

Development and Validation of the School Refusal Evaluation Scale¹ for Adolescents

Marie Gallé-Tessonneau,¹ PhD and Kamel Gana,² PR

¹Research Center INSERM U1219, Bordeaux Population Health, University of Bordeaux and, ²Department of Psychology, University of Bordeaux

All correspondence concerning this article should be addressed to Marie Gallé-Tessonneau, PhD, Department of Psychology, Bâtiment H, University of Bordeaux, 3 ter place de la Victoire, 33076 Bordeaux Cedex, France. E-mail: marie.galle-tessonneau@hotmail.fr.

¹SCREEN is the acronym of SChool REfusal EvaluatiON. The scale is available from the first author.

Received December 19, 2017; revisions received July 12, 2018; accepted July 12, 2017

Abstract

Objective School refusal (SR) is considered as a clinical form of emotionally based school absenteeism related to the experience of strong negative emotions when attending school. However, there are no psychometrically sound measures specifically designed to assess SR. The aim of the present research was to develop and validate a multidimensional self-report measure of SR: the SChool REfusal EvaluatiON Scale (SCREEN) for adolescents. **Methods** Two studies were performed to develop this scale. Study 1 used content analysis on 42 semi-directed interviews to identify and organize the indicators of SR and to generate items. The item pool was then reviewed by experts to construct a pilot scale. Study 2 examined the psychometric properties of this pilot scale of the SCREEN in a community sample of 649 French adolescents (age: 10–16 years) and a clinical sample of 31 adolescents. **Results** The initial SCREEN structure was identified via exploratory factor analysis, and the resultant model was validated using confirmatory factor analysis. In its final version (an 18-item, four-factor measurement model), the SCREEN reliably measures four interrelated aspects of SR. Data suggest good reliability and validity of scores on this SR measure. Results of the receiver operating characteristic (ROC) curve analysis indicated good accuracy of the SCREEN (sensitivity = .94 and specificity = .88). **Conclusions** The SCREEN offers a new and valid measure to assess SR in research and clinical settings.

Key words: school absenteeism; school attendance problem; school phobia; school refusal; school refusal assessment scale; school refusal evaluation; school refusal scale; truancy.

Introduction

Beyond the acquisition of knowledge, school plays an important role in the development of social, social-emotional, and civil skills (Kearney & Graczyk, 2014). However, it also hosts the expression of adolescents' psychosocial disorders, notably school violence, school burnout, truancy, academic anxiety, and school refusal (SR). The reasons behind school absenteeism are manifold, encompassing family-related, health-related, emotion-related factors (Inglés, González-Maciá, García-Fernández, Vicent, & Martínez-Monteagudo, 2015; Maynard, Salas-Wright, Vaughn,

& Peters, 2012), and school-related factors such as school environment, school climate, and class climate (Havik, Bru, & Ertesvåg, 2015a; Hendron & Kearney, 2016; Simons, Hwang, Fitzgerald, Kielb, & Lin, 2010).

Although it is difficult to estimate, SR is considered to affect 1–5% of school-aged children (Egger et al., 2003; Havik et al., 2015a; King & Bernstein, 2001; Nair et al., 2013), boys and girls being almost equally concerned (King & Bernstein, 2001), with a higher prevalence during adolescence than childhood (Gallé-Tessonneau, Doron, & Grondin, 2017).

School absenteeism has several consequences, including poor school performance, increased risk for unemployment, impaired social functioning, and mental health problems in adulthood (Havik et al., 2015a; Kearney, 2008; McCune & Hynes, 2005; McShane, Walter, & Rey, 2004). The early identification of SR is crucial because a positive prognosis depends on early clinical care (Okuyama, Okada, Kuribayashi, and Kaneko, 1999). The absence of instruments allowing early screening of SR can cause delays in treatment delivery. This is detrimental to the prospect of school return; therefore, the present study sought to establish a novel tool to assess SR in adolescents, namely, the School REfusal EvaluationN Scale (SCREEN).

Although further research should be undertaken on the development of specific measurement tools for the early detection of SR in children, this study primarily focuses on adolescents given that SR is more prevalent in adolescence than childhood (Gallé-Tessonau et al., 2017). The following sections address issues of SR conceptualization and measurement and provide rationale for the development of a new SR scale (Holmbeck & Devine, 2009).

Conceptual Clarification

It is noteworthy to mention that the first step of scale construction, as provided by DeVellis (2011) in his “Guidelines in scale development,” is “Step 1: Determine clearly what it is you want to measure” (p. 73).

At least three meaningful distinctions have been drawn between the four key concepts used to refer to school avoidance: school phobia, school truancy, SR, and school refusal behavior (SRB).

First of all, although school phobia and SR are often used interchangeably to refer to irrational fear of attending school (Kearney, 2008), conceptual differences emerged between these concepts (see Kahn and Nursten, 1962). Since then, the term SR was recommended because it takes into consideration the heterogeneity of the causes of the phenomenon, and because it is a broader and more inclusive concept than school phobia, which is being used less frequently in recent research literature (Kearney, 2008). Even Johnson (1957), who coined the term “school phobia” in 1941 (Johnson, Falstein, Szurek, & Svendsen, 1941), recognized in 1957 that “school phobia is a misnomer” (p. 307). Second, some researchers made a distinction between SR and SRB (see Heyne, Gren Landell, Melvin, & Gentle-Genitty, 2018). SR refers the experience of strong negative emotions when a young person is at school or faced with the prospect of going to school (Berg, 1997), while “school refusal behavior is an umbrella term that subsumes constructs such as truancy, school refusal, and school phobia” (Kearney, 2008, p. 452). Third, a distinction between school truancy and

SR has long been drawn. SR is differentiated from school truancy on the grounds that the former is related to the experience of strong negative emotions when attending school or faced with the prospect of going to school, and the latter is characterized by poor motivation for school or a negative attitude toward school, and a tendency to seek more pleasurable activities outside of school during school time (Heyne et al., 2018). Thus, in the present study, we conceptualize SR as a clinical form of emotionally based school nonattendance related to the experience of strong negative emotions when attending school. These negative emotions are often the source of the absenteeism at school, which occurs through, inter alia, being often tardy to school, visiting the school infirmary frequently and/or the school office, and calling parents to leave school and return home.

Clinical Manifestations of SR

Although SR is not classified as an independent diagnostic category in the international classification systems (i.e., DSM-5, ICD-10), its clinical criteria have been well documented for a long time. Berg, (1992, 1997; Berg et al., 1969) presented five clinical criteria of SR: (1) Reluctance, severe difficulty in attending school, or frequent refusal to go to school; (2) Seeking the comfort and security of home, preferring to remain close to parental figures, especially during school hours; (3) Displaying evidence of emotional upset and distress when the adolescent is faced with the prospect of having to go to school, although this may take the form of unexplained physical symptom (including dizziness, headaches, nausea, back pain); (4) Absence of serious antisocial tendencies, apart from possible aggressiveness when attempts are made to force school attendance; (5) The problem is not hidden to parents, who have made reasonable efforts to ensure the presence of the child in school throughout the history of the problem.

There is an absolute consensus that the psychological symptoms of SR are often accompanied by complaints of somatic symptoms (Bernstein et al., 1997; Egger et al., 2003; Havik, Bru, & Ertesvåg, 2015b; Nair et al., 2013). Often these symptoms are present on mornings before school, and sometimes they occurred the nights before. They disappear when the adolescent is allowed to stay at home.

Measurement of SR

To our knowledge, there are no psychometrically sound measures specifically designed to assess SR. However, there are three instruments used currently to capture some constructs close to SR: (1) the “school phobia” subscale of the Screen for Child Anxiety Related Emotional Disorders (SCARED, Birmaher et al., 1997), (2) the School Refusal Assessment Scale-Revised

(SRAS-R; Kearney, 2002, 2006), which is a revised version of the School Refusal Assessment Scale (Kearney & Silverman, 1993), and (3) the Reasons of School Non-Attendance Scale, recently developed by Havik et al., (2015a) to assess functional types of reasons underlying school nonattendance.

A serious limitation of these scales is that they were not developed specifically to assess SR. Their items were often generated from theory, without empirical validation which can be problematic in particular with a broad construct such as SR. Because a substantial body of literature underlined that distinction between SR and other related constructs, such as truancy, is important (Havik et al., 2015b, 2015a; Heyne et al., 2018), we believe that it is crucial to use an appropriate approach to develop a measure for accurately assessing SR.

Purpose of the Present Research

The aim of the present research was the construction and the validation of a self-report measure of SR. To have an appropriate approach of SR assessment, it seems crucial to adopt a multisource information gathering process. Along with the theoretical-conceptual approach, an empirical approach of the phenomenon was initiated in this research. Indeed, students themselves (Gregory & Purcell, 2014), child psychiatry professionals, school professionals (including school nurses), family physicians, and pediatricians have all been involved in defining the construct and elaborating its indicators and items. A multistep approach was used (Gallé-Tessonneau, Grondin, Koleck, & Doron, 2018). Study 1 describes the initial development, and Study 2 the validation of the SCREEN for adolescents.

This research received ethical approval from the independent Committee of the Protection of Persons.

Study 1. Scale Construction

This study was designed to identify and organize the indicators of SR, and to develop the novel scale. A two-stage approach was followed: (1) identify relevant manifestations of SR using the content analysis method, and (2) generating an item pool, examining the content validity of these items using experts reviewing, and then scale construction.

Stage 1: Identify Relevant Manifestations of SR Using Content Analysis Method Participants

Forty-two participants from the French region of Aquitaine were derived from four groups: (a) middle school professionals ($n = 17$; e.g., teachers, school director, and school nurse), (b) middle school students attending public school ($n = 15$; 11–16 years old, $M = 13.7$ years, $SD = 1.3$), (c) care professionals ($n = 4$; i.e., psychiatrists and psychologists) working with adolescents displaying SR, and (d) adolescents

displaying SR ($n = 6$; 11–16 years old; $M = 13.6$ years, $SD = 1.7$). Adolescents displaying SR were recruited by psychiatrists according to Berg's criteria of SR.

Material and Procedure

Semi-structured interviews were conducted using an interview guide based on a literature review of absenteeism and indicators of SR (Kearney, 2008; King & Bernstein, 2001) as well as advice from experts and researchers (e.g., What do you think are the behaviors of adolescents displaying SR at school? According to you, how can we detect these students earlier? How would you describe a student displaying SR? According to you, what are the first signs of SR?). These interviews were preparatory pilot-tested with two school professionals and four middle school students. All the interviews were conducted by an experienced psychologist. The interviews were recorded and lasted between 15 and 45 min.

Content Analysis

A corpus of 42 interviews was analyzed with a two-step thematic content analysis. First, an inductive content analysis was performed on 70% of the corpus, and second, a deductive content analysis was conducted on the remaining 30%. The categories were created by consensus between the first author of this article and one postgraduate psychology student. A dual coding, performed on 20% of the corpus revealed a good inter-coder reliability, as indicated by Cohen's kappa coefficient ($\kappa = .84$).

Results

Total 4 themes and 10 subthemes were identified through the content analysis: (1) "school" (subthemes: "going to school," "inter-class time," "in-class time," and "absenteeism"), (2) "daily and family life" ("daily life at home" and "youth and parent responses"), (3) "health" ("medical issues" and "psychological vulnerabilities"), and (4) "socialization outside home and family" ("lack of peer interactions" and "lack of community interactions").

Stage 2: Item Generation, Item Selection by Experts (Content Validity), and Scale Construction Item Generation

Based on the verbatim records of the model generated on Stage 1, word frequency analysis allowed identification of the 10 most frequent words in each subtheme. These were used as indicators to generate an item pool. Thus, a set of 64 items was created and worded in the first person.

Item Selection and Content Validity of the Item Pool

As recommended by DeVellis (2011), to maximize the content validity of the scale, the item pool was

reviewed by experts. A panel of 11 care professionals (psychiatrists and psychologists) and 11 adolescents displaying SR (age: 11–16 years old; $M = 14.3$, $SD = 1.6$) was asked to rate (0–10) each item with respect to its relevance for the SR construct. Experts were also invited to comment on these items and to make suggestions.

The items that were at the 60th percentile or higher of the rating were retained (statistical approach): items with a rating above 7.2 by the professionals and above 6.63 by the adolescents. Thus, 35 items were selected (20 items by both professionals and adolescents and 15 by one of the two groups). In addition, we ensured that all subcategories were represented by at least one item (theoretical approach; eight items).

The experts' comments were analyzed: five items were reworded for better comprehension and one item was removed as redundant with another item. Six items were reversed to maximize the comprehension. In total, 42 items were retained.

Items' Clarity and Comprehensibility of the 42-Item Pilot Scale

The 42-item pilot scale using a five-point Likert, from 1 (*not at all like me*) to 5 (*much like me*), format response was administered to 12 students (11 years old, first year of middle school). They were asked to examine the clarity and the comprehensibility of the instruction of the scale, the response format, and each item formulation. Identifying respondents' cognitive burden while reading and understanding the item helps improving its clarity (Tourangeau, Rips, & Rasinski, 2000).

Our interviews and probing questions revealed that comprehension of the instructions of the scale and the clarity of the items were good (no issues were raised by participants on the wording of the items or on scale instruction). Thus, the 42-item pilot scale of School REfusal EvaluationN for adolescents (SCREEN) was subjected to the validation process in the next step.

Study 2. Initial Validation of the SCREEN

This study was designed to examine the psychometric properties of the SCREEN by (1) exploring its factor structure, through an exploratory factor analysis (EFA) in a sample of middle school students. EFA allows the removal of redundant items and provides a statistical examination of the structure of the item pool; (2) cross-validating, through a confirmatory factor analysis (CFA), its factor structure in a sample of middle school students; (3) analyzing its items and assessing the reliability of its score (using composite reliability index, CRI); (4) examining its construct validity (i.e., convergent, divergent validity), using correlation coefficient and coefficient of determination, r^2 ,

(Steiger & Ward, 1987) to determine the level of overlap between SCREEN scores and scores in other measures tapping constructs close to SR, and (5) establishing a cutoff value to determine a clinically significant score (using receiver operating characteristic, ROC, curve).

Participants and Procedure

Community Sample

Total 624 adolescents (58% girls) 10–16 years old ($M = 12.4$ years, $SD = 1.3$), from six French public middle schools, participated in this study. With regard to family characteristics, the sample is constituted of two-parent families (68.5%), families with separated parents (29.5%), and families with a deceased parent (2%). Half of the participants' mothers (53%) had a level of education above the bachelor's degree (9% no diploma, 16% below the bachelor's degree, 15% bachelor's degree, and 7% unknown). Fathers of the participants in our sample came from various socio-professional categories, particularly the middle (42%) and upper (31%) classes (1% retirement, 6% unemployment, 7% laborer, and 13% unknown). Adolescents took part on a voluntary basis after their parents had previously given explicit written consent. The questionnaire was group administered in a classroom setting in the six public high schools. The questionnaire was completed on paper form or on a computer. As for other studies, Birnbaum (2000) found there was no significant difference between paper or computer form for all the measures used in this study.

This sample had been randomly divided into two subsamples (DeVellis, 2011): (1) the first subsample ($n = 420$, 58% girls, age: 10–16 years old, $M = 12.2$ years, $SD = 1.2$) was used to explore the factor structure and item analysis of the SCREEN; (2) the second subsample ($n = 204$, 58% girls, age: 10–16 years old, $M = 12.7$ years, $SD = 1.3$) was used to cross-validate the factor structure (CFAs). To split the sample, we used the "rule of 10," referring to the minimum sample size in factor analysis, which should be at least 10 observations per item in the measure being analyzed (i.e., 42 items * 10 = 420 participants in the first subsample).

The construct validity of the SCREEN and ROC curve analyses were conducted on the total sample ($n = 624$ adolescents; age: 10–16 years old; $M = 12.4$, $SD = 1.3$).

Clinical Sample

Total 31 adolescents (15 boys and 16 girls) 10–16 years old ($M = 13.3$ years, $SD = 1.5$) displaying SR. They were recruited in five mental health centers according to Berg's criteria of SR (independent criterion). Adolescents took part on a voluntary basis after

their parents had previously given explicit written consent. They completed the protocol at home or in the mental health-care center. The clinical group was recruited to run ROC curve analysis.

Each participant (in both community sample groups and clinical sample) filled out the same protocol.

Measures

In addition to some demographic variables (e.g., age, gender, and school grade), the participants filled out the following measures.

The SCREEN. A 42-item pilot scale developed to assess SR (Study 1). Each item was rated on a five-point Likert scale ranging from 1 (*not at all like me*) to 5 (*much like me*), with higher scores indicating severe SR.

The Screen for Child Anxiety Related Emotional Disorders-R. Only the “school phobia” dimension (four items) of this scale was used in the present study (Muris, Merckelbach, Brakel, and Mayer, 1999; Muris, Merckelbach, Schmidt, and Mayer, 1998). Each item was rated on a three-point scale, where 0 = *almost never*, 1 = *some time*, and 2 = *often*, with higher scores indicating higher school phobia. The Screen for Child Anxiety Related Emotional Disorders-R (SCARED-R) displayed good validity and reliability estimates in the general population and clinical population (Crocetti, Hale, Fermani, Raaijmakers, & Meeus, 2009; Muris, Merckelbach, Schmidt, and Mayer, 1998). Cronbach’s alpha reliability for our study was .66, which is similar to those obtained in others studies (Cronbach’s alpha from .66 to .70, Bouvard et al., 2013; Crocetti et al., 2009). This scale was used to assess the convergent validity of SCREEN scores. Thus, we hypothesized that SCREEN scores would be highly related to school phobia.

The SRAS (SRAS Kearney & Silverman, 1993) is a self-report measure containing 16 items that assess the function(s) of adolescent’s absenteeism. The SRAS determines the relative strength of four functions of SRB (see Introduction). Each item was rated on a seven-point scale ranging from 0 (*never*) to 6 (*always*). Coefficients alpha in this sample were, respectively, .74, .73, .55, and .57 for each dimension, which is close to European results (Cronbach’s alpha .66, .67, .67, .29, Brandibas, Jeunier, Gaspard, & Fourasté, 2001). This scale was used to assess the convergent validity of SCREEN scores. Thus, we hypothesized that SCREEN scores would be highly related to the first three functions of school absenteeism and less related to the fourth one. Despite the fact that “school phobia” subscale in SCARED-R and SRAS are not targeted specifically at assessing SR, they are the closest measures, likely to capture some aspects of SR.

Therefore, we hypothesized positive but not perfect correlations between these scales and SCREEN scores.

The Child Behavior Checklist-Youth Self-Report (CBCL-YSR; Achenbach, 1991) is a 112-item self-report measure widely used to assess emotional and behavioral problems (i.e., clinical syndromes) among adolescents of age 11–18 years (Van Meter et al., 2014). We used a computerized scoring program to obtain scores for the following syndrome scales: withdrawn, somatic complaints, anxious/depressed, social problems, thought problems, attention problems, delinquent behavior, and aggressive behavior. Also, the CBCL-YSR provides scores for the total behavior problems. Each item was rated on a three-point scale ranging from 0 (*not true*) to 2 (*very true or often true*). Because of computer scoring program standardized for age and sex, internal consistency is not available, but Cronbach’s coefficient alpha for the French YSR scales ranged from .83 to .92 (Wyss, Voelker, Cornock, & Hakim-Larson, 2003). This scale was used to assess construct validity. Because SR, as we conceive it, is an emotional and behavioral problem, we hypothesized that SCREEN scores would be particularly strongly related to total behavior problems as well as anxious-depressed and internalizing problems.

Absenteeism. Absenteeism was assessed in two ways: absenteeism *at* school and absenteeism *from* school. To assess absenteeism *at* school, adolescents were asked to report the number of times in the past 3 weeks they (1) went to the school infirmary, (2) went to the school office, (3) called parents to leave school and return home, and (4) delay. It was a quantitative subjective measure. To assess absenteeism *from* school, for each participant, we recorded the total number of half-days they were absent within the past 3 weeks. The data were provided by the school administration. Because SR cannot be reduced to absenteeism, we hypothesized that SCREEN scores would be modestly related to absenteeism.

Data Analysis

The first subsample ($n = 420$) was used to explore the initial version of the SCREEN (42 items) through a principal axis factor analysis with oblique rotation. Before performing the factor analysis, we checked the factorability of the items. Bartlett’s test of sphericity (χ^2 (861, $N = 420$) = 6,012.852, $p < .01$) and the Kaiser–Meyer–Olkin test of sampling adequacy ($KMO = .88$) indicated that our data were suitable for factor analysis. Several criteria were applied to determine the number of factors that should be retained: (1) the eigenvalues criterion (i.e., eigenvalues > 1), (2) the “proportion of variance accounted for” criterion (i.e., minimum 5% of the variance), (3) the “number of items per factor” criterion (i.e., at least four items per factor, with significant factor loading

Table I. Factor Loading of SCREEN Items (A Final four-factor Solution)

Item (number)	F1	F2	F3	F4
–I can't explain why I can't go to school (3)	.61	.19	.00	–.01
–When I get to school, I don't feel well when it comes time to go into the building (4)	.72	–.01	–.02	–.02
–I feel like I have a mental block when it comes to going to school, like I won't be able to (7)	.75	–.04	–.02	.08
–I'm afraid to go to school (11)	.77	–.03	.09	–.04
–On the way to school, I don't feel well (18)	.77	–.06	–.02	.01
–I tell my parents that I don't want to go to school and I want to stay at home (2)	.28	.52	–.06	.07
–In the morning, I don't want to go to school (8)	.09	.78	–.09	–.02
–I have a hard time going back to school after school vacation (12)	–.22	.74	.13	.00
–It's hard for me to go back to school on Mondays (16)	–.03	.69	.05	–.04
–I'm afraid of what others in my class think of me (1)	–.04	.05	.78	–.04
–I'm scared of doing a bad job in class (5)	.02	.12	.45	.17
–I'm fragile and sensitive (10)	–.04	–.13	.53	.07
–In the classroom, I feel good, I feel comfortable ^a (14)	.12	.14	.43	–.09
–I feel very comfortable with my classmates at school ^a (15)	.14	–.18	.53	–.05
–I'm absent more often this year than last year (6)	–.12	.02	.02	.57
–I often go to the school infirmary or administration office because I don't feel well (9)	.16	–.03	–.02	.54
–I'm frequently absent because I don't feel well (13)	–.05	–.05	.00	.57
–I often ask that my parents be called to come pick me up when I'm at school (17)	.04	.02	–.07	.51
Explained variance (%)	16	11	9	8
Factor intercorrelations				
Anxious anticipation		.49*	.49*	.40*
Difficult transition			.29*	.32*
Interpersonal discomfort				.21*

Note. $N = 420$; $*p < .05$. F1: Anxious anticipation; F2: Difficult transition; F3: Interpersonal discomfort; F4: School avoidance.

^aReversed items.

The English version of the SCREEN was obtained through a translation and back-translation procedure as recommended by the International Test Commission guidelines.

$\geq .40$), and (4) parallel analysis scree plots test (break point). Bootstrap method was used to improve the model's reliability. Items were selected if factor loading $> .40$ and at least a difference of $.20$ when cross-loadings among factors occurred.

The second subsample ($n = 204$) was used to cross-validate through a CFA the solution obtained through the EFA.

The reliability of the scores obtained by our participants was assessed using CRI (Raykov, 1997). For validity correlation, correlations $< .30$ were considered small, correlations between $.30$ and $.49$ were considered medium, and correlations of $\geq .50$ were considered large (Cohen, 1988).

The clinical sample was used to determine clinically significant score on the SCREEN, through a ROC curve analysis.

Results

EFA and Item Analysis of the 42-Item Pilot Scale of the SCREEN for Adolescents

As indicated in Table I, a four-factor solution was retained accounting for 44% of the item variance after five iterations. This solution is composed of 18 items (Table I), which factor loadings ranged from $.43$ to $.78$. As expected, factors are significantly intercorrelated (correlations ranged from $.21$ to $.49$; see Table I). The first factor (five items) reflects *Anxious*

anticipation when the adolescent is faced with the prospect of having to go to school. Indeed, the items refer to being afraid to go college, the path and the difficulty of returning to the school, and the inability to put into words this anxiety. The second factor (four items) represents *Difficult transition* between home and school. The items refer to the separation from home and relational issues between child and parents about it. The third factor (five items) reflects the adolescent's *Interpersonal discomfort* in relationships at school. The items refer to the adolescent's malaise in their relations with others (at school) and to the impact of these relations on his self-perception. The fourth factor (four items) represents *School avoidance*, through two types of behaviors: absenteeism *from* school and absenteeism *at* school (avoidance behavior such as visits to the infirmary or call parents).

Cross-Validation of the Factor Structure: CFA

A CFA was used to cross-validate the factor structure obtained through the EFA (i.e., an 18-item, four-factor model; $n = 204$). CFA requires multivariate normality of the sample data distribution. Normality of the items of the SCREEN was examined using the skewness and Kurtosis scores. As seen in Table II, the univariate skewness values were all significant, and except for five items, the univariate kurtosis values were also significant. Thus, the multivariate normality was affected: Mardia's coefficient was equal to 258.42

Table II. Item Analysis of the SCREEN

Item number	Mean (SD)	Corrected item–test correlation	Skewness	Kurtosis
Item 3	1.35 (0.873)	.649	2.715	6.813
Item 4	1.31 (0.872)	.492	3.213	9.845
Item 7	1.24 (0.708)	.491	3.518	13.020
Item 11	1.18 (0.615)	.542	3.979	16.647
Item 18	1.19 (0.561)	.531	3.946	18.934
Item 2	1.48 (0.909)	.527	2.023	3.470
Item 8	1.89 (1.26)	.520	1.371	0.713 ^a
Item 12	3.57 (1.44)	.421	−0.455	−1.274
Item 16	2.44 (1.40)	.434	0.664	−0.879
Item 1	2.09 (1.30)	.402	1.004	−0.229 ^a
Item 5	2.25 (1.20)	.375	0.646	−0.613 ^a
Item 10	2.11 (1.24)	.366	0.919	−0.245 ^a
Item 14	2.38 (1.27)	.497	0.722	−0.486 ^a
Item 15	2.10 (1.27)	.363	1.040	−0.027 ^a
Item 6	1.44 (1.00)	.248	2.417	4.992
Item 9	1.29 (0.640)	.388	2.294	4.798
Item 13	1.19 (0.517)	.485	3.049	9.439
Item 17	1.19 (0.577)	.414	3.362	12.635

Note. ^aNonsignificant.

indicating significant ($p < .001$) violation of normality. Consequently, we opted for MLMV estimator available in *Mplus* 7 software (Muthén & Muthén, 2010) to assess the fit of the measurement model underlying the SCREEN scores. The chi-square statistic was reported here, along with the comparative fit index (CFI), the Tucker–Lewis index (TLI), the standardized root mean squared residual (SRMR), and the root mean square error of approximation (RMSEA) and its 90% confidence interval (90% CI). A nonsignificant χ^2 is indicative of perfect fit. The CFI and TLI vary along a 0-to-1 continuum in which values $>.90$ and $.95$ are typically taken to reflect acceptable and excellent fits to the data, respectively (Marsh, Hau, & Wen, 2004). RMSEA values $<.05$ suggest a good model fit, and $<.08$ an acceptable model fit (Browne & Cudeck, 1993). The lower limit should be close to 0, while the upper limit should be $<.08$, and an SRMR value of $\leq .08$ indicates good model fit (Hu & Bentler, 1999).

The four-correlated factor model yielded mixed results ($\chi^2 (129, N=202) = 164.164, p = .020$; RMSEA = .037 [90% CI = .016, .053]; SRMR = .066; CFI = .878; TLI = .856). The modification indices suggested that the CFI and TLI can be improved by freeing the covariance between some measurement errors, notably those of a pair of items from *Difficult transition* factor. As underlined by Rubio and Gillespie (1995), the similarity of item format and item redundancy may contribute to the covariance between the error terms of the two items. A model including this correlated error terms fitted the data well ($\chi^2 (128, N=202) = 148.81, p = .100$; RMSEA = .028 [90% CI = .000, .046]; SRMR = .061; CFI = .928; TLI = .914).

All the items loaded significantly on their respective factors, and they yielded coefficient values of .40 or higher, ranging from .42 to .84 (Brown, 2015). The correlations between the four factors were statistically significant (range from .26 [between school avoidance and interpersonal discomfort] to .70 [between anxious anticipation and difficult transition]).

Reliability and Item Analysis

The reliability of the scores obtained by our participants was assessed using CRI (Raykov, 1997): the CRI values were .70 for the *Interpersonal discomfort* dimension, .70 for *School avoidance* dimension, .80 for the *Anxious anticipation* dimension, .77 for the *Difficult transition* dimension, and .92 for the total score scale. The corrected item-total correlations ranged from $r = .25$ (Item 6) to $r = .65$ (Item 3) (Table II). The deletion of any item was not likely to improve the reliability of the scale. Thus, the final 18-item version of the SCREEN will be subjected to construct validity examination.

Validity Correlations

Construct validity was examined by computing Pearson correlations between, on the one hand, the subscales of the SCREEN and the total SCREEN scores and, on the other hand, the scores on the SCARED, SRAS, CBCL-YSR, and absenteeism. As expected, (see Table III), SCREEN total scores (as well as scores on each subscale) were significantly and positively related to school phobia ($r = .64, p = .000$), to functions of SRB ($r = .59, p = .000$), and psychopathology (r -values range from .26 to .62, $p < .001$). The strong correlation between the SCREEN and the CBCL-YSR’s dimensions total behavior problem ($r = .62, p = .000$), anxious-depressed ($r = .53, p = .000$),

Table III. Correlations Between the SCREEN and the Measures Used in This Study

	Anxious anticipation	Difficult transition	Interpersonal discomfort	School avoidance	SCREEN total scores	r^2
SCARED (school phobia dimension)	.47***	.56***	.38***	.27***	.64***	.41
SRAS						
Avoidance of school-related stimuli	.48***	.33***	.55***	.13*	.58***	.34
Escape from aversive social-evaluative situations	.40***	.18***	.59***	.12*	.50***	.25
Pursuit of care from significant others	.37***	.45***	.30***	.22**	.51***	.26
Pursuit of tangible reinforcers outside of school	.06	.17**	.01	.05	.12*	.01
SRAS total score	.44***	.42***	.45***	.19***	.59***	.35
CBCL-YSR						
Withdrawn	.34***	.20***	.45***	.16**	.44***	.19
Somatic complaints	.31***	.29***	.29***	.26***	.42***	.18
Anxious/depressed	.40***	.21***	.59***	.16**	.53***	.28
Social problems	.28***	.16**	.39***	.14*	.35***	.12
Thought problems	.23***	.18***	.31***	.12***	.32***	.10
Attention problems	.31***	.35***	.41***	.22***	.50***	.25
Delinquent behavior	.14*	.28***	.11	.19***	.26***	.07
Aggressive behavior	.19***	.30***	.23***	.16*	.36***	.13
Internalizing problems	.42***	.29***	.57***	.22***	.57***	.32
Total problems	.40	.40	.52	.25	.62***	.38
Absenteeism at school						
School infirmary	.20***	.13	.10	.35***	.20***	.04
Ask to back home	.18***	.17**	.10	.35***	.27***	.07
Visit to the school office	.15*	.22***	.14*	.29***	.28***	.08
Delay	.11	.20***	.05	.24***	.19***	.04
Absenteeism from school						
Total of absence	.06	.12	.08	.34***	.19**	.04

Note. $N = 624$; *** $p < .001$; ** $p < .01$; * $p < .05$; p -adjusted values (Holm's method); SCARED = Screen for Child Anxiety Related Emotional Disorders; SRAS = School Refusal Assessment Scale; CBCL-YSR = Child Behavior Checklist-Youth Self-Report.

and internalizing problems ($r = .57, p = .000$) is consistent with the conception of SR as an emotional and behavioral problem. SR is linked to psychopathological distress in teenagers but cannot be reduced to the psychopathological disorders measure by the CBCL-YSR.

It is noteworthy that neither SCARED scores ($r = .06, p = .99$) nor SRAS scores ($r = .12, p = .42$) correlated with absenteeism from school as recorded by school administration, while there was a significant association between SCREEN scores and absenteeism ($r = .19, p = .018$).

Overlap With School Phobia (SCARED) and Functions of SRB (SRAS)

Although expected, the strong correlations between the SCREEN and the SCARED's school phobia subscale ($r = .64$) on one side, and functions of SRB ($r = .59$) on the other side, raise questions of redundancy between these measures. If the SCREEN and school phobia or functions of SRB (SRAS) are indicators of the same underlying construct, then the amount of variance these scales share in common could be high (Steiger & Ward, 1987). Thus, coefficients of determination were computed by squaring the correlations (r^2) to see the shared variance between the SCREEN and each of these measures. As suggested by Muus, Williams, and Ringsberg (2007), an r^2 value $< .50$ is

considered appropriate, as higher r^2 values might suggest substantial overlap between measures. As shown in Table III, the SCREEN and school phobia subscale (SCARED) shared in common 41% of the variance, while the SCREEN and functions of SRB (SRAS) shared 35% of the variance.

Clinically Significant Scores on the SCREEN

ROC curve analysis was conducted using the clinical sample ($n = 31$) SCREEN scores as standard criterion. The clinical group had on average 61.5 (minimum–maximum = 34–85; $SD = 13.5$), while the school group ($n = 624$) had on average 31.4 (minimum–maximum = 18–83; $SD = 9$). The ROC curve analysis showed good proprieties (the area under the curve = .96). The cutoff score of 41 in the SCREEN had the highest validity in predicting clinically relevant SR. The sensitivity and specificity for this cutoff value were .94 and .88, respectively.

Discussion

SR has severe consequences on adolescent's psychological functioning as well as on his/her family functioning. Our purpose in the present research was to develop a multidimensional self-report measure of SR that is currently lacking in the literature. To this aim, two studies were conducted. The resultant SCREEN

for adolescents (11–16 years old) is intended to be usable in research and clinical settings.

The SCREEN is an 18-item scale. The initial SCREEN structure was identified via EFA on an item pool with items based on theoretically guided and empirically generated aspects of SR. The resultant four-factor measurement model was then validated using CFA. Thus, the SCREEN reliably measures four inter-related aspects of SR. First, the anxiety of school attendance and the fear of being confronted with the school, which corresponds in clinical practice to reluctance in attending school or frequent refusal to go to school. This dimension is highly related to the “Avoidance of school-related stimuli” subscale of the SRAS, which is a function of SRB. Second, the adolescent–parent relationships in SR, referring in clinical practice to the difficulty to deal with the transition and the separation with home and parents for adolescents displaying SR; this dimension is moderately associated with the “pursuit of care from significant others” subscale of the SRAS, which is another function of school absenteeism. Third, the interpersonal relationship difficulties at school for adolescents displaying SR and the impact of these difficulties on their self-perception. This dimension is highly related to withdrawn, anxious-depressed, and social problems syndromes as assessed with the CBCL and “Escape from aversive social-evaluative situations” subscale of the SRAS. Fourth, the avoidance of school refers to adolescent difficulties in coping with emotional and psychic distress and failing to regulate them. It is expressed through the absenteeism (punctual and iterative) and somatic manifestations. This dimension is related to nursing visits and total number of absence.

The reliability of the SCREEN total scores was good. As expected, validity correlations revealed that the SCREEN total scores were correlated with both the SRAS and the school phobia as assessed with the SCARED. However, the coefficients of determination showed no substantial overlap between these measures, suggesting that SCREEN scores were capturing specific information compared with SCARED scores on one hand and SRAS scores on the other. Simply put, the SCREEN has sufficient uniquely reliable variance to warrant separate interpretation from either the school phobia subscale of the SCARED or the functions of SRB as assessed with the SRAS. In addition, unlike SRAS and SCARED scores, SCREEN scores were correlated with absenteeism from school as recorded by school administration. This correlation was weak, indicating that the SCREEN scores capture more than just the absenteeism phenomenon or truancy. All these findings support the incremental convergent validity of the SCREEN scores (Haynes & Lench, 2003).

Also, results of the ROC curve indicated good accuracy of the SCREEN scores. Using a cutoff score of 41

(sensitivity = .93 and specificity = .88), the SCREEN can be prudently used as a screening tool to detect SR. In comparison, Birmaher et al. (1997) have reported a sensitivity of .67 and a specificity of .58 for the “school phobia” dimension (four items) of the SCARED (for a cutoff score of 3).

Because the association between the SCREEN scores and the total psychopathological problems (CBCL-YSR) was strong, we suggest including SCREEN scores as a complement to the screening and diagnosis of psychopathologic disorders. This is even more important given the complexity and the heterogeneity of SR (Fremont, 2003).

The SCREEN could fill the gap in SR measurement, and then could be helpful in various contexts. First, in clinical settings, it could be used by pediatricians, psychiatrists, psychologists, and family physicians to screen SR. Owing to the importance of somatic complaints in SR, family physicians as well as pediatricians, who are essential first-line professionals in this process, need a psychometrically sound measure of SR, because early detection of SR could hasten treatment and return to school (Fremont, 2003). Second, in research settings, the SCREEN could promote research in a field considered by the National Association of School Psychologists, as “one of the major challenges facing educators in the 21st century.”

This study has a number of limitations. First, our clinical sample was small, and thus, our results should be considered as preliminary. A cross-validation in large clinical and more diverse samples is needed, especially for chronic health condition populations. Second, the SCREEN is a self-report measure, which is sensitive to number of bias, such as social desirability. The impact of social desirability on each item response was not estimated in the present study. Future SCREEN studies in which a social desirability measure is used are needed. Recall here that the validation of any measure is an ongoing and cumulative process allowing instrument refinement (Holmbeck & Devine, 2009). Investigating the utility of the SCREEN in assessing responsiveness to treatment is needed. Cross-cultural validation of the SCREEN in other languages is an interesting and challenging project. We are ready to contribute to any relevant initiative with this aim. Measurement tools and measurement invariance are the foundation for international empirical research. SR is a universal phenomenon (SR is consubstantial with school), but what about the universality of its clinical manifestations? This could be a relevant avenue of future research.

Compliance With Ethical Standards

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or

national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

Acknowledgments

The authors are especially indebted to Sabine Broc (University of Westminster) who graciously edited this manuscript. The authors would like to thank David Heyne (Leiden University) for his helpful comments on an earlier draft of this manuscript.

Conflict of interest: None declared.

References

- Achenbach, T. M. (1991). *Manual for Child Behavior Checklist 4-18, 1991 Profile*. Burlington, VT: Department of Psychiatry, University of Vermont.
- Berg, I. (1992). Absence from school and mental health. *The British Journal of Psychiatry: The Journal of Mental Science*, 161, 154–166.
- Berg, I. (1997). School refusal and truancy. *Archives of Disease in Childhood*, 76, 90–91.
- Berg, I., Nichols, K., & Pritchard, C. (1969). School Phobia—Its classification and relationship to dependency. *Journal of Child Psychology and Psychiatry*, 10, 123–141.
- Bernstein, G. A., Massie, E. D., Thuras, P. D., Perwien, A. R., Borhardt, C. M., & Crosby, R. D. (1997). Somatic symptoms in anxious-depressed school refusers. *Journal of the American Academy of Child & Adolescent Psychiatry*, 36, 661–668.
- Birmaher, B., Khetarpal, S., Brent, D., Cully, M., Balach, L., Kaufman, J., & Neer, S. (1997). The Screen for Child Anxiety Related Emotional Disorders (SCARED): Scale construction and psychometric characteristics. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36, 545–553.
- Birnbaum, M. H. (2000). *Psychological experiments on the Internet*. San Diego, CA: Academic Press.
- Bouvard, M., Roulin, J.-L., & Denis, A. (2013). The French version of the Screen for Child Anxiety Related Emotional Disorders-Revised (Scared-R): Factor structure, convergent and divergent validity in a sample of teenagers. *Psychologica Belgica*, 53, 3.
- Brandibas, G., Jeunier, B., Gaspard, J.-L., & Fourasté, R. (2001). Evaluation des modes de refus de l'école: Validation française de la SRAS (School Refusal Assessment Scale). *Psychologie et Psychométrie*, 22, 45–58.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. London: Guilford Publications.
- 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.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edn.). Hillsdale, NJ: Routledge.
- Crocetti, E., Hale, W. W., Fermani, A., Raaijmakers, Q., & Meeus, W. (2009). Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED) in the general Italian adolescent population: A validation and a comparison between Italy and The Netherlands. *Journal of Anxiety Disorders*, 23, 824–829.
- DeVellis, R. F. (2011). *Scale development: Theory and applications*. London: Sage.
- Egger, H. L., Costello, J. E., & Angold, A. (2003). School refusal and psychiatric disorders: A community study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 42, 797–807.
- Fremont, W. P. (2003). School refusal in children and adolescents. *American Family Physician*, 68, 1555–1560.
- Gallé-Tessonneau, M., Doron, J., & Grondin, O. (2017). Des critères de repérage aux stratégies de prise en charge du refus scolaire: Une revue de littérature internationale systématique. From diagnosis criteria to therapeutic strategies of school refusal: An international systematic review. *Pratiques Psychologiques*, 23, 1–19.
- Gallé-Tessonneau, M., Koleck, M., & Doron, J. (2018). Considérations méthodologiques pour la construction de questionnaires: L'exemple de la School REfusal EvaluationN (SCREEN). Methodological considerations for scale design. The example of the School REfusal EvaluationN (SCREEN) survey. *Annales Médico-Psychologiques*. Advance online publication. doi: 10.1016/j.amp.2017.03.029.
- Gregory, I. R., & Purcell, A. (2014). Extended school non-attenders' views: Developing best practice. *Educational Psychology in Practice*, 30, 37–50.
- Havik, T., Bru, E., & Ertesvåg, S. K. (2015). Assessing reasons for school non-attendance. *Scandinavian Journal of Educational Research*, 59, 316–336.
- Havik, T., Bru, E., & Ertesvåg, S. K. (2015). School factors associated with school refusal- and truancy-related reasons for school non-attendance. *Social Psychology of Education*, 18, 221–240.
- Haynes, S. N., & Lench, H. C. (2003). Incremental validity of new clinical assessment measures. *Psychological Assessment*, 15, 456–466.
- Hendron, M., & Kearney, C. A. (2016). School climate and student absenteeism and internalizing and externalizing behavioral problems. *Children and Schools*, 38, 109–116.
- Heyne, D., Gren Landell, M., Melvin, G., & Gentle-Genitty, C. (2018). Differentiation between school attendance problems: Why and how? *Cognitive and Behavioral Practice*. Advance online publication. doi: 10.1016/j.cbpra.2018.03.006.
- Holmbeck, G. N., & Devine, K. A. (2009). Editorial: An Author's checklist for measure development and validation manuscripts. *Journal of Pediatric Psychology*, 34, 691–696.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6, 1–55.
- Inglés, C. J., González-Maciá, C., García-Fernández, J. M., Vicent, M., & Martínez-Monteaudo, M. C. (2015). Current status of research on school refusal. *European Journal of Education and Psychology*, 8, 37–52.

- Johnson, A. M. (1957). School phobia. *American Journal of Orthopsychiatry*, 27, 307–309.
- Johnson, A. M., Falstein, E. I., Szurek, S. A., & Svendsen, M. (1941). School phobia. *American Journal of Orthopsychiatry*, 11, 702–711.
- Kahn, J. H., & Nursten, J. P. (1962). School refusal: A comprehensive view of school phobia and other failures of school attendance. *American Journal of Orthopsychiatry*, 32, 707–718.
- Kearney, C. A. (2002). Identifying the function of school refusal behavior: A revision of the school refusal assessment scale. *Journal of Psychopathology and Behavioral Assessment*, 24, 235–245.
- Kearney, C. A. (2006). Confirmatory factor analysis of the school refusal assessment scale-revised: Child and parent versions. *Journal of Psychopathology and Behavioral Assessment*, 28, 139–144.
- Kearney, C. A. (2008). School absenteeism and school refusal behavior in youth: A contemporary review. *Clinical Psychology Review*, 28, 451–471.
- Kearney, C. A., & Graczyk, P. (2014). A Response to Intervention model to promote school attendance and decrease school absenteeism. *Child and Youth Care Forum*, 43, 1–25.
- Kearney, C. A., & Silverman, W. K. (1993). Measuring the function of school refusal behavior: The school assessment scale. *Journal of Clinical Child Psychology*, 22, 85–96.
- King, N. J., & Bernstein, G. A. (2001). School refusal in children and adolescents: A review of the past 10 years. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40, 197–205.
- Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, 11, 320–341.
- Maynard, B. R., Salas-Wright, C. P., Vaughn, M. G., & Peters, K. E. (2012). Who are truant youth? Examining distinctive profiles of truant youth using latent profile analysis. *Journal of Youth and Adolescence*, 41, 1671–1684.
- McCune, N., & Hynes, J. (2005). Ten year follow-up of children with school refusal. *Irish Journal of Psychological Medicine*, 22, 56–58.
- McShane, G., Walter, G., & Rey, J. M. (2004). Functional outcome of adolescents with school refusal. *Clinical Child Psychology and Psychiatry*, 9, 53–60.
- Muris, P., Merckelbach, H., Brakel, A. V., & Mayer, B. (1999). The revised version of the screen for child anxiety related emotional disorders (SCARED-R): Further evidence for its reliability and validity. *Anxiety, Stress, & Coping*, 12, 411–425.
- Muris, P., Merckelbach, H., Schmidt, H., & Mayer, B. (1998). The revised version of the Screen for Child Anxiety Related Emotional Disorders (SCARED-R): Factor structure in normal children. *Personality and Individual Differences*, 26, 99–112.
- Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide* (6th edn.). Los Angeles, CA: Muthén & Muthén.
- Muus, I., Williams, L. S., & Ringsberg, K. C. (2007). Validation of the Stroke Specific Quality of Life Scale (SS-QOL): Test of reliability and validity of the Danish version (SS-QOL-DK). *Clinical Rehabilitation*, 21, 620–627.
- Nair, M. K. C., Russell, P. S. S., Subramaniam, V. S., Nazeema, S., Chembagam, N., Russell, S. . . Charles, H. (2013). ADad 8: School Phobia and Anxiety Disorders among adolescents in a rural community population in India. *Indian Journal of Pediatrics*, 80, 171–174.
- Okuyama, M., Okada, M., Kuribayashi, M., & Kaneko, S. (1999). Factors responsible for the prolongation of school refusal. *Psychiatry and Clinical Neurosciences*, 53, 461–469.
- Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential Tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research*, 32, 329–353.
- Rubio, D. M., & Gillespie, D. F. (1995). Problems with error in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 2, 367–378.
- Simons, E., Hwang, S., Fitzgerald, E. F., Kielb, C., & Lin, S. (2010). The impact of school building conditions on student absenteeism in Upstate New York. *American Journal of Public Health*, 100, 1679–1686.
- Steiger, J. H., & Ward, L. M. (1987). Factor analysis and the coefficient of determination. *Psychological Bulletin*, 101, 471–474.
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response* (1st edn). Cambridge, UK; New York, NY: Cambridge University Press.
- Van Meter, A., Youngstrom, E., Youngstrom, J. K., Ollendick, T., Demeter, C., & Findling, R. L. (2014). Clinical decision-making about child and adolescent anxiety disorders using the achenbach system of empirically based assessment. *Journal of Clinical Child and Adolescent Psychology*, 43, 552–565.
- Wyss, C. A., Voelker, S. L., Cornock, B. L., & Hakim-Larson, J. (2003). Psychometric properties of a French-Canadian translation of Achenbach's youth self-report. *Canadian Journal of Behavioural Science*, 35, 67–71.