University of Reading

The risk-taking and self-harm inventory for adolescents: Development and psychometric evaluation

Article

Accepted Version

Vrouva, I., Fonagy, P., Fearon, P. R. M. and Roussow, T. (2010) The risk-taking and self-harm inventory for adolescents: Development and psychometric evaluation. Psychological Assessment: A Journal of Consulting and Clinical Psychology, 22 (4). pp. 852-865. ISSN 1040-3590 doi: 10.1037/a0020583 Available at https://centaur.reading.ac.uk/17507/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.
Published version at: http://dx.doi.org/10.1037/a0020583
To link to this article DOI: http://dx.doi.org/10.1037/a0020583

Publisher: American Psychological Association

Publisher statement: 'This article may not exactly replicate the final version published in the APA journal. It is not the copy of record.'

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

University of Reading

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

The Risk-Taking and Self-Harm Inventory for Adolescents (RTSHIA): Development and Psychometric Evaluation

Ioanna Vrouva, University College London

Peter Fonagy, University College London

Pasco Fearon, University of Reading

Trudie Roddouw, North East London Foundation NHS Trust

This study reports on the development and psychometric evaluation of the Risk-Taking (RT) and Self-Harm (SH) Inventory for Adolescents (RTSHIA), a self-report measure designed to assess adolescent RT and SH in community and clinical settings. 651 young people from secondary schools in England ranging in age from 11.6 up to 18.7 years and 71 young people referred to mental health services for SH behavior in London between the ages of 11.9 and 17.5 completed the RTSHIA along with standardized measures of adolescent psychopathology. Two factors emerged from the principal axis factoring and RT and SH were further validated by a confirmatory factor analysis as related, but different constructs, rather than elements of a single continuum. Inter-item and test-retest reliability were high for both components (Cronbach's alpha = .85 and .93, r_{tt} = .90 and .87) and considerable evidence emerged in support of the measure's convergent, concurrent and divergent validity. The findings are discussed with regard to potential utility of the RTSHIA for research and clinical purposes with adolescents.

Keywords: risk-taking, self-harm, adolescence, screening measure, psychometric evaluation

Risk-taking (RT) and self-harm (SH) occur across the lifespan, but many types of RT and SH typically appear for the first time in adolescence (Boyer, 2006; Yates, 2004). There is no consensus about what constitutes RT and SH. Most conceptualizations of RT, including the one adopted in this study, refer to 'engagement in behaviors that are associated with some probability of undesirable results' (Boyer, 2006, p. 291). These behaviors are closely related to personality traits such as sensation seeking (Zuckerman, 1979) and to poor judgment due to adolescents' psychosocial immaturity (Steinberg, 2007).

SH covers a wide spectrum of behaviors, furthermore the intentionality, destructiveness and conventionality of thoughts and acts of SH sit on a broad continuum (Turp, 2003). One of the most debated issues is whether SH is an expression of suicidality or a means of coping (Simeon & Favazza, 2001). Consequently, some researchers use definitions that exclude behaviors with conscious suicidal intent, and use terms such as 'non-suicidal self-injury' (NSSI) to describe intentional, direct damage to one's body tissue without suicidal intent (Pattison & Kahan, 1983). Other researchers use more inclusive definitions that encompass a wide range of behaviors, including both suicidal acts and different types of NSSI such as cutting or burning, and consider it important to assess for suicidal behaviors in studies on SH (Hickey, Hawton, Fagg, & Weitzel, 2001). This is because suicidal intent may not always be reliably measured (Fliege et al., 2006), and because NSSI is one of the strongest predictors of suicidal behaviors (Nock, Joiner, Gordon, Lloyd- Richardson, & Prinstein, 2006; Zahl & Hawton, 2004). In this paper we use the term SH in its broader meaning to indicate culturally unacceptable behavior that involves direct and deliberate infliction of physical harm to one's body, regardless of the presence of suicidal intent and in the absence of a pervasive developmental disorder (Simeon & Hollander, 2001; Walsh, 2008; Yates, 2004).

In our opinion, it is important to measure RT and SH simultaneously, because they are clinically, empirically and conceptually linked. They both derive from the young person's wish to momentarily experience something that is subjectively perceived to be essential and desirable, regardless of consequences. Both behaviors are subject to peer influence (Steinberg, 2008). Both seem to be a manifestation of developmental vulnerability, as they frequently emerge in adolescence (Boyer, 2006; Yates, 2004) and happen in an impulsive state of mind (Steinberg, 2007; Webb, 2002) in which cognitive control and reflectivity are weak, emotional and stress reactivity are high and decision making and consequence appraisal are poor. Both have been associated with a state of increased arousal (Howard, Yan, Ling, & Min, 2002; Gratz, 2003) that can expose the individual to various degrees of risk, leading subsequently to various degrees of harm. Furthermore, there is substantial comorbidity as many types of adolescent RT are linked to psychiatric disorders, such as substance abuse and conduct disorders (APA, 2000), which are also associated with raised rates of SH (Evans, Hawton, & Rodham, 2004). Adolescents with psychiatric disorders, in particular, frequently engage in multiple risk behaviors, including SH (Lescano et al., 2007). Moreover, RT and SH are defining characteristics of a well-established clinical syndrome, borderline personality disorder (BPD; APA, 2000).

The multi-factorial nature of SH has been raised by several experts. For instance, Turp (2003) suggests a continuum model, which relates high visibility manifestations of SH to virtually invisible ones, with an array of gradations in-between, which include RT and other more indirect forms of SH (or potentially SH) behavior. Similarly, Claes and Vandereycken (2007) contend that SH may include indirect modes of self-destructive behaviors, such as particular habits of eating, drinking and smoking. Abnormal eating behaviors and negative body image have been significantly associated with SH and suicidality in several studies, therefore several SH measures contain eating disorder (ED) items (Sansone & Sansone, 2002). Likewise, Walsh (2008) regards RT as a form of indirect SH, together with EDs, substance abuse and medication discontinuance, arguing that the self-destructive intent of the individual engaging in indirect SH is

usually ambiguous and at times "more of omission than commission" (p.25). Some forms of SH are less clear than others. However, a recurrent pattern of such behaviors should be viewed in some contexts as potentially life-threatening, hence early identification is vital (Walsh, 2008).

Conversely, there are noticeable differences between RT and SH. Firstly, SH can hardly be seen as typical or adaptive adolescent behavior, whereas RT has been considered normative and adaptive and linked with developmental tasks such as autonomy and exploration (Lavery, Siegel, Cousins, & Rubovits, 1993). Consequently, its position in the developmental normality-psychopathology continuum differs. Secondly, whereas the mental state context of SH is that of emotional distress, RT is linked with a variety of moods, including euphoric ones (Steinberg, 2004). Thirdly, the presence of peers increases RT (Steinberg, 2008), whereas SH usually takes place in solitude. Fourthly, whereas in SH the physical damage is direct and intentional, in RT the physical damage is a likely *consequence* of the behavior but not its original conscious *aim*.

It follows from the above that at their extremes, RT and SH are clearly distinct. However, there is a range of behaviors and attitudes which fall under the banner of SH, and inhabit the grey areas between indifference, risky behavior and explicit intent to harm or even kill oneself. Assessing various forms of SH and organizing information about the whole spectrum with a single instrument has practical and conceptual advantages. Combining these dimensions makes it easier and more convenient for young people to complete. Clinically, there is value in capturing the ways in which adolescents may harm themselves in order to screen for risk, identify comorbidity and obtain a comprehensive picture. Assessment of the wider presentation in the general population of young people is also important, in order to better understand when and how these different constructs are related and also to build the evidence base concerning the wider spectrum of potentially self-destructive behaviors. A common measurement framework would enable different research groups to collect relevant data and advance knowledge more consistently. Statistically, the use of a single measure consisting of items sharing the same response format enables the investigation of the factor structure of several RT and SH behaviors

irrespective of differences in item distribution and item categorization properties that might otherwise influence the quality of the factor solution (Bernstein & Teng, 1989).

Widely-used self-report measures, such as the Adolescent Risk-Taking Scale (Alexander et al., 1990), the Involvement Scale of the Risk-Involvement Ratings (Lavery, et al., 1993), Arnett's (1996) questionnaire focusing on reckless behavior, and the Adolescent Risk-Taking Questionnaire (Gullone, Moore, Moss, & Boyd, 2000) were developed to assess aspects of adolescent RT. However, few reliable and valid self-report measures of SH have been developed specifically for adolescents, despite the limitations of applying adult conceptualizations of SH to adolescent behavior (Pfeffer, Jiang, & Kakuma, 2000).

Several researchers have used one or two items that allow participants to define SH as they understand it (e.g. Hawton, Rodham, Evans, & Weatherall, 2002; Ross & Heath, 2002). Although such operational definitions sample the entire domain of behaviors that participants feel are self-destructive, the use of such measures may not provide sufficient cues for participants to recollect all relevant kinds of SH (Lundh, Karim, & Quilisch, 2007) and has been shown to provide false positives when followed up by an interview (e.g. Ross & Heath, 2002).

The Functional Assessment of Self-Mutilation (FASM; Lloyd, Kelley, & Hope, 1997), the Deliberate Self-Harm Inventory (DSHI; Gratz, 2001; Lundh et al., 2007) and the Suicidal Behaviors Questionnaire-Revised (Osman et al., 2001) have been used in studies of both adults and adolescents. However, the first two assess only non-suicidal SH, whereas the latter focuses on suicidality and does not capture the wider range of behaviors that are self-destructive in the absence of clear suicidal intent, thus limiting the scope of investigated behaviors. Similarly, scales of suicidal tendency or potential in adolescents, e.g. the Multi-Attitude Suicide Tendency Scale for Adolescents (Orbach et al., 1991) and the Child-Adolescent Suicidal Potential Index (CASPI; Pfeffer et al., 2000) were designed to assess attitudes, affects, coping skills and risk propensity rather than to gather detailed behavioral data on SH.

The Self-Injurious Thoughts and Behaviors Interview (SITBI; Nock, Holmberg, Photos, & Michel, 2007) uniformly and comprehensively assesses a broader range of self-injurious thoughts and behaviors than any measures previously reported in the literature (Nock et al., 2007). It assesses the presence, frequency and characteristics of five types of SH constructs: suicidal ideation, suicide plans, suicide gestures, suicide attempts, and NSSI. However, in many research applications, such as in investigations aimed at identifying mechanisms and patterns in large samples, it may not be feasible to administer interview measures.

Few measures conceptualized RT and SH as closely related constructs, thus containing items referring to both behaviors simultaneously. The Self-Injury Inventory (Zlotnick et al., 1996) and the Self-Harm Inventory (SHI; Sansone, Wiederman, & Sansone 1998) were validated with adult clinical samples, and may therefore lack the necessary developmental specificity to measure RT and SH in adolescents. The Adolescent Risk Inventory (Lescano et al., 2007) assesses RT and SH behaviors simultaneously and was developed specifically for use with adolescents. It consists of behavior and attitude items reflecting three domains: sex risk, abuse/SH, and acting out. A shortcoming of this measure is that has only been validated with adolescents in psychiatric treatment, thus its utility as a screening tool for community adolescents is unknown. Moreover, its focus is on risk behavior more generally, resulting in the inclusion of only three items relevant to direct SH (thinking of cutting, cutting and attempting suicide).

Despite the significant implications of early screening, there is a notable dearth of self-report measures of the multifactorial elements of SH for adolescents. Thus, there is a need for a comprehensive tool in order to make quick but valid assessments of relevant behaviors among youth in community and clinical settings. Moreover, we need to learn more about the structure of RT and SH and understand the links among various constructs. For example, behaviors such as unsafe sex, substance abuse and disordered eating, seem to overlap significantly with both RT and SH. In response to this need, we included items that probe more generally into the broader spectrum of RT and SH, rather than limiting our questions to those concerning suicidal behavior

and self-mutilation. The Risk Taking and Self-Harm Inventory for Adolescents (RTSHIA), designed to overcome the aforementioned limitations, is described below, followed by a preliminary investigation of its psychometric properties.

Method

Participants

Data from two studies were used in the present analysis: a survey that took place as part of a doctoral research program between 2006 and 2008 and an ongoing randomized clinical trial (RCT) investigating treatment efficacy for self-harming (SH) adolescents in North East London. In total, 722 young people (651 community adolescents and 71 adolescents referred to outpatient treatment for SH behavior) participated in the study. A sample of 589 adolescents participated in the doctoral survey and were recruited from Grade 6 up to Grade 12 in nine high schools in London. Four of the high schools were girls' schools and the majority of participants (82.7%) were female. Their mean age was 15.3 (SD = 2.1, range 11.6-18.7). A subsample of 101 participants (60.4% female) from the doctoral survey completed the RTSHI twice within a period of three months. Their mean age was 15.2 (SD = 2.2, range 11.6-18).

A sample of 133 adolescents participated in the RCT study and fell into three distinct subsamples: a clinical SH group (n = 53), referred to child and adolescent mental health services (CAMHS) for SH behavior, a clinical non-SH group (n = 18), referred to CAMHS for mental health problems other than SH, and a community sample (n = 62), drawn from high schools. RCT participants were recruited from Grade 7 up to Grade 12 and the majority (67.7%) were female. Their mean age was 15.8 (SD = 1.5, range 11.9-18.4). All participants in the clinical SH group are administered the RTSHIA at 3, 6, 9 and 12 months from intake. Three-month follow-up data were available for 36 participants from the RCT. Of these participants, 27 completed the RTSHIA at both 3 and 6 months and 18 participants provided data at 3, 6, and 9 months.

No participant dropped out of the study. On average, each item was answered by 98.4% of the respondents, and 91.12% of respondents answered all items. Out of the 651 community

adolescents, 59 (9.1%) did not complete all items. Data from these participants were discarded and only data from the remaining 592 community adolescents were used when performing the exploratory factor analyses, as the statistical program Factor does not handle missing values.

Measures

Description of the RTSHIA

In developing the RTSHIA, we created and selected items that tap into RT and SH in both direct and indirect ways. Item 10 about alcohol misuse was selected from a study on adolescent injuries and substance abuse (Chiolero & Schmid, 2002). In items 15-19, we used the wording of items found in the DSHI (Gratz, 2001; Lundh et al., 2007). In items 7, 25 and 26 we used the wording of items found in the SHI (Sansone et al., 1998). Items 17 and 23 were partly taken from the FASM (Lloyd et al., 1997).

The 12 RT-related items ranged from mild behaviors such as smoking tobacco and taking chances while doing one's hobbies, to serious risk-taking, such as participating in gang violence and putting oneself at risk of sexual abuse. The 22 SH-related items ranged again from milder behaviors such as picking at wounds and pulling one's hair out to more serious SH such as taking an overdose and trying to commit suicide. Moreover, we used an item (item 34) developed by Lundh et al. (2007), which enquires about SH leading to hospitalization or to an injury severe enough to require medical treatment.

Most SH items were about self-mutilation (e.g. cutting, burning, biting, scratching one's skin, etc), followed by a question about the part(s) of the body that were deliberately injured, if applicable, as employed in the Self-Injury Questionnaire-Treatment Related (Claes & Vandereycken 2007), developed for adults. Three items were about disordered eating (starving oneself, eating too much and using laxatives), two items about self-demeaning behavior (staying in a friendship or a relationship with somebody who repeatedly hurt one's feelings, and trying to make oneself suffer by thinking horrible things about oneself) and two items were about SH

ideation, with or without suicidal intent. The items were ordered gradually in terms of severity and expected frequency (from the milder and more frequent to the more serious and rarer). Most items contained the word "intentionally" and four items ended with the phrase "to hurt yourself" or "to hurt or punish yourself". This was added in order to indicate specific deliberation, when the behavior did not in and of itself unambiguously imply self-harm (items 20 and 25-27).

The items were on a 4-point Likert scale, referring to life-long history and could be answered by selecting "never", "once", "more than once", or "many times" , as employed by Lundh et al. (2007). There was also a free-response question about other self-destructive behaviors not mentioned in the inventory which participants might have engaged, and another question asking whether participants knew somebody well who had deliberately injured him/herself. At the end, we asked participants to comment briefly on why in their opinion some young people harm themselves. The RTSHIA items are given in the Appendix.

Other measures

The survey and the RCT employed different test batteries that were developed to suit the particular aims of each research project. Both included the RTSHIA, a series of questions designed to gather demographic information, and the following measures:

a) the Short Mood and Feelings Questionnaire (SMFQ; Angold et al., 1995) is a 13-item, 3-point Likert scale designed to assess core depressive symptomatology and to screen youth in epidemiological studies. The SMFQ has been found to correlate well to the Children's Depression Inventory (r = .67; Angold et al., 1995) and to discriminate well between clinical and nonclinical samples (Angold et al., 1995; Costello & Angold, 1988). Cronbach's alpha in our study was .89.

b) the Borderline Personality Features Scale for Children (BPFS-C; Crick, Murray-Close, & Woods, 2005), is a 24-item, 5-point Likert scale designed to assess BPD traits in juveniles. Over the course of a year, children's borderline personality features as assessed by the BPFS-

C tracked together with indicators (i.e., cognitive sensitivity, emotional sensitivity, friend exclusivity, and aggression) of borderline pathology in childhood (Crick et al., 2005). Cronbach's alpha in our study was .88.

c) Some participants in the survey and all participants in the RCT also completed the Millon Adolescent Clinical Inventory (MACI; Millon, 1993), a 160-item True/False inventory designed to assess a wide range of DSM-IV related psychological disturbances. The internal consistency of the MACI subscales ranges from .73 to .91 (Millon, 1993). Moreover, good criterion and concurrent validity has been reported in a study of hospitalized adolescents (Pinto & Grilo, 2004).

Participants in the RCT also completed the Adolescent Dissociative Experiences Scale (A-DES; Armstrong, Putnam, Carlson, Libero, & Smith, 1997), an 11-point Likert scale screening measure for pathological dissociation during adolescence. Data from several studies have yielded Spearman-Brown and Cronbach's alpha reliability coefficients of .90 or higher (Keck-Seeley, Perosa, & Perosa, 2004), and the A-DES has been found to discriminate well between adolescents with dissociative disorders and adolescents with general psychiatric disorders, and between adolescents who reported both physical and sexual abuse and adolescents who reported no abuse (Armstrong et al., 1997). In our study, Cronbach's alpha was .96.

Finally, RCT participants were assessed using the Childhood Interview for DSM-IV Borderline Personality Disorder (CI-BPD; Zanarini, 2007), a semi-structured interview developed to assess DSM-IV-TR (APA, 2000) BPD in latency-age children and adolescents. It consists of nine domains/symptoms, which correspond to the nine DSM-IV-TR diagnostic criteria for BPD and are rated by the interviewer as absent (level 0), mild or moderate (level 1) and serious or severe (level 2). When 5 or more symptoms are rated at level 2, the diagnosis of BPD is warranted. The scale can be used both categorically and dimensionally. In this study, the alpha for the dimensional scale was .90, and the inter-rater agreement (ICC) was .95.

Data Analyses

The statistical programs SPSS V.16, MPlus V. 5.2 (Muthén and Muthén, 1998-2007) and Factor (Lorenzo-Seva & Ferrando, 2006) were used to analyze the data.

Determination of factor structure. In order to examine the factor structure that underpins the RTSHIA, we intended to cross-validate our analyses by conducting both exploratory and confirmatory factor analyses. Therefore the nonclinical sample (survey and RCT community data) was randomly split into two parts, to avoid problems of capitalizing on chance fluctuations in the sample. The data derived from the first subgroup (age: $M = 15.5$, $SD = 2.0$ years) were used for a series of ULS (unweighted least squares) Principal Axis Factoring (PAF, also known as Principal Factors Analyses). Several items' distributions were asymmetric, therefore ULS factor analysis was performed on the polychoric correlation matrix. Oblique rotation (direct oblimin) was used because the factors were assumed to be related to each other rather than independent. Delta was set to 0.

The number of expected factors was not specified prior to running the analysis. In order to determine the right number of factors to retain, we performed a parallel analysis using marginally bootstrapped samples (PA-MBS; Lattin, Carroll, & Green, 2003), considered to be one of the most accurate factor retention techniques (Fabrigar, Wegener, MacCallum, & Strahan, 1999). It involves the generation of several correlation matrices of random variables based on the same distribution, sample size and number of variables in the actual data set, but with a zero population correlation matrix. Parallel analysis with bootstrapped samples uses samples in which the number of variables, the sample size and the marginal kurtosis are kept at their original level, which assures (at the 95% confidence level) that the derived factors determine the distribution of eigenvalues.

To investigate the reliability of these factors, a CFA was carried out using MPlus with a different subsample ($n = 331$). It was composed of the remaining 296 participants of the 592 nonclinical sample that answered all items and 35 nonclinical participants who left a maximum of

6 questions unanswered (age: M = 15.4, SD = 2.0 years). Analyses of univariate skewness and kurtosis indicated significant departures from normality, implying that data were multivariate non-normal. Therefore, WLSMV (weighted least-squares means and variance adjusted) estimation was used to examine model fit. This method uses weighted least-square parameter estimates from the diagonal of the weight matrix and has been recommended for multivariate non-normal data and for categorical/ordinal variables on the basis of simulation studies (Muthén, du Toit, & Spisic, 1997).

First, an independence model was examined. It assumes that all of the items are uncorrelated and is the standard control in CFA. Second, a one-factor model, which presupposes that all of the items pertain to the same factor, was assessed. Third, a two-factor orthogonal model was examined. This model assumes the pattern of loadings that emerged from the PAF, except the factors are designated as independent of one another. Finally, a two-factor oblique model was scrutinized which, unlike the orthogonal model, allows inter-correlation among the factors. Four indices were invoked to assess the efficacy of these models, including the Comparative Fit Index (CFI; Bentler, 1990), the Tucker-Lewis Index (TLI; Tucker & Lewis, 1973), the Root Mean Square Error Of Approximation (RMSEA; Steiger, 1990) and the Weighted Root Mean Square Residual (WRMR; Yu & Muthén, 2002). The absolute fit of the model can be represented by the χ^2/df ratio, which should be less than 5.0. (Bollen, 1989). The efficacy of the derived models was further examined across gender, age and nonclinical and clinical participants.

Distribution of scores. Once the factor structure of the instrument was identified, we report descriptive statistics (M, SD) for the RT and the SH subscale scores as well as the total RTSHIA scale for the entire sample.

Internal consistency. Cronbach's alpha coefficient and corrected item-total correlations were computed.

Test-retest reliability. Test-retest reliability (Pearson r) was calculated for a subsample of the doctoral survey participants who completed the items initially and 3 months later.

Validity. The validity of the RTSHIA was assessed in three ways. Firstly, the RT and SH scales were correlated (Pearson *r*) with theoretically related measures, such as the SMFQ, the MACI subscales, the A-DES and the CI-BPD in order to establish construct (convergent) validity. Given the limitations of using Hotelling's *t* test to compare related correlation coefficients (Steiger, 1980), we used Steiger's *z* to compare the correlation coefficients between the RTSHIA scales and relevant standardized measures. In these analyses, Bonferroni corrections were applied to ensure a comparison-wise error rate of .05. Secondly, we repeated the correlations between the RTSHIA and key validity measures, after controlling for SMFQ total scores (using partial correlations), in order to examine the measure's discriminatory power and establish divergent validity. The SMFQ was chosen based on depression's high comorbidity with SH behavior in adolescence (Evans et al., 2004) and the high degree of overlap between depressive affect and SH observed in the present study. Thirdly, we used independent samples *t* tests to compare the SH group obtained from the RCT to a randomly selected (though age, gender and ethnicity matched) subsample obtained from the school survey, in order to establish concurrent validity.

Ability to detect change. Next, we examined the correlations between the difference scores of the RTSHIA derived scales and the SMFQ, both of which were readministered to subsamples of the RCT SH participants at 3, 6, and 9 months after intake.

Discriminatory ability. Receiver Operating Characteristic (ROC) curves were employed to identify multiple pairs and provide a graphical display of the relationship between test sensitivities (proportion of true positives correctly identified) and specificities (proportion of true negatives correctly identified) at multiple cutoff scores for the SH scale score. The capacity of the RTSHIA to differentiate between the RCT SH clinical group and an age, gender and ethnicity matched community group was therefore assessed.

Results

Determination of Factor Structure: Exploratory Factor Analysis

Before performing the PAF, two items (21 and 26) were discarded. Their item reliability indices were below 0.15, and more than 96% of respondents answered *Never* to those items resulting in excessive positive skewness (6.8 and 5.2 respectively) and kurtosis (45.1 and 25.1 respectively). Although also asymmetric, the remaining items were answered positively (*Once*, *More than Once* or *Many Times*) by at least 5% of the sample.

Two items (1, 2) in which the communality was below 0.3 were discarded. The remaining 29 items were subjected to another ULS factor analysis. Two factors emerged from this analysis, however items 8, 9 and 24 had moderate loadings on both factors (.49 & .31, .30 & .52 & .51 and .31 respectively) and were removed in order to improve factor simplicity.

A final ULS factor analysis (oblique rotation) was carried out with the remaining 26 items. According to this procedure, the advised number of extracted factors was two. The inter-factors correlation was .55. Together the first two factors explained 60.6% of the variance. The component matrix revealed a dominant first factor (accounting for 49.8% of the variance) with 18 out of 26 items having rotated loadings reaching .56 or higher and a second factor (accounting for 10.8% of the variance) with the remaining 8 items having loadings of .49 or higher.

Table 1 displays the pattern matrix that emerged from this analysis. The final column relates to the CFA and will be addressed later. To enhance readability, coefficients that do not exceed .3 are omitted. The two factors are respectively designated as *self-harming* (SH) and *risk-taking* (RT) *behaviors.* The structure matrix (available on request) revealed a similar configuration. The two emerged factors will be delineated below.

[TABLE 1 ABOUT HERE]

Determination of Factor Structure: Confirmatory Factor Analysis

To investigate the reliability of these factors, a CFA was carried out using MPlus with the other subsample. Table 2 presents the fit indices associated with each model. The two-factor oblique model, assuming that RT and SH are different, but related constructs appeared to fit the

data better, and its indices were close to the recommended criteria of a reasonable model. The CFI and the TLI were larger than .95 and considered acceptable. The RMSEA exceeded the recommended cutoff value of .05 (Browne & Cudeck, 1993) or .06 (Hu & Bentler, 1998). However, it has been suggested that precise numerical cutoff points for RMSEA should be used with caution (Chen, Curran Bollen, Kirby, & Paxton, 2008). The WRMR exceeded the recommended value of 1. The correlation between the RT and SH scales was $r = .44$.

[TABLE 2 ABOUT HERE]

The standardized coefficients associated with each item are presented in the last column of Table 1. All of these coefficients exceed 0.4, providing support for the efficacy of the model. Although not all indices surpassed the criteria of an acceptable two factor oblique model, no definite criteria exist to determine precise cutoffs, and "interpretation of fit indices has to take into account a number of measures as well as the nature of the data and the model under examination" (Heubeck, 2000, p. 443). Given the heterogeneity of the data and the theoretical coherence of the emerged factors, on the whole the model was considered to be supported.

Factor Structure Across Gender, Age and Samples of Nonclinical and Clinical Participants

In order to investigate the efficacy of the model across gender and age, separate CFAs were carried out for girls ($n = 555$) and boys ($n = 142$) and for early (Grades 6 and 7, $n = 179$), middle (Grades 8 and 9, $n = 171$) and late adolescents (Grades 10, 11 and 12, $n = 347$). Moreover, CFAs were carried out for nonclinical ($n = 627$) and clinical participants ($n = 70$), including 52 SH and 18 non-SH clinical participants. The results of these analyses are presented in Table 2. Most indices were close to the recommended cutoff values. However, replication with larger samples would allow higher levels of certainty regarding the underlying two-factor structure.

Distribution of Scores

We first aggregated the responses given to all retained 26 items by both the doctoral survey and the RCT participants into total scores. The mean was 7.69 (*SD* = 7.93), the median 6, the mode 1, and the range 0 - 62. We next aggregated the responses given to the 8 RT items, again for the whole sample. The mean was 3.24 (*SD* = 3.30), the median 2, the mode 0, and the range 0 - 22. Then, we aggregated the responses given to the 18 SH items, again for the whole sample. The mean was 4.47 (*SD* = 5.84), the median 2, the mode 0, and the range 0 - 45. All three distributions were positively skewed and leptokurtic.

Reliability

The RT and SH scales were computed by adding the Likert-type ratings given to the 8 and 18 items loading on each component respectively. The next set of analyses investigates the reliability of the two scales. The alpha for the RT scale (*N* = 707) was high at .85 and the item reliability indices ranged from .45 (item 4) to .69 (item 12). The alpha for the SH scale (*N* = 675) was also high at .93 and the corrected item-total correlations ranged from .33 (item 23) to .78 (item 32). Three-month test-retest reliability (Pearson *r*) for the RT and SH subscales was .90 and .87 respectively.

Validity

The correlations between the RT and SH subscales and the relevant measures are presented in Table 3.

[TABLE 3 ABOUT HERE]

As evident in Table 3, both subscales correlated significantly and in the expected direction with the SMFQ and the BPFS-C and most MACI scales. The highest correlations observed were those between RT and the MACI substance abuse scale and the SH and the MACI suicidal tendency scale. Although both RT and SH correlated positively with the SMFQ, the BPFS-C, the MACI doleful, self-demeaning, borderline and identity diffusion scales and the number of CI-

BPD interview criteria, the correlations were significantly larger for the SH scale. Moreover, SH correlated highly (medium effect size, i.e. $r > .5$; Cohen, 1992) and positively with the MACI introversive, self-devaluation, childhood abuse and depressive affect scales, and negatively with the MACI conforming scale. Besides the MACI substance abuse scale, RT correlated highly ($r > .5$) and positively with the MACI unruly scale, and negatively with the MACI anxious feelings scale. The correlations between RT and MACI unruly, forceful, substance abuse, delinquent predisposition, impulsive propensity and anxious feelings scales were significantly stronger compared to the correlations between these MACI scales and SH. Interestingly, the MACI social insensitivity scale correlated positively with RT and negatively with SH.

Next, we conducted partial correlation analyses to examine the unique association between the SH scale and key validity measures, after partialling out the SMFQ scores. The results revealed that the SH scale remained significantly correlated with the BPFS-C (partial $r = .26$, $p < .001$), the MACI suicidal tendency scale and borderline scales (partial $r = .32$, $p < .001$ and partial $r = .15$, $p = .027$, respectively) and the number of CI-BPD interview criteria (partial $r = .46$, $p < .001$). However, all associations decreased in magnitude.

Our next step was to compare the SH group of the RCT and an age and gender equivalent group drawn from the community. The SH group scored significantly higher than the community matched group on both the RT and the SH scales, $t(65.46) = 4.27$, $p < .001$, significant at the Bonferroni corrected 0.0083 level, and $t(66.09) = 7.5$, $p < .001$, significant at the Bonferroni corrected 0.0083 level, respectively (unequal variances assumed).

The SH group scored higher than the non-SH clinical group on RT and SH scales, although the difference was statistically significant for only the SH scale, $t(66) = 0.52$, $p = .61$, n.s., and $t(65.18) = 8.29$, $p < .001$, (unequal variances assumed), significant at the Bonferroni corrected 0.0083 level. Table 4 presents the means and standard deviations obtained by each group.

[TABLE 4 ABOUT HERE]

Ability to Detect Change

The magnitude of all correlations between the difference scores (baseline value minus value at 3, 6 and 9 months) of the RT, the SH and the SMFQ was medium or large. In particular, differences in RT and SMFQ correlated significantly ($p < .05$) at 3 and 6 months ($r = .42$ and $r = .39$ respectively), but not at 9 months ($r = .33$, $p = .23$, *n.s.*). Differences in SH and SMFQ correlated significantly at 3 ($r = .48$, $p < .01$) 6 ($r = .40$, $p < .05$) and 9 months ($r = .88$, $p < .001$). This suggests that the ratings of these measures of psychopathology were changing in a similar direction and magnitude over time.

Discriminatory ability.

Using ROC curves (Figure 1), we selected a cutoff value that maximized the sum of sensitivity (proportion of true positives correctly identified) and specificity (proportion of true negatives correctly identified) and minimized error. This gave a value of 7, with a sensitivity of .94 and a specificity of .72, yielding a positive predictive value of 76.2%, correctly identifying 48 of the 51 RCT SH participants (true positives), and 38 of the 53 community participants (true negatives). Depending on factors such as the severity of relative cost of errors, and the prevalence of the phenomenon, this score may be adjusted accordingly.

[FIGURE 1 ABOUT HERE]

Discussion

The major purpose of this study was to investigate the structure of RT and SH behaviors in adolescence, and to develop a reliable and valid measure of adolescent RT and SH. This article described the construction of the RTSHIA, and examined its psychometric properties. Two factors emerged from our data, referred to as RT and SH behaviors. Both factors had high internal consistency and test-retest reliability, and sufficient convergent, concurrent and divergent validity.

The RT factor is composed of eight items; one item concerning illegal activities, one item concerning school noncompliance, one item concerning staying out late at night without parental knowledge, one item about participation in violence, one sexual RT item (having many sexual partners within a short period of time), and three substance abuse related items (smoking tobacco, using excessive amounts of alcohol and using illicit drugs).

The SH factor consisted of a number of "direct" SH items, including various forms of self-mutilation (e.g cutting, burning, biting, picking at wounds, etc) overdosing and attempting suicide. Moreover, this factor also included an item regarding the presence of self-demeaning thoughts, an item about staying in relationships or friendships with people who repeatedly hurt one's feelings, and two items enquiring about the presence of self-harming or suicidal thoughts.

The finding that suicidality-and self-mutilation-related items loaded on the same component provides some support to the multifactorial nature of SH. This is in accordance with literature in which SH is defined at the behavioral level with no reference to intent (Hawton et al., 2002), because suicidal acts and self-destructive behavior may often be interchangeable in nature.

The association between the RT and SH factors and the loading of SH thoughts and psychological SH items on the same factor is in accordance with how young people themselves seem to understand and experience SH. For example, in single-item SH measures, young people have been reported to positively endorse the item on the basis of having engaged in RT behaviors and mentally hurting themselves (Ross & Heath, 2002). Moreover, young people responded in our open-ended question about other types of SH by mentioning a broad range of SH phenomena beyond the strict definition of self-inflicted tissue damage.

The fact that RT and SH-related items did not load on the same factor points to the possibility that the functions of RT and SH are distinct, and that they are likely to correspond to different vulnerabilities and coping mechanisms. In particular, the patterns of correlations of RT and SH with measures of psychopathology suggest a degree of overlap, but they also provide

support for RT as a mainly externalizing problem, linked with aggression and impulsivity, and SH as a mainly internalizing problem, linked with depressive affect and self-devaluation.

The above inter-correlated two-factor structure was demonstrated via the CFA and the factorial invariance of each solution across gender and age groups. Although the goodness of fit indices were not always within the strictly acceptable range, when considered together they were adequately close to support the two-factor model, which was further demonstrated by the strength of the factor loadings. The two-factor model is consistent with empirical findings and theoretical proposals about RT and SH being closely related, albeit not identical.

It is notable that the factor analysis resulted in a total of seven items originally included in the scale being discarded for statistical reasons. For example, three items were discarded because they loaded moderately on both factors. Unprotected sex appeared to be linked with both sexual RT and indirect SH in the form of self-devaluation. Analyses investigating possible moderation effects of gender, age or sample type (clinical vs. nonclinical) did not reveal any differences. Given that unsafe sex is a major risk behavior among adolescents, it is advisable that the item is retained in the RTSHIA. As it pertains to both RT and SH, its inclusion in the total RTSHIA scale would not affect the high internal consistency of the overall scale.

Putting oneself at risk of sexual abuse and deliberately inhaling something harmful, excluding cigarette smoke or drugs or swallowing something inedible also loaded positively on both factors. Like unsafe sex, putting oneself at risk of sexual abuse may be pertinent to both RT and SH. However, the formulation of the latter item may have caused the dual loading and therefore needs sharpening. Some respondents may have not taken into account the phrase "excluding cigarette smoke or drugs" and endorsed the item for both drug-related behaviors such as glue sniffing as well as for direct SH behaviors such as swallowing bleach or inedible objects.

Besides finding a meaningful factor structure for the RTSHIA items, both RTSHIA scales had very high internal consistency. It is interesting that if all 26 items had been grouped under the same single factor, internal consistency would have remained very high (.92). Very

good evidence for test-retest reliability was also found, albeit in a sample group which did not cover the complete age range of our participants.

The RTSHIA demonstrated promising convergent validity as evidenced by moderate to strong correlations with validated measures of depressive affect, borderline personality traits and suicidality. The RT factor shared about 50% variance with the MACI substance abuse scale, whereas SH shared about 50% variance with the MACI suicidality scale and the CI-BPD. The size of these associations suggests that while the RTSHIA overlapped somewhat with these measures, it appeared to be statistically independent and therefore a measure of unique constructs.

Furthermore, even after controlling for the effects of depressive affect, the associations between SH and both BPD traits and suicidality remained significant, albeit decreased. This provides initial evidence for the divergent validity of the RTSHIA. Moreover, the RTSHIA and especially the SH scale was shown to have validity as a measure of change. Difference scores on the RTSHIA were correlated with difference scores on the SMFQ, a self-report measure of depressive affect, which is the primary target of interventions with individuals who self-harm.

The RTSHIA may represent an improvement over existing measures of RT and SH for several reasons. First, it provides a wealth of information about the full spectrum of potentially self-destructive behaviors as well as other very significant information, such as the localization of injuries in the body. As a self-report measure, it has the benefit of standardized administration, wording and scoring (Osman et al., 2002) and can be faster, more economical to administer and score and more efficient as it can be completed in groups (Eyman & Eyman, 1990). Furthermore, the quality of data produced by self-report measures has been compared to that obtained through clinical interviews (Shain, Naylor & Alessi, 1990). People may be more comfortable admitting to sensitive thoughts and acts when they are asked to circle a response or write a brief explanation instead of providing a verbal report, which may be influenced by interpersonal reactions to interviewers (Gutierrez, Osman, Barrios & Kopper, 2001). Moreover, assuring participants of the confidentiality and anonymity of self-reports seems to be easier (Gutierrez et al., 2001). In any

case, there may be few alternatives to self-report data when requesting personal and sensitive information from young people (Osman et al., 2002).

Another advantage of the RTSHIA is that its data can be used flexibly in a variety of ways. As the RTSHIA not only identifies individuals who are at risk but also reports on specific risk behaviors, some items (e.g. regarding substance abuse or suicidal attempts) can be analyzed separately and be used to alert clinicians to the potential for significant risk that requires further assessment, depending on the research or clinical focus, and to identify targets for intervention. As a provisional guide, we found in this study that a score of 7 or greater maximizes the predictive power of the measure (sum of sensitivity and specificity) in differentiating clinical from non-clinical individuals. A score above the cutoff should prompt the researcher or clinician to perform a more extensive assessment for SH and gather more detailed information. Moreover, we found that a minority (3 out of 51 young people) in our clinical SH sample scored below the cutoff on the SH scale. Because not all individuals who self-harm engage in multiple methods, it is preferable to err on the side of caution and answers on individual items of importance should be considered together with the total score on the RT and SH scales.

Furthermore, the RTSHIA was developed for use with young people and is composed of behaviors consistent with the adolescent perspective. Despite the existence of several RT measures developed for adolescents (Gullone et al., 2000), the majority of SH measures were not developed and have not been validated with adolescents. In our view, an adolescent-relevant instrument is essential because despite considerable overlap, we should not assume that behaviors defined by adults as risky or self-destructive have the same function in adolescents.

Another advantage is that the RTSHIA provides an assessment of the frequency of SH behavior (an aspect that is likely to have important clinical implications), as opposed to simply recoding the presence or absence of the behavior. Moreover, the study included a large and highly diverse sample of participants drawn from a wide range of age groups. In addition, the use of RTSHIA was validated in a clinical population. Finally, there are robust psychometric data

establishing the measure's validity and reliability with participants from the whole range of secondary education, namely from Grade 6 up to Grade 12.

Several limitations must be acknowledged before conclusively recommending adoption of the RTSHIA for research and clinical purposes. First, the RTSHIA is primarily a self-report measure, and therefore dependent upon respondents' comprehension of items, concentration ability and openness. Despite reassurance with regard to anonymity and confidentiality, some of the participants may have hesitated to give personal information to the adult researchers, possibly seen as authority figures (Fox & Hawton, 2004). Perhaps a lie scale could rule out invalid responses more easily. In addition, some respondents, particularly the younger adolescents, may have felt that certain items were not applicable to them. While we appreciate the implications of asking young people some personal questions prematurely, (e.g. items regarding sexual behavior), all of the items of the RTSHIA were endorsed positively by at least one respondent in each year group, starting from Grade 6. Clearly, the topic should be introduced sensitively and informatively, so that the possibility of adverse reactions is minimized.

Second, the wording of the questions does not allow the discrimination between individuals who are currently involved in RT and SH versus those who did so in the past but not any more. However, the wording of the items is readily adaptable to the timeline of interest (e.g. *Over the last month…*). Another shortcoming related to the wording is that the phrase "to hurt or punish yourself" has a quite complex functional meaning. A less specific phrase e.g. "to hurt yourself or feel bad" appears more appropriate and could easily replace the previous formulation in items 25 and 27. Furthermore, the presence of NSSI cannot be determined with a single item. For example, it is impossible to determine whether a respondent who reported cutting him/herself, taking an overdose and attempting suicide was intending to kill him/herself when (s)he cut. However, in the absence of suicidal intent, the RTSHIA does contain a sufficient number of items that can be used to infer whether NSSI has ever been present.

Another limitation concerns external validity and the generalizability of the findings. The samples, although diverse, were not all selected to be representative of the broader adolescent population (most data were collected in London), and the ways in which participants were recruited made it impossible to examine selective participation effects. Moreover, the present study suffers from self-selection at the school level, resulting in a disproportionately large number of girls and adolescents of 17-18 years. In addition, the sample size of the retest study was small and did not cover the complete age range of our participants.

Finally, the CFA resulted in dropping seven items, which has inevitably resulted in a less comprehensive measure than originally planned. The RT scale in particular, being eight items in length, cannot be sufficiently comprehensive. It is important to emphasize that by no means does the RTSHIA enquire about all possible forms of RT and SH. However, both scales have been shown to work, having good psychometric properties and overall appearing reliable and valid.

In light of the aforementioned limitations, it is important to note that the results of this study are preliminary, and replication with different samples is necessary to ensure the robustness of the findings and the validity of the measure for other populations (Haynes, Richard, & Kubany, 1995). One useful area for future research would be to use or advance the RTSHIA to distinguish reliably between persistent RT and SH as a psychopathological phenomenon, and a temporary experimentation with some behaviors as a developmental phenomenon without any severe or long-term consequences. The potential of the measure to function as a brief screening tool for emerging BPD, a disorder in which the co-occurrence of RT and SH may be more prominent, is also worth investigating.

In summary, the incremental contribution of the RTSHIA to the SH literature is twofold. Firstly, it responds to a need for brief questionnaire-based instruments in addition to more in-depth interview-based ones. Secondly, by assessing a broader range of *potentially* self-destructive behaviors that are not directly self-injurious, it captures the wider presentation of the ways in which adolescents may harm themselves. However, as it has been pointed out previously (Nock et

al., 2007), the evidence-based assessment of self-harm should rely on multiple assessment methods. It would be useful in the future to validate the RTSHIA against interview measures such as the SITBI. It might turn out that they tap into different 'truths' about self-harm and they are complimentary in predicting criteria outcomes but it is also likely that there is additional value in obtaining interview data beyond self-report questionnaires. In such cases, the RTSHIA can be used as a primary screening measure that provides basic information about several risk-taking and self-harm behaviors. More detailed information about each behavior may be obtained by administering the SITBI (Nock, et al., 2007) or other, more focused measures, especially in clinical settings, where more in-depth information is required.

For the present time, the RTSHIA seems well-positioned to fill the void in adolescent RT and SH instrumentation and appears to have the potential as a reliable, valid, straightforward and flexible measure that can be used as a multifocused screening tool for the identification of a wide range of problem behaviors and thoughts in adolescence and as a tool for the assessment of symptomatic change and therapeutic effectiveness in treating young people who self-harm. Although it may be early to draw conclusions about the utility of the scale, the evidence exists to recommend the RTSHIA as a psychometrically sound, comprehensive tool with the potential for both research and clinical applications, worthy of further empirical investigation.

Acknowledgements

We are grateful to all participating young people and their teachers. We thank the UCL Department of Psychology undergraduate and postgraduate students, Dr Claire Wright, and the NELFT staff Katarina Kovacova and Eleana Stathopoulou for their support and assistance with data collection.

References

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision*. Washington, DC: American Psychiatric Association.

Alexander, C. S., Kim, Y. J., Ensminger, M., Johnson, K. E., Smith, B. J., & Dolan, L. J. (1990). A measure of risk-taking for young adolescents: Reliability and validity assessments. *Journal of Youth and Adolescence*, *19*, 559-569.

Angold, A., Costello, E. J., Messer, S. C., Pickles, A., Winder, F., & Silver, D. (1995) The development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research*, 5, 237 - 249.

Armstrong, J., Putnam, F. W., Carlson, E., Libero, D., & Smith, S. (1997). Development and validation of a measure of adolescent dissociation: The Adolescent Dissociative Experience Scale. *Journal of Nervous and Mental Disease*, *185*, 491-497.

Arnett, J. (1996). Sensation seeking, aggressiveness, and adolescent reckless behavior. *Personality and Individual Differences*, *20*, 693-702.

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238-246.

Bernstein, I. H., & Teng, G. (1989). Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization. *Psychological Bulletin, 105,* 467-477.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Boyer, T.W. (2006). The development of risk-taking: A multi-perspective view. *Developmental Review*, *26,* 291–345.

Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In: Bollen, K. A. & Long, J. S. (Eds.) *Testing Structural Equation Models*. pp. 136–162. Beverly Hills, CA: Sage.

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An Empirical Evaluation of the Use of Fixed Cutoff Points in RMSEA Test Statistic in Structural Equation Models. *Sociological Methods Research, 36,* 462-494.

Chiolero, A., & Schmid, H. (2002). Repeated self-reported injuries and substance use among young adolescents: The case of Switzerland. *Social and Preventive Medicine, 47*(5), 289-297.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Claes, L., & Vandereycken, W. (2007). The Self-Injury Questionnaire-Treatment Related (SIQ-TR): Construction, reliability, and validity in a sample of female eating disorder patients. In P.M.

Goldfarb (Ed.), *Psychological Tests and Testing Research Trends*. New York: Nova Science Publishers.

Crick, N. R., Murray-Close, D., & Woods, K. (2005). Borderline personality features in childhood: A short-term longitudinal study. *Development and Psychopathology, 17,* 1051 1070.

Evans, E., Hawton, K., & Rodham, K. (2004). Factors associated with suicidal phenomena in adolescents: a systematic review of population-based studies. *Clinical Psychology Review, 24,* 957-979.

Eyman, J.R., & Eyman, S.K. (1990). Suicide risk and assessment instruments. In P. Cimbolic & D.A. Jobes (Eds.), *Youth suicide: Issues, assessment, and intervention* (pp. 9–32). Springfield, IL: Thomas.

Fabrigar, L. R.,Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*, 272-299.

Fliege, H., Kocalevent, R. D., Walter, O. B., Beck, S., Gratz, K. L., Gutierrez, P. M., & Klapp, B. F. (2006). Three assessment tools for deliberate self-harm and suicide behavior: Evaluation and psychopathological correlates. *Journal of Psychosomatic Research, 61*, 113-121.

Fox, C., & Hawton, K. (2004). *Deliberate Self-Harm in Adolescence*. Jessica Kingsley: London.

Gratz, K. L. (2001). Measurement of deliberate self-harm: Preliminary data on the deliberate self-harm inventory. *Journal of Psychopathology and Behavioral Assessment, 23*(4), 253−263.

Gratz, K. L. (2003). Risk factors and functions of deliberate self-harm: An empirical and conceptual review. *Clinical Psychology Science and Practice, 10,* 192–205.

Gullone, E., Moore, S. Moss, S. & Boyd, C. (2000). The Adolescent Risk-Taking Questionnaire: Development and Psychometric Evaluation. *Journal of Adolescent Research, 15,* 231-250.

Gutierrez, P.M., Osman, A., Barrios, F.X., & Kopper, B.A. (2001). Development and Initial Validation of the Self-Harm Behavior Questionnaire. *Journal of Personality Assessment, 77*(3), 475-490.

Hawton, K., Rodham, K., Evans, E., & Weatherall, R. (2002). Deliberate self-harm in adolescents: self report survey in schools in England. *BMJ, 325,* 1207-1211.

Haynes, S. N., Richard, D. C. S. & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, *7,* 238–247.

Heubeck, B. G. (2000). Cross-Cultural Generalizability of CBCL Syndromes Across Three Continents: From the USA and Holland to Australia. *Journal of Abnormal Child Psychology, 28*, 439–450.

Hickey, L., Hawton, K., Fagg, J., & Weitzel, H. (2001) Deliberate self-harm patients who leave the accident and emergency department without a psychiatric assessment. A neglected population at risk of suicide. *Journal of Psychosomatic Research*, *50,* 87-93.

Howard, R., Yan, T. S., Ling, L. H., & Min, T. S. (2002). Risk taking and metamotivational state.

Personality and Individual Differences, 32, 155-165.

Hu, L. & Bentler, P.M. (1998). Fit indices in covariance structure modelling: Sensitivity to underparameterized model misspecification, *Psychological Methods, 3*, 424–453.

Keck-Seeley, S., Perosa, S., & Perosa, L. (2004). A validation study of the Adolescent Dissociative Experiences Scale. *Child Abuse and Neglect, 28,* 755-769.

Lattin, J., Carroll, D.J., & Green, P.E. (2003). *Analyzing multivariate data.* Duxbury Press.

Lavery, B., Siegel, A.W., Cousins, J. H., & Rubovits, D. S. (1993). Adolescent risk-taking: An analysis of problem behaviors in problem children. *Journal of Experimental Child Psychology*, *55*, 277-294.

Lescano, C.M., Hadley, W.S., Beausoleil, N. I., Brown, L.K., D'eramo, D., & Zimskind, A. (2007). A Brief Screening Measure of Adolescent Risk Behavior. *Child Psychiatry and Human Development, 37,* 325–336.

Lloyd, E. E., Kelley, M. L., & Hope, T. (1997, April). *Self-mutilation in a community sample of adolescents: Descriptive characteristics and provisional prevalence rates.* Poster session presented at the annual meeting of the Society for Behavioral Medicine, New Orleans, LA.

Lorenzo-Seva, U., & Ferrando, P.J. (2006). FACTOR: A computer program to fit the exploratory factor analysis model. *Behavioral Research Methods, Instruments and Computers, 38*(1), 88-91.

Lundh, L.G., Karim, J., & Quilisch, E. (2007). Deliberate self-harm in 15-year-old adolescents: A

pilot study with a modified version of the Deliberate Self-Harm Inventory. *Scandinavian Journal of Psychology, 48,* 33–41.

Millon, T. (1993). *The Millon Adolescent Clinical Inventory*. Minneapolis, MN: National Computer Systems, Inc.

Muthén, L.K., & Muthén, B.O. (1998-2007). *MPlus User's Guide*. Fifth Edition. Los Angeles, CA: Muthén & Muthén.

Muthén, B., du Toit, S.H.C. & Spisic, D. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modelling with categorical and continuous outcomes. Unpublished technical report.

Nock, M. K., Holmberg, E. B., Photos, V. I. , & Michel, B. D. (2007). Self-Injurious Thoughts and Behaviors Interview: Development, Reliability and Validity in an Adolescent Sample. *Psychological Assessment*, *19*(3), 309-317.

Nock, M. K., Joiner, T. E., Gordon, K. H., Lloyd-Richardson, E., & Prinstein, M. J. (2006). Non-suicidal self-injury among adolescents: Diagnostic correlates and relation to suicide attempts. *Psychiatry Research, 144*, 65-72.

Orbach I., Milstein, I., Har-Even, D., Apter, A., Tiano, S. & Elizur, A (1991). A Multi-Attitude Suicide Tendency Scale for Adolescents. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 398-404.

Osman, A., Bagge, C. L., Gutierrez, P. M., Konick, L. C., Kopper, B. A., & Barrios, F. X. (2001). The Suicidal Behaviors Questionnaire-Revised (SBQ-R): Validation with clinical and non-clinical samples. *Assessment, 8,* 445-455.

Osman, A., Barrios, F.X., Gutierrez, P.M., Wrangham, J.J., Kopper, B.A., Truelove, R.S., et al., (2002). The Positive and Negative Suicide Ideation (PANSI) Inventory: Psychometric evaluation with adolescent psychiatric inpatient samples. *Journal of Personality Assessment, 79,* 512–530.

Pattison, E.M., & Kahan, J. (1983). The deliberate self-harm syndrome. *Americal Journal of Psychiatry, 140,* 867–72.

Pfeffer, C. R., Jiang, H. & Kakuma, T. (2000). Child-Adolescent Suicidal Potential Index (CASPI): a screen for risk for early onset suicidal behavior. Psychological Assessment, *12*(3), 304-18.

Ross, S. & Heath, N. (2002). A study of the frequency of self-mutilation in a community sample of adolescents. *Journal of Youth and Adolescence*, *31*, 67–77.

Sansone, R.A., & Sansone, L.A. (2002). Assessment Tools for Self-Harm Behavior Among Those with Eating Disorders. *Eating Disorders, 10,* 193-203.

Sansone, R.A., Wiederman, M.W., & Sansone, L. A. (1998). The Self-harm Inventory (SHI): development of a scale for identifying self-destructive behaviors and borderline personality disorder. *Journal of Clinical Psychology, 54,* 973– 983.

Shain, B.N., Naylor, M., & Alessi, N. (1990). Comparison of self-rated and clinician-rated measures of depression in adolescents. *American Journal of Psychiatry, 147,* 793-795.

Simeon, D., & Favazza, A. R. (2001). Self-injurious behaviors: Phenomenology and assessment. In D. Simeon & E. Hollander (Eds.), *Self-injurious behaviors: Assessment and treatment* (pp. 1-28). Washington, DC: American Psychiatric Press.

Simeon, D., & Hollander, E. (Eds.). (2001). *Self injurious behaviors: Assessment and treatment.* Washington, DC: American Psychiatric Press.

Steiger, J.H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87,* 245-251.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25,* 173–180.

Steinberg, L. (2004). Risk taking in adolescence: what changes, and why? *Annals of the New York Academy of Sciences, 1021,* 51-8.

Steinberg, L. (2007). Risk-taking in adolescence: New perspectives from brain and behavioral science. *Current Directions in Psychological Science, 16,* 55-59.

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review, 28*, 78-106.

Tucker, L.R. & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*(1), 1-10.

Turp, M. (2003). *Hidden Self-Harm. Narratives from Psychotherapy*. Kingsley, London.

Walsh, B.W. (2008). *Treating Self-Injury: A Practical Guide.* Guilford Press, New York.

Webb, L. (2002). Deliberate self-harm in adolescence: a systematic review of psychological and psychosocial factors. *Journal of Advanced Nursing, 38*(3), 235-244.

Yates, T.M. (2004). The developmental psychopathology of self-injurious behavior: Compensatory regulation in posttraumatic adaptation. *Clinical Psychology Review, 24* (1), 35-74.

Yu, C.-Y. & Muthén, B. (2002). Evaluation of model fit indices for latent variable models with categorical and continuous outcomes. Technical report.

Zahl D, & Hawton, K. (2004). Repetition of deliberate self-harm and subsequent suicide risk: long-term follow-up study of 11583 patients. *British Journal of Psychiatry, 185,* 70– 5.

Zanarini, M.C. (2007, October). *Childhood Prevalence of Borderline Personality Disorder.* Paper presented at the 54th Annual Meeting of the American Academy of Child and Adolescent Psychiatry. Boston, MA.

Zlotnick, C., Shea, M. T., Pearlstein, T., Simpson, E., Costello, E., & Begin, A. (1996). The Relationship Between Dissociative Symptoms, Alexithymia, Impulsivity, Sexual Abuse, and Self-Mutilation. *Comprehensive Psychiatry, 37*(1), 12-16.

Zuckerman, M. (1979). Sensation seeking: Beyond the optimal level of arousal. Hillsdale, NJ: Lawrence Erlbaum Associates.

Table 1

Factor Pattern Matrix Rotated to the Direct Oblimin Criterion, $\delta = 0$ ($n = 296$)

RTSHIA Item	Factor 1 Self- harming	Factor 2 Risk-taking	h^2	CFA
Item 3		.806	.636	.683
Item 4		.490	.328	.432
Item 5		.639	.464	.856
Item 6		.711	.567	.712
Item 7		.626	.540	.703
Item 10		.649	.511	.817
Item 11		.936	.833	.867
Item 12		.741	.561	.837
Item 13	.809		.620	.875
Item 14	.619		.650	.789
Item 15	.807		.531	.740
Item 16	.699		.594	.661
Item 17	.651		.491	.702
Item 18	.764		.594	.798
Item 19	.725		.714	.717
Item 20	.723		.388	.519
Item 23	.674		.419	.424
Item 25	.813		.591	.723
Item 27	.687		.457	.710
Item 28	.566		.381	.559
Item 29	.829		.610	.761
Item 30	.720		.633	.800
Item 31	.790		.659	.814
Item 32	.834		.674	.827
Item 33	.778		.763	.917
Item 34	.729		.734	.796
Trace	1.92	0.42	2.34	
% of variance	49.8	10.8	60.6	

Note. RTSHIA = Risk-Taking and Self-Harm Inventory for Adolescents; Percentage variance is post-rotation; because here there were 26 measured variables, percentage of variance is trace divided by 26 times 100 (or trace times 26). h^2 = communality coefficient. The last column

presents the standardized coefficients that emerged from the Confirmatory Factor Analysis (CFA).

Table 2

Goodness-of-Fit Indices generated by the Confirmatory Factor Analysis (CFA) of the Risk-Taking and Self-Harming Items using data from Half of the Sample (n = 331) for four different models and relevant indices generated by the CFA of the Risk-Taking and Self-Harming Items for the two-Factor Oblique Model across Gender, Age and Samples of Nonclinical and Clinical Participants

Model type using data from half of the sample (*n* = 331)	χ^2	***df***	***CFI***	***TLI***	***RMSEA***	***WRMR***
Independence	2393.684	50				
One factor	576.217	67	.783	.838	0.152	2.197
Two-factor orthogonal	336.985	53	.879	.886	0.127	2.008
Two-factor oblique	185.663	74	.952	.968	0.068	1.187
Two-Factor Oblique Model for each Sub-Group	χ^2	***df***	***CFI***	***TLI***	***RMSEA***	***WRMR***
Clinical & Nonclinical Girls (*n* = 555)	284.734	89	.958	.983	.063	1.215
Clinical & Nonclinical Boys (*n* = 142)	93.241	36	.963	.971	.106	1.178
Clinical & Nonclinical Early adolescents (Grades 6 & 7) (*n* = 179)	91.195	46	.916	.952	.074	1.050
Clinical & Nonclinical Middle adolescents (Grades 8 & 9) (*n* = 171)	149.616	46	.943	.960	.115	1.295
Clinical & Nonclinical Late adolescents (Grades 10, 11 & 12) (*n* = 347)	176.348	66	.967	.982	.069	1.131
Nonclinical participants (*n* = 627)	281.624	95	.954	.974	.056	1.269
Clinical participants 52 SH and 18 non-SH, (*n* = 70)	105.590	30	.917	.945	.190	1.235

Notes. CFI = Comparative Fit Index, TLI = Tucker-Lewis Index, RMSEA = Root Mean Square Error of Approximation Index, WRMR= Weighted Root Mean Square Residual.

The chi-square and degrees of freedom are adjusted to obtain a correct *p*-value with weighted least-squares means and variance adjusted estimation (WLSMV).

Table 3

Pearson Correlations and Steiger's z scores comparing the correlation coefficients between the Risk-Taking and Self-Harm scales and the SMFQ, the BPFS-C, the MACI scales, the A-DES and the Childhood Interview for DSM-IV Borderline Personality Disorder (CI-BPD) Criteria

Measures	RT	SH	Steiger's *z*
Shortened Mood and Feelings Questionnaire (SMFQ)	.251**	.610**	-10.08(**)
Borderline Personality Features Scale for Children (BPFS-C)	.273**	.576**	-8.54(**)
MACI introversive scale	.026	.514**	-.7.86(**)
MACI inhibited scale	-.083	.406**	-7.58(**)
MACI doleful scale	.255**	.589**	-6.24(**)
MACI submissive scale	-.465**	-.333**	-2.20(*)
MACI dramatizing scale	-.038	-.494**	7.30(**)
MACI egotistic scale	-.072	-.469**	6.33(**)
MACI unruly scale	.552**	.161*	6.51(**)
MACI forceful scale	.436**	.109	5.18(**)
MACI conforming scale	-.431**	-.552**	2.17(*)
MACI oppositional scale	.354**	.481**	-2.15(*)
MACI self-demeaning scale	.235**	.478**	-3.98(**)
MACI borderline scale	.354**	.554**	-3.49(**)
MACI identity diffusion scale	.221**	.503**	-4.65(**)
MACI self-devaluation scale	.138*	.516**	-6.18(**)
MACI body disapproval scale	.023	.357**	-5.15(**)
MACI sexual discomfort scale	-.497**	-.330**	-2.82(*)
MACI peer insecurity scale	-.089	.349**	-6.71(**)
MACI social insensitivity scale	.229**	-.243**	7.17(**)
MACI family discord scale	.426**	.283**	2.33(*)
MACI childhood abuse scale	.256**	.550**	-4.99(**)
MACI eating dysfunctions scale	.042	.408**	-5.72(**)
MACI substance abuse scale	.707**	.491**	4.43(**)
MACI delinquent predisposition scale	.441**	-.115	8.69(**)
MACI impulsive propensity scale	.444**	.231**	3.45(**)
MACI anxious feelings scale	-.559**	-.168*	-6.54(**)
MACI depressive affect scale	.056	.509**	-7.30(**)
MACI suicidal tendency scale	.206*	.698**	-8.99(**)
Adolescent Dissociative Experiences Scale (A-DES)	.197	.483**	-3.01(**)

Number of CI-BPD interview criteria at level 2 (serious or severe)	.323**	.748**	-6.01(**)

Notes. MACI: Millon Adolescent Clinical Inventory. For the above correlations, *n* was 598 & 572 for the SMFQ, 616 & 594 for the BPFS-C, 231 & 224 for the MACI scales, 98 & 96 for the A-DES and 119 & 117 for the CI-BPD interview criteria at level 2 and the RT and SH subscales respectively.

* Correlation (or correlation difference) is significant at the $p < 0.05$ level. ** Correlation (or correlation difference) is significant at the Bonferroni corrected $p < 0.0016$ level.

Table 4

Number of participants (n), means and SDs on the RT and SH scales among the study's groups

	Risk-Taking			**Self-Harming**		
Group	***n***	***M***	***SD***	***n***	***M***	***SD***
RCT SH group	50	6.42	5.03	51	15.78	9.84
Age, gender & ethnicity matched community group	52	3.17	2.13	53	4.77	4.07
Clinical non-SH control group	18	5.72	4.43	18	3.17	2.75
Remaining community group	587	2.90	2.98	553	3.44	4.26

Note. The sample size varies slightly due to missing cases.

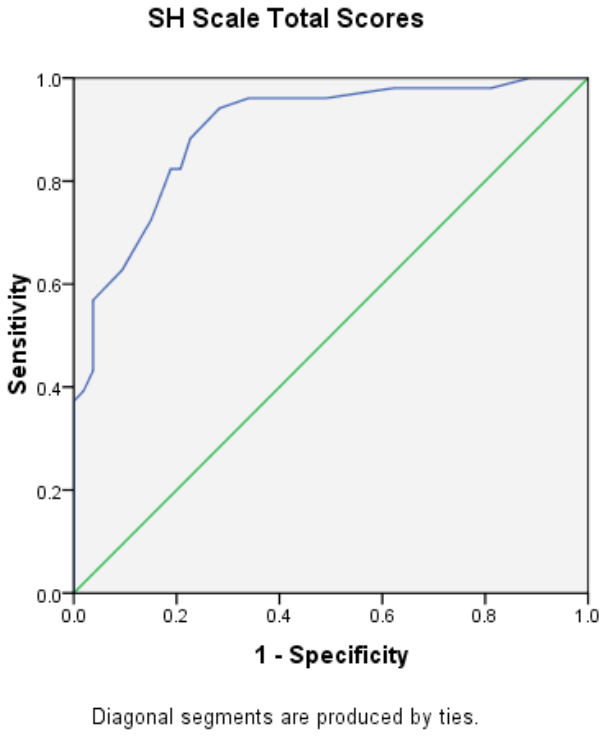

Figure 1. Receiver Operating Characteristic (ROC) plot of the relationship between scores on the Self-Harming scale of the RTSHIA and clinical group membership (RCT group referred to mental health services for self-harming behavior).

Note. RTSHIA: Risk-Taking and Self-Harm Inventory for Adolescents

Copy of the original RTSHIA measure

Instructions

This questionnaire asks about a number of different things that young people sometimes do. Please do not be concerned if some statements seem unusual. They are included to provide us with greater understanding and knowledge about these behaviors and the best way to help young people.

- Please complete this questionnaire on your own.
- If a statement is not applicable to you, please circle *Never*.
- You do not have to answer any questions that you prefer not to answer.
- Please try to answer as truthfully as possible.
- All your answers are kept strictly confidential.

1. Have you ever taken chances while doing your hobbies (e.g. not wearing your helmet and other safety gear, riding risky stances on your skateboard, etc)?	Never	Once	More than once	Many times
2. Have you ever deliberately crossed the road dangerously or driven recklessly (e.g. raced, did not fasten your seatbelt, drove while intoxicated or drunk)?	Never	Once	More than once	Many times
3. Have you ever put yourself in a risky situation (such as classroom cheating, traveling without a valid ticket, shoplifting etc) knowing that you may get caught?	Never	Once	More than once	Many times
4. Have you ever been suspended (i.e. punished with exclusion) or dropped out of school?	Never	Once	More than once	Many times
5. Have you ever stayed out late at night, without your parents knowing where you are?	Never	Once	More than once	Many times
6. Have you ever participated in gang violence, physical fights or held a weapon?	Never	Once	More than once	Many times
7. Have you ever been promiscuous (i.e. had many sexual partners within a short period of time)?	Never	Once	More than once	Many times
8.*Have you ever had sex avoiding precautions against sexually transmitted diseases or pregnancy?	Never	Once	More than once	Many times
9. Have you ever put yourself at risk of sexual abuse?	Never	Once	More than once	Many times
10. Have you ever had so much alcohol that you were really drunk?	Never	Once	More than once	Many times

11. Have you ever used drugs (such as marijuana, cocaine, LSD etc)?	Never	Once	More than once	Many times
12. Have you ever smoked tobacco?	Never	Once	More than once	Many times

Please say yes to the following questions only if you did the behavior intentionally, or on purpose, to hurt yourself. Circle *Never* if you did something only accidentally (e.g. you tripped and banged your head on accident).

13. Have you ever intentionally cut your skin?	Never	Once	More than once	Many times
14. Have you ever intentionally burned yourself with a hot object (such as a cigarette)?	Never	Once	More than once	Many times
15. Have you ever intentionally bitten yourself, to the extent that you broke the skin?	Never	Once	More than once	Many times
16. Have you ever intentionally banged your head against something, hit or punched yourself, to the extent that you caused a bruise to appear?	Never	Once	More than once	Many times
17. Have you ever intentionally prevented wounds from healing or picked at areas of your body to the point of drawing blood?	Never	Once	More than once	Many times
18. Have you ever intentionally scraped, scrubbed or scratched your skin to the point of breaking your skin or drawing blood?	Never	Once	More than once	Many times
19. Have you ever intentionally rubbed a sharp object (such as sandpaper) or dripped anything toxic (such as acid) onto your skin?	Never	Once	More than once	Many times
20. Have you ever exercised an injured part of your body intending to hurt yourself?	Never	Once	More than once	Many times
21. Have you ever deliberately broken a bone in your body either by making yourself fall or in another way?	Never	Once	More than once	Many times

22. Please choose A or B

…...…A. I've never deliberately injured myself

…...…B. I have at least once deliberately injured myself

If you answered B, which body parts did you deliberately injure?

Please tick one (or more) of the following options.

..…..... Torso, belly, buttocks …….Hands, arms, fingers, nails

……. ..Head …….Legs, feet, toes

…...…..Face …….Other (please write)…….

…...…..Neck

23. Have you ever intentionally pulled your hair out?	Never	Once	More than once	Many times
24. Have you ever deliberately inhaled something harmful	Never	Once	More	Many

(excluding cigarette smoke or drugs) or swallowed something inedible?			than once	times
25. Have you ever starved yourself to hurt or punish yourself?	Never	Once	More than once	Many times
26. Have you ever used laxatives to hurt or punish yourself? *Laxative*: a drug that makes you go to the toilet	Never	Once	More than once	Many times
27. Have you ever forced yourself to eat too much to hurt or punish yourself?	Never	Once	More than once	Many times
28. Have you ever stayed in a friendship or a relationship with somebody who repeatedly hurt your feelings on purpose?	Never	Once	More than once	Many times
29. Have you ever tried to make yourself suffer by thinking horrible things about yourself?	Never	Once	More than once	Many times
30. Have you ever taken an overdose? (i.e. taken an excessive amount of medication without having been prescribed this dosage)	Never	Once	More than once	Many times
31. Have you ever seriously thought about harming a part of your body?	Never	Once	More than once	Many times
32. Have you ever seriously thought about killing yourself?	Never	Once	More than once	Many times
33. Have you ever tried to kill yourself?	Never	Once	More than once	Many times
34. Have you ever intentionally hurt yourself in any of the above mentioned ways so that it led to hospitalization or injury severe enough to require medical treatment?	Never	Once	More than once	Many times

35. Have you engaged in any other self-destructive behaviors not asked about in this questionnaire? If yes, please describe below

……

36. Please choose A or B

……A. I know no one well who has deliberately injured himself/herself

……B. I know someone well who has deliberately injured himself/herself

37. If you answered B, why do you think he/she did this?

……

38. If you answered A or B, why do you think some young people harm themselves?

……

Thank you very much for completing this questionnaire

Notes. The items with their numbers in **heavy type** were retained in the final RTSHIA. Item **8*** loaded equally on both the Risk-Taking and the Self-Harm factors and was therefore excluded

from the relevant scales in order to improve factor simplicity. However, as it is a major adolescent risk behavior, it is recommended that the item is retained in the RTSHIA.