Comprehensive Psychiatry 88 (2019) 39–48

Contents lists available at ScienceDirect

Comprehensive Psychiatry

journal homepage: www.elsevier.com/locate/comppsych

Is a brief self-report version of the Columbia severity scale useful for screening suicidal ideation in Chilean adolescents?

D. Núñez [a,b,*], V. Arias [c,d], P. Méndez-Bustos [e], A. Fresno [a]

[a] Faculty of Psychology, Universidad de Talca, Avenida Lircay s/n, Talca, Chile
[b] Millennium Nucleus to Improve the Mental Health of Adolescents and Youths, Imhay, Chile
[c] Faculty of Psychology, Department of Personality, Assessment and Psychological Treatment, Universidad de Salamanca, Avenida de la Merced, 109, CP 37005, Salamanca, Spain
[d] Institute of Community Integration (INICO), Universidad de Salamanca, Avenida de la Merced, 109, CP 37005, Salamanca, Spain
[e] School of Psychology, Universidad Católica del Maule, Avenida San Miguel 3605, Talca, Chile

ARTICLE INFO

ABSTRACT

Background: Given the high rates of suicide in the adolescent population and the reluctance of this population to seek help, developing proactive and effective strategies to timely detect individuals at high risk for suicide in non-clinical contexts is a worldwide recognized need. A series of brief self-report questionnaires have been developed for this purpose, however there are few studies providing evidence on their capability to accurately classify suicidal risk levels in specific populations. One of the instruments frequently used to evaluate suicide risk is the Columbia-Suicide Severity Rating Scales (C-SSRS). The goal of this study is to provide psychometric evidence about the accuracy of the Suicidal Ideation subscale (SI) of the C-SSRS to classify suicidal risk levels in a sample of Chilean adolescents using Item Response Theory (IRT).
Methods and materials: Through the two parameter logistic model (2-PLM), we analyzed the capability of a self-report questionnaire addressing suicidal ideation (SI) to differentiate and classify participants according to their SI severity levels. We tested two main parameters: difficulty (localization) and discriminating power of 6 items extracted and adapted from the Columbia-Suicide Severity Rating Scales (C-SSRS). We administered this questionnaire to a general sample of 1645 adolescents aged 13 to 18.
Results: Our results show that the items differentiate symptoms addressing suicidal thoughts according to their severity, providing an accurate classification of the SI risk level.
Conclusions: These findings support the usage of the C-SSRS in Chilean adolescents. Further research is needed to test its predictive value in different populations.

1. Introduction

Suicide-related behaviors are common among school-age adolescents. Suicide is the second leading cause of death among youth, ages 13–18 [1]. It has been conceptualized as a continuum developing from suicidal ideation (SI), suicide plans, suicide attempts and complete suicide [2]. However, suicidal behavior and suicidal ideation are currently thought as distinct but correlated phenomena [3,4]. SI, defined as thinking about, considering, or planning suicide [5], is highly prevalent in adolescents [6,7], and it is regarded as a predictor of future suicidal attempts and psychiatric crises in this population [8–13].

The general adolescents' reluctance to seek professional help [14], the decreased help-seeking among adolescents with SI [15], and the low capacity of parent and teachers to recognize suicide warning signs [16] indicate that proactive and effective staging-based programs aimed to timely detect and refer subjects at risk for suicide are needed [17–19]. These programs frequently use screening questionnaires, which should confidently rule out individuals with no evident risk (i.e., low false negatives) [20]. In non-clinical settings (i.e., schools), these screening tools should be used in the context of multi-level screening programs [21] addressing a broader range of psychological symptoms, other risk factors and warning signs for suicide [22], and using complementary strategies such as training community members as gatekeepers [23,24].

Several brief, easy-to-use self-report questionnaires addressing SI have proliferated for screening purposes [25] which are frequently used in educational institutions [26–30], clinical settings [31], and emergency departments [32–36]. Despite their relatively expanded usage and the recognized need for developing shorter instruments aimed at effectively and reliably distinguishing individuals who are at high risk for suicide from those who are not [20], some issues preventing a clear recommendation of screening measures for SI have been recently highlighted. For instance, studies analyzing their

* Corresponding author at: Faculty of Psychology, Avenida Lircay s/n, Talca, Chile.
E-mail address: dnunez@utalca.cl (D. Núñez).

https://doi.org/10.1016/j.comppsych.2018.11.002
0010-440X/

diagnostic accuracy are surprisingly scant [37]; there are no standard measures being used in population-based studies [25] and more diverse population representation and non-English versions in studies is required [39].

One of the instruments frequently used to screen SI is the Columbia-Suicide Severity Rating Scale (C-SSRS) [40]. When used in adolescent students, it has shown good sensitivity and reasonable specificity; however, a second-stage evaluation has been recommended to reduce the burden of its low specificity [41]. As described by Scott et al. [19], the thresholds can be modified to reduce the screen-positive population. The C-SSRS has been reported as a useful tool to predict patients' returns to emergency departments because of suicide-related complaints [33], and as a reliable and valid instrument to assess SI and behavior in daily clinical practice and research settings [10,31]. Despite these findings, the psychometric evidence about the C-SSRS is scarce and not conclusive [42] and additional research assessing its capability to accurately identify subjects at risk for suicide is needed [43–45].

Literature strongly encourages the development of systematic and standardized suicide screening [46] based on target populations' needs and characteristics [20] and shows that patients seem to be more honest when answering self-report scales about SI [47]. In this context, the present study was aimed to provide psychometric evidence on the accuracy of the SI subscale of the C-SSRS to classify suicidal risk levels using the "item response theory" (IRT). [48]. First, we determined the relation between a latent trait (i.e., suicidal ideation) and its observable manifestation (a single item's response) [49]. Then, we tested the capability of the scale to differentiate subjects according to their SI severity levels, analyzing the scale at the single item level through two main parameters: difficulty (localization) and discriminating power of each item. In addition, we assessed the accuracy of the subscale by computing both the item and the test information functions (TIF). Furthermore, we addressed the external validity of the scale by analyzing associations between depression and SI, which have been previously observed in adolescents [7].

2. Methods and materials

2.1. Participants, procedure and measures

We conducted a cross-sectional study with 1776 adolescents recruited between April and August 2015 in secondary schools in the city of Talca, Chile. A total of 131 adolescents did not entirely complete the questionnaires. These subjects were not considered for the analysis. We conducted the analyses with a final sample of 1645 Chilean young adolescents (mean age = 16.03, SD = 1.42, women = 54.1%). The only inclusion criterion was that the students voluntarily agreed to participate in the study. We tested whether the 131 excluded subjects differ from included participants on demographic and clinical characteristics. Given that 13 subjects did not respond any item, they were excluded from this analysis. Using the rest of the sample we tested if the patterns of the missing data were associated to demographic variables (gender and age) and clinical severity, addressed by the scores on depression, anxiety and stress (DASS-21). We conducted this analysis through the Little's test [50], which contrasts the null hypothesis positing that there is no systematic relationships between the missing data and the other variables ($p < .05$). In all cases, the Little's test was not significant ($p \geq .05$). Therefore, the probability of missing data, in this sample, does not depend on gender, age or clinical severity.

We assessed SI through the Suicidal Ideation subscale (SI) of the Columbia–Suicide Severity Rating Scale [40]. Instead of conceptualizing suicidal ideation and suicidal behavior as a unidimensional construct existing along a continuum, the C-SSRS differentiates suicidal ideation from suicidal behavior. The former is addressed by two subscales: severity (subscale 1) and intensity (subscale 2); the latter is addressed by two subscales: behavior (actual, aborted, interrupted and nonsuicidal self-injurious (subscale 3), and lethality of attempts (subscale 4).We adapted the first subscale for being used as a self-report questionnaire. Given that the last item of the original scale is a compound question asking about active planning ("Have you started to work out or worked out the details of how to kill yourself?"), and intentions ("Do you intend to carry out this plan?"), we did split this item into two different questions (questions 5 and 6; Table 1). Therefore, severity of SI was finally rated on a 6-point ordinal scale in which 1 = wish to be dead, 2 = nonspecific active suicidal thoughts, 3 = thoughts about how to commit suicide, 4 = suicidal thoughts and general intentions, 5 = suicidal thoughts with detailed plan, and 6 = intentions to conduct plans. Frequency of SI was addressed by asking participants when these thoughts happened: ever in life (SI_L) and/or during last month (SI_M).The questionnaire should take between 5 and 8 min to complete.

We computed the internal reliability of these two subscales. Because of the binary nature of the data, we included the Kuder-Richardson coefficient (KR20). Given that the Cronbach's alpha, and consenquently also the KR20, have been criticized because the lack of precision to reflect the true parameter, or its dependency of the number of variables [51–53], we additionally estimated the model-based reliability (McDonald's coefficient omega) [54], computed on the standardized factorial loads of the two-parameter logistic model. Both scales showed adequate internal consistency ($KR20_L = 0.808$, $KR20_M = 0.838$), and model-based reliability ($\omega_L = 0.968$; $\omega_M = 0.972$).

We additionally addressed depression, anxiety and stress by the Depression, Anxiety, and Stress Scales (DASS-21) [55] previously adapted and validated with Chilean adolescents [56].

This study was approved by the Bioethics Committee of Universidad de Talca, Chile. We conducted the study in those public schools who agreed to participate after meetings with directive committees. Researchers participated in different parents' meetings to present the research project. The project was explained to both the directive committee and students. After its approval and once written informed consents were obtained from both, the adolescents and their caregivers, the participants completed the questionnaires that were administered in a classroom setting by trained psychologists.

2.2. Data analysis

2.2.1. Unidimensionality and local dependence

In order to demonstrate the existence of a clearly dominant dimension [57], we tested two basic requirements of IRT analysis: unidimensionality and local independence. Whereas the former addresses whether all items are measuring a single characteristic (latent variable), the latter evaluates if the items are independent of each other after adjusting the causal effect of the latent variable.

Table 1
Parameters of the SI_M scale.

Item		α_i	s.e. (α_i)	β_i	s.e. (β_i)
1	Have you wished you were dead or wished you could go to sleep and not wake up?	2.65	0.30	1.24	0.08
2	Have you actually had any thoughts of killing yourself?	5.26	0.86	1.61	0.08
3	Have you been thinking about how you might do this?	4.68	0.64	1.73	0.08
4	Have you had these thoughts and had some intention of acting on them?	5.53	1.04	1.71	0.08
5	Have you started to work out or worked out the details of how to kill yourself?	4.54	0.56	2.01	0.09
6	Do you intend to carry out this plan?	3.89	0.62	2.19	0.13

We examined dimensionality through a parallel analysis [58,59] based on a tetrachoric correlation matrix. Additionally, we computed the fit of two models by confirmatory factor analysis (CFA). Model 1 (M1) hypothesizes that the twelve items of the scale are organized into a single dimension of SI, and model 2 (M2) hypothesizes the existence of two correlated but empirically separable dimensions (SI throughout life, and SI during the last month). In both models residual correlations corresponding to pairs of items with identical wording were released.

We tested the conditional independence criterion by computing the modified indexes (MI) and the standardized expected parameter change indexes (SEPC). Local dependence is negligible if MI values are <10, and if SEPC values are not higher than 0.2 [60]. Second, we inspected the standardized local dependence (LD-χ^2) values of the matrix of expected and observed response frequencies for each item (χ2 values >10 suggest violation of local independence) [61].

2.2.2. Parameter estimation

In clinical scales, the localization parameter (β) is interpretable as the relative symptom severity (i.e., the point in the latent variable, theta, where a person has a probability of 50% of endorsing the symptom). The 2-PLM model, besides the item localization, estimates a discrimination parameter (α) for each item, assuming that the items are not equally accurate in different areas of the latent variable. The 2-PLM model specifies a person's probability of being classified into the lower category ("no", or absent symptom) as opposed to being classified into the higher category ("yes", or present symptom) as a function of the interaction among the degree to which a person presents the measured trait (suicidal ideation), the localization parameter (severity), and its discriminant capacity.

2.2.3. Measurement accuracy

First, to evaluate the measurement accuracy of each item in different levels of the latent construct, we tested the trait range θ by the item information function (IFF). The higher information value indicates a larger precision of the items for a certain zone of the theta continuum. Second, we computed the test information function (TIF), which addresses how well a test estimates SI severity. Therefore, it represents the combined value of the information function of all items. As in IFF, a higher information value indicates a higher precision of the scale. Third, to evaluate the relation between the subject's scores and the latent SI scores, we inspected the Test Characteristic Curve (TCC). Fourth, we computed the sample distribution in the latent trait by computing the estimated a posteriori scores.

Finally, to ensure a fair gender comparison, we computed the differential item functioning (DIF) by the Wald test, and when necessary, we estimated the DIF size according to the effect size indexes suggested by Meade [62].

2.2.4. External validity

We divided the sample into three groups according to their position in the latent variable (group 1 ≤ 1 SD; group 2 = between 1 and 2 SD; and group 3 ≥ 2 SD). Then, we compared the latent scores for depression, anxiety and stress, estimated by the Graded Response Model [63]. For each comparison (effect sizes; Cohen's *d*) we used group 1 as a reference.

The IRT models were estimated with IRTPRO 4.0 [64]. The factorial models were estimated using MPlus 7.0 [65] using weighted least squares with mean and variance adjusted (WLSMV).

3. Results

3.1. Unidimensionality and local dependence

As shown in Fig. 1, the parallel analysis suggests that a single component for each subscale is plausible since the random eigenvalues from component 3 surpassed those obtained from the real data.

The CFA supported these results. Model 1 (1 dimension, 12 symptoms) yielded acceptable fit indices (RMSEA = 0.075; CFI = 0.976; TLI = 0.968; DF = 48). Model 2 (2 correlated factors) showed substantially better fit indices (RMSEA = 0.0036; CFI = 0.994; TLI = 0.992; DF = 47), without a significant reduction of parsimony (one degree of freedom). Correlation between factors was 0.75 and the average variance extracted for SI_M and SI_L was 0.83 and 0.81, respectively. Therefore, both factors can be regarded as empirically separable dimensions. Both MI and SEPC values were respectively lower than 10 and 0.2 (for 56 of the 60 possible contrasts). Finally, the χ2 values for each subscale were lower than 10 (range = 0.1–6.6), reaching the local independency requirement.

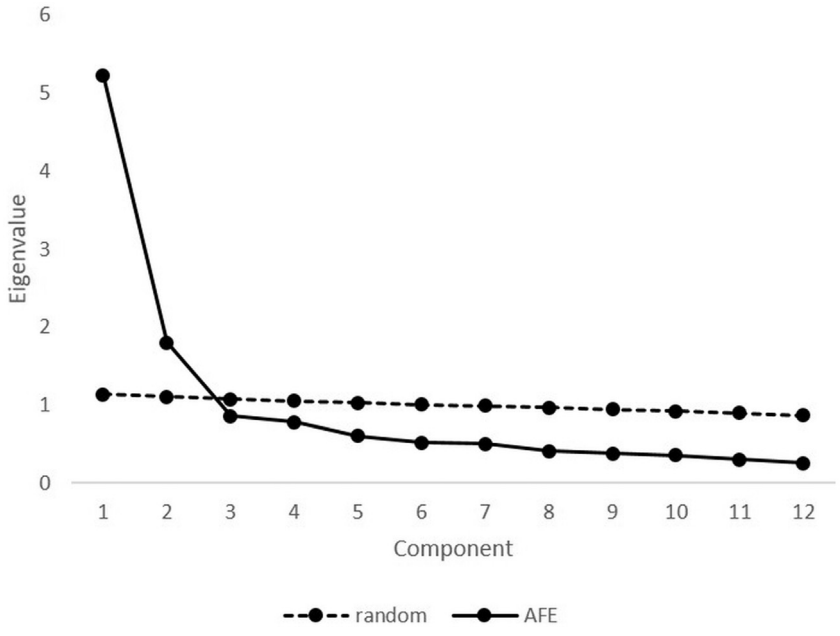

Fig. 1. Parallel analysis.

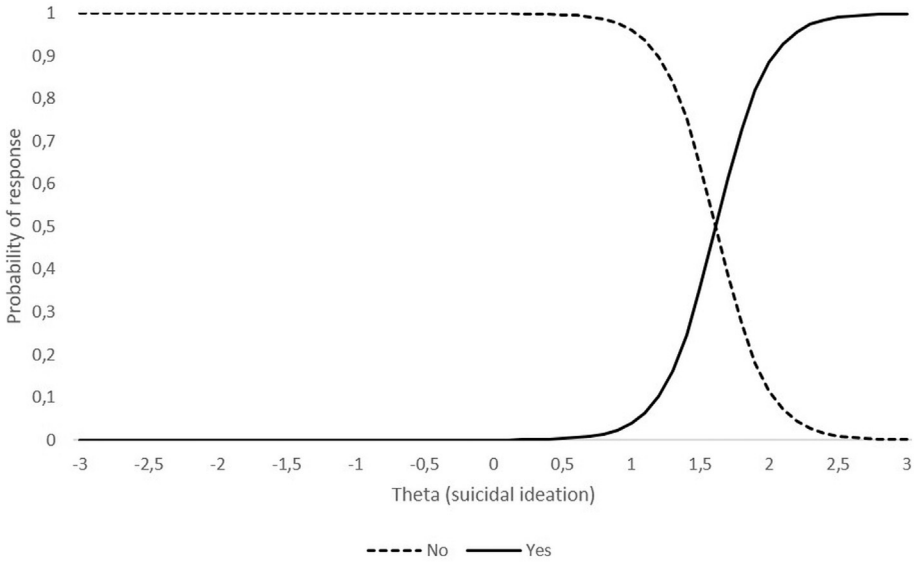

Fig. 2. Category characteristic curves, item 2, SI_M scale.

3.2. Fitting data to the two-parameter model (2-PLM)

Both M_2 and RMSEA values for the SI_M scale were ($M_2 = 35$; $GL = 9$; $p = .0001$; $RMSEA = 0.04$) and for the SI_L scale were ($M_2 = 95$; $GL = 9$; $p = .0001$; $RMSEA = 0.06$). In terms of absolute fit, both M_2 statistics were significant. However, the low associated RMSEA values suggest that this misfit may be due to a limited amount of unmodeled error [66]. Most of the items had non-significant S-χ^2 values.

3.3. Parameters estimation

3.3.1. Item localization

As an example, Fig. 2 shows the category characteristic curves for item 2 of the SI_M scale ("Have you actually had any thoughts of killing yourself?"). The y-axis represents the latent variable θ ($M = 0$; $SD = 1$). For each item, two curves are drawn, which represent the probability (x-axis) of being located in each of the response categories. Thus, for being placed within the "Yes" category, the suicide ideation level should be relatively high (approximately 1.6 standard deviations above the mean). By contrast, the most probable answer (100%) of an adolescent located on the mean ($\theta = 0$) would be "No".

3.3.2. Item information functions

The item information functions (IIFs) are depicted in Figs. 3 and 4. For instance, in Fig. 3, item 2 (SI_M scale) starts providing information from $\theta = 1.0$, which drastically decreases from $\theta = 1.8$, being this range where the item is most productive for the measure. By contrast, within the lesser and higher severity ranges, the item does not discriminate between subjects with different trait levels.

Two additional comments deserve to be mentioned. First, in both scales the information provided by item 1 ("Have you wished you

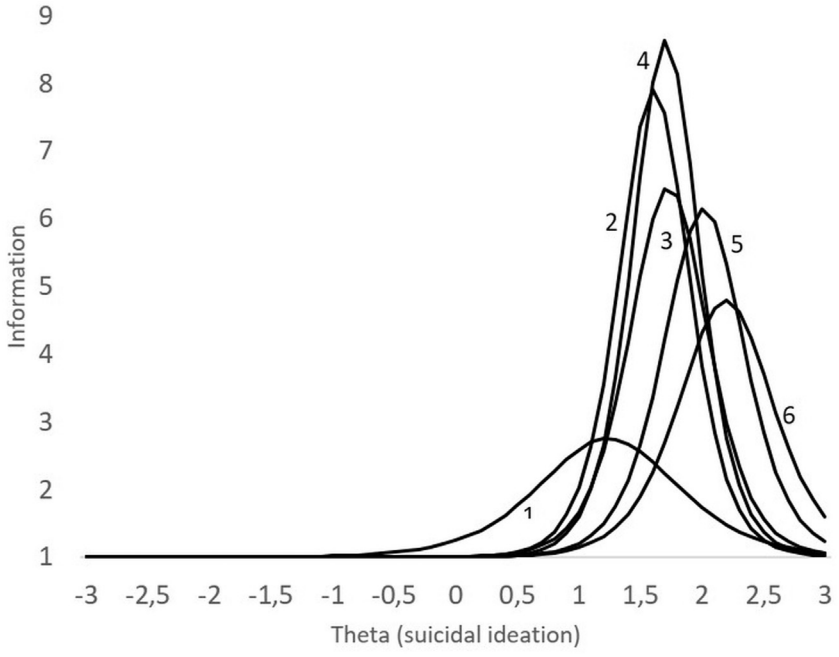

Fig. 3. Item information function, SI_M scale.

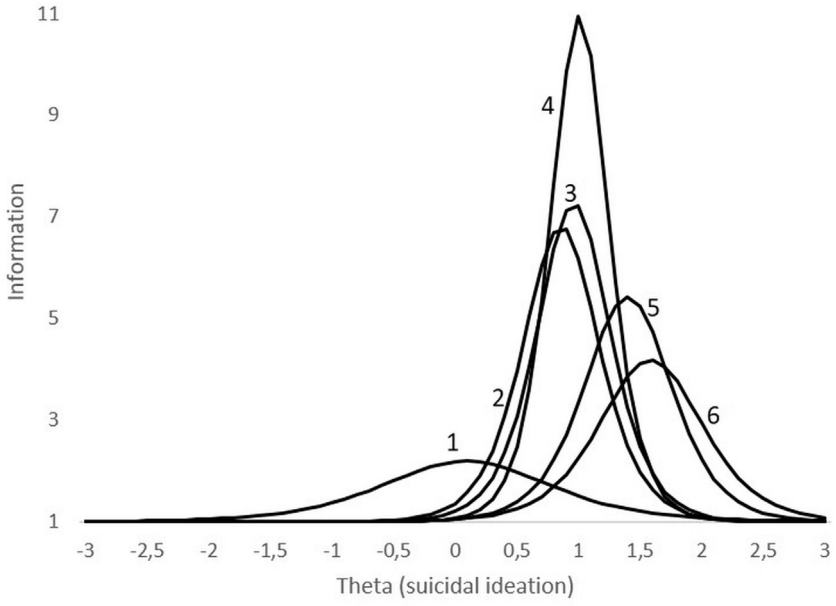

Fig. 4. Item information function, SI_L scale.

were dead or wished you could go to sleep and not wake up?") is substantially lower than the rest of the items. In addition, in the SI_L scale, its peak of information is close to the mean, suggesting a low diagnostic value in this sample (it does not provide information in clinically relevant levels of SI). Second, considering the informational peaks, the items can be grouped into three severity clusters. For the SI_M scale, the first cluster is represented by item 1, whose peak is between 1 and 1.5 SD above the mean. The second cluster (items 2, 3 and 4) contributes with maximal information approximately between +1.5 and +2 SD. The content of this cluster is associated to a general SI and non-elaborated suicidal plans. Finally, the cluster 3 (items 5 and 6), shows its major discriminant capability at very severe levels (2 SD above the mean) and represents a detailed suicidal planning and the intention to carry out the plan. A similar pattern is observed for the SI_L scale, with a shift to less severe levels of the variable.

The parameters α_i and β_I are shown in Tables 1 and 2. Item 1 obtained the lower discriminant power, particularly in the SI_L scale. These results are in line with the information provided by each item. As depicted in Figs. 3 and 4, item 1 showed the poorest contribution in both scales, particularly in the SI_L scale.

According to the Baker's classification [67], we conclude that the discrimination parameters have been very high for all items of both scales, which is usual for clinical scales [68].

Regarding the β_i parameters, we observed a clear difference in symptom severity, with a range from the mean to +1.6 SD for the SI_L scale and from +1.24 to +2.19 for the SI_M scale. For the SI_L scale, the β parameter of item 1 was 0.09, which means that a person located at the mean of the latent variable (and consequently within a region without diagnostic value) will have a high probability to give a positive answer. By contrast, the β parameter of the same symptom in the SI_M scale was 1.24, which indicates that a positive response to the item represents a relatively high SI level. On the other side, severity transitions between the symptom thresholds are reasonably reduced, revealing that the symptoms homogeneously cover a wide range of the latent variable, without information gaps between consecutive items. The exception was given by item 1 of the SI_L scale ($\beta_i = 0.09$) with respect to the next item of this scale ($\beta_i = 0.86$). This difference implies nearly 1 SD between one item and the next one.

3.4. Measurement accuracy

3.4.1. Test information function

Figs. 5 and 6 depict the test information function (TIF) and the measurement error distribution of the SI_M scale and the SI_L scale, respectively. The solid line represents the TIF, equivalent to the combined value of the information functions of the six items shown in Figs. 3 and 4. In the case of the SI_M scale, the TIF values are higher for θ values between +1.2 and +2.5 SD. The lowest measurementerror values correspond to the highest information test values. As expected, the lower the standard error, the higher the accuracy (reliability) with which the scale provides information about the latent trait. Consequently, the SI_M scale is most accurate within high SI levels, and does not provide enough information (it does not discriminate subjects according to their SI level) either in intermediate or low SI levels (>2.5 SD), where the error measurement exceeds the information provided by the test. We observed similar results for the SI_L scale, but slightly closer to lower severity SI levels, making this scale more accurate within the range of +0.5 and +1.5 SD.

Table 2
Parameters of the SI_L scale.

Item		α_i	s.e. (α_i)	β_i	s.e. (β_i)
1	Have you wished you were dead or wished you could go to sleep and not wake up?	2.18	0.17	0.09	0.05
2	Have you actually had any thoughts of killing yourself?	4.82	0.43	0.86	0.05
3	Have you been thinking about how you might do this?	5.01	0.47	0.96	0.05
4	Have you had these thoughts and had some intention of acting on them?	6.31	0.74	1.01	0.05
5	Have you started to work out or worked out the details of how to kill yourself?	4.21	0.61	1.40	0.05
6	Do you intend to carry out this plan?	3.57	0.52	1.59	0.06

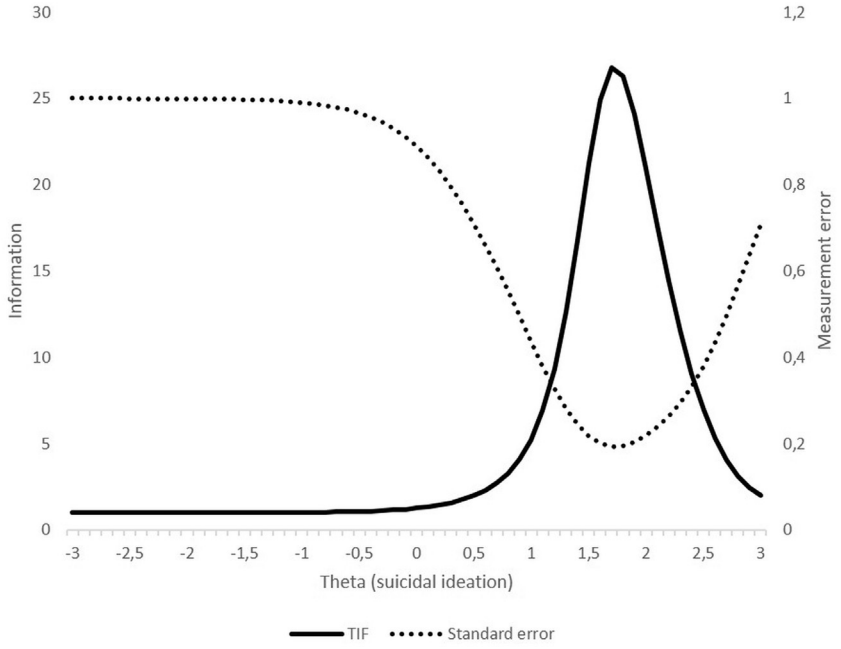

Fig. 5. Test information function, and measurement error distribution, SI_M scale.

Fig. 7 shows the Test Characteristic Curve (TCC), which illustrates the correspondence between the subject's scores (x axis) and the latent SI scores (y axis). Finally, we computed the sample distribution by estimated a posteriori scores (Table 3). In the SI_M scale, most of the sample (82.1%) obtained the minimal score (equivalent to a direct score = 0) and a low proportion of subjects obtained high (6.3%) or very high (1.8%) scores. In the SI_L scale, 49.5% obtained a minimal score, 24.3% was located between the mean and 1 SD, and 15.3% and 4.4% obtained high and very high scores, respectively.

3.5. Additional validity evidence

Fig. 8 shows the effect sizes (Cohen's *d*) for each comparison using group 1 as a reference. Group 2 (high SI level in the SI_M scale, panel A) showed significantly more depressive symptoms compared to group 1, with a large effect size ($d = 1.44$), a difference even greater when comparing group 3 with group 1 (very high SI level) ($d = 2.77$). The rest of comparisons were significant, with high or very high effect sizes. The group with highest SI levels consistently showed higher anxiety and stress scores. We observed a similar pattern for the SI_L scale (panel B), but the differences were much more moderate, being depression the variable showing the greatest difference with respect to the group with lower SI.

3.6. Gender-based differential item functioning analysis

DIF analysis did not reveal evidence supporting a differential item functioning. Chi-square values were between 0 ($p = .90$) and 1.6 (p

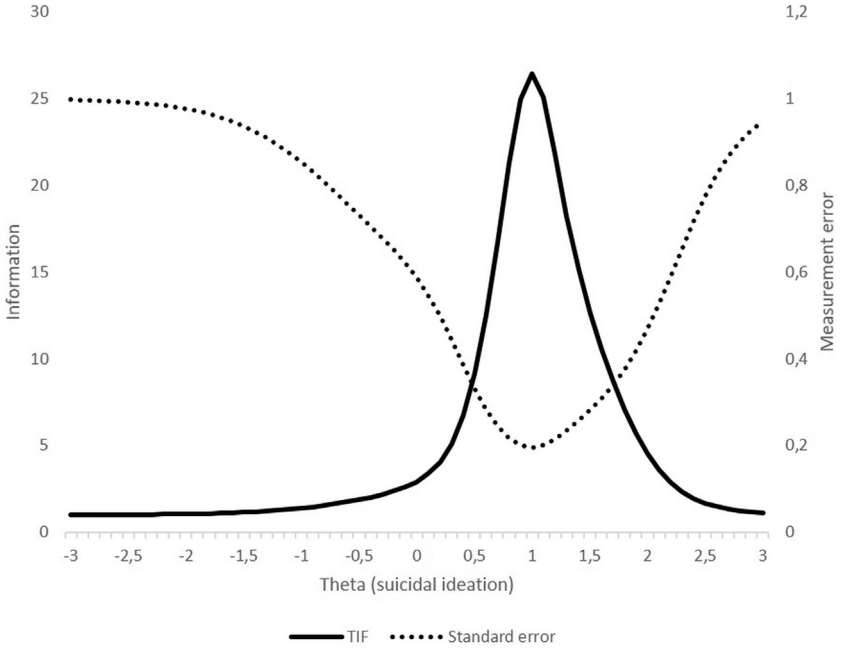

Fig. 6. Test information function, and measurement error distribution, SI_L scale.

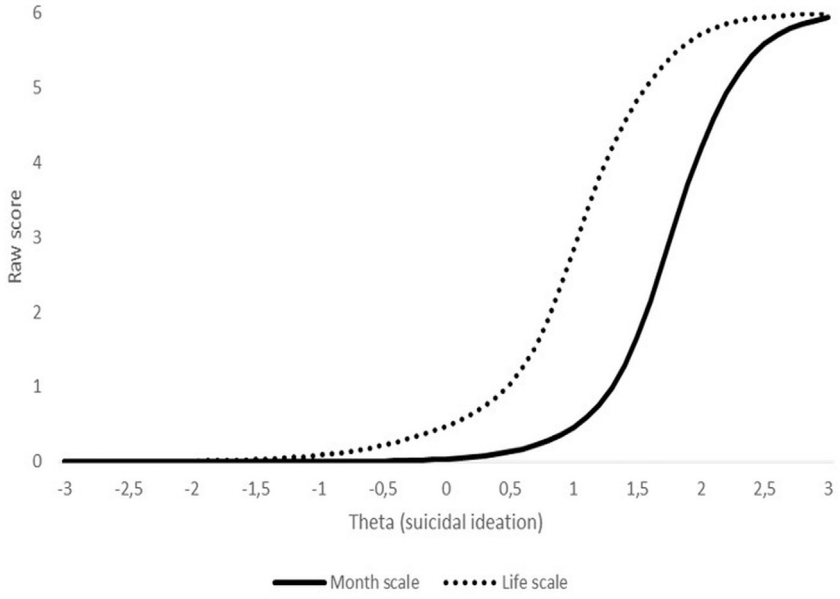

Fig. 7. Test characteristic curve.

= .45), and between 0.1 (p = .90) and 3.4 (p = .18) for SI_L and SI_M scales, respectively. Consequently, we did not compute additional iterations or DIF effect size estimations. Finally, we observed significant differences in latent means, with higher values for women in both subscales (SI_M = 0.64 SD; SI_L = 0.74 SD).

4. Discussion

We tested the accuracy of the SI sub-scale of the Columbia-Suicide Severity Rating Scale to differentiate participants according to their SI severity levels in a large and general sample of adolescents aged 13 to 18. We analyzed the difficulty and the discrimination power of each item, and the accuracy with which the scale provides information about the latent trait. The item information function analysis yielded a very high ability for the items to differentiate severity levels of SI, which can be grouped into three clusters. The test information function analysis showed that the scale provides very accurate information about SI, particularly within high SI levels. According to Baker's classification [67], discrimination parameters were very high. Overall, these findings suggest that the scale is a suitable instrument to both screen SI and to accurately classify and differentiate at-risk groups. Moreover, our results support the usage of the scale for clinical and research purposes [31] and agree with previous research showing that it is a useful instrument to efficiently detect SI in adolescents [39,40]. Nevertheless, as we slightly modified the scale, direct comparisons with our results should be made with caution.

We addressed SI by six items asking about the presence of suicidal thoughts ever in life (SIL) and/or during the last month (SIM). Although we observed a high correlation between SI_M and SI_L, the data are better represented by a two-factor model. Consequently, SI_M and SI_L can be regarded as empirically separable dimensions. First, the best fit was obtained by the two-factor model; second, there is discriminant validity (the variance explained by the items is higher than the variance explained by the between factor correlations); and third, the factors show a different relational pattern with one criterion (Fig. 8). This could mean that SI_L might be a predisposing factor to SI_M, but (for a screening) it is not a unique requisite to classify a subject as being at risk. Therefore, at a first glance, adolescents with a high score in SI_L and a low score in SI_M, could not be considered as needing an in depth evaluation or immediate clinical assistance. However, because the associations between SI_L and the current levels of emotional distress depicted in Fig. 8, and because the evidence revealing that the past SI has been shown as an important predictor of eventual death by suicide [69], subjects with this profile should be comprehensively assessed.

Regarding the general functioning of the scale, we observed a clear floor effect for both SIL and SIM scales, and consequently, most of subjects obtained the minimum score (zero). The scale is productive only within a small range of the variable, particularly in moderate-high (SIL) and high-very high (SI_M) areas, with an effectiveness range of 1 SD in both cases, which is an expectable result for a clinical scale applied to a general sample. All items showed a high discriminative capability. The exception was item 1 (SI_L scale), whose informational peak was located in a non-clinically relevant area (the mean area of the variable). It was affirmatively responded by 25% of subjects. Given the probable high proportion of false positives generated by this item, it can be regarded as an inaccurate measure to detect suicidal ideators. For this reason, a "yes" response in the SI_L scale could be interpreted as a normative adolescent phenomenon rather than a clinical symptom. Because of its low discrimination ability, it is a probable multidimensional item and a "yes" response should be considered only if the following items representing higher SI severity levels are affirmatively answered. According to both severity and peak information, the items (both scales) are organized into the following three groups: group 1 = the general idea of being dead (item 1); group 2 = the elaboration of specific ideas about suicide (items 2, 3 and 4); and group 3 = specific intentions and detailed plans (items 5 and 6). These groups represent different risk levels (i.e., low, moderate, and high risk), which provide useful information either to characterize response patterns or to adopt different prevention strategies and behaviors (i.e., supervision and follow up in case of low-moderate risk, in-depth assessment and specific interventions in case of high risk). This grouping pattern was similar for both scales, but in the case of SI_L, the items were slightly moved towards less severity areas. As we previously stated, we did

Table 3
Distribution of adolescents according to their EAP scores.

EAP score	Month scale	Life scale
Min	82.10%	49.50%
<1 SD	9.90%	24.30%
1–2 SD	6.30%	15.30%
>2 SD	1.80%	4.40%

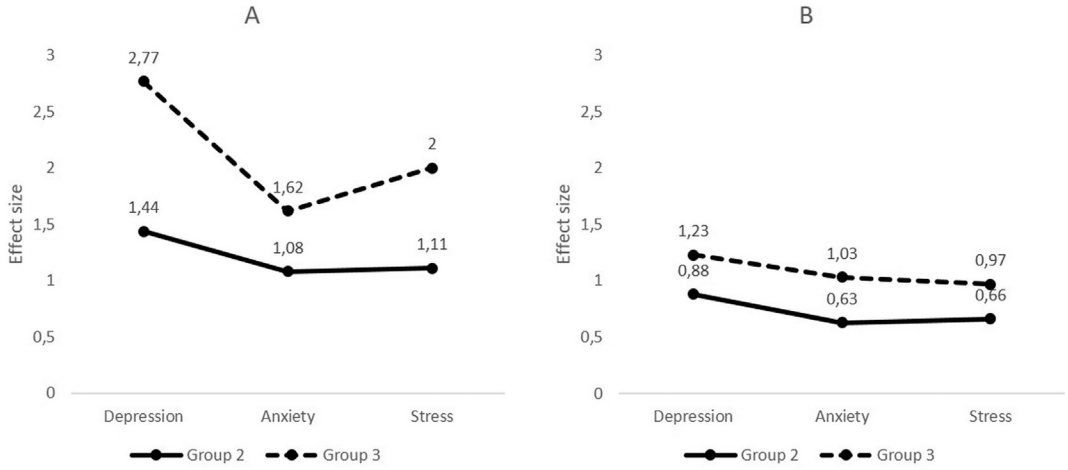

Fig. 8. Between-group comparison of latent scores for depression, anxiety and stress.

split the last question of the original scale into two items (items 5 and 6). Given that the items representing higher SI severity levels did provide more accurate information about SI, and because these items address specific and different aspects of SI (active planning and intentions), we recommend keeping these items separated if the scale is used as a self-report.

Overall, these findings suggest that the scale not only distinguishes between high and low-risk subjects, but it also allows classifying subjects according to their differential risk level, which is a relevant aspect discussed nowadays [33,43]. Despite the evidence suggesting that suicidal attempts could be prevented by identifying individuals with SI [70,71], the relationship between SI and death by suicide has been questioned [72], and further research is needed to understand the progression of SI and the specific pathways by which it may lead to suicide attempts [4,73]. As posited by Burke and Alloy [74], providing specific categories of ideators could lead to reduce false positive variables distinguishing between ideators and attempters, one of the current and critical issues in the field. Within this ideation-to-action framework [75], our results reveal that the current self-report version of the C-SSRS can accurately classify SI. This could be helpful not only for future studies analyzing SI, but also for strategies aimed to properly and timely detect individuals needing a more thorough assessment. As clearly stated by Silverman & Berman, this assessment must not be only focused on SI questions [45].

Proportion of subjects classified into one of the proposed categories fits with evidence showing that, in the general population, the prevalence of SI during a lifetime is 18.49% [76], with a range of 2.1–18.5% [42]. Additionally, our data support previous studies revealing that the lifetime prevalence of SI was 4.7% in adolescents aged 15 to 19 [77] and 5% in adolescents aged 14 to 19 in the previous year [78]. Concerning the gender of participants, we did not observe a differential functioning between men and women. Therefore, the scale seems to be a fair screening tool. It does not present gender biases; the items do not need changes and it is not necessary to generate separate rates. Moreover, the scale estimates the average differences between males and females in a precise and unbiased manner. Finally, we observed higher SI levels in females, which supports previous evidence [7,79].

We analyzed the external validity of the scale. Our results showed that higher SI levels are positively related with substantially higher emotional distress levels (depressive, anxiety and stress symptoms), addressed by the Depression, Anxiety and Stress Scales (DASS-21) [55], which support a good concurrent validity. Moreover, the greater differences in depressive symptoms are generated by higher SI levels, which is proof of discriminant validity. Although these differences are less for SI_L, they remain significant, supporting that a higher level of SI is a factor risk for psychopathology and poor functioning in adolescence [80], being particularly associated to depressive symptoms [7,16,81,82].

5. Limitations

Some limitations deserve to be mentioned. First, we did not address suicidal attempts. Therefore, we cannot either compare the functioning of the scale in subjects with suicidal ideation and subjects with suicidal attempts or draw conclusions about the external validity of SI in these two groups. Second, we did not evaluate convergent validity. Despite the lack of a gold standard to study suicide ideations [42], future studies analyzing this psychometric property could use the Beck Scale of Suicidal Ideation (BSSI), a comprehensive self-report measuring severity of SI. According to Batterham et al. [25] it shows considerable evidence of psychometric robustness in adults. It has also been used in adolescents [83] showing adequate psychometric properties in this population [84]. Third, and despite we addressed external validity, we used a limited range of variables, and further research analyzing variables suggested by recent theories is needed [3,75]. Fourth, because our cross-sectional design, we cannot test whether SI predicts future suicide attempts in our sample.

6. Conclusions

In summary, through the item response theory, our results demonstrate that the self-report version of the C-SSRS is a useful tool to address SI in adolescents. It properly distinguishes between subjects with and low risk and accurately classifies them according to their differential risk level. These findings support the usage of the C-SSRS in Chilean adolescents.

References

[1] Centers for Disease Control and Prevention. Web-based injury statistics query and reporting system, National Center for Injury Prevention and Control. CDC; 2013 http://www.cdc.gov/injury/wisqars/leadingcauses.html.
[2] Sveticic J, De Leo D. The hypothesis of a continuum in suicidality: a discussion on its validity and practical implications. Ment Illn 2012;4:73–8.
[3] Klonsky D, May D. The three-step theory (3ST): a new theory of suicide rooted in the "ideation-to-action" framework. Int J Cogn Ther 2015;8(2):114–29.
[4] May A, Klonsky D. What distinguishes suicide attempters from suicide ideators? A meta-analysis of potential factors. Clin Psychol Sci Prac 2016;23:5–20.
[5] Klonsky ED, May A, Saffer B. Suicide, suicide attempts, and suicidal ideation. Annu Rev Clin Psychol 2016;12:307–30.
[6] Auerbach RP, Tsai B, Abela JRZ. Temporal relationships among depressive symptoms, risky behavior engagement, perceived control, and gender in a sample of adolescents. J Res Adolesc 2010;20(3):726–47.

[7] Nock M, Green J, Hwang I, McLaughlin K, Sampson N, Zaslavsky A, et al. Prevalence, correlates, and treatment of lifetime suicidal behavior among adolescents. JAMA Psychiat 2013;70(3):300–10.
[8] Czyz EK, King CA. Longitudinal trajectories of suicidal ideation and subsequent suicide attempts among adolescent inpatients. J Clin Child Adolesc Psychol 2015; 44(1):181–93.
[9] Evans E, Hawton K, Rodham K, Deeks J. The prevalence of suicidal phenomena in adolescents: a systematic review of population-based studies. Suicide Life Threat Behav 2005;35(3):239–50.
[10] Horwitz AG, Czyz E, King CA. Predicting future suicide attempts among adolescent and emerging adult psychiatric emergency patients. J Clin Child Adolesc Psychol 2015;44:751–61.
[11] Miranda R, De Jaegere E, Restifo K, Shaffer D. Longitudinal follow-up study of adolescents who report a suicide attempt: aspects of suicidal behavior that increase risk of a future attempt. Depress Anxiety 2014;31:19–26.
[12] Thompson R, Proctor L, English D, Dubowitz H, Narasimhan S, Everson M. Suicidal ideation in adolescence: examining the role of recent adverse experiences. J Adolesc 2012;35(1):175–86.
[13] Scott L, Pilkonis P, Hipwell A, Keenan K, Stepp S. Non-suicidal self-injury and suicidal ideation as predictors of suicide attempts in adolescent girls: a multi-wave prospective study. Compr Psychiatry 2015;58:1–10.
[14] Gulliver A, Griffiths K, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. BMC Psychiatry 2010; 10:113.
[15] Rickwood D, Deane F, Wilson C, Ciarrochi J. Young people's help-seeking for mental health problems. Aust e-J Adv Ment Health 2005;4(3):1–34.
[16] Joffe BI, van Lieshout RJ, Duncan L, Boyle M. Suicidal ideation and behavior in adolescents aged 12–16 years: a 17-year follow-up. Suicide Life Threat Behav 2014;44:497–509.
[17] Cotter P, Kaess M, Corcoran P, Parzer P, Brunner R, Keeley H, et al. Help-seeking behaviour following school-based screening for current suicidality among European adolescents. Soc Psychiatry Psychiatr Epidemiol 2015;50(6):973–82.
[18] Husky M, Kaplan A, McGuire L, Flynn L, Chrostowski C, Olfson M. Identifying adolescents at risk through voluntary school-based mental health screening. J Adolesc 2011;34:505–11.
[19] Scott M, Wilcox H, Huo Y, Turner J, Fisher P, Shaffer D. School-based screening for suicide risk: balancing costs and benefits. Am J Public Health 2010;100(9): 1648–52.
[20] Boudreaux E, Horowitz L. Suicide risk screening and assessment designing instruments with dissemination in mind. Am J Prev Med 2014;47(3S2):S163–9.
[21] Hilt LM, Tuschner RF, Salentine C, Torcasso G, Nelson KR. Development and initial psychometrics of a school-based screening program to prevent adolescent suicide. Pract Innov 2018;3(1):1–17.
[22] King Ch, Arango A, Foster C. Emerging trends in adolescent suicide prevention research. Curr Opin Psychol 2018;22:89–94.
[23] Erbacher T, Singer J, Poland S. Suicide in schools. A practitioner's guide to multilevel prevention, assessment, intervention and postvention. New York: Routledge; 2015.
[24] Stone DM, Holland K, Bartholow B, Crosby AE, Davis S, Wilkins N. Preventing suicide: a technical package of policies, programs, and practices. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention; 2017.
[25] Batterham P, Ftanou M, Pirkis J, Brewer J, Mackinnon A, Beautrais A, et al. Systematic review and evaluation of measures for suicidal ideation and behaviors in population-based research. Psychol Assess 2015;27(2):501–12.
[26] Gould M, Marrocco F, Hoagwood K, Kleinman M, Altschuler E. Service use by at-risk youth after school-based suicide screening. J Am Acad Child Adolesc Psychiatry 2009;48(12):1193–201.
[27] Horowitz L, Ballarda E, Pao M. Suicide screening in schools, primary care and emergency departments. Curr Opin Pediatr 2009;21(5):620–7.
[28] Kaess M, Brunner R, Parzer P, Carli V, Apter A, Balazs Bobes J, et al. Risk-behaviour screening for identifying adolescents with mental health problems in Europe. Eur Child Adolesc Psychiatry 2014;23(7):611–20.
[29] King C, Eisenberg D, Zheng K, Czyz E, Kramer A, Horwitz A, et al. Online suicide risk screening and intervention with college students: a pilot randomized controlled trial. J Consult Clin Psychol 2015;83(3):630–6.
[30] Scott M, Wilcox H, Schonfeld I, Davies M, Hicks R, Turner JB, et al. School-based screening to identify at-risk students not already known to school professionals: the Columbia suicide screen. Am J Public Health 2009;99:334–9.
[31] Al-Halabí S, Sáiza P, Buróna P, Garrido M, Benabarr A, Jiménez E, et al. Validación de la versión en español de la Columbia-Suicide Severity Rating Scale (Escala Columbia para Evaluar el Riesgo de Suicidio). Rev Psiquiatr Salud Ment (Barc) 2016;9(3): 134–42.
[32] Brown G, Currier G, Jager-Hyman S, Stanley B. Detection and classification of suicidal behavior and nonsuicidal self-injury behavior in emergency departments. J Clin Psychiatry 2015;76(10):1397–403.
[33] Gipson PY, Agarwala P, Opperman KJ, Horwitz AG, King CA. Columbia-suicide severity rating scale (C-SSRS): predictive validity with adolescent psychiatric emergency patients. Pediatr Emerg Care 2015;31:88–94.
[34] King C, O'Mara R, Hayward C, Cunningham R. Adolescent suicide risk screening in the emergency department. Acad Emerg Med 2009;16:1234–41.
[35] Patel A, Watts C, Shiddell S, Couch K, Smith A, Moran M, et al. Universal adolescent suicide screening in a pediatric urgent care center. Arch Suicide Res 2017. https://doi.org/10.1080/13811118.2017.1304303.
[36] Ronquillo L, Minassian A, Vike GM, Wilson MP. Literature-based recommendations for suicide assessment in the emergency department: a review. J Emerg Med 2012;43:836–42.
[37] Runeson B, Odeberg J, Pettersson A, Edbom T, Adamsson I, Waern M. Instruments for the assessment of suicide risk: a systematic review evaluating the certainty of the evidence. PLoS One 2017;12(7):e0180292. https://doi.org/10.1371/journal.pone.0180292.
[39] Kreuze E, Lamis D. A review of psychometrically tested instruments assessing suicide risk in adults. OMEGA J Death Dying 2017;77(1):36–90.
[40] Posner K, Brown G, Stanley B, Brent D, Yershova K, Oquendo M, et al. The Columbia–suicide severity rating scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry 2011;168: 1266–77.
[41] Shaffer D, Scott M, Wilcox H, Maslow C, Hicks R, Lucas C, et al. The Columbia suicide screen: validity and reliability of a screen for youth suicide and depression. J Am Acad Child Adolesc Psychiatry 2004;43(1):71–9.
[42] Ghasemi P, Shaghaghi A, Allahverdipour H. Measurement scales of suicidal ideation and attitudes: a systematic review article. Health Promot Perspect 2015;5(3): 156–68.
[43] Interian A, Chesin M, Kline A, Miller R, St. Hill L, Latorre M, et al. Use of the Columbia-suicide severity rating scale (C-SSRS) to classify suicidal behaviors. Arch Suicide Res 2017;22(2):278–94.
[44] Giddens JM, Sheehan KH, Sheehan DV. The Columbia-suicide severity rating scale (C–SSRS): has the gold standard become a liability? Innov Clin Neurosci 2014;11: 66–80.
[45] Silverman M, Berman A. Suicide risk assessment and risk formulation part I: a focus on suicide ideation in assessing suicide risk. Suicide Life-Threat Behav 2014;44(4): 420–31.
[46] Horowitz L, Ballarda E, Pao M. Suicide screening in schools, primary care and emergency departments. Curr Opin Pediatr 2009;21(5):620–7.
[47] Yigletu H, Tucker S, Harris M, Hatlevig J. Am Psychiatr Nurses Assoc 2004;10:9–15.
[48] Embretson SE, Reise SP. Item response theory for psychologists. Mahwah, NJ: Lawrence Erlbaum; 2000.
[49] Fraley C, Waller N, Brennan C. An item response theory analysis of self-report measures of adult attachment. J Pers Soc Psychol 2000;78(2):350–65.
[50] Little RJ. A test of missing completely at random for multivariate data with missing values. J Am Stat Assoc 1988;83(404):1198–202.
[51] Sijtsma K. On the use, the misuse, and the very limited usefulness of Cronbach's alpha. Psychometrika 2009;74(1):107–20.
[52] Raykov T, Dimitrov DM, Asparouhov T. Evaluation of scale reliability with binary measures using latent variable modeling. Struct Equ Model 2010;17 (2):265–79.
[53] Dunn TJ, Baguley T, Brunsden V. From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. Br J Psychol 2014;105(3): 399–412.
[54] McDonald RP. In: Mahwah NJ, editor. Test theory: a unified treatment. L. Erlbaum Associates; 1999.
[55] Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the depression anxiety stress scales (DASS) with the Beck depression and anxiety inventories. Behav Res Ther 1995;33(3):335–43.
[56] Román F, Vinet EV, Alarcón AM. Escalas de Depresión Ansiedad Estrés (DASS-21): Adaptación y propiedades psicométricas en estudiantes secundarios de Temuco. Revista Argentina Clínica Psicológica 2014;23(2):179–90.
[57] Hambleton RK, Swaminathan J. Item response theory: principles and aplications. Boston, MA: Kluwer; 1985.
[58] Horn JL. A rationale and test for the number of factors in factor analysis. Psychometrika 1965;32:179–85.
[59] Timmerman ME, Lorenzo-Seva U. Dimensionality assessment of ordered polytomous items with parallel analysis. Psychol Methods 2011;16(2):209.
[60] Saris WE, Satorra A, van der Veld WM. Testing structural equation models or detection of misspecifications? Struct Equ Model 2009;16(4):561–82.
[61] Thissen D. IRTPRO User's Guide. Lincolnwood, IL: Scientific Software International; 2011.
[62] Meade AW. A taxonomy of effect size measures for the differential functioning of items and scales. J Appl Psychol 2010;95(4):728–43.
[63] Samejima F. A new family of models for the multiple-choice item. University of Tennessee, Knoxville, TN: Office of Naval Research Report; 1979 (79-4).
[64] Cai L, du Toit SCH, Thissen D. IRTPRO: flexible, multidimensional, multiple categorical IRT modeling [computer software]. Chicago, IL: Scientific Software International; 2011.
[65] Muthén BO, Muthén LK. Mplus 7 base program. Los Angeles, CA: Muthén & Muthén; 2012.
[66] Cai L, Thissen D, du Toit SCH. IRTPRO for windows [computer software]. Lincolnwood, IL: Scientific Software International; 2015.
[67] Bake F. The basic of item response theory: College Park: ERIC clearinghouse on assessment and evaluation. University of Maryland; 2001.
[68] Reise SP, Bonifay WE, Haviland MG. Scoring and modeling psychological measures in the presence of multidimensionality. J Pers Assess 2013;95:129–40.
[69] Beck AT, Brown GK, Steer RA, Dahlsgaard KK, Grisham JR. Suicide ideation at its worst point: a predictor of eventual suicide in psychiatric outpatients. Suicide Life-Threat Behav 1999;29:1–9.
[70] Panagioti M, Gooding P, Tarrier N. A prospective study of suicidal ideation in posttraumatic stress disorder: the role of perceptions of defeat and entrapment. J Clin Psychol 2015;71(1):50–61.
[71] Rueter M, Holm K, McGeorge C, Conger R. Adolescent suicidal ideation subgroups and their association with suicidal plans and attempts in young adulthood. Suicide Life-Threat Behav 2008;38(5):564–75.
[72] Wintersteen M, Berman AL, Silverman MM. What is meant by "increased risk for suicide?" [letter to the editor]. JAMA Pediatr 2013;167:675–6.

[73] Dhingra K, Boduszek D, O'Connor RC. Differentiating suicide attempters from suicide ideators using the integrated motivational-volitional model of suicidal behavior. J Affect Disord 2015;186:211–8.
[74] Burke T, Alloy L. Moving toward an ideation-to-action framework in suicide research: a commentary on May and Klonsky (2015). Clin Psychol 2016;23(1): 26–30. https://doi.org/10.1111/cpsp.12134.
[75] Klonsky D, May A. Differentiating suicide attempters from suicide ideators: a critical frontier for suicidology research. Suicide Life-Threat Behav 2014;44(1):1–5.
[76] Lee J, Lee M, Liao S, Chang C, Sung S, Chiang H, et al. Prevalence of suicidal ideation and associated risk factors in the general population. Formos Med Assoc 2010;109 (2):138–47.
[77] Fuller-Thomson E, Hamelin G, Stephen JR, Granger S. Suicidal ideation in a population-based sample of adolescents: implications for family medicine practice. ISRN Fam Med 2013. https://doi.org/10.5402/2013/282378.
[78] Begum B, Rahman AKM, Rahman F, Soares J, Reza H, Macassa G. Prevalence of suicide ideation among adolescents and young adults in rural Bangladesh. Int J Ment Health 2017;46(3):177–87.
[79] McKinnon B, Gariépy G, Sentenac M, Elgar F. Adolescent suicidal behaviours in 32 low- and middle-income countries. Bull World Health Organ 2016;94: 340–350F.
[80] Reinherz H, Tanner J, Berger S, Beardslee W, Fitzmaurice G. Adolescent suicidal ideation as predictive of psychopathology, suicidal behavior, and compromised functioning at age 30. Am J Psychiatry 2006;163:1226–32.
[81] Lasgaard M, Goossens L, Elklit A. Loneliness, depressive symptomatology, and suicide ideation in adolescence: cross-sectional and longitudinal analyses. J Abnorm Child Psychol 2011;39:137–50.
[82] Roberts RE, Roberts CR, Chen YR. Suicidal thinking among adolescents with a history of attempted suicide. J Am Acad Child Adolesc Psychiatry 1998;3: 1294–300.
[83] Kim YJ, Moon SS, Lee JH, Kim JK. Risk factors and mediators of suicidal ideation among Korean adolescents. Crisis 2018;39(1):4–12.
[84] Holi M, Pelkonen M, Karlsson L, Kiviruusu O, Ruuttu T, et al. Psychometric properties and clinical utility of the scale for suicidal ideation (SSI) in adolescents. BMC Psychiatry 2005;5:6. https://doi.org/10.1186/1471-244X-5-8.