This is the author version of an article published as:


© 2020. This is an Accepted Manuscript of an article published by Taylor & Francis in Journal of Clinical Child & Adolescent Psychology on 20/02/2020, available online https://doi.org/10.1080/15374416.2020.1716364

It is deposited under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.
A measure of safety behaviors for use with young people: The Subtle Avoidance Measure for Youth

Cindy Chapman
Ronald M Rapee

Centre for Emotional Health, Department of Psychology,
Macquarie University

Corresponding author: Ron Rapee, Centre for Emotional Health, Department of Psychology, Macquarie University, Sydney NSW 2109, Australia; Ph: +61 2 98508032
Email: ron.rapee@mq.edu.au

Running head: Measuring safety behaviors in youth
Abstract

Objective: Safety behaviors have been found to undermine successful exposure in the treatment of anxiety disorders for both adults and children. Although reliable measures of safety behaviors have been developed for use with adults, no such measure has been developed specifically for pediatric populations. In light of this limitation, the current study aimed to develop and validate a measure of the use of safety behaviors suitable for children: The Subtle Avoidance Measure for Youth (SAMY).

Methods: Clinical (n = 174) and community (n = 138) young people, aged 7-13 years, provided data.

Results: Both exploratory and confirmatory factor analyses supported a three-factor solution of the SAMY, which reflected checking behaviors, behaviors related to image management, and behaviors related to physical protection. The SAMY and its subscales demonstrated strong internal consistency, test-retest reliability, construct validity, and the ability to discriminate between clinical and community participants.

Conclusions: Given its solid psychometric properties, the SAMY will prove useful for both research and clinical work with anxious young people.

Key words: Anxiety, safety behavior, assessment, psychometric, avoidance
Introduction

Exposure-based treatments are the “gold-standard” for the reduction of anxiety disorders (Rapee, Schniering, & Hudson, 2009). One crucial factor that has been found to hinder effective exposure is avoidance during exposure tasks, which often occurs through the use of safety behaviors, sometimes also referred to as subtle avoidance (Furukawa et al., 2009; Lovibond, Mitchel, Minard, Brady, & Menzies, 2009). These behaviors comprise cognitive or behavioral strategies that are employed to reduce the perceived risk of feared outcomes occurring and seemingly protect the individual from perceived harm. They can include active, overt behaviors (e.g., frequently washing hands), as well as subtle restriction of behaviors (e.g., avoiding eye contact) (Salkovskis, 1991). While the use of safety behaviors may bring short-term relief to an anxious individual, theoretical models suggest that prolonged use is problematic because it interrupts threat disconfirmation and can paradoxically increase the likelihood of negative outcomes (Deacon & Maack, 2008; Powers, Smits, & Telch, 2004). Consistent with this view, research with adult populations has shown that use of safety behaviors can maintain anxiety and weaken the impact of exposure-based treatments (Furukawa et al., 2009; Kim, 2005; Lovibond et al., 2009; Moscovitch et al., 2013; Taylor & Alden, 2010).

Given the relevance of safety behaviors for the maintenance and treatment of anxiety disorders, surprisingly little attention has been paid to their reliable and valid assessment. At present, there are only a handful of published measures of safety behaviors for use among adults. Most of these scales have been focussed on behaviors of relevance to a specific anxiety disorder including social anxiety disorder (Subtle Avoidance Frequency Examination (SAFE) (Cuming et al., 2009); Social Phobia Safety Behaviors Scale (SPSBS) (Burato, Crippa, & Loureiro, 2009), panic disorder (Texas Safety Maneuver Scale (TSMS) (Kamphuis & Telch, 1998), and fears of storms (Storm-Related Safety Behavior Scale (SRSBS) (Krause,
MacDonald, Goodwill, Vorstenbosch, & Antony, 2018). One measure, the Safety Behavior Assessment Form (SBAF) (Goodson, Haeffel, Raush, & Hershenberg, 2016), includes coverage of a range of safety behaviors of relevance across several disorders. The SBAF is a 41-item self-report questionnaire that was developed largely from research within a military population. It contains three factors: monitoring and vigilance behaviors; bodily and health related behaviors; and escaping and evading of social situations and has shown solid psychometric properties. More recently, another measure assessing a range of safety behaviors among adults, the Subtle Avoidance Measure (SAM), has shown two factors: Tangible outcome management and social image management (Soo & Rapee, 2019).

In contrast to the adult literature, far less is known about the use of safety behaviors among children and adolescents. Nevertheless, some studies have shown similar relevance of safety behaviors to the maintenance of anxiety in young populations as among adults. For example, youth with social anxiety disorder aged 8-13 years have been shown to engage in greater use of safety behaviors than non-clinical controls (Kley, Tuschen-Caffier, & Heinrichs, 2012) and use of safety behaviors has been shown to predict social anxiety among participants aged 11-14 years (Hodson, McManus, Clark, & Doll, 2008) and 14-17 years (Qasmieh et al., 2018; Szollos et al, 2019; Thomas et al., 2015). Among older adolescents, safety behaviors predict constructs that are theoretically related to social anxiety such as fears of negative and positive evaluation (Szollos et al., in press; Lipton, Augenstein, Weeks, & De Los Reyes, 2014) and predict social anxiety above any variance predicted by depression (Qasmieh et al., 2018). Further, the use of safety behaviors during exposure has been associated with significantly reduced efficacy for treatment of child anxiety (Hedtke, Kendall, & Tiwari, 2009).

Given these theoretically important findings and the important role of safety behaviors in theories of anxiety, it is surprising that no specifically child-focused measures of
safety behaviors for clinically anxious youth currently exist. As a result, researchers who evaluate safety behaviors in younger people have used measures originally designed for adults (e.g., Kley et al., 2012; Szollos et al., in press; Thomas et al., 2015). However, studies of the psychometric properties of adult safety behaviors measures among younger populations are rare and have mostly focused specifically on social anxiety. In two studies with older adolescents aged 14-17 years, the SAFE demonstrated good psychometric properties as a construct related to social anxiety (Qasmieh et al., 2018; Thomas, Daruwala, Goepel, & De Los Reyes, 2012). Another study adapted an adult measure of safety behaviors related to trauma for youth who had experienced a traumatic event (Alberici et al., 2018). The resulting 22-item scale showed solid psychometric properties.

Theoretically, it has been suggested that use of safety behaviors will be closely related to the content of threat expectations (Salkovskis 1991; Salkovskis et al., 1996). Hence, it is likely that people with different forms of anxiety (e.g., social anxiety, panic disorder) will make use of different safety strategies. As noted above, research and assessment of safety behaviors in both adults and adolescents has typically focused on one specific form of anxiety, with most of this work focusing on social anxiety (e.g., Furukawa et al., 2009; Kley et al., 2012; Qasmieh et al., 2018; Taylor & Alden, 2010). A few broader measures of safety behaviors for use with adults have begun to appear (e.g., Goodson et al., 2016; Soo & Rapee, 2019). Exploratory factor analyses of these measures have commonly demonstrated two or three factors: a factor related to social concerns and image management; a factor related to physical protection and health; and a factor related to vigilance and checking (Goodson et al., 2016; Soo & Rapee, 2019). Diagnostically, it might be predicted that these factors will relate differently to specific anxiety disorders: image management to social anxiety disorder; physical protection to panic disorder or separation anxiety disorder; and vigilance to obsessive compulsive disorder or generalized anxiety disorder. However, empirical
evaluation of this issue is rare, largely due to the lack of research in this area that makes use of a range of anxiety disorders.

While there is value in validating adult measures with younger populations, they carry the limitation of not being specifically developed with these populations in mind, and thus questions remain about their developmental sensitivity. Safety behaviors may be especially developmentally restrained. For example, it is possible that younger people may rely more on use of mobile technology or transitional objects while adults might rely more on carrying medications or using alcohol. Therefore, there is likely to be strong value in developing a measure of safety behaviors specifically for younger populations and developed from items generated by young people and their therapists. In addition, given high prevalence of comorbidity among young anxious populations (Rapee et al., 2009), relying on separate measures of safety behaviors that are relevant to specific forms of fear limits their clinical and theoretical utility. Further, although there has been some examination of the role of safety behaviors in social anxiety (as noted above), evaluation of this issue among youth with other anxiety disorders (e.g., separation anxiety, generalized anxiety) has been far less extensive, perhaps partly due to the lack of relevant measures of safety behaviors. In addressing these gaps in the literature, this paper aimed to develop and validate a new, broad measure of safety behaviors for use with young people, which we called the Subtle Avoidance Measure for Youth (SAMY).

The measure made use of young people with a range of anxiety disorders and their therapists to generate items addressing the full range of subtle behaviors and strategies that anxious young people use to feel better when they are anxious. Given the breadth of items and the range of different anxiety disorders that the youth reported, we predicted that the items would “subdivide” into two or more conceptually distinct factors. Based on the limited prior research with adults (Goodson et al., 2016; Soo & Rapee, 2019), we predicted a factor
reflecting socially-relevant safety behaviors and a factor reflecting protection from physical threats; and possibly a factor reflecting more general checking or vigilance behaviors. Diagnostically, we expected that children with social anxiety disorder would score highest on socially-relevant items, while children with separation anxiety disorder would score highest on items related to physical protection. If a third factor relevant to checking behaviors emerged, we predicted that it would characterize children with obsessive compulsive concerns. We had no strong predictions about children with generalized anxiety disorder due to the wide range of concerns that they report, but given the strong links between generalized anxiety and social anxiety disorder, we predicted that they would show the greatest similarities to children with social anxiety. Finally, given the relevance of safety behaviors for maintenance of anxiety disorders, we predicted stronger relationships between the overall safety measure and measures of anxiety than measures of unrelated constructs such as conduct disorder or hyperactivity. Some research has suggested that depressed individuals also engage in safety behaviors (Moulds, Kandris, Williams, & Lang, 2008) and we therefore expected depression to show a moderate relationship with the overall measure.

Method

Initial Item Generation

The initial items for the SAMY were generated using a multi-step process to incorporate feedback from both clinicians in the field and families with anxious children. Items were first derived by modifying items from existing measures of safety behaviors, specifically the Subtle Avoidance Frequency Examination (SAFE; Cuming et al., 2009) and the Modified Social Behavior Questionnaire for Children (M-SBQ-C; Kley et al., 2012). As existing measures at the time focused largely on social anxiety, additional items were generated from research in the broader anxiety literature and interviews with clinicians.
experienced in the assessment and treatment of child and adolescent anxiety disorders. Clinicians were asked to consider both cognitive and behavioral safety behaviors that they may have encountered in their practice.

These initial steps produced a preliminary pool of 35 items. Semi-structured interviews with anxious children and their parents were then conducted to obtain further feedback and input for the questionnaire. Five parents and their children (7-11 years) who were involved in treatment for their anxiety disorder were interviewed separately with a semi-structured interview. The interview was adapted and revised from a survey used by Kley et al. (2012) and involved open-ended questions regarding the child’s potential use of any overt or subtle strategies to reduce anxiety in various situations. A sample prompt was “Can you think of anything that you do (your child does) so that no one notices your fear?”

The children and parents also provided feedback about ambiguity and interpretation of the existing items. These interviews resulted in revision, removal and addition of several items. The eventual list of items was reviewed by three university staff and five clinical psychology doctoral students who had experience working with anxious children, resulting in a final list of 37 items for psychometric testing.

Participants

Participants for the study comprised 312 youth (162 male, 150 female), aged from 7 to 13 years, M = 9.55, S.D. = 1.87. The participants were recruited from two samples, a clinical sample and a community sample.

The clinical group consisted of 174 children (91 males, 83 females) who sought treatment at the Centre for Emotional Health at Macquarie University in Sydney, Australia for an anxiety disorder between 2014 and 2016. All clinical participants met criteria for one or more anxiety disorders according to DSM-IV criteria – the most common primary
disorders being Generalized Anxiety Disorder (44.9%), Social Anxiety Disorder (21.0%), and Separation Anxiety Disorder (13.0%). Diagnoses were made by trained clinicians at our Centre following structured interview with the Anxiety Disorders Interview Schedule for DSM-IV, Child version (Silverman & Albano, 1996) as part of the clinical trials for which they were recruited. Most clinical children came from European/Australian backgrounds (89%); attended public schools (61%), and lived with two biological parents (88%), who were mostly employed (80%).

The community group consisted of 138 children and adolescents (71 males, 67 females). This population was recruited through Girl Guides and Scouts NSW units across Sydney. They were not screened for mental disorders and were included if they fell within the age range and volunteered to participate.

The clinical (52.3% male) and community (51.4% male) samples did not differ significantly on sex distribution, $\chi^2(1, N=312) = .02, p = .881$, nor on age, (clinical, $M = 9.51$ years, $SD = 1.89$; control, $M = 9.62$ years, $SD = 1.85$), $t(310) = 0.52, p = .606$. The clinical group also scored significantly higher than the community group on symptoms of anxiety (Clinical – $M = 37.61$, S.D. = 16.09; Community – $M = 27.98$, S.D. = 15.58), $t(274) = 5.05, p<.001$.

Measures

All measures were child self-report to reduce any confounding between assessment of validity and informant disagreement.

37-item SAMY. All participants completed the 37-item SAMY by rating the frequency with which they would use the potential safety behaviors when they felt anxious on a five-point (0-4) Likert rating scale from never to always. Higher scores indicated greater use of the safety behaviors.
Spence Children’s Anxiety Scale (SCAS) (Spence, 1998). The SCAS comprises 38 items, scored 0-3, to assess anxiety symptoms among young people. The measure has six subscales reflecting generalized anxiety, social phobia, separation anxiety, panic/agoraphobia, obsessive compulsive disorder, and physical injury fears. The total scale and subscales have demonstrated internal consistency of 0.60 to 0.92, and test-retest reliability of 0.45 to 0.60 in both child and adolescent populations (Spence, 1998; Spence, Barrett, & Turner, 2003). Internal consistency for the total scale in the current sample was good (alpha = .90).

Children’s Automatic Thoughts Scales (CATS) (Schniering & Rapee, 2002). The CATS is a 40-item measure (rated 0-4) of negative self-statements reported by youth that contains four subscales of physical threat, social threat, personal failure, and hostility. The total scale and subscales have demonstrated internal consistency ranging from 0.85 to 0.95, and test-retest reliability ranging from 0.66 to 0.79 (Schniering & Rapee, 2002). Internal consistencies across the total and subscales for the current sample were strong (alphas = .91-.96).

Short Mood and Feelings Questionnaire (SMFQ) (Angold, Costello, Messer, & Pickles, 1995). The SMFQ is a measure of recent depressive symptoms for youth aged 6-17 years. It contains 13 items rated 0 to 2. The SMFQ has demonstrated good internal consistency (α = 0.85) and the ability to successfully discriminate between children in psychiatric and control groups (Angold et al., 1995). Internal consistency was strong for the current sample (alpha = .95).

Clinical participants only.

As a requisite of their participation in treatment, the clinical group also completed two additional measures.
**Strengths and Difficulties Questionnaire (SDQ)** (Goodman, 1997). The SDQ contains 25 items rated 0-2 and provides five subscales (conduct problems, hyperactivity, emotional symptoms, peer problems and prosocial behavior). In the current study, only the conduct problems, hyperactivity, and emotional symptoms subscales were used. The SDQ has demonstrated sound internal consistency (α = 0.73) and the ability to discriminate between children in the community and those attending a mental health clinic (Goodman, Meltzer, & Bailey, 1998). Internal consistency for the three relevant subscales in the current sample ranged from poor to adequate (alphas = .74 (emotional difficulties); .49 (hyperactivity); .28 (conduct problems)).

**Child Anxiety Life Interference Scale (CALIS)** (Lyneham et al., 2013). The CALIS contains 9 items rated from 0 to 4 that describe life interference and impairment associated with anxiety. It has demonstrated good internal consistency, (α = .80), moderate to strong positive correlations with measures of self-reported internalising symptoms (e.g., SCAS, SDQ-emotional symptoms), and sensitivity to treatment change. In the current sample, internal consistency was solid (alpha = .84).

**Procedure**

Approval for this study was granted by the Macquarie University Human Research Ethics Committee and all participants and their parents/caregivers gave informed consent for participation in the study. Families in the clinical group were initially screened by telephone and then completed questionnaires and diagnostic assessment prior to entering treatment or external referral. For the community population, the experimenter visited individual Girl Guides and Scouts NSW units throughout Sydney, NSW and provided an information session about the study and instructions on accessing the information page, consent form and questionnaires online. One hundred and thirteen community participants also completed the
same battery of questionnaires approximately 12-13 weeks after their initial completion. Families were reimbursed with a $15 gift voucher for their time and participation in the study.

**Statistical Analyses**

Following examination of item distributions, the factor structure of the SAMY was evaluated using exploratory factor analysis (EFA) and items that did not fit clearly were removed. The sample was randomly divided into two subsamples in order to initially conduct the EFA (n = 168) and subsequently replicate the factor structure using confirmatory factor analysis (CFA) (n = 144) (Fabrigar, Wegener, MacCallum, & Strahan, 1999). The total sample (N=312) was then utilized for psychometric analyses including reliability and validity, with clinical-only or control-only analyses used where relevant. Construct validity was determined by comparisons against anxiety-related measures and comparison between clinical and community participants. Discriminant validity was tested by comparison against unrelated constructs (measures of externalizing). SPSS version 23 was used for the EFA and psychometric analyses, and AMOS version 22 was used to conduct the CFA.

**Results**

**Item and sample characteristics.**

The original set of 37 items of the SAMY is described in Appendix A. Means for the 37 items in the EFA subsample ranged from 0.57 to 1.96, standard deviations ranged from 0.84 to 1.44, and the full range of the scale was used for all items. Skewness of individual items ranged from .07 to 1.68 and kurtosis ranged from -1.36 to 2.85. Given the high positive kurtosis value of one item (29. “Hide your face”, kurtosis = 2.85) which was above the cut-off value of 2, it was deleted from the item pool (West, Finch, & Curran, 1995).
An inter-item correlation matrix was produced and no items correlated too strongly ($r \geq \pm 0.8$) with each other (Field, 2013). Correlation coefficients between each item and the total questionnaire score were calculated, and as all items demonstrated item-total correlations greater than 0.3, no items were excluded from the scale based on this recommendation (Kline, 2014).

**Exploratory factor analysis (first subsample).** The remaining 36 items were subjected to a principal axis factor analysis with a Promax rotation and the interpretability of the factors was examined. Promax rotation was chosen as correlation among factors was expected and oblique rotation methods are highly recommended for this reason (Costello & Osborne, 2005).

The initial solution identified eight factors with eigenvalues greater than 1 (Kaiser criterion) and this solution explained 52.92% of the total variance. Examination of the scree plot suggested a break after the third or fourth factor. Hence, both three and four-factor analyses were explored. The three-factor solution, explaining 41.51% of total variance, was chosen as the most parsimonious and interpretable.

A threshold of .32 was implemented as a minimum loading for an item to be retained (Tabachnick & Fidell, 2001). This cut-off resulted in the removal of three items (11. “Play or fidget with something”, 17. “Avoid answering the phone or door” and 26. “Say that you are sick and/or visit school sick bay”) as they did not sufficiently load onto any factors. These items were removed and the EFA was recalculated. All items loaded onto one factor at greater than .32. However, three items (items 14, 15, and 30) cross-loaded onto additional factors and were removed. The EFA was recalculated on the remaining 30 items. An additional item (21. “Practice sentences you’ll say to others in your mind”) now failed to load sufficiently on any factor and was removed. In addition, two items (24. “Call or text parents frequently” and 19. “Imagine you are somewhere else”) were conceptually dissimilar from
the other items in the factors they loaded on and showed the weakest loadings on their respective factors (.32 and .37, respectively) and thus were removed.

From this process, a three-factor solution with 27 items offered a theoretically sound and parsimonious representation of the data, with the factors accounting for 46.18% of total variance. Factor loadings for the final 27-item SAMY are presented in Table 1. Examination of the items that loaded onto each of the three factors suggested relatively coherent categories of safety behaviors represented by each. The first factor appeared to reflect safety behaviors related to excessive checking tendencies and vigilance (labelled ‘Checking’). The second factor focused clearly on safety behaviors related to restrictive and active behaviors associated with social situations and protection from negative evaluation (labelled ‘Image management’). Items on the third factor reflected safety behaviors related to physical help and protection through external agents (labelled ‘Physical protection’).

**Confirmatory factor analysis (second subsample).** To confirm the factor structure of the 27-item SAMY, a CFA was conducted on the second half of the randomly split sample (n = 144). Multivariate tests of normality were conducted through AMOS and revealed significant positive skewness and kurtosis in the data. This was expected given the sample characteristics (i.e., inclusion of nonclinical individuals) and the nature of the questionnaire. In light of this, a bootstrap procedure was employed to reduce potential biases resulting from deviations from multivariate normality (Efron & Tibshirani, 1994). The Bollen-Stine bootstrap (500 samples; Bollen & Stine, 1992) was used to obtain a corrected p-value for the $\chi^2$ statistic. Model fit was also evaluated on the basis of a range of goodness-of-fit indices, including the Comparative Fit Index (CFI), (Bentler, 1990) the Incremental Fit Index (IFI) (Bollen, 1989), and the Root Mean Square Error of Approximation (RMSEA) (Steiger, 1990). Lastly, correlated error variances (i.e., those that represent non-random measurement
error) within the same factor were permitted, justified on the basis of content or wording (Brown, 2014).

The resulting model demonstrated reasonable fit; CFI = .89, IFI = .89, RMSEA = .07. The bootstrap $p$ value for the $\chi^2$ statistic was significant, $\chi^2 (309) = 516.47, p < .001$. However, the $\chi^2$ statistic is sensitive to sample size and can detect significant differences despite well-fitting models, and thus the relative $\chi^2$ statistic ($\chi^2 / df$) is often recommended as it is adjusted for sample size. The relative $\chi^2$ statistic for this model was 1.67, indicating good fit. Figure 1 presents the path diagram with the standardized estimates obtained from the CFA.

**Internal consistency.** Internal consistency of the 27-item scale and each of the three factors was calculated using Cronbach’s coefficient alpha on the full sample. The total scale demonstrated good internal consistency (α = .91), as did the three factors (checking α = .92, image management α = .84, physical protection α = .83).

**Discriminant validity.** Discriminant validity was evaluated by examining the ability of the SAMY to discriminate between clinically anxious and community participants. For clinical utility, participants were dichotomized into two age groups: younger children (7-9 years) and older children (10-13 years). Four, three-way ANOVAs were conducted to examine the effect of clinical status, age group, and sex on total SAMY scores as well as each subscale. For the total score on the SAMY, there was a significant main effect of clinical status, $F(1,304) = 61.15, p < .001, \eta^2 p = .167$, no significant main effect of age group, $F(1,304) = 1.13, p = .289, \eta^2 p = .004$, and no significant main effect of sex, $F(1,304) = 0.39, p = .533, \eta^2 p = .001$. There were no significant two or three-way interactions, all $F$’s < 1.50, all $p$’s > .26. A similar pattern of results was shown in the checking subscale - there was a
significant main effect of clinical status, $F(1,304) = 59.35, p < .001, \eta^2_p = .163$, no significant main effect of age group, $F(1,304) = 1.53, p = .217, \eta^2_p = .005$, and no significant main effect of sex, $F(1,304) = 0.31, p = .578, \eta^2_p = .001$. There were no significant two or three-way interactions, all $F$’s $< 0.50$, all $p$’s $> .48$. Similarly, for the image management subscale, there was a significant main effect of clinical status, $F(1,304) = 15.04, p < .001, \eta^2_p = .047$, no significant main effect of age group, $F(1,304) = 1.20, p = .274, \eta^2_p = .004$, and no significant main effect of sex, $F(1,304) = 0.11, p = .739, \eta^2_p < .001$. There were no significant two or three-way interactions, all $F$’s $< 1.35$, all $p$’s $> .25$. However, a different pattern emerged for the physical protection subscale. There was a significant main effect of clinical status, $F(1,304) = 37.26, p < .001, \eta^2_p = .109$, a significant main effect of age group, $F(1,304) = 7.90, p = .005, \eta^2_p = .025$, and a significant main effect of sex, $F(1,304) = 6.30, p = .013, \eta^2_p = .020$. There was also a significant interaction between clinical status and child sex, $F(1,304) = 5.35, p = .021, \eta^2_p = .017$, with girls reporting somewhat higher scores than boys among community participants, but not among clinical participants. The remaining interactions were not significant, all $F$’s $< 2.10$, all $p$’s $> .15$. Mean scores are presented in Table 2.

As an additional examination of relationships with specific diagnoses, scores on the three subscales of the SAMY were compared between participants diagnosed with a particular anxiety disorder, to those without that disorder only among clinical participants. First, three anxiety disorders, social anxiety disorder (SAD), separation anxiety disorder (SEP), and generalized anxiety disorder (GAD), were dummy-coded to indicate their presence or absence anywhere in the diagnostic profile. Scores on the subscales of the SAMY were then compared between these newly created groups (e.g., clinical participants with social anxiety disorder vs those without) using one-way ANOVAs. Data are presented in Table 3.
Participants with GAD scored significantly higher on the SAMY Checking subscale compared to those without, $F(1, 172) = 4.78, p = 0.030, \eta^2_p = .027$. However, the groups did not differ significantly, on SAMY image management, $F(1, 172) = 3.61, p = 0.059, \eta^2_p = .021$, nor in SAMY physical protection, $F(1, 172) = 1.29, p = 0.258, \eta^2_p = .007$. Participants with SAD did not score significantly higher on the SAMY image management subscale than those without, $F(1, 172) = 2.70, p = .102, \eta^2_p = .015$. The groups also did not differ significantly on SAMY checking, $F(1, 172) = 0.44, p = 0.834, \eta^2_p < .001$, nor in SAMY physical protection, $F(1, 172) = 0.14, p = 0.712, \eta^2_p = .001$. Finally, participants with SEP scored significantly higher on the SAMY physical protection subscale than those without, $F(1, 172) = 5.36, p = .022, \eta^2_p = .030$. The groups did not differ significantly on SAMY checking, $F(1, 172) = 0.23, p = 0.634, \eta^2_p = .001$, nor on SAMY image management, $F(1, 172) < 0.01, p = 0.978, \eta^2_p < .001$.

**Construct validity.** To examine convergent and divergent validity, the scores of the SAMY and its subscales were correlated with other administered measures. Results broadly were as expected (see hypotheses) and are presented in Table 4. In general, the largest (moderate to strong) correlations were shown against the measure of anxiety symptoms, the SCAS. The lowest correlations (nonsignificant) were shown against externalising scales including the hyperactivity and conduct problems subscales of the SDQ. Intermediate relationships were shown with the measure of depression (SMFQ).

As an additional evaluation of unique relationships, a multiple regression (entering all variables simultaneously) was conducted in which the SAMY total score was the dependent variable and predictors included the primary measure of anxiety (SCAS total), SMFQ, CALIS, and SDQ hyperactivity and conduct problems. The regression accounted for 10.7%
of variance in SAMY total scores, $F(1,140) = 16.77, p<.001$, and SCAS total was the only significant predictor, $B = 0.363$, S.E. = 0.09.

Correlations between the SAMY subscales and the subscales of the SCAS and CATS also demonstrated expected patterns. Image management showed its strongest relationships with SCAS social anxiety and CATS social threat, while physical protection showed its strongest relationships with SCAS separation anxiety and physical injury and CATS physical threat. The checking subscale showed its strongest relationships with the GAD, OCD, and physical injury subscales of the SCAS and interestingly, with the personal failure subscale of the CATS.

**Test-retest reliability.** Retest reliability was assessed approximately 12-13 weeks after the initial data collection with a subsample of community participants who returned a second completion of the measure (n = 113). Test-retest reliability was $r = .835, p < .001$ for the total scale, $r = .746, p < .001$ for the checking subscale, $r = .876, p < .001$ for the image management subscale, and $r = .900, p < .001$ for the physical protection subscale.

**Discussion**

The aim of the present study was to develop and validate a measure of safety behaviors specifically for use within pediatric populations. The resultant questionnaire, the Subtle Avoidance Measure for Youth (SAMY) comprised 27 items that loaded onto three clear factors, reflecting subtle avoidance and safety behaviors around checking, image management, and physical protection. Overall, the SAMY demonstrated promising characteristics and should fill an important niche of relevance to research and treatment of anxiety and related disorders among young people.
Examination of the factor structure of the SAMY with an unconstrained EFA provided evidence for a three-factor solution that reflected three different, but related, types of safety behaviors: checking, image management, and physical protection. This structure was broadly supported through a CFA, although some of the fit indices were slightly below optimum, suggesting the possibility that alternate factor structures may be identified in future. Nonetheless, the identified factors make good, face valid sense and it is unlikely that alternate factor structures will differ too extensively from the current one. Theoretically, it has been argued that safety behaviors should be related to the specific threat expectations of individuals (Salkovskis, 1991), perhaps predicting a myriad of idiosyncratic safety behaviors. However, similar factor structures to the one identified here have also been shown within studies with adult samples (e.g., Goodson et al., 2016; Soo & Rapee, 2019). An image management factor has been most commonly identified and the majority of research has focused on safety behaviors of direct relevance to social anxiety disorder (Burato et al., 2009; Cuming et al., 2009). Some evidence has also supported identification of safety behaviors related to physical threat (Kamphuis & Telch, 1998). Two attempts to create broad measures of safety behaviors for use with adults have demonstrated factor structures that are similar to the current one (Goodson et al., 2016; Soo & Rapee, 2019). Our measure with adults showed a two-factor structure reflecting image management (similar to the current measure) and tangible outcome management (including both checking behaviors and physical protection) (Soo & Rapee, 2019). The measure developed by Goodson et al. (2016) showed a very similar structure to the current one, reflecting evasion of social situations (image management), health related behaviors (physical protection), and monitoring and vigilance (checking). It appears that there is some degree of “substructure” between safety behaviors and this structure may be generally consistent across development. Naturally, future research
will be needed to directly compare the factor structure underlying safety behaviors across the developmental spectrum.

In general, the SAMY showed solid reliability. Internal consistency of both the total score and each subscale demonstrated substantial homogeneity among items within the safety behavior construct. Of particular relevance to both clinical use and research, the scale showed good test-retest reliability over a period of three months, demonstrating that the questionnaire provides a stable measure of safety behaviors in children and adolescents. Test-retest reliability was assessed only with control participants, given that clinical participants underwent treatment, and therefore we cannot be certain that similar retest reliability will be shown in clinical populations. However, the demonstrated reliability was similar to levels shown by symptom measures of child-reported anxiety (Arendt, Hougaard, & Thastum, 2014).

Discriminant validity was supported by the ability of the SAMY and its subscales to discriminate between clinically anxious and community participants. Further, the measure demonstrated good convergent and divergent validity, with stronger associations to anxiety-related measures and minimal relationships with measures of externalizing. Moderate relationships were shown with measures of depression, reflecting the suggestion that people suffering depression also engage in safety behaviors (Moulds et al., 2008) and with the broader measure of emotional symptoms. It was interesting to note the low to moderate relationship between the total SAMY score and the CALIS, which is a measure of the impact and life interference caused by anxiety. Logically, two competing hypotheses could be predicted for this relationship. On the one hand, engagement in safety behaviors is related with anxiety symptoms and might itself lead to life impairment, predicting a strong correlation between these measures. On the other hand, judicious use of safety behaviors might allow greater engagement with fear-provoking situations (Rachman, Radomsky, &
Shafran, 2008), thereby reducing life impairment. Perhaps the most likely prediction is that some safety behaviors (e.g., “avoid asking questions”) will directly impact functioning while the more subtle behaviors (e.g., “always carry a mobile”) might facilitate engagement in fear-provoking situations and reduce the impact of anxiety on functioning. Future research that focuses in more detail on the functions and impact of specific safety behaviors will help to untangle these ideas.

Interestingly, not only did the overall measure show the expected relationships demonstrating its validity, but theoretically predicted relationships were also shown for the separate factors (subscales). As predicted, each of the subscales discriminated logically between specific forms of anxiety. Image management was closely related to social anxiety, especially within the correlations with the SCAS and also with the social threat subscale of the CATS. This finding is consistent with a large body of research among both adults and youth that have focused on social anxiety and its relationship with socially-relevant safety behaviors (Moscovitch et al., 2013; Qasmieh et al., 2018; Taylor & Alden, 2010). Given the strong social-threat focus within generalized anxiety, it is not surprising that generalized anxiety also showed strong associations with image management. The checking subscale showed similar patterns. As predicted, it was most strongly related to generalized anxiety and to obsessive compulsive symptoms as assessed through the SCAS. While generalized anxiety is characterized by a broad range of worries, among youth it is perhaps most strongly grounded in social threat concerns and perfectionism. Therefore, the associations between generalized anxiety and both image management and checking safety behaviors is not surprising. Empirically, depression has shown its strongest associations with generalized anxiety within evaluations of the factor structure of mental disorders (Lahey, Krueger, Rathouz, Waldman, & Zald, 2017). Therefore, it was also not surprising to see significant relationships between image management, checking, and both symptoms of depression and
thoughts related to personal failure. At least one study with adults has noted that depressed individuals also report some use of safety behaviors, particularly related to avoiding or managing instructive memories (e.g., distracting oneself, blocking the memory or using alcohol) (Moulds et al., 2008). It is possible that a broader coverage of safety strategies might provide additional items that reflect a more “depressogenic” set of safety behaviors, but these items would be unlikely to emerge in a young sample such as ours. In contrast to the first two subscales, physical protection showed a slightly different pattern of relationships. The physical protection subscale was most strongly related to separation anxiety and to some extent with the physical injury subscale of the SCAS. It was also the only subscale that showed differences according to age and sex – higher scores among younger children and girls, especially among community participants. These differences likely reflect the content of the items – several items referred to use of parents as safety cues. Among adult measures, questionnaires assessing physical-threat relevant safety behaviors include items describing use of substances or reliance on a partner (e.g. Kamphuis & Telch, 1998). Future developmental research would benefit from closer examination of age-related changes in use of physical protection safety behaviors.

There were several strengths and limitations to this study. We sampled both clinically anxious and community populations to determine generalizability of results. The initially identified factor structure was replicated in a separate sample. However, although the sample size allowed for an item to participant ratio of 1:5 and was considered sufficient for exploratory factor analysis (Floyd & Widaman, 1995; Gorsuch, 1997), this ratio is in fact the minimum recommendation, with some authors advocating a ratio of 1:10 (Nunnally & Bernstein, 1994). Future validation studies would benefit from larger sample sizes to ensure the factorial stability of the SAMY. The internal consistency of the measure of divergent validity (SDQ) was quite poor in the current sample and it was also only completed by the
clinical participants. Therefore, the data on divergent validity need to be taken with caution and replication is required. Furthermore, both samples were comprised of children from middle socio-economic backgrounds in Australia and the community controls were recruited from scouting/guides organizations, which are unlikely to represent the typical Australian population. To increase generalizability of findings, replication in other geographical areas and socio-demographic groups and potentially in other cultures is recommended. Retest reliability was also only evaluated in the community sample and future research needs to evaluate retest reliability among clinical participants. Finally it would be valuable for future research to evaluate changes in use of safety behaviors following treatment and more importantly, to see whether treatment change is partially mediated by changes in use of safety behaviors among youth.

Despite these limitations the current study provided initial demonstration of the ability to reliably measure safety behaviors among young people. Given the solid psychometric properties of the SAMY and its subscales, the measure has the potential to inform treatment planning by identifying the types and extent of safety behaviors used by anxious young people presenting for treatment. Given the habitual and instinctive nature of safety behaviors, many individuals may struggle to identify their (or their children’s) behaviors without systematic prompting. Thus, the SAMY could be implemented as a component of psychoeducation and subsequent treatment planning. The measure can also be used throughout treatment to monitor possible threats to in vivo exposure although the ability of SAMY scores to predict response to exposure is yet to be tested.

Given the central role and clinical relevance of safety behaviors in the development and treatment of anxiety, there has not been sufficient research dedicated to the assessment of these behaviors among young people. This study has filled a critical gap in the literature and has provided the groundwork for the valid assessment of safety behaviors in children and
adolescents. While further work will help to refine and validate the measure, the SAMY is a promising tool for use in both clinical and research contexts.
Acknowledgement:

This research was supported by National Health & Medical Research Council grant APP1027556 to R Rapee. The authors would like to thank Dr Lauren McLellan for facilitating data collection.
References


doi:http://dx.doi.org/10.1016/j.jbtep.2008.08.003


DOI: 10.1016/j.janxdis.2013.09.008


**Table 1:**

Factor Loadings for the 27-Item Subtle Avoidance Measure for Youth in the Exploratory Factor Analysis Sample (n=168).

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>18. Return permission slips the day after receiving them</td>
<td>.843</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Check items in your bag to make sure you haven't forgotten things</td>
<td>.812</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25. Excessively study for tests and exams</td>
<td>.794</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36. Double-check instructions to make sure you’ve heard correctly</td>
<td>.789</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Check with parents/guardians or teacher</td>
<td>.762</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Always stick to rules</td>
<td>.702</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Check and re-check homework/school work</td>
<td>.695</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Keep an eraser for correcting mistakes</td>
<td>.695</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Always ask questions to make sure of things</td>
<td>.641</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Have back ups (e.g., extra pencils in exams, pack extra clothes in case you need to change)</td>
<td>.561</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32. Stand or sit in a way so people won’t notice you much (e.g., behind people; keep still; look busy)</td>
<td>.859</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31. Speak really fast and avoid pauses</td>
<td>.665</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33. Do anything to try not to shake (e.g., hold things tightly, clench fists)</td>
<td>.657</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28. Avoid asking questions (e.g., in class)</td>
<td>.650</td>
<td></td>
<td></td>
</tr>
<tr>
<td>37. speak softly/ in a whisper</td>
<td>.603</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27. Worry about or spend a long time on how you look</td>
<td>.531</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35. Don’t say what you really want to</td>
<td>.489</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Always carry a mobile</td>
<td>.426</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Look down/avoid eye contact</td>
<td>.425</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34. Shut your eyes</td>
<td>.394</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Keep quiet or try not to say much</td>
<td>.354</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Stand close to parent’s side/ hold on to parent’s clothes</td>
<td>.782</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. Sleep with a family member</td>
<td>.778</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Make sure parents/ guardians are always close by</td>
<td>.755</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. Use a night light, lamp, fan, etc at night</td>
<td>.679</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. Have your parents complete tasks on your behalf (e.g., saying things for you, collecting things for you)</td>
<td>.578</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Carry, wear, or have something special nearby (e.g., lucky charm, bracelet, blanket, toy)</td>
<td>.437</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2:
Mean scores between younger and older, clinical and community girls and boys on the total SAMY and its subscales.

<table>
<thead>
<tr>
<th></th>
<th>7-9 years</th>
<th>10-13 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Clinical</td>
<td>Community</td>
</tr>
<tr>
<td>Males</td>
<td>(n=54)</td>
<td>(n=31)</td>
</tr>
<tr>
<td>Females</td>
<td>(n=35)</td>
<td>(n=40)</td>
</tr>
<tr>
<td>Total SAMY 1</td>
<td>44.87</td>
<td>25.10</td>
</tr>
<tr>
<td></td>
<td>(17.88)</td>
<td>(9.58)</td>
</tr>
<tr>
<td>Checking</td>
<td>21.48</td>
<td>11.58</td>
</tr>
<tr>
<td></td>
<td>(9.76)</td>
<td>(8.38)</td>
</tr>
<tr>
<td>Image management</td>
<td>13.06</td>
<td>8.81</td>
</tr>
<tr>
<td></td>
<td>(6.82)</td>
<td>(6.04)</td>
</tr>
<tr>
<td>Physical protection</td>
<td>10.33</td>
<td>4.71</td>
</tr>
<tr>
<td></td>
<td>(5.11)</td>
<td>(4.83)</td>
</tr>
</tbody>
</table>

1. SAMY – Subtle Avoidance Measure for Youth
Table 3:
Mean scores (and standard deviations) on the Subtle Avoidance Measure for Youth subscales for clinical participants according to their dichotomized diagnostic groups.

<table>
<thead>
<tr>
<th></th>
<th>GAD</th>
<th>SAD</th>
<th>SEP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes (n=90)</td>
<td>No (n=84)</td>
<td>Yes (n=97)</td>
</tr>
<tr>
<td>SAMYch</td>
<td>20.93 (9.81)</td>
<td>17.90 (8.33)</td>
<td>19.34 (9.54)</td>
</tr>
<tr>
<td>SAMYim</td>
<td>14.27 (7.26)</td>
<td>12.14 (7.49)</td>
<td>14.06 (8.04)</td>
</tr>
<tr>
<td>SAMYpp</td>
<td>10.10 (5.15)</td>
<td>9.20 (5.29)</td>
<td>9.54 (5.44)</td>
</tr>
</tbody>
</table>

Note: GAD – Generalized anxiety disorder; SAD – Social anxiety disorder; SEP – Separation anxiety disorder; SAMYch – SAMY checking subscale; SAMYim – SAMY image management subscale; SAMYpp – SAMY physical protection subscale.
Table 4:

Correlations between the SAMY (total score and subscales) and Other Administered Measures.

<table>
<thead>
<tr>
<th></th>
<th>Total SAMY</th>
<th>Checking</th>
<th>Image management</th>
<th>Physical protection</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SCAS (n = 276)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>.57**</td>
<td>.41**</td>
<td>.51**</td>
<td>.44**</td>
</tr>
<tr>
<td>Separation</td>
<td>.36**</td>
<td>.16**</td>
<td>.21**</td>
<td>.58**</td>
</tr>
<tr>
<td>Social</td>
<td>.39**</td>
<td>.20**</td>
<td>.61**</td>
<td>.09</td>
</tr>
<tr>
<td>GAD</td>
<td>.44**</td>
<td>.35**</td>
<td>.39**</td>
<td>.29**</td>
</tr>
<tr>
<td>OCD</td>
<td>.45**</td>
<td>.44**</td>
<td>.31**</td>
<td>.28**</td>
</tr>
<tr>
<td>Panic/Agoraphobia</td>
<td>.42**</td>
<td>.27**</td>
<td>.45**</td>
<td>.28**</td>
</tr>
<tr>
<td>Physical injury</td>
<td>.41**</td>
<td>.39**</td>
<td>.20**</td>
<td>.37**</td>
</tr>
<tr>
<td><strong>CATS (n = 275)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>.47**</td>
<td>.30**</td>
<td>.53**</td>
<td>.26**</td>
</tr>
<tr>
<td>Physical threat</td>
<td>.41**</td>
<td>.26**</td>
<td>.36**</td>
<td>.39**</td>
</tr>
<tr>
<td>Social threat</td>
<td>.33**</td>
<td>.09</td>
<td>.61**</td>
<td>.09</td>
</tr>
<tr>
<td>Personal failure</td>
<td>.42**</td>
<td>.40**</td>
<td>.38**</td>
<td>.14*</td>
</tr>
<tr>
<td>Hostile intent</td>
<td>.33**</td>
<td>.21**</td>
<td>.34**</td>
<td>.22**</td>
</tr>
<tr>
<td><strong>SMFQ (n = 276)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>.26**</td>
<td>.19**</td>
<td>.28**</td>
<td>.12</td>
</tr>
<tr>
<td><strong>SDQ (n = 142)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional Symptoms</td>
<td>.37**</td>
<td>.30**</td>
<td>.35**</td>
<td>.22**</td>
</tr>
<tr>
<td>Conduct Problems</td>
<td>-.03</td>
<td>-.15</td>
<td>.15</td>
<td>-.07</td>
</tr>
<tr>
<td>Hyperactivity</td>
<td>.05</td>
<td>-.02</td>
<td>.16</td>
<td>-.03</td>
</tr>
<tr>
<td><strong>CALIS (n = 142)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>.22**</td>
<td>.09</td>
<td>.27**</td>
<td>.19**</td>
</tr>
</tbody>
</table>

*p<.05  **p<.01

Note: SAMY – Subtle Avoidance Measure for Youth; SCAS – Spence Children’s Anxiety Scale; GAD – Generalized Anxiety Disorders (subscale); OCD – Obsessive Compulsive Disorder (subscale); CATS – Children’s Automatic Thoughts Scale; SMFQ – Short Mood and Feelings
Questionnaire; SDQ – Strengths and Difficulties Questionnaire; CALIS – Children’s Anxiety Life Interference Scale.
Figure 1. Confirmatory factor analysis path diagram of the three-factor model of the 27-item Subtle Avoidance Measure for Youth. All factor loadings are significant at $p < .001$.

Note: SAMY – Subtle Avoidance Measure for Youth
Appendix A

Full set of tested items from the Subtle Avoidance Measure for Youth (SAMY)

1. Carry, wear, or have something special nearby (e.g., lucky charm, bracelet, blanket, toy)
2. Always carry a mobile
3. Stand close to parent’s side/hold onto parent’s clothes
4. Make sure parents/guardians are always close by
5. Look down/avoid eye contact
6. Always ask questions to make sure of things
7. Check with parents/guardians or teacher
8. Keep an eraser for correcting mistakes
9. Have back ups (e.g., extra pencils in exams, pack extra clothes in case you need to change)
10. Check items in your bag to make sure you haven’t forgotten things
11. Play or fidget with something (e.g., crack knuckles, twirl hair, bite nails)
12. Keep quiet or try not to say much
13. Always stick to rules
14. Rehearse what you’ll say or do in your head
15. Have your parents check homework
16. Check and re-check homework/school work
17. Avoid answering phone or door
18. Return permission slips the day after receiving them
19. Imagine you are somewhere else
20. Use a night light, lamp, fan, etc at night
21. Practice sentences you’ll say to others in your mind
22. Sleep with a family member
23. Have your parents complete tasks on your behalf (e.g., saying things for you, collecting things for you)
24. Call or text parents frequently to check
25. Excessively study for tests and exams
26. Say that you are sick/unwell and/or visit school sick bay
27. Worry about or spend a long time on how you look
28. Avoid asking questions (e.g., in class)
29. Hide your face (e.g., behind people; with a hat)
30. Try to think about other things
31. Speak really fast and avoid pauses
32. Stand or sit in a way so people won’t notice you much (e.g., behind people; keep still; look busy)
33. Do anything to try not to shake (e.g., hold things tightly, clench fists)
34. Shut your eyes
35. Don’t say what you really want to
36. Double-check instructions to make sure you’ve heard correctly
37. speak softly/ in a whisper