

Assessment tools of non-suicidal self-injury in adolescents 1990-2016: a systematic review

Yolanda Viridiana Chávez-Flores (<https://orcid.org/0000-0003-0613-167X>)¹

Carlos Alejandro Hidalgo-Rasmussen (<https://orcid.org/0000-0002-5287-2076>)²

Libia Yanelli Yanez-Peñúnuri (<https://orcid.org/0000-0003-4682-5123>)³

Abstract *The purpose of this systematic review was to identify the instruments created or adapted to assess non-suicidal self-injury (NSSI) among adolescents. The Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) methodology was used. Two individual reviewers analyzed the psychometric properties of instruments published in English or Spanish from 1990 to 2016 considering standardized quality criteria. The PsycINFO, PubMed, ISI Web of Knowledge, Scopus, SciELO, ScienceDirect, and EBSCO databases were consulted. Eighteen studies that created or adapted 11 instruments were selected. Most were developed in the United States or Canada, and none were developed in Latin America. Several studies presented no evidence of the psychometric properties of their instruments. Seven of the 18 studies obtained at least one positive score. The Alexian Brothers Urge to Self-Injure Scale (ABUSI) and the Impulse, Self-harm, and Suicide Ideation Questionnaire for Adolescents (ISSIQ-A) obtained the highest positive scores. The limitation of this study is that only seven databases were employed for the literature search in English and Spanish. The reporting of the psychometric properties of NSSI instruments among adolescents should be improved, and adaptations to Latin American countries should be developed for international comparisons.*

Key words *Self-destructive behavior, Adolescent, Surveys and questionnaires, Cross-cultural comparison, Review*

¹ Escuela de Ciencias de la Salud Valle de las Palmas, Universidad Autónoma de Baja California. Calzada Universidad 14418, Parque Industrial Internacional. 22427 Tijuana B.C.

México. carlos.hidalgo@academicos.udg.mx

² Centro de Investigación en Riesgos y Calidad de Vida, Departamento de Promoción, Preservación y Desarrollo de la Salud, Universidad de Guadalajara. Centro de Estudios Avanzados, Universidad de Playa Ancha. Cd. Guzman Jalisco Mexico.

³ Departamento de Ciencias Sociales, Unidad Regional Norte, Universidad de Sonora. Caborca Sonora México.

Non-suicidal self-injury (NSSI) was included in the Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-5), as a condition that requires greater research¹. NSSI is defined as the deliberate self-inflicted destruction of the body or skin tissue without the intention of dying¹. NSSI behaviors include self-cutting, burning, skin rubbing, hitting, biting, and others².

Internationally, the prevalence of NSSI varies based on the studied population, age group, and criteria used to operationally define this construct. The rates of NSSI among adults during the last 6 months were estimated at 4% for the general population and 21% for the clinical population³, whereas those for adolescents over the last 12 months increase to 16.1-18.0% for the general population⁴ and 60% for the clinical population⁵. The onset age of NSSI is between 12 and 15 years^{6,7}. The risk of this behavior is higher in adolescents^{8,9} in both the general and clinical populations, which suggests the need for assessment instruments that are validated and appropriate for this age group.

The comprehensive and precise study of NSSI allows, in clinical practice, to support decision making regarding diagnosis and treatment plans. Furthermore, in research field it supports theoretical development and favors agreement among researchers for the denomination and operationalization of the construct¹⁰. Because of the incidence of NSSI behaviors and health risks associated among adolescents and adults, the number of studies and the development of various evaluation tools (including interviews and self-report tools) have increased¹¹ since the 1990s. To ensure that the data provided by these reports are useful for the diagnosis and treatment of NSSI, evaluating the psychometric quality of these instruments is paramount to generate evidence about their validity (content, criteria, and construct) and reliability (internal, consistency, and reproducibility) as well as to ensure that evidence is reported according to standardized criteria^{12,13}. Likewise, the use of exploratory or confirmatory factor analyses is required to understand the nature and quantity of the underlying factors of the variables included in an instrument¹⁴ and to provide cross-cultural validity when using instruments from different languages or countries^{13,15,16}. Cross-cultural validity is understood as how the performance of the items in a translated or culturally adapted instrument reflects the performance of the items in the original version of the instrument¹⁷.

Systematic reviews attempt to collect all of the empirical evidence that fits previously specified eligibility criteria to answer a specific research question¹⁸. Although systematic reviews and studies that analyze the psychometric properties of the instruments that evaluate NSSI have been published^{10,19,20}, they were not specifically performed with regard to adolescents. Data regarding this age group is necessary given that vulnerability to NSSI behaviors is greatest during puberty and adolescence. Therefore, the objective of this research was to execute a systematic review to identify the assessment instruments for to assess NSSI among adolescents, created or adapted from 1990 to 2016 as well as to analyze their psychometric properties.

Method

The systematic review protocol applied was approved and registered by the Research and Postgraduate Coordination of the university in charge of this study.

Search strategies

Using the methodology guidelines of Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA; <http://www.prisma-statement.org/>), an author of the present study conducted (from October 28 to November 28, 2016) a comprehensive search of studies published in English and Spanish between January 1990 and October 2016 and indexed by the PsycINFO, PubMed, ISI Web of Knowledge, Scopus, SciELO, ScienceDirect, and EBSCO databases.

The search terms were divided into three groups using the Boolean operators “AND” and “OR”. The terms in English were (“self-injury OR self-injurious behavior OR self-mutilation OR non-suicidal self-injury * OR NSSI OR self-poisoning * OR self-harm * OR deliberate self-harm * OR parasuicide OR deliberate self-injury * OR deliberate self-determination OR self-aggression * OR autoaggression * OR autoaggressive “*) AND (“instrument * OR measure * OR questionnaire * OR scale * OR assessments * OR inventory * OR psychometric * OR validity OR reliability OR psychometric properties OR factor analysis) AND (“Adolescence OR youth OR teenager OR teens OR adolescents”). The terms in Spanish were (“autolesi* OR comportamientos autolesivos OR autoagresión OR lesiones autoprovocadas OR automutilación OR autodaño

OR autolesión* OR lesiones autoinflingidas OR autodestructivo OR autoflagelación”) AND (“instrumento* OR cuestionario* OR escala* OR inventario* OR entrevista* OR autoreporte OR pruebas psicométricas OR validación OR confiabilidad OR propiedades psicométricas OR análisis factorial”) AND (“adolescencia OR adolescentes OR pubertad”).

Subsequently, another author performed a directed search to verify that the initial search identified internationally recognized instruments. In addition, both authors independently conducted a manual search based on the reference lists of the identified studies to add those not included to the main search. Additional data were requested from two authors because one published article did not include sufficient data concerning the instrument to evaluate its possible inclusion in this study.

Selection of studies

Based on the outcome of the main, directed and manual searches (which were performed independently), the research team agreed select studies based on the following inclusion and exclusion criteria. Studies that developed, adapted, or evaluated the psychometric properties of instruments or domains of NSSI with regard to adolescents (10 to 19 years of age) were included. Studies that used the term “self-injury” to evaluate suicidal behavior or those in which the evaluation of psychometric properties could not be performed because of insufficient data or the authors’ lack of response were excluded. To reduce bias with regard to the search and study selection, the definition of self-injury was reviewed in each article, and its theoretical foundation was identified to ensure that NSSI was evaluated.

Data extraction

Two of the authors of this study separately used an electronic form previously designed and piloted by the research team to collect the data. Data from each study concerning the characteristics of the instrument (i.e., name, author, country, year of publication, dimensions, and number of items), the study characteristics (i.e., population type, sample size, average age, and sex distribution), and reported evidence of the psychometric properties of each instrument (considering the criteria of Terwee et al.^{12,13}) were extracted.

Evaluation of the studies

One of the authors evaluated the quality and bias risk of the selected studies using Downes et al.’s²¹ AXIS tool. To analyze the psychometric characteristics of the instruments, two authors independently evaluated the selected studies according to the standardized criteria proposed by Terwee et al.¹² on a four-point scale (positive, intermediate, poor, or unavailable) and nine psychometric characteristics (content validity, internal consistency, criterion validity, construct validity, agreement, reliability, responsiveness, floor-ceiling effect, and interpretability). Although these criteria exclude the cross-cultural validity of the instruments, it was added in a complementary manner based on the Agreement-based Standards for the Selection of Health Measurement INstruments (COSMIN) tool (www.cosmin.nl) created by Terwee et al.¹³ that evaluates the evidence presented in cross-cultural validation studies as excellent, good, fair, or poor. Ten psychometric characteristics were evaluated. The rating provided by each reviewer was analyzed; disagreements were discussed by a third party to reach agreement. The evaluation of the psychometric characteristics performed by the two independent reviewers showed an almost perfect agreement (Kappa = .86).

Results

The main search identified 282 studies, and the manual and directed searches identified 29, for a total of 311. A total of 32 duplicate studies were eliminated. Based on the inclusion and exclusion criteria, 207 studies were eliminated, and 72 were selected for a full-text review. Based on the full review, 54 studies were excluded. The main grounds for rejection were the lack of psychometric characteristics for any instrument (71%), the evaluated instruments were created based on a population of adults or young adults (23%), or that the evaluation method included fewer than three items (5.7%). Finally, 18 studies were included in the systematic review to evaluate the psychometric properties (Figure 1).

Table 1 shows the characteristics of the studies selected for evaluation. Of these 18 studies, 11 were instruments for evaluate NSSI in adolescents: the Self-Injurious Thoughts and Behaviors Interview (SITBI)²² with one adaptation²³; the Self-Harm Behavior Questionnaire (SHBQ)²⁴; the Functional Assessment of Self-Mutila-

tion (FASM)⁶ with three adaptations²⁵⁻²⁷; the Self-Injury Motivation Scale, Adolescent version (SIMS-A)²⁸; the Alexian Brothers Assessment of Self-Injury (ABASI)²⁹; the Alexian Brothers Urge to Self-Injure Scale (ABUSI)³⁰, four adaptations of the instrument Ottawa Self-Injury Inventory (OSI)³¹⁻³⁴; the Risk-Taking and Self-Harm Inventory for Adolescents (RTSHIA)³⁵; the Repetitive Non-Suicidal Self-Injury Questionnaire (R-NS-SI-Q)³⁶; the 6-item measure of Prinstein³⁷; and the Impulse, Self-harm, and Suicide Ideation Questionnaire for Adolescents (ISSIQ-A)³⁸. With its publication in 2007, the SITBI²² became the first instrument intended to measure NSSI among adolescents. The SITBI is an interview-format instrument with 169 items that was developed in the United States using a general population of adolescents and young adults. Of the 18 included studies, eleven (61.1%) were performed in the United States, and four (22.2%) were performed in Canada. Eight (44.4%) were adapted cross-culturally. No instruments created in or adapted for Latin America were found. Of the total studies, six (33.3%) were performed with clinical populations, ten (55.5%) were performed with students, one (5.6%) was performed with a general

population, and one (5.6%) was performed with both patients and students. Regarding the use of a factor analysis, three studies (16.6%) conducted an exploratory factor analysis, eight (44.4%) conducted a confirmatory factor analysis, and seven (38.8%) did not report such an analysis.

Regarding the quality and risk of bias of the selected studies, all listed clear objectives. Concerning their methods, all showed appropriate designs and clearly defined samples with a sampling frame representing the researched population as well as risk factors and outcome variables appropriate for their objectives. Only 33% justified their sample sizes, 28% ensured a random selection, 39% took action to address non-respondents, 78% evaluated risk factors and outcome variables using previously tested or published instruments, 78% specified what was applied to determine the significance or precision estimates, and 94% described the methodology to replicate them. Regarding the results, all studies described the basic data and delivered the results of the analyses described in the methods. Although 61% provided information about non-respondents, 44% provided insufficient information about a non-response bias. Approximately 83% of the studies had outcomes that were internally consistent. Limitations were indicated in all of the studies' discussions, and 94% of the results and authors' discussions were justified. No study reported conflicts of interest, and 94% secured ethical approval or the consent of the participants.

Results of the evaluation of the psychometric properties

Table 2 shows the studies that created instruments that evaluate NSSI among adolescents and their psychometric properties. Of the seven instruments, only the ISSIQ-A³⁸ obtained positive content validity scores, whereas five instruments (71.4%) obtained poor ratings. When evaluating internal consistency, two of the seven instruments (28.5%), the ABUSI³⁰ and ISSIQ-A³⁸, obtained positive scores, and two instruments presented no data (28.5%). Five instruments (71.4%) provided data on criteria validity, but none obtained positive ratings. Regarding construct validity, five instruments (71.4%) obtained a positive score, and two instruments (28.5%) had no available data. Concerning reliability, two instruments (28.5%), the ABUSI³⁰ and SITBI²², obtained positive scores, whereas the ABASI²⁹ (14.2%) obtained poor scores, and the remaining

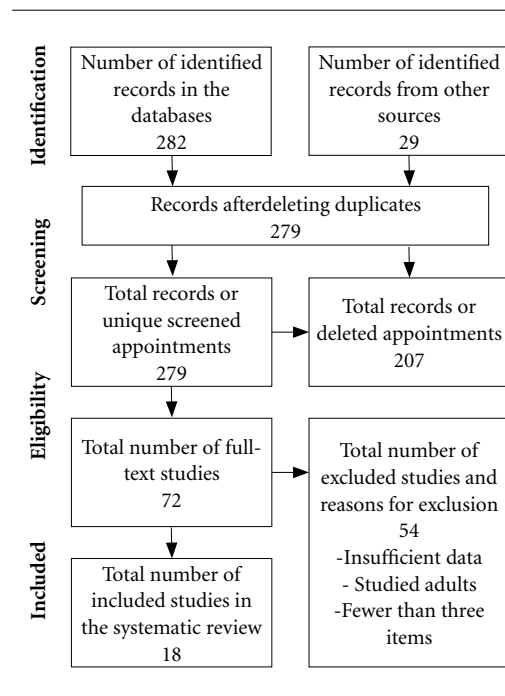


Figure 1. Description of the search and selection process of the included studies in the systematic review.

Table 1. Description of the NSSI instruments and sample characteristics.

Instrument	Author(s), year	Country of creation and/or adaptation	n	Type of participant	Age, M (SD)	No. Items	Dimensions	Factor Analysis
ABASI	Washburn et al., 2015	USA ^c	511	P	17.3 (6.2)	28	Criteria A, B, C and D of the DSM-5	EFA
ABUSI	Washburn et al., 2010	USA ^c	36	P	18.7 (7.5)	5	Unidimensional	EFA
FASM	Nock & Prinstein, 2004	USA ^c USA ^a	108	P	14.8 (1.4)	22	Methods, frequency and self-injury function	CFA
	Leong et al., 2014 (C-FASM)	USA ^c Ch ^a	825	St	11.4 (1.3)			
	Calvete et al., 2015 (FASM-E)	USA ^c S ^a	1864	St	15.3 (2.0)			
	Dahlstrom et al., 2015 (FASM)	USA ^c Sw ^a	3097	St	15-17*			
ISSIQ-A	Carvalho et al., 2015	P ^c	1722	St	16.7 (1.3)	56	Impulsivity, self-injury, risk behaviors, suicidal ideation	CFA
R-NSSI-Q	Manca et al., 2014	I ^c	383	St	23.3 (4.0)	38	Frequency and non-suicidal self-injury methods	NR
			251	R	16.4 (1.7)			
			953	St	16.2 (1.5)			
RTSHIA	Vrouva et al., 2010	En ^c	651	P	15.3 (2.1)	27	Self-injury and risk taking	CFA
				St	15.2 (2.2)			
				St	15.8 (1.5)			
SHBQ	Muehlenkamp et al., 2010	USA ^c USA ^a	1386	St	15.4 (1.4)	22	NSSI behaviors, suicidal ideation, suicide threat and past suicide attempts	CFA
SIMS-A	Swannell et al., 2008	USA ^c Au ^a	38	P	M 15.7 (1.0) H 16.1 (0.9)	22	Emotions Regulation, communicating/influencing others, punishment/excitement and psychosis/emptiness of insight	EFA

it continues

instruments presented no data. Only two instruments (28.5%), the R-NSSI-Q³⁶ and RTSHIA³⁵, presented evidence of responsiveness and obtained a positive score. None of the instruments provided data concerning agreement or of floor or ceiling effects. Five instruments (71.4%) provided data on interpretability, but none obtained a positive score.

Of the 11 studies that described an instrument adaptation (Table 3), only the SHBQ²⁴ obtained a positive score for internal consistency.

Six instrument adaptations (54.5%) obtained poor scores, and four adaptations (36.3%) presented no data. With regard to criteria validity, only the C-FASM²⁵ provided evidence and obtained an intermediate score. Regarding construct validity, only two instruments (18.1%), the C-FASM²⁵ and SHBQ²⁴, obtained positive scores, and the SITBI-G²³ (9.0%) scored poorly. Concerning reliability, only the SITBI-G²³ presented evidence and obtained a poor score. None of the instruments adaptations provided evidence of

Table 1. Description of the NSSI instruments and sample characteristics.

Instrument	Author(s), year	Country of creation and/or adaptation	n	Type of participant	Age, M (SD)	No. Items	Dimensions	Factor Analysis
SITBI	Nock et al., 2007	USA ^c	94	PG	17.1 (1.9)	169	Suicidal ideation, suicidal planning, suicidal gestures, suicidal attempt, and NSSI	NR
	Fischer et al., 2014 (SITBI-G)	USA ^c G ^a	111	P	15.4 (1.7)			NR
OSI	Plener et al., 2009	C ^c G ^a	665	E	14.8 (0.7)	21	Self-injury methods, suicidal behavior, functions, coping strategies, and addiction	NR
	Csorba et al., 2010	C ^c Hu ^a	427	E	16.7 (NR)	37	Impulse, behaviors, feelings, and behavior climate	NR
	Rodav et al., 2014	C ^c Is ^a	275	E	14.8 (1.4)	34	Frequency, methods, and functions of NSSI	NR
	Nixon et al., 2015	C ^c C ^a	94	P	15.7 (1.5)	31	Internal emotional regulation, social influence, external emotional regulation, and search for sensations	CFA
6 Item-Measure	Prinstein et al., 2008	USA ^c	148	E	13.5 (0.8)	6	NR	NR

Note: ABASI, Alexian Brothers Assessment of Self-Injury; ABUSI, Alexian Brothers Urge to Self-Injure Scale; FASM, Functional Assessment of Self-Mutilation; ISSIQ-A, Impulse, Self-harm and Suicide Ideation Questionnaire for Adolescents; R-NSSI-Q, Repetitive Non-Suicidal Self-Injury Questionnaire; RTSHA, Risk-taking and self-harm inventory for adolescents; SHBQ, Self-Harm Behavior Questionnaire; SIMS-A, Self-Injury Motivation Scale Adolescent. Au, Australia; Ch, China; S, Spain; USA, United States of America; I, Italy; En, England; P, Portugal; Sw, Sweden. P, Patients; R, Residents in therapeutic community; St, Students; CFA, Confirmatory factor analysis; EFA, Exploratory factor analysis. * Age range is presented; c Country of creation; a Country of adaptation.

agreement, responsiveness or floor or ceiling effects. Regarding interpretability, five instrument adaptations (45.5%) presented evidence and obtained intermediate scores.

Only the FASM²⁵ adaptation in China obtained a positive evaluation regarding cross-cultural validity, and the remaining adaptations (87.5%) were evaluated as poor. These poor evaluations were for not describing or poorly describing the translator's expertise regarding the tested construct; for not clarifying whether there was more than one translator, whether they worked independently, or how they resolved their differences; or for not clarifying whether the translation was reviewed by a committee involving the researchers or the creators of the original instrument. Some studies did not pilot the translated instrument or perform a confirmatory factor analysis.

Discussion

This systematic review identified the NSSI assessment instruments created or adapted for adolescents between 1990 and 2016 as well as analyzed the outcomes of their psychometric properties. A total of 18 studies were identified; seven created instruments, and 11 adaptations. Seven of the 18 studies obtained at least one positive score. The ABUSI³⁰ and the ISSIQ-A³⁸ obtained the greatest number of positive scores, fulfilling three of the nine evaluation criteria. However, no evidence regarding the psychometric properties of the various instruments was found using Terwee et al.'s^{12,13} evaluation criteria, and no data were presented regarding agreement or floor or ceiling effects.

Because NSSI is a widespread phenomenon reported across different countries⁴, it is associa-

Table 2. Analysis of the psychometric properties of studies reporting the creation of instruments to assess NSSI among adolescents considering the criteria of Terwee et al.^{1,2}

Instrument	Content Validity	Internal Consistency	Criteria Validity	Construct Validity	Agreement	Reliability	Responsiveness	Floor-Ceiling Effect	Interpretability	Positive Scores Total
ABASI	-	-	?	0	0	-	0	0	?	0
ABUSI	-	+	?	+	0	+	0	0	?	3
ISSIQ-A	+	+	-	+	0	0	0	0	?	3
R-NSSI-Q	?	0	?	+	0	0	+	0	?	2
RTSHIA	-	-	?	+	0	0	+	0	?	2
SITBI	-	0	0	+	0	+	0	0	0	2
6-item measure	-	-	0	0	0	0	0	0	0	0

Note: ABASI, Alexian Brothers Assessment of Self-Injury; ABUSI, Alexian Brothers Urge to Self-Injure Scale; ISSIQ-A, Impulse, Self-harm and Suicide Ideation Questionnaire for Adolescents; R-NSSI-Q, Repetitive Non-Suicidal Self-Injury Questionnaire; RTSHIA, Risk-taking and self-harm inventory for adolescents; SITBI, Injurious Thoughts and Behaviors Interview. Criteria: + positive; ? intermediate; - poor; 0 unavailable.

Table 3. Analysis of the psychometric properties of the adaptations of the instruments evaluating NSSI among adolescents considering the criteria of Terwee et al.^{1,2}

Instrument	Content Validity	Internal Consistency	Criteria Validity	Construct Validity	Agreement	Reliability	Responsiveness	Floor-Ceiling Effect	Interpretability	Positive Scores Total
FASM	EU	-	0	0	0	0	0	0	0	0
C-FASM	Ch	-	?	+	0	0	0	0	?	1
FASM-E	E	-	0	0	0	0	0	0	?	0
FASM	S	-	0	0	0	0	0	0	0	0
SHBQ	EU	+	0	+	0	0	0	0	?	2
SIMS-A	Au	-	0	0	0	0	0	0	0	0
SITBI-G	A	0	0	-	0	-	0	0	0	0
OSI	A	0	0	0	0	0	0	0	?	0
OSI	Hu	0	0	0	0	0	0	0	0	0
OSI	Is	0	0	0	0	0	0	0	?	0
OSI	C	-	0	0	0	0	0	0	0	0

Note: FASM, Functional Assessment of Self-Mutilation; SHBQ, Self-Harm Behavior Questionnaire; SIMS-A, Self-Injury Motivation Scale Adolescent; SITBI, Injurious Thoughts and Behaviors Interview; OSI, Ottawa Self-Injury Inventory. G, Germany; Au, Australia; C, Canada; Ch, China; S, Spain; USA, United States of America; Hu, Hungary; Is, Israel; Sw, Sweden. Criteria: + positive; ? intermediate; - poor; 0 unavailable.

ted with negative well-being^{39,40}, higher probability of suicide^{41,43}, and a higher prevalence of NSSI has been reported among adolescents compared with adult populations, and that the present review identified 11 instruments, more than half of which were created over the last 7 years (as well as seven of eight adaptations), suggest that this field of study is still incipient. No instruments created for or adapted to Latin America were identified, despite evidence that NSSI is a common phenomenon in these populations⁴⁴⁻⁴⁷. Eleven instruments were created or adapted in the United States; only the FASM^{6,25-27}, the SITBI^{22,23}, and the OSI³¹⁻³⁴ were adapted for application among the adolescents of several countries to allow for international comparisons³¹. Therefore, it is necessary to make progress toward culturally adapting the instruments with the best psychometric properties for several regions of the world.

Regarding the target population of the analyzed instruments, seven were originally designed for adolescents, and four were designed for adults. It is necessary that the instruments applied to adolescents be designed specifically for this age group; if they were created for adults, then a semantic adaptation should be made because adolescents might interpret and answer questions from an instrument in a different way because of developmental differences compared with adults⁴⁸.

Of the evaluated instruments, those with the greatest evidence of validity and reliability were the ABUSI³⁰ and ISSIQ-A³⁸. These instruments evaluated the different dimensions of the NSSI. The ABUSI³⁰ evaluates criteria A, B, C, and D of the DSM-5, whereas the ISSIQ-A³⁸ is useful for determining impulsivity, NSSI, risky behaviors, and the intrapersonal or interpersonal functions of the NSSI. More dimensions are addressed in other instruments that might be useful in clinical practice given the relevant information they provide; however, it is necessary to provide greater evidence of validity and reliability. Seven of the 18 analyzed studies did not report the use of factor analysis methods. These methods provide evidence of the exploration or confirmation of the dimensions evaluated by an instrument.

Eleven of the 18 analyzed instruments were created for or adapted to general or school populations; therefore, it is important to evaluate the psychometric properties of the instruments across different types of vulnerable populations (e.g., clinical populations), which would enable researchers to establish cut-off points and broaden the analysis of psychometric characteristics

such as responsiveness and specificity; however, one possible difficulty would be accessing this population, partially because of the low prevalence of NSSI or the lack of seeking help among those who show these behaviors.

Assessing the psychometric properties of the instruments is important because of its relevance regarding the evaluation and treatment of NSSI. Regarding the internal consistency of the creations, the ABUSI³⁰, the ISSIQ-A³⁸, and the adaptation of the SHBQ²⁴ obtained positive scores. This criterion was met by only three instruments. From our perspective, this finding suggests that difficulties (e.g., those regarding content validity) were present during previous phases of the construction of the instruments or their adaptations. Thus, instruments with internal consistency issues also scored low on content validity. The application of Cronbach's alpha to evaluate the internal consistency of the instruments was criticized because this statistic improves as the number of items in the instrument increases^{49,50}. The evidence for content validity was weak because only one of the seven created instruments, the ISSIQ-A³⁸, obtained positive scores. The participation of experts and the target population in focus groups is critical for the items' final outcomes (and the domains to which they contribute) to be included in the instrument. Although most of the analyzed studies did not report satisfactory evidence with regard to the two evaluated criteria, they contribute to an understanding of how the scales behave.

No instrument obtained positive scores regarding criteria validity because there are few gold standards in psychology. Therefore, the need to create standardized evaluation criteria and updates of the psychometric properties of the instruments is evident because the criteria of Terwee *et al.*¹² were elaborated from a medical context. Regarding construct validity, obtained positive scores the instruments created ISSIQ-A³⁸, ABUSI³⁰, R-NSSI-Q³⁶, RTSHIA³⁵, and SITBI²², and those adapted the C-FASM²⁵ and SHBQ²⁴. Because this type of validity helps prove that a specific construct is being evaluated, it is an essential feature among instruments designed to evaluate the set of indicators that comprise or relate to a specific diagnostic criterion. Based on all of the analyzed studies, it is noteworthy that this value was not reported by more than half. This fact might be related to the shortage of validated instruments that evaluate NSSI among adolescents, through which a convergent validity analysis could be performed.

Only the created instruments the ABUSI³⁰ and SITBI²² obtained positive reliability scores. Terwee et al.¹² distinguished between the relative and absolute measurement of reliability, which he called reliability and agreement, respectively. No study reported having calculated this latter condition, which shows that an analysis like the minimum important change (MIC), which describes the certainty that the magnitude of a change is not the result of chance, is not widely used among researchers in this field even though its use has increased, most likely because it requires more dissemination and training for this community of researchers to become familiar with it, compared with traditional analysis Kappa. Although reliability is much less complex to calculate, presents difficulties because of the need for multiple measurements of the same participants, so it was only evaluated by four instruments: the ABASI²⁹, ABUSI³⁰, and SITBI²² as well as the adaptation SITBI-G²³. Therefore, we also think that the journals where these studies are published should play a primary role in requiring that certain important analyses are executed by the researchers to guarantee the quality of the creation or adaptation of an instrument.

Positive scores in responsiveness were only reported for the created instruments R-NSS-Q³⁶ and RTSHIA³⁵. Previous systematic reviews such as that concerning self-injury instruments for adults¹⁹ have argued that the importance of reporting responsiveness has received little attention. It is necessary to insist on the importance of evaluating these properties for the creation or adaptation of instruments because, rather than being an exclusive failure of the NSSI instruments aimed at adolescents, it is a general weakness of the studies and reviews of the creation or adaptation of instruments.

Our results showed that of the seven studies that culturally adapted an instrument, only one adaptation, the C-FASM²⁵, obtained a good evaluation, and six adaptations obtained poor evaluations. This result occurred because the process of translating the instrument and verifying its adaptation through a review by a committee, pilot study, or confirmatory factor analysis was not described or only briefly described. Thus, the instruments evaluated in this review were not culturally adapted based on established standards. In this regard, Leong et al.²⁵ mentioned that the translation of psychological instruments without empirical validation is common because the psychometric properties are not exhaustively verified or the procedure is not

described in detail. It cannot be assumed that a translated version of an instrument has the same psychometric properties as the original version; therefore, the validity of a translated instrument should be evaluated before applying it to another country or culture.

The lack of high scores regarding the evaluation of an instrument does not necessarily denote poor psychometric properties; rather, it represents an absence of evidence because no available data were found for any of the nine evaluated criteria for any of the instruments. One relevant aspect is that several instruments showed evidence of their psychometric characteristics, although they differed from Terwee et al.'s¹² criteria. Therefore, the evidence was considered as insufficient. In the case of the RTSHIA³⁴, reliability was evaluated using Pearson's *r* coefficient, which is not an incorrect procedure⁵¹. In the case of the SITBI-G²³, inter-rater reliability was evaluated by agreement; although this method delivers excellent Kappa coefficients, it is not possible to evaluate it using standardized criteria. In addition, the considered evaluating criteria were designed for self-report instruments; therefore, because the SITBI^{22,23} is an interview, it was not possible to evaluate its internal consistency or perform a factor analysis.

Regarding the current review's limitations, this search was conducted using only seven databases and two languages; therefore, subsequent studies might seek to broaden this search. It was not possible to collect data from one study, even after the corresponding author was contacted. Future studies should address important practical aspects when selecting an evaluation instrument, such as the time to respond according to the number of items, the required resources for qualifications, or the training needed for its application⁵².

One challenge for Ibero-American researchers is to quickly become familiar with the progress toward measuring the properties of instruments and their evaluation techniques to generate knowledge. Likewise, it is important that the knowledge and techniques required to evaluate the properties of instruments in developing countries (e.g., most of Latin America), especially those that are least likely to be reported in scientific articles, be promoted in research and training centers. If the researchers know and value the important role of instrument adaptations, they will achieve intercultural comparability and a greater understanding of the studied phenomenon, in this case the NSSI, even if they

are developed and designed in other countries out of necessity.

The use of different nomenclatures to refer to the NSSI, the different ways of managing it, and the lack of adequately standardized and validated assessment instruments might trigger potential biases in the estimates of NSSI⁴, which limits our knowledge of its incidence and prevalence⁵³. This review found that little or insufficient evidence regarding the psychometric properties of the ins-

truments evaluating NSSI among adolescents is available, which justifies greater research because it is during this developmental stage that higher prevalence of NSSI is observed, in both general and clinical populations. This outcome is consistent with that regarding the NSSI instruments for adults, where a lack of developed instruments and insufficient evidence of psychometric properties is also observed. This lack of data hinders the evaluation of NSSI using standardized methods¹⁰.

Collaborations

YV Chávez-Flores worked on the design, analysis, interpretation, drafting, review and approval of the final version of the article. CA Hidalgo-Rasmussen worked on the conception, design, analysis, interpretation, drafting, review and approval of the final version of the article. LY Yanez-Peñúñuri worked on the analysis of the studies, interpretation, review and approval of the final version of the article.

Acknowledgments

The author expresses gratitude to Dr. Karina Franco Paredes for her orientation and support with regard to the methodological revision of this study.

The main author received grant from the National Council of Science and Technology. The funding source played no role in conducting this study. No conflicts of interest are present.

References

- Asociación Americana de Psiquiatría. *Manual Diagnóstico y Estadístico de los Trastornos Mentales DSM-5*. México: APA; 2014.
- Zetterqvist M, Lundh LG, Dahlström Ö, Svedin CG. Prevalence and function of non-suicidal self-injury (NSSI) in a community sample of adolescents, using suggested DSM-5 criteria for a potential NSSI disorder. *J Abnorm Child Psychol* 2013; 41(5):759-773.
- Briere J, Gil E. Self-mutilation in clinical and general population samples: Prevalence, correlates, and functions. *Am J Orthopsychiatry* 1998; 68(4):609-620.
- Muehlenkamp JJ, Claes L, Havertape L, Plener PL. International prevalence of adolescent non-suicidal self-injury and deliberate self-harm. *Child Adolesc Psychiatry Ment Health* 2012; 6(1):10.
- Kaess M, Parzer P, Mattern M, Plener PL, Bifulco A, Resch F, Brunner R. Adverse childhood experiences and their impact on frequency, severity, and the individual function of nonsuicidal self-injury in youth. *Psychiatry Res* 2012; 206(2-3):265-272.
- Nock MK, Prinstein MJ. A functional approach to the assessment of self-mutilative behavior. *J Consult Clin Psychol* 2004; 72(5):885-890.
- Muehlenkamp JJ, Gutiérrez PM. Risk for Suicide Attempts Among Adolescents Who Engage in Non-Suicidal Self-Injury. *Arch Suicide Res*. 2007; 11(1):69-82.
- Hawton K, Van Heeringen K. *The International Handbook of Suicide and Attempted Suicide*. London: Wiley; 2008.
- Klonsky ED. Non-suicidal self-injury in United States adults: prevalence, sociodemographics, topography and functions. *Psychol Med* 2011; 41(9):1981-1986.
- Klonsky ED, Lewis SP. Assessment of Nonsuicidal Self-Injury. In: *The Oxford Handbook of Suicide and Self-Injury*. New York: Oxford University Press; 2014.
- Asociación Española de Psiquiatría del Niño y el Adolescente (AEPNYA). *Autolesiones*. Madrid: Protocolos; 2008.
- Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, de Vet HCW. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; 60(1):34-42.
- Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Qual Life Res* 2012; 21(4):651-657.
- Fabrigar L, MacCallum R, Wegener D, Strahan E. Evaluating the use of Exploratory Factor Analysis in Psychological Research. *Psychol Methods* 1999; 4(3):272-299.
- Carvajal A, Centeno C, Watson R, Martínez M, Sanz Rubiales Á. ¿Cómo validar un instrumento de medida de la salud? *An Sist Sanit Navar* 2011; 34(1):63-72.
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000; 25(24):3186-3191.
- Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HCW. International consensus on taxonomy, terminology, and definitions of measurement properties for health related patient reported outcomes: results of the COSMIN study. *J Clin Epidemiol* 2010; 63:737-745.
- Green S, Higgins PT, Alderson P, Clarke M, Mulrow DC, Oxman DA. *Cochrane Handbook: Cochrane Reviews: Ch 1: Introduction*. In: *Cochrane Handbook for Systematic Reviews of Interventions* 2011; 6:3-10.
- Latimer S, Meade T, Tennant A. Measuring engagement in deliberate self-harm behaviours: psychometric evaluation of six scales. *BMC Psychiatry* 2013; 13(1):4.
- Borschmann R, Hogg J, Phillips R, Moran P. Measuring self-harm in adults: A systematic review. *Euro Psychiatry* 2012; 27(3):176-180.
- Downes MJ, Brennan ML, Williams HC, Dean RS. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ Open* 2016; 6:e011458.
- Nock MK, Holmberg EB, Photos VI, Michel BD. Self-Injurious Thoughts and Behaviors Interview: Development, reliability, and validity in an adolescent sample. *Psychol Assess* 2007; 19(3):309-317.

23. Fischer G, Ameis N, Parzer P, Plener PL, Groschwitz R, Vonderlin E, Kaess M. The German version of the self-injurious thoughts and behaviors interview (SITBI-G): a tool to assess non-suicidal self-injury and suicidal behavior disorder. *BMC Psychiatry* 2014; 14(1):265.
24. Muehlenkamp JJ, Cowles ML, Gutierrez PM. Validity of the self-harm behavior questionnaire with diverse adolescents. *J Psychopathol Behav Assess* 2010; 32(2):236-245.
25. Leong CH, Wu AMS, Poon MMY. Measurement of perceived functions of non-suicidal self-injury for chinese adolescents. *Arch Suicide Res* 2014; 18(2):193-212.
26. Calvete E, Orue I, Aizpuru L, Brotherton H. Prevalence and functions of non-suicidal self-injury in Spanish adolescents. *Psicothema* 2015; 27(3):223-228.
27. Dahlstrom O, Zetterqvist M, Lundh LG, Svedin CG. Functions of nonsuicidal self-injury: exploratory and confirmatory factor analyses in a large community sample of adolescents. *Psychol Assess* 2015; 27(1):302-313.
28. Swannell S, Martin G, Scott J, Gibbons M, Gifford S. Motivations for Self-Injury in an Adolescent Inpatient Population: Development of a Self-Report Measure. *Australas Psychiatry* 2008; 16(2):98-103.
29. Washburn JJ, Potthoff LM, Styer DM. Assessing DSM-5 Nonsuicidal Self-Injury Disorder in a Clinical Sample. *Psychol Assess* 2015; 27(1):31-41.
30. Washburn JJ, Juzwin KR, Styer DM, Aldrige D. Measuring the urge to self-injure: Preliminary data from a clinical sample. *Psychiatry Res* 2010; 178(3):540-544.
31. Plener PL, Libal G, Keller F, Fegert JM, Muehlenkamp JJ. An international comparison of adolescent non-suicidal self-injury (NSSI) and suicide attempts: Germany and the USA. *Psychol Med* 2009; 39(9):1549.
32. Csorba J, Dinya E, Ferencz E, Páli E, Nagy E, Horváth A, Vados M. A study of Hungarian adolescent outpatients suffering from self-injurious behaviour. *Psychiatr Danub* 2010; 22(1):39-45.
33. Rodav O, Levy S, Hamdan S. Clinical characteristics and functions of non-suicide self-injury in youth. *Eur Psychiatry* 2014; 29(8):503-508.
34. Nixon MK, Levesque C, Preyde M, Vanderkooy J, Cloutier PF. The Ottawa Self-Injury Inventory: Evaluation of an assessment measure of nonsuicidal self-injury in an inpatient sample of adolescents. *Child Adolesc Psychiatry Ment Health* 2015; 9:26.
35. Vrouva I, Fonagy P, Fearson PRM, Roussov T. The risk-taking and self-harm inventory for adolescents: Development and psychometric evaluation. *Psychol Assess* 2010; 22(4):852-865.
36. Manca M, Presaghi F, Cerutti R. Clinical specificity of acute versus chronic self-injury: Measurement and evaluation of repetitive non-suicidal self-injury. *Psychiatry Res* 2014; 215(1):111-119.
37. Prinstein MJ, Nock MK, Simon V, Aikins JW, Cheah CSL, Spirito A. Longitudinal trajectories and predictors of adolescent suicidal ideation and attempts following inpatient hospitalization. *J Consult Clin Psychol* 2008; 76(1):92-103.
38. Carvalho CB, Nunes C, Castilho P, da Motta C, Caldeira S, Pinto-Gouveia J. Mapping non-suicidal self-injury in adolescence: Development and confirmatory factor analysis of the impulse, self-harm and suicide ideation questionnaire for adolescents (ISSIQ-A). *Psychiatry Res* 2015; 227(2-3):238-245.
39. Nock MK, Joiner TE, Gordon KH, Lloyd-Richardson E, Prinstein MJ. Non-suicidal self-injury among adolescents: Diagnostic correlates and relation to suicide attempts. *Psychiatry Res* 2006; 144(1):65-72.
40. Ross S, Heath NL, Toste JR. Non-suicidal self-injury and eating pathology in high school students. *Am J Orthopsychiatry* 2009; 79(1):83-92.
41. Whitlock J, Muehlenkamp J, Eckenrode J, Purington A, Baral AG, Barreira P, Kress V. Nonsuicidal self-injury as a Gateway to suicide in Young adults. *J Adolesc Health* 2013; 52:4.
42. Grandclerc S, De Labrouhe D, Spodenkiewicz M, Lachal J, Moro MR. Relations between nonsuicidal self-injury and suicidal behavior in adolescence: A systematic review. *PLoS ONE* 2016; 11:4.
43. Asarnow JR, Porta G, Spirito A, Emslie G, Clarke G, Wagner KD, Brent DA. Suicide attempts and non-suicidal self-injury in the treatment of resistant depression in adolescents: Findings from the TOR-DIA study. *J Am Acad Child Adolesc Psychiatry* 2011; 50(8):772-781.
44. Benjet C, González-Herrera I, Castro-Silva E, Méndez E, Borges G, Casanova L, Medina-Mora ME. Non-suicidal self-injury in Mexican Young adults: Prevalence, associations with suicidal behavior and psychiatric disorders, and DSM-5 proposed diagnostic criteria. *J Affect Disord* 2017; 215:1-8.
45. Arcoverde RL, Amazonas MCLDA. *Autolesão deliberada: relatos em comunidades virtuais*. Recife: Universidade Federal de Pernambuco; 2011.
46. Thyssen LS, Camp IV. Non-suicidal Self-Injury in Latin America. *Salud Ment* 2014; 37(2):153-157.
47. Albores-Gallo L, Méndez-Santos JL, García A, Delgadillo GY, Chávez-Flores CI, Martínez OI. Autolesiones sin intención suicida en una muestra de niños y adolescentes de la ciudad de México. *Actas Esp Psiquiatr* 2014; 42(4):159-168.
48. Gutierrez PM, Osman A. *Adolescent suicide: An integrated approach to the assessment of risk and protective factors*. DeKalb: Northern Illinois University Press; 2008.
49. Boyle GJ. Does item homogeneity indicate internal consistency or item redundancy in psychometric scales? *Pers Individ Dif* 1991; 12(3):291-294.
50. Shevlin M, Miles JN, Davies MN, Walker S. Coefficient alpha: a useful indicator of reliability? *Pers Individ Dif* 2000; 28(2):229-237.
51. Vaz S, Falkmer T, Passmore AE, Parsons R, Andreou P. The Case for Using the Repeatability Coefficient When Calculating Test-Retest Reliability. *PLoS ONE* 2013; 8:9.
52. Keszei AP, Novak M, Streiner DL. Introduction to health measurement scales. *J Psychosom Res* 2010; 68(4):319-323.
53. Butler AM, Malone K. Attempted suicide v. non-suicidal self-injury: Behaviour syndrome or diagnosis? *Br J Psychiatry* 2013; 202(5):324-325.

Article submitted 21/05/2017

Approved 18/12/2017

Final version submitted 20/12/2017

