Development of a scale battery for rapid assessment of risk and resilience

Tyler M. Moore\textsuperscript{a,b,1,*}, Lauren K. White\textsuperscript{a,b,1}, Ran Barzilay\textsuperscript{a,b}, Monica E. Calkins\textsuperscript{a,b}, Jason D. Jones\textsuperscript{a,b}, Jami F. Young\textsuperscript{a,b}, Ruben C. Gur\textsuperscript{a,b}, Raquel E. Gur\textsuperscript{a,b}

	extsuperscript{a} Department of Psychiatry, Brain Behavior Laboratory, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

\textsuperscript{b} Lifespan Brain Institute, Department of Child and Adolescent Psychiatry and Behavioral Sciences, Children’s Hospital of Philadelphia, Philadelphia, PA 19104, USA

\textsuperscript{*} Corresponding author at: Brain Behavior Laboratory, Perelman School of Medicine, University of Pennsylvania, 3700 Hamilton Walk, Philadelphia, PA 19104, USA.

\textsuperscript{1} These authors contributed equally.

\textsuperset{\textsuperscript{+}}This work was supported by the Lifespan Brain Institute (LiBI); NIMH grants MH089983, MH117014, and MH096891; and the Dowshen Program for Neuroscience.

\textsuperset{\textsuperscript{E-mail addresses}}{tymoore@pennmedicine.upenn.edu, tymoore@upenn.edu (T.M. Moore).}

\textsuperset{\textsuperscript{1}}These authors contributed equally.

https://doi.org/10.1016/j.psychres.2020.112996

Received 7 April 2020; Received in revised form 8 April 2020; Accepted 8 April 2020

Available online 18 April 2020

0165-1781/ © 2020 Elsevier B.V. All rights reserved.

ARTICLE INFO

Keywords:
Resilience
Environmental risk
Trauma
Social environment
Scale abbreviation

ABSTRACT

It is critical to understand the factors that increase risk for development of psychiatric disorders as well as promote resilience against disorders. The current study describes the development of a brief tool for risk/resilience assessment that takes a broad perspective of “risk” and “resilience” to characterize the phenomena, and assesses multiple factors that span intrapersonal, interpersonal, and wide-ranging external contexts. We administered twelve scales (212 items) to a diverse population comprising help-seeking and community participants ($N = 298; 46\%$ female) in the greater Philadelphia area. We used exploratory item-factor analysis to determine how items cluster across scales. After determining that a seven-factor solution was optimal, computerized adaptive testing (CAT) simulation was run to determine what would happen if the seven full-form factors were administered adaptively. These results were used to select items for short-form scales, producing seven final scales (items = 47). Validity was assessed by relating short-form scores to demographics, clinical diagnoses, scales, and criteria; these relationships were also compared to the relationships found with the original scales. Almost all effects detected by the twelve original scales were detected by the substantially abbreviated short-forms. The abbreviated battery shows promise for rapid assessment of multiple risk and resilience parameters, a necessity in large-scale studies.

1. Introduction

The strong genetic component of many psychiatric disorders is well established (Alemany et al., 2019) and the contribution of multiple environmental influences is also recognized (Kessler et al., 2010). Understanding the factors that increase the risk for the development of psychiatric disorders as well as promote resilience against the disorders is critical. To date, there is no standard method to assess risk and resilience across the lifespan. Establishing such an assessment can be beneficial to developmental psychopathology and has the potential to advance translational science. The current paper describes the development of a risk and resilience scale battery, including aggregation of well-established scales and, using empirical methods, selection of the optimal items among them.

Risk and resilience are multifaceted processes, thus characterizing and delineating their role in developmental psychopathology is complex. In particular, the term “resilience” has taken on multiple meanings and definitions over the last few decades (Masten and Barnes, 2018; Luthar et al., 2000). On the one hand, resilience can be thought of as a developmental process in which an individual achieves healthy or adaptive development despite exposure to risk factors. In this view of resilience, research focuses on identifying protective factors that promote adaptive outcomes in the face adversity. Notably, many factors can be thought of as both protective (e.g., positive family support) and risk (e.g., the absence of positive family support) factors. On the other hand, resilience has been conceptualized as a multi-faceted trait or characteristic of the individual that promotes positive development in the absence of adversity and/or adaptive responding to challenges (Luthar et al., 2000). In this view of resilience, research has focused on capturing what that trait is and on how those with high levels of the resilience trait develop or respond to challenges relative to those low on the resilience trait. Despite a great amount of interest in risk and resilience in developmental psychopathology research, the field lacks a standard assessment that measures multiple processes and captures the multifaceted, multidirectional nature of risk and resilience factors.
The current study describes the development of a risk and resilience assessment that takes a broad perspective of “risk” and “resilience” to better characterize the dynamic processes, and assesses multiple factors that span intrapersonal, interpersonal, and wide-ranging external contexts. Notably, it is not clear how the link between risk and resilience factors and the status of developmental disorders change across the lifespan. Critical to this endeavor is a standard assessment of risk and resilience that can be administered across different developmental epochs, from childhood and adolescence to adulthood.

In the current study, we first describe the development of a risk and resilience battery, present scale results and factor analyses from data collected on children, adolescents, and adults. Lastly, for data reduction and dissemination of the battery, we present results from computerized adaptive testing, which selects optimal items from all available scale items. This resulted in a short-form risk and resilience battery that can be administered across a wide age span in both community and clinical samples.

2. Methods

2.1. Participants

The sample included 298 individuals (46% females) who presented to the Lifespan Brain Institute (LBI) of Penn Medicine and Children’s Hospital of Philadelphia for sequential research assessments. These participants were recruited to take part in several ongoing studies examining neuropsychiatric disorders across development. Participants were recruited through the Department of Child and Adolescent Psychiatry and primary care offices at the Children’s Hospital of Philadelphia and advertising in community outlets. A large subset of the cohort (N = 140) was obtained as part of an ongoing community-based longitudinal study of youths from the Philadelphia Neurodevelopmental Cohort (Satterthwaite et al., 2016; Calkins et al., 2017). As such, the sample is heterogeneous, with a moderately wide age range (8–35 years; Mean = 18.72, SD = 5.03) and various psychiatric disorders. The race/ethnicity of the sample was: Caucasian = 32%; African American = 55%; Asian = 4%. Clinical diagnostic information for Phenotypes and eXposures (Hamilton et al., 2011) and the nine item Patient Health Questionnaire-9 (PHQ-9) (Kroenke et al., 2001) were used to assess self-reported depression symptoms. The seven item Promis Pediatric Anxiety (PPA) (Irwin et al., 2010) and the 41-item Screen for Child Anxiety Related Disorders (SCARED) (Birmaher et al., 1997) were used to assess self-reported anxiety symptoms.

In addition to self-report scales, the majority of participants (N = 250) also underwent a computer based semi-structured clinical interview with modules based on the KSADS and Structured Interview for Prodromal Symptoms (SIPS version 4.0; Calkins et al., 2017). The K-SADS modules provided a standardized assessment of DSM-IV axis 1 psychopathology (i.e., mood disorders, ADHD, anxiety disorders, OCD, PTSD, suicide ideation). The SIPS modules assessed psychosis spectrum symptoms. Collateral interviews were also conducted for participants age 8–18; only collateral interviews were conducted for participants 8–10 years of age. After the clinical assessments, information was aggregated across proband and collateral reports and medical records, if available, and consensus diagnoses were made by a team of clinical psychologists and psychiatrists. Three clinical consensus ratings on participant functioning were also given for each participant: Global Assessment of Functioning (GAF) (McGlashan et al., 2001), Global Function: Social Scale (GF: Social), and Global Function: Role Scale (Cornblatt et al., 2007).

2.2. Risk and resilience battery

The current battery is a composite of multiple, well-established questionnaires that were selected to assess various factors related to both risk and resilience. The scales were chosen by a team of experts, consisting of developmental psychologists, clinical psychologists, and adult and child and adolescent psychiatrists. Through consensus, the team decided on multiple domains that spanned intrapersonal (e.g., emotion regulation) and interpersonal factors (e.g., family relationships), as well as broader contexts (e.g., neighborhood safety). Priority was given to questionnaires that were well suited for a wide age range. Whenever possible, open access scales from PhenX (consensus measures for Phenotypes and eXposures) (Hamilton et al., 2011) and PROMIS (Cella et al., 2007) were chosen. See Table 1 for description of each scale (Cella et al., 2007; Hamilton et al., 2011; Earls et al., 2005; Mujahid et al., 2007; Forman et al., 1997; Fisher et al., 2000; Tiet et al., 1998; Betts et al., 2015; Furman and Buhrmester, 1985; Wagnild and Young, 1993; Liebenberg et al., 2013; Ebesutani et al., 2012; Kaufman et al., 2016; Ellis and Rothbart, 2005; Mynard and Joseph, 2000). The risk and resilience battery was computerized and administered on a laptop or tablet. The battery administered was modified according to participant age (see Table 1); age restrictions were placed only on younger participants. All individual scales (full-form) were scored using unit-weighted means (basic mean scores). For the factor scores, each participant received a score calculated by summing responses on all items within each factor and dividing by the total possible points for the participant on that factor. This was done even when a single score mixes across items with different numbers of response categories. Polytomous items therefore add more variance to the scores than do dichotomous items.

2.3. Clinical assessment

Participants were asked to fill out several self-report clinical scales to measures depression, anxiety, and psychosis spectrum symptoms. The seven item Promis depression Scale (PDS) (Pilkonis et al., 2011) and the nine item Patient Health Questionnaire-9 (PHQ-9) (Kroenke et al., 2001) were used to assess self-reported depression symptoms. The seven item Promis Pediatric Anxiety (PPA) (Irwin et al., 2010) and the 41-item Screen for Child Anxiety Related Disorders (SCARED) (Birmaher et al., 1997) were used to assess self-reported anxiety symptoms.

In addition to self-report scales, the majority of participants (N = 250) also underwent a computer based semi-structured clinical interview with modules based on the KSADS and Structured Interview for Prodromal Symptoms (SIPS version 4.0; Calkins et al., 2017). The K-SADS modules provided a standardized assessment of DSM-IV axis 1 psychopathology (i.e., mood disorders, ADHD, anxiety disorders, OCD, PTSD, suicide ideation). The SIPS modules assessed psychosis spectrum symptoms. Collateral interviews were also conducted for participants age 8–18; only collateral interviews were conducted for participants 8–10 years of age. After the clinical assessments, information was aggregated across proband and collateral reports and medical records, if available, and consensus diagnoses were made by a team of clinical psychologists and psychiatrists. Three clinical consensus ratings on participant functioning were also given for each participant: Global Assessment of Functioning (GAF) (McGlashan et al., 2001), Global Function: Social Scale (GF: Social), and Global Function: Role Scale (Cornblatt et al., 2007).

2.4. Statistical analysis

Analyses proceeded in three steps:

1. Factor analysis to determine item clustering (which items load on which factors).
2. For each factor identified above, estimate item response theory (IRT) (Reise and Moore, 2012) item parameters via the Graded Response Model (Samejima, 1969).
3. Simulate computerized adaptive testing (CAT) sessions to determine overall quality of items (used in selecting items for the short-form).

The first step in the factor analysis was to determine the (empirical) optimal number of factors to extract, which was done using a combination of the minimum average partial method (Velicer, 1976), parallel analysis (Horn, 1965) with Glorfeld correction (Glorfeld, 1995), and subjective evaluation of the scree plot (Cattell, 1966). Exploratory factor models were then estimated using least-squares extraction and oblimin rotation.

Unidimensional IRT Graded Response Models (GRMs) were estimated for each (sub-)scale resulting from the factor analysis above, and item fit was examined via the signed chi-square test (Orlando and Thissen, 2000) extended for polytomous items (Kang and Chen, 2008). Following the procedures in Moore et al. (2015), item parameter estimates (discriminances and difficulty thresholds) were then inputted to
Table 1
Description and summary information for the Twelve Original Scales.

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Name of Scale</th>
<th>Abbreviation</th>
<th>Domain</th>
<th>Age range</th>
<th># of items</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effortful control</td>
<td>Early Adolescent Temperament Questionnaire – Effortful Control Scale¹</td>
<td>EATQ</td>
<td>Intrapersonal</td>
<td>8+</td>
<td>16</td>
<td>3.46 (0.59)</td>
</tr>
<tr>
<td>Emotion regulation</td>
<td>Difficulties in Emotion Regulation Scale²</td>
<td>DERS</td>
<td>Intrapersonal</td>
<td>8+</td>
<td>18</td>
<td>2.11 (0.73)</td>
</tr>
<tr>
<td>Positive/negative affect</td>
<td>Positive and Negative Affect Schedule for Children³</td>
<td>PANAS</td>
<td>Intrapersonal</td>
<td>8+</td>
<td>5/5</td>
<td>2.32 (1.07)</td>
</tr>
<tr>
<td>Resilience</td>
<td>The Child &amp; Youth Resilience Measure⁴</td>
<td>CYRM</td>
<td>Intrapersonal</td>
<td>8+</td>
<td>12</td>
<td>2.54 (0.40)</td>
</tr>
<tr>
<td>Resilience Scale</td>
<td>Resilience Scale¹</td>
<td>RS</td>
<td>Intrapersonal</td>
<td>16+</td>
<td>25</td>
<td>5.28 (3.11)</td>
</tr>
<tr>
<td>Positive/negative relationship</td>
<td>Network of Relationships Inventory⁶</td>
<td>NRI</td>
<td>Intrapersonal</td>
<td>8+</td>
<td>24/26</td>
<td>3.72 (0.73)</td>
</tr>
<tr>
<td>Peer victimization/bullying</td>
<td>Multidimensional Peer Victimization Scale¹</td>
<td>MPVS</td>
<td>Intrapersonal</td>
<td>8+</td>
<td>21</td>
<td>0.26 (0.36)</td>
</tr>
<tr>
<td>Adverse experiences</td>
<td>Adverse Life Experiences Scale⁷</td>
<td>ALES</td>
<td>Broader context</td>
<td>8+</td>
<td>25</td>
<td>0.17 (0.13)</td>
</tr>
<tr>
<td>Racial discrimination</td>
<td>Adolescent Discrimination Distress Index⁹</td>
<td>ADDI</td>
<td>Broader context</td>
<td>11+</td>
<td>15</td>
<td>0.32 (0.49)</td>
</tr>
<tr>
<td>Discrimination</td>
<td>Everyday Discrimination Scale¹</td>
<td>EDS</td>
<td>Broader context</td>
<td>11+</td>
<td>10</td>
<td>1.54 (0.66)</td>
</tr>
<tr>
<td>Neighborhood</td>
<td>Neighborhood Safety and Crime¹</td>
<td>NSC</td>
<td>Broader context</td>
<td>11+</td>
<td>3</td>
<td>2.58 (1.16)</td>
</tr>
<tr>
<td>Neighborhood</td>
<td>Neighborhood Community Cohesion¹</td>
<td>NCC</td>
<td>Broader context</td>
<td>11+</td>
<td>5</td>
<td>2.78 (0.79)</td>
</tr>
</tbody>
</table>


Firestar (Choi, 2009) to simulate CAT sessions—i.e. simulate what would have happened if the scale items had been administered adaptively. The frequencies with which items were used in the simulations were then used to select the best (most informative) items for the short-form.

With the short-forms selected, the next steps were: (1) compare group differences found on the full original scales to group differences found using the short-forms, and (2) compare correlations of full original scales with clinical validity scales and criteria to the same correlations for the short-form. The groups used for #1 were based on demographics (age, sex, and race); the validity criteria used for #2 were the self-report clinical scales (PHQ-9, PDS, SCARED, PSA, and the PRIME) and clinical assessments of function (GAF, GF: Social, and GF: Role). Note that, because the primary goal of these analyses was to compare test forms rather than make substantive scientific claims about the causes/correlates of risk and resilience, corrections for multiple comparisons were not applied. If applied, the correction would be larger for the fourteen full-forms than for the seven short-forms, which would be inappropriate because the comparison here is between two batteries meant to be administered in their entirety. Analyses assess convergent validity of the short-form factors are reported in Supplemental Materials. Differences in short-form scales across diagnostic categories were also examined and reported in the Supplemental Materials.

3. Results

3.1. Factor analysis and item calibration

The minimum average partial and parallel analysis methods suggested 13 and 27 factors, respectively—both clear over-extractions confirmed by the fact that their rotated solutions (not shown) include factors comprising very narrow item content (e.g. attitude toward one specific family member). Subjective evaluation of the scree plot—visually determining the point at which successively plotted eigenvalues begin to form a linear trend—suggested 10 factors. We thus first settled on the 10-factor solution as empirically optimal, and Supplementary Table S1 shows the results. The ten factors could be given these conceptual labels, respectively: trait resilience, peer victimization, difficulties in emotion regulation, positive family relationships, positive relational and community resources factors distress captured by the Child and Youth Resilience Measure (CYRM), discrimination-related distress as measured by the Adolescent Discrimination Distress Index (ADDI), life events captured by the Adverse Life Experiences Scale (ALES), negative friend/sibling relationships, negative parent relationships, and negative neighborhood-level characteristics.

The ten-factor solution was problematic for our purposes in two ways. First, many of the factors comprise items from only a single scale, meaning such factors do not contribute much new information beyond the fact that those items correlate well within their own scale. Second (and relatedly), given the end goal of creating a small number of short-forms (one per factor), ten factors are simply too many. Thus, our decision regarding the number of factors to extract was heavily influenced by the practical problem of measuring as many constructs as possible as quickly as possible without being redundant. The ten short-forms that would come from the ten-factor solution shown in Supplementary Table S1 would cover overlapping (redundant) phenomena. Extracting successively larger numbers of factors, we found that the two-factor solution comprised negative social relationships (NSR) and Self-Reliance; the three-factor solution comprised NSR, Self-Reliance, and positive
Exploratory factor analysis solution of the 212 risk and resilience items.

<table>
<thead>
<tr>
<th>Item</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
<th>F5</th>
<th>F6</th>
<th>F7</th>
</tr>
</thead>
<tbody>
<tr>
<td>RS item 23</td>
<td>0.77</td>
<td>0.04</td>
<td>−0.14</td>
<td>−0.03</td>
<td>−0.03</td>
<td>0.13</td>
<td>−0.09</td>
</tr>
<tr>
<td>RS item 10</td>
<td>0.75</td>
<td>−0.01</td>
<td>0.02</td>
<td>0.04</td>
<td>0.04</td>
<td>0.10</td>
<td>−0.08</td>
</tr>
<tr>
<td>RS item 2</td>
<td>0.75</td>
<td>0.17</td>
<td>0.00</td>
<td>−0.04</td>
<td>−0.07</td>
<td>0.08</td>
<td>−0.22</td>
</tr>
<tr>
<td>RS item 17</td>
<td>0.73</td>
<td>0.03</td>
<td>−0.18</td>
<td>−0.06</td>
<td>−0.01</td>
<td>0.12</td>
<td>0.00</td>
</tr>
<tr>
<td>RS item 13</td>
<td>0.73</td>
<td>−0.02</td>
<td>0.09</td>
<td>−0.11</td>
<td>0.07</td>
<td>0.22</td>
<td>−0.02</td>
</tr>
<tr>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
</tr>
<tr>
<td>MPVS item 13</td>
<td>0.05</td>
<td>0.84</td>
<td>0.11</td>
<td>−0.05</td>
<td>−0.14</td>
<td>0.02</td>
<td>0.13</td>
</tr>
<tr>
<td>MPVS item 3</td>
<td>−0.03</td>
<td>0.83</td>
<td>0.00</td>
<td>0.05</td>
<td>0.04</td>
<td>0.00</td>
<td>0.13</td>
</tr>
<tr>
<td>MPVS item 11</td>
<td>−0.03</td>
<td>0.83</td>
<td>0.03</td>
<td>0.08</td>
<td>0.09</td>
<td>−1.12</td>
<td>0.02</td>
</tr>
<tr>
<td>MPVS item 4</td>
<td>−0.02</td>
<td>0.81</td>
<td>−0.06</td>
<td>−0.03</td>
<td>0.08</td>
<td>−0.03</td>
<td>0.15</td>
</tr>
<tr>
<td>MPVS item 14</td>
<td>0.00</td>
<td>0.76</td>
<td>−0.06</td>
<td>0.07</td>
<td>0.05</td>
<td>0.08</td>
<td>−0.02</td>
</tr>
<tr>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
</tr>
<tr>
<td>DERS item 13</td>
<td>0.06</td>
<td>0.03</td>
<td>0.74</td>
<td>0.02</td>
<td>−0.05</td>
<td>−0.09</td>
<td>0.15</td>
</tr>
<tr>
<td>DERS item 11</td>
<td>−0.04</td>
<td>0.08</td>
<td>0.71</td>
<td>−0.01</td>
<td>−0.01</td>
<td>−0.07</td>
<td>0.06</td>
</tr>
<tr>
<td>DERS item 10</td>
<td>−0.14</td>
<td>−0.05</td>
<td>0.68</td>
<td>−0.02</td>
<td>0.14</td>
<td>0.08</td>
<td>−0.09</td>
</tr>
<tr>
<td>DERS item 16</td>
<td>−0.15</td>
<td>−0.03</td>
<td>0.66</td>
<td>0.07</td>
<td>0.04</td>
<td>0.09</td>
<td>0.00</td>
</tr>
<tr>
<td>DERS item 8</td>
<td>−0.01</td>
<td>0.13</td>
<td>0.65</td>
<td>0.00</td>
<td>−0.03</td>
<td>−0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
</tr>
<tr>
<td>NRI item on admiration with Parent 1</td>
<td>0.08</td>
<td>0.03</td>
<td>0.04</td>
<td>0.74</td>
<td>−0.11</td>
<td>0.02</td>
<td>0.05</td>
</tr>
<tr>
<td>NRI item on lasting relationship with Sibling</td>
<td>0.05</td>
<td>−0.06</td>
<td>0.04</td>
<td>0.70</td>
<td>−0.02</td>
<td>0.13</td>
<td>−0.11</td>
</tr>
<tr>
<td>NRI item on caring relationship with Sibling</td>
<td>0.12</td>
<td>−0.10</td>
<td>−0.02</td>
<td>0.66</td>
<td>0.08</td>
<td>0.05</td>
<td>−0.06</td>
</tr>
<tr>
<td>NRI item on admiration with Sibling</td>
<td>0.04</td>
<td>−0.05</td>
<td>0.03</td>
<td>0.66</td>
<td>−0.05</td>
<td>0.18</td>
<td>−0.17</td>
</tr>
<tr>
<td>NRI item on lasting relationship with Parent1</td>
<td>0.08</td>
<td>0.14</td>
<td>0.10</td>
<td>0.65</td>
<td>−0.08</td>
<td>0.00</td>
<td>−0.11</td>
</tr>
<tr>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
</tr>
<tr>
<td>NRI item on relationship nerves with Parent2</td>
<td>0.05</td>
<td>0.02</td>
<td>0.04</td>
<td>0.07</td>
<td>0.67</td>
<td>0.02</td>
<td>−0.06</td>
</tr>
<tr>
<td>NRI item on arguments with Parent2</td>
<td>−0.05</td>
<td>−0.05</td>
<td>0.15</td>
<td>0.06</td>
<td>0.64</td>
<td>−0.03</td>
<td>−0.11</td>
</tr>
<tr>
<td>NRI disagree_Parent2</td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
<td>−0.03</td>
<td>0.63</td>
<td>0.08</td>
<td>−0.09</td>
</tr>
<tr>
<td>NRI item on relationship nerves with Parent1</td>
<td>0.02</td>
<td>0.07</td>
<td>0.00</td>
<td>−0.10</td>
<td>0.62</td>
<td>0.02</td>
<td>0.07</td>
</tr>
<tr>
<td>NRI item on hassles with Parent2</td>
<td>−0.01</td>
<td>0.09</td>
<td>0.06</td>
<td>0.00</td>
<td>0.62</td>
<td>−0.14</td>
<td>−0.04</td>
</tr>
<tr>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
</tr>
<tr>
<td>ADDI item 14</td>
<td>0.13</td>
<td>−0.18</td>
<td>0.10</td>
<td>0.12</td>
<td>−0.14</td>
<td>0.71</td>
<td>0.10</td>
</tr>
<tr>
<td>ADDI item 8</td>
<td>0.12</td>
<td>0.03</td>
<td>0.09</td>
<td>−0.04</td>
<td>0.02</td>
<td>0.67</td>
<td>−0.05</td>
</tr>
<tr>
<td>ADDI item 7</td>
<td>0.11</td>
<td>0.32</td>
<td>−0.02</td>
<td>−0.13</td>
<td>−0.03</td>
<td>0.55</td>
<td>0.00</td>
</tr>
<tr>
<td>ADDI13</td>
<td>0.01</td>
<td>0.14</td>
<td>0.19</td>
<td>0.14</td>
<td>0.07</td>
<td>0.53</td>
<td>0.17</td>
</tr>
<tr>
<td>NSC item 3</td>
<td>−0.04</td>
<td>−0.11</td>
<td>−0.06</td>
<td>−0.03</td>
<td>0.17</td>
<td>0.51</td>
<td>−0.05</td>
</tr>
<tr>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
</tr>
<tr>
<td>ALES item 13</td>
<td>0.02</td>
<td>0.17</td>
<td>0.01</td>
<td>−0.07</td>
<td>0.03</td>
<td>−0.01</td>
<td>0.73</td>
</tr>
<tr>
<td>ALES item 17</td>
<td>0.07</td>
<td>−0.14</td>
<td>0.12</td>
<td>−0.04</td>
<td>0.17</td>
<td>0.05</td>
<td>0.59</td>
</tr>
<tr>
<td>ALES item 16</td>
<td>0.09</td>
<td>−0.01</td>
<td>0.15</td>
<td>−0.09</td>
<td>0.02</td>
<td>0.18</td>
<td>0.59</td>
</tr>
<tr>
<td>ALES item 15</td>
<td>0.11</td>
<td>0.09</td>
<td>0.06</td>
<td>0.14</td>
<td>0.08</td>
<td>−0.03</td>
<td>0.57</td>
</tr>
<tr>
<td>ALES item 14</td>
<td>−0.06</td>
<td>−0.05</td>
<td>0.21</td>
<td>0.09</td>
<td>0.12</td>
<td>0.21</td>
<td>0.52</td>
</tr>
<tr>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
</tr>
</tbody>
</table>

Note. Factor extraction method = least squares; rotation = oblimin; inter-factor correlations not shown; highest absolute loading for each item is bolded; items shown are only the top five highest loading items on that factor; F = factor; rs = Resilience Scale; mpvs = Multidimensional Peer Victimization Scale; DERS = Difficulties in Emotion Regulation Scale; NRI = Network of Relationships Inventory; ADDI = Adolescent Discrimination Distress Index; ALES = Adverse Life Experiences Scale; NSC = Neighborhood Safety and Crime.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Full factor</th>
<th>High precision (α = 0.91)</th>
<th>Moderate precision (α = 0.80)</th>
<th>Minimum acceptable precision (α = 0.70)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1</td>
<td>49</td>
<td>7.3</td>
<td>3.0</td>
<td>2.4</td>
</tr>
<tr>
<td>Factor 2</td>
<td>28</td>
<td>19.7</td>
<td>14.2</td>
<td>10.8</td>
</tr>
<tr>
<td>Factor 3</td>
<td>35</td>
<td>13.6</td>
<td>4.8</td>
<td>2.5</td>
</tr>
<tr>
<td>Factor 4</td>
<td>30</td>
<td>12.5</td>
<td>4.2</td>
<td>2.8</td>
</tr>
<tr>
<td>Factor 5</td>
<td>27</td>
<td>11.8</td>
<td>4.5</td>
<td>2.8</td>
</tr>
<tr>
<td>Factor 6</td>
<td>31</td>
<td>12.1</td>
<td>3.6</td>
<td>2.0</td>
</tr>
<tr>
<td>Factor 7</td>
<td>12</td>
<td>12.0</td>
<td>11.4</td>
<td>10.6</td>
</tr>
<tr>
<td>Total</td>
<td>212</td>
<td>89.0</td>
<td>45.7</td>
<td>33.9</td>
</tr>
</tbody>
</table>

Table 2 shows abbreviated results of the 7-factor exploratory factor analysis, and Supplementary Table S2 shows the full results. Factor 1 comprises items from the Resiliency scale, CYRM, Positive and Negative Affect Schedule for Children (PANAS), Difficulties in Emotion Regulation Scale (DERS), Early Adolescent Temperament Questionnaire – Effortful Control Scale (EATQ), and Network of Relationships Inventory (NRI), with the 19 highest-loading items all from the Resiliency scale. Factor 1 captures Self-Reliance. Factor 2 comprises items from the Multidimensional Peer Victimization Scale (MPVS), ADDI, Everyday Discrimination Scale (EDS), and ALES, with the 17 highest-loading items all from the MPVS. Factor 2 captures peer-victimization. Factor 3, which captures difficulties with emotion regulation, comprises items from the DERS, PANAS, EATQ, EDS, and ALES, with the 13 highest-loading items all from the DERS. Factor 4 comprises items from the NRI, split into two separate factors; and the six-factor solution comprised the same factors as the five-factor solution, except that negative family relationships and negative environment (i.e., neighborhoods/discrimination) split into two separate factors. Finally, the seven-factor solution (used here) was chosen because it comprised the same factors as the six-factor solution, plus an important seventh factor capturing stressful life events experienced by the individual (as opposed to negative characteristics of the persons themselves).

Social relationships; the four-factor solution comprised Self-Reliance, positive family relationships, peer-victimization/difficulties with emotion regulation, and general negative environment (i.e., family relationships/neighborhoods/discrimination/life stressors); the five-factor solution comprised the same four factors as the four-factor solution, except that victimization and difficulties in emotion-regulation
Table 4
Items composing the risk and resilience short forms.

<table>
<thead>
<tr>
<th>Abbreviated short form</th>
<th>47 items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale</td>
<td>Items</td>
</tr>
</tbody>
</table>

Factor 1: self-reliance
RS When I’m in a difficult situation, I can usually find my way out of it.
RS I am determined.
RS My belief in myself gets me through hard times.

Factor 2: peer victimization
MPVS Called me names
MPVS Tried to make my friends turn against me
MPVS Made fun of me for some reason
MPVS Took something of mine without permission
MPVS Swore at me
MPVS Refused to talk to me
EDSE You are called names or insulted.
MPVS Made fun of me because of my appearance
MPVS Tried to get me into trouble with my friends
EDSE People act as if they’re better than you.
MPVS Hurt me physically in some way
MPVS Sent me a nasty text
MPVS Said something mean about me on a social networking site
MPVS Stole something from me

Factor 3: emotion dysregulation
DERS When I’m upset, I have difficulty controlling my behaviors
DERS When I’m upset, I have difficulty focusing on other things
DERS When I’m upset, I have difficulty concentrating
DERS When I’m upset, I believe that I will end up feeling very depressed
DERS When I’m upset, I have difficulty getting work done
DERS When I’m upset, I have difficulty controlling my behaviors

Factor 4: positive relationships
NRI How much do you and [Parental Fig. 1] treat you like you’re admired and respected?
NRI How sure are you that this [Parental Fig. 1] relationship will last no matter what?
NRI How sure are you that this [Sibling] relationship will last no matter what?
NRI How much does [Sibling] really care about you?

Factor 5: negative relationships
NRI How much do you and [Parental Fig. 1] get annoyed with each other’s behavior?
NRI How much do you and [Parental Fig. 1] disagree and quarrel?
NRI How much do you and [Parental Fig. 1] hassle or nag one another?
NRI How much do you and [Parental Fig. 1] get on each other’s nerves?
NRI How much do you and [Parental Fig. 2] get on each other’s nerves?

Factor 6: neighborhood danger
NSC My neighborhood is safe from crime.
NSC Violence is not a problem in my neighborhood.
NSC I feel safe walking in my neighborhood, day or night.
NCC People in my neighborhood can be trusted.

Factor 7: stressful events
ALES One of the parents went to jail
ALES Family moved
ALES Parents got into trouble with the law
ALES Parents got divorced
ALES One parent was away from home more often
ALES Someone in the family was arrested
ALES Got new stepmother or father
ALES Parent got a new job
ALES Attended a new school
ALES Get seriously sick or injured
ALES Got new brother or sister
ADDI You were given a lower grade than you deserved.

ADDI = Adolescent Discrimination Distress Index; ALES = Adverse Life Experiences Scale; NSC = Neighborhood Safety and Crime; *the originator of the Resilience Scale (Wagnild and Young, 1993; Wagnild, personal communication, November-December 2019) does not condone the use of the Resilience Scale in any form except the full, 25-item form.

CYRM, and ALES, with the 20 highest-loading items all positively-valenced items (reflecting positive states or outcomes) from the NRI. Factor 4 captures positive relationships with family and friends (especially family). Factor 5, which captures negative relationships with family and friends (especially family), comprises items from the NRI and ALES, with the 24 highest-loading items all negatively-valenced items from the NRI. Factor 6 compromises items from the ADDI, Neighborhood Safety and Crime (NSC), and Neighborhood Community Cohesion (NCC), CYRM, ALES, and EDS, with the top 12 items being ADDI, the neighborhood scales (NSC and NCC), and one negatively-loaded item from the CYRM. Factor 6 captures negative environments (high discrimination and crime). Factor 7 compromises all ALES items, except for one ADDI item. Factor 7 captures stressful life events, primarily concerning family. Raw Cronbach’s alpha for factors 1–7 were 0.95, 0.90, 0.93, 0.92, 0.91, 0.83, and 0.68, respectively. Cronbach’s alpha based on polychoric correlation matrices (more appropriate here) were 0.96, 0.95, 0.95, 0.94, 0.93, 0.90, and 0.85.

Correlations among the factors (not shown) were weak (maximum 0.22 between factors 2 and 5). The low inter-factor correlations suggest that the factors are conceptually distinct, meaning a model that included a factor explaining variance across all items (bifactor model) would be inappropriate.

With the seven factors above established, the items in each factor were calibrated in seven separate IRT GRMs. Supplementary Table S3 shows the item parameter estimates, and Supplementary Table S4 shows the item fit statistics (p-values corrected within-sub-scale using false detection rate; Benjamini and Hochberg, 1995). All items achieve acceptable fit except for three items (asking about non-family friends) in the “Positive Relationships” factor. None of these three items was selected for the final abbreviated scales.

3.2. CAT simulation

Table 3 shows the number of items necessary (47 in total) to achieve three levels of measurement precision on each of the seven factors (new scales; see Table 4 for list of items). By far, the factor that can be abbreviated the most is Factor 1 (Self-Reliance), which can be reduced by 49.0–7.3 = 41.7 items (85%) and still retain high precision. Factor 7 (Stressful Life Events) showed the opposite: even if we are willing to accept the minimum precision in Table 3, Factor 7 can be shortened by only 1.4 items (12%). The final full-battery (total item) lengths for high, medium, and minimum precision were 89.0, 45.7, and 33.9, respectively.

3.3. Validity analyses

Fig. 1 shows the results of group comparisons on the fourteen full-form scales originally selected for the Risk and Resilience battery. For sex (top graph), the scales detected significant differences in (1) perceived neighborhood safety/cohesion (NSC & NCC; worse in females), and (2) everyday discrimination (EDS; worse in females). For race (middle graph), the scales detected significant differences in (1) perceived neighborhood safety/cohesion (NSC & NCC; much better in White race), (2) difficulties in emotion regulation (DERS; worse in White race), (3) some aspects of Self-Reliance (CYRM; higher in “white” race), and (4) discrimination distress (ADDI; much lower in White race). For age (bottom graph), the scales detected significant differences in (1) difficulties in emotion regulation (DERS; worse in children), (2) peer victimization (MPVS; higher in children), (3) adverse life experiences in last year (ALES; higher in children), and (4) negative
Sex
- Male
- Female

Race
- White
- Black
- Other

Age Group
- 8 - 17 years
- 18+ years

(caption on next page)
relationships (NRI negative; more negative relationships in children).

Supplementary Fig. S1 shows the results of sex comparisons on the seven short-form scales obtained from CAT-simulation. Consistent with Fig. 1, females show significantly more perceived neighborhood danger (and lower cohesion). The significant sex difference in everyday discrimination seen in Fig. 1 was not seen when using the short-forms, but note that the short-form comprising discrimination-related phenomena (“Victimization”) shows a sex difference approaching significance.

Fig. 2 shows the results of race comparisons on the seven short-form scales. The significant associations with emotion regulation (DERS), neighborhood danger (NCC and NSC), and discrimination (ADDI) from Fig. 1 are captured in Fig. 2 by the “Emotion Dysregulation”, “Neighborhood Danger”, and “Stressful Events” factors, respectively. The significant race difference on the CYRM (sub-set of Self-Reliance) in Fig. 1 was not apparent in any factor in Fig. 2.

Fig. 3 shows the results of age group comparisons on the seven short-form scales. The significant age group differences in Fig. 1 for emotion dysregulation (DERS), adverse life events (ALES), and peer victimization (MPVS) are captured in Fig. 3 by the “Emotion Dysregulation”, “Victimization”, and “Stressful Events” short-forms, respectively. The significantly worse relationships (NRI) for younger participants seen in Fig. 1 was not seen in Fig. 3. Conversely, one of the associations detected by the short-forms (higher Self-Reliance among older participants) was not detected when using the individual scales.

Finally, we tested the relationship of the risk and resilience measures with clinical scales that reflect level of function and levels of self-reported depression, anxiety, and psychosis spectrum symptoms. Supplementary Fig. S2 shows the Pearson correlation coefficients for the relationships of full- and short-form scales with clinical validity criteria (symptoms and function). The cutoff for statistical significance at this sample size (familywise error uncorrected) is ± 0.12, meaning most correlations in Figure S2 are significant. As expected, both the strongest (DERS) and weakest (ADDI for function ratings; NSC for clinical symptoms) relationships are seen in the full-forms, with the short-forms showing effect sizes mostly in the middle. Four measured constructs show no noticeable difference between the full- and short-forms: Positive Family (full “NRI Pos” compared to short “Positive Family”), Negative Family (full “NRI Neg” compared to short “Negative Family”), Victimization (full “MPVS” compared to short “Victimization”), and Neighborhood Danger (full “NSS” and “NSC” compared to short “Neighborhood Danger”). Stressful Events (full “ALES” compared to short “Stressful Events”) shows differences (lower magnitude) only in the Cornblatt scales. Finally, Self-Reliance and Emotion Dysregulation do show noticeable differences between full- and short-forms, in the range expected by the substantial decrease in items.

4. Discussion

As the field of developmental neuropsychiatry evolves, it is recognized that granular characterization of the individual’s intrapersonal and environmental phenotypes is essential to understanding the biological mechanisms that underlie risk or resilience to develop serious psychiatric conditions (Cathomas et al., 2019). Therefore, it is critical to
include thorough measurement of environmental and dispositional risk/resilience factors in studies that investigate brain and behavior (Southwick and Charney, 2012). A major challenge for research in developmental psychopathology is the need to distinguish between factors that confer risk (or protection), such as familial conflicts vs. school bullying. One child might be relentlessly bullied at school but come home to a supportive family, while another child might experience no problems at school but come home to an abusive or neglectful family; otherwise, we are not in a position to know which of these two children is at higher risk for mental illness. Perhaps the first compelling example of this ambiguity was the work of Sameroff et al. (1987), who found that multiple social-familial-environmental factors predicted childhood IQ, yet dozens of combinations of those factors produced the same predictions—i.e. there was no specific aggregation of risk factors that predicted IQ.

Further complicating the research—arguably the motivation behind the research—is that children differ in how they respond to adverse events and environments, meaning even if we had a thorough understanding of the relative harms of bullying and family stress, we would still be able to make only rough predictions about which of the two children in the above example will suffer more long-term harm. Thus, to understand the phenomena of risk and resilience, measurement tools need to capture both broad information about the environment and relevant information about the individual. The problem with broad measurement is that inter-item correlations will tend to be lower (compared to a narrow construct) (Crocker and Algina, 1986), requiring many more items to achieve acceptable measurement precision. However, thorough, lengthy measurement is simply not compatible with the current era of large-scale genomic and international population studies. The purpose of the present study was to develop a battery of risk and resilience measures that are thorough enough to account for the complexity of social-environmental phenomena, yet brief enough to be used in large samples, where risk and resilience might not be the primary subject of study.

We found that the twelve Risk & Resilience scales (212 items total) could be adequately summarized by seven factors, and that the seven resulting sub-scales could be abbreviated substantially (47 items; 22% of total). A series of validation analyses revealed that the individual scales and, more importantly, the short-from factor scores were significantly related to a series of clinical criterion (clinical function ratings, self-report symptoms, and clinical diagnoses) in both healthy and patient populations. Of interest to future research, the magnitude of relations between the risk and resilience factors and the clinical outcomes varied, suggesting that certain areas of risk/reliance might differentially protect or increase risk for certain maladaptive outcomes. Although there are no studies to which to compare the present study directly, our approach and results are consistent with previous research. While some studies focus exclusively on the external (environmental) risk factors (Kipke et al., 2007; Whipple et al., 2010), most psychiatric and psychological studies involving assessment of risk (Nikulina et al., 2011; Brown et al., 1998; Frissen et al., 2015; Dubowitz et al., 2002; Dupéré et al., 2007; Kupersmidt et al., 1995; Hill et al., 1999; Singh et al., 2010; Sameroff et al., 1987; Bifulco et al., 1994), (1) draw the distinction between neighborhood-level and family-level phenomena, and (2) acknowledge the importance of interactions among these factors and individual-level factors. Indeed, all studies cited above found differential effects of neighborhood and familial risk factors.

Despite its strengths, this study has some notable limitations. First, the sample (N = 292) was relatively small given the complexity of the analyses. However, sample size recommendations for factor analysis vary widely, from as few as 50 (Barrett and Kline, 1981) to as many as 400 (Aleamoni, 1976). Using the standards of Comrey and Lee (2013), our N is between “Fair” (N = 200) and “Good” (N = 300); nonetheless, future work using larger samples with similar batteries are needed to investigate the generalizability of the present findings. Second, we did not explore item bias (differential item functioning; DIF), which occurs when groups of interest (e.g. race, sex, etc.) do not have equal probabilities of endorsement even when holding overall trait level constant.
We did not have the sample size necessary to explore DIF (see Zwick, 2012 for review), but because the constructs explored here are heavily influenced by variables such as race and sex, it is critical that future studies explore DIF on these scales. Thirdly, some scales used in the present battery were not originally designed for use in a wide age range. Supplemental analyses found scales intended for younger groups to be reliable in youth and adults. Younger participants were not given the scales designed for older participants, resulting in missing data for some items. Finally, the recommended short-forms assessing family relationships include some items that ask about siblings, and the study participant might not have any siblings (missing data). Measurement error will therefore be slightly lower (better) for individuals who do have at least one sibling. Despite these limitations, the current study presents a valid, brief way to assess a wide range of risk and resilience factors across the lifespan. Further testing and validation of this assessment battery will help move the field of developmental psychopathology forward, especially in light of growing effort to characterize the biological substrates of risk and resilience (Cathomas et al., 2019).

Declaration of Competing Interest
None.

Supplementary materials

References
Cron, J., Algina, J., 1986. Introduction to classical and modern test theory. ERIC.

McGlashan, T.H., Miller, T.J., Woods, S.W., Rosen, J.L., Hoffman, R.E., Davidson, L., 2001. Structured Interview For Prodromal Syndromes. PRIME Research Clinic, Yale School of Medicine, New Haven, CT.


