, Oberfield RA, Halamandaris PV, Bremner JD. The intellectual performance of traumatized children and adolescents with or without posttraumatic stress disorder. J Abnorm Psychol. 2006;115:332-340.
40. Saltzman W, Layne CM, Pynoos RS, Olafson E, Kaplow JB, Boat B. Trauma and grief component therapy for adolescents: a modular approach to treating traumatized and bereaved youth. Cambridge, UK: Cambridge University Press; 2017.
41. Briggs EC, Fairbank JA, Greeson JK, *et al*. Links between child and adolescent trauma exposure and service use histories in a national clinic-referred sample. Psychol Trauma. 2013;5:101-109.
42. Briggs EC, Greeson JK, Layne CM, *et al*. Trauma exposure, psychosocial functioning, and treatment needs of youth in residential care: preliminary findings from the NCTSN core data set. J Child Adolesc Trauma. 2012;5:1-15.
43. Kaplow JB, Layne CM, Rolon-Arroyo B. Evidence-based assessment in the aftermath of disasters: towards a best-practice model for evaluating hurricane-exposed youth. APSAC Advisor. 2018;30:6-20.
44. Greeson JKP, Briggs EC, Layne CM, *et al*. Traumatic childhood experiences in the 21st century: broadening and building on the ACE studies with data from the national child traumatic stress network. J Interpers Violence. 2014;29:536-556.
45. Layne CM, Greeson JKP, Kim S, *et al*. Links between trauma exposure and adolescent high-risk health behaviors: findings from the NCTSN core data set. Psychol Trauma. 2014;6(Suppl 1):S40-S49.
46. Stover CS, Hahn H, Im JJ, *et al*. Agreement of parent and child reports of trauma exposure and symptoms in the early aftermath of a traumatic event. Psychol Trauma. 2010;2:159-168.