Psychological Assessment
2007, Vol. 19, No. 3, 309–317

1040-3590/07/$12.00 DOI: 10.1037/1040-3590.19.3.309

Self-Injurious Thoughts and Behaviors Interview: Development, Reliability, and Validity in an Adolescent Sample

Matthew K. Nock, Elizabeth B. Holmberg, Valerie I. Photos, and Bethany D. Michel
Harvard University

The authors developed the Self-Injurious Thoughts and Behaviors Interview (SITBI) and evaluated its psychometric properties. The SITBI is a structured interview that assesses the presence, frequency, and characteristics of a wide range of self-injurious thoughts and behaviors, including suicidal ideation, suicide plans, suicide gestures, suicide attempts, and nonsuicidal self-injury (NSSI). This initial study, based on the administration of the SITBI to 94 adolescents and young adults, suggested that the SITBI has strong interrater reliability (average $\kappa = .99$, $r = 1.0$) and test–retest reliability (average $\kappa = .70$, intraclass correlation coefficient = .44) over a 6-month period. Moreover, concurrent validity was demonstrated via strong correspondence between the SITBI and other measures of suicidal ideation (average $\kappa = .54$), suicide attempt ($\kappa = .65$), and NSSI (average $\kappa = .87$). The authors concluded that the SITBI uniformly and comprehensively assesses a wide range of self-injury-related constructs and provides a new instrument that can be administered with relative ease in both research and clinical settings.

Keywords: suicide, self-injury, assessment, reliability, validity

Although impressive advances have been made in the study of self-injurious thoughts and behaviors (SITB) over the past several decades (Hawton & van Heeringen, 2000; Jacobs, 1999; Maris, Berman, & Silverman, 2000), the rates of SITB in the general population have remained virtually unchanged (Kessler, Berglund, Borges, Nock, & Wang, 2005). This may be due in part to a lack of clarity and consistency in the way SITB are measured across research studies and clinical settings. Science and practice in this area will advance most rapidly with the availability of measures that clearly and consistently assess these behaviors.

A review of all existing measures of self-injury-related constructs is beyond the scope of this article; however, a brief overview of prior work and a summary of key limitations will help to place the current work in context. Given the importance of SITB and the problems in the measurement of these constructs, the National Institute of Mental Health (NIMH) recently commissioned two systematic reviews of suicide assessment measures available for use with children/adolescents (Goldston, 2000) and adults/older adults (Brown, 2000). These comprehensive reviews identified numerous, valuable measures currently used to assess self-injury-related constructs; however, in doing so they also highlighted several important limitations. One significant limitation is that many of the measures currently available do not use clear and specific definitions of the SITB being assessed. Although operational definitions for distinct types of SITB have been outlined (O'Carroll, Berman, Maris, & Moscicki, 1996), many measures fail to adhere to these definitions, and many do not differentiate among different SITB constructs (e.g., some measures classify all self-injurious behaviors as "parasuicide" or "suicide attempts," regardless of whether the individuals intended to die from their behavior). The use of broad and inconsistent definitions makes it difficult to compare results from different studies and also can obscure important differences in the data and lead to erroneous conclusions (see Linehan, 1997, for a review). For instance, a recent reanalysis of data from the National Comorbidity Survey (Kessler et al., 1994) revealed that requiring intent to die in the definition of a *suicide attempt* reduced the U.S. lifetime prevalence of self-reported suicide attempts from 4.6% to 2.7% and exposed important differences between those with intent to die and those who engaged in self-injury without such intent (Nock & Kessler, 2006). Yet most measures that include questions about suicide attempts do not specifically assess the presence of intent to die (Nock, Wedig, Janis, & Deliberto, in press).

A related concern is that many of the measures used to assess SITB do not provide clear, objective data about the presence and frequency of the SITB in question, but instead provide data on somewhat arbitrary scales that may be difficult for clinicians and researchers to interpret (Blanton & Jaccard, 2006; Kazdin, 2006). For instance, the fact that a patient scored a "7" on a measure of the severity of suicidal ideation may be less readily useful to most clinicians than knowing that an individual has thought about killing herself daily for the past month or that she has made three suicide attempts in the past year. Data reported in nonarbitrary metrics such as the presence, frequency, and characteristics on different types of SITB, although basic, will likely be of value for clinical and research purposes.

Matthew K. Nock, Elizabeth B. Holmberg, Valerie I. Photos, and Bethany D. Michel, Department of Psychology, Harvard University.

This research was supported by grants from the Milton Fund and Talley Fund of Harvard University to Matthew K. Nock. We are grateful to the members of the Laboratory for Clinical and Developmental Research for their assistance with this work and to the adolescents and families who participated in this study.

Correspondence concerning this article should be addressed to Matthew K. Nock, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138. E-mail: nock@wjh.harvard.edu

A second major limitation is that virtually all existing measures, even those that include clear and specific definitions of SITB, assess only a limited range of SITB. More specifically, most assess suicidal ideation; a smaller number examine suicide attempts; and few assess other constructs such as suicide plans, suicide gestures (i.e., leading someone to believe one wants to die by suicide when there is no intention of doing so),[1] and nonsuicidal self-injury (NSSI; i.e., direct, deliberate self-injury in which there is no intent to die). When these latter constructs are assessed, it typically is done using a single item. The narrow focus of prior measures limits the availability of information about the less studied constructs, introduces further difficulties in making comparisons across studies (given variability in assessment procedures across studies), and precludes the ability to examine relations *among* self-injury-related constructs. Assessing relations among different SITB is important because it can help psychologists to better understand when and how these different constructs are related, and also because earlier work suggests that milder forms of SITB are often among the best predictors of more severe SITB. For instance, the presence of a suicide plan and the presence of NSSI are both associated with an increased risk of suicide attempt (Kessler, Borges, & Walters, 1999; Nock, Joiner, Gordon, Lloyd-Richardson, & Prinstein, 2006).

The need for more comprehensive methods of assessing SITB has been highlighted over the course of decades (Hirschfeld & Blumenthal, 1986; Reynolds, 1990) and remains an important task for improving evidence-based assessment in this area (Joiner, Walker, Pettit, Perez, & Cukrowicz, 2005). Indeed, the development of measures that uniformly and comprehensively examine the presence, frequency, and characteristics of a broad range of SITB would enable researchers and clinicians to more carefully examine the relations among these different types of SITB, and to test the relations between each type of SITB and related constructs; and it would facilitate valid comparisons across research studies and clinical settings.

The purpose of the current study was to develop a comprehensive measure of a wide range of SITB that can be easily used by researchers and clinicians, and to conduct a preliminary evaluation of this tool. Toward this end, we report here on the Self-Injurious Thoughts and Behaviors Interview (SITBI), a brief interview-based measure that uniformly assesses the presence, frequency, and characteristics of (a) suicidal ideation, (b) suicide plans, (c) suicide gestures, (d) suicide attempts, and (e) NSSI. The characteristics of SITB assessed by the SITBI include age of onset, methods, severity, functions, precipitants, experience of pain, use of alcohol and drugs during SITB, impulsiveness, peer influences, and self-reported future probabilities for each type of SITB. This article describes the SITBI and presents preliminary data on its psychometric properties, including descriptive statistics, interrater reliability, and interinformant agreement for quantitative items; test–retest reliability for the presence and frequency of each type of SITB assessed over a 6-month period; and construct validity via relations between the SITBI and other measures of self-injury-related constructs, including suicidal ideation, suicide attempts, and NSSI.

Method

Participants

A total of 94 (female $n = 73$; male $n = 21$) adolescents and young adults (age in years: $M = 17.1$, $SD = 1.9$, range 12–19) were recruited via announcements posted in local psychiatric clinics and newspapers and on community bulletin boards and Internet message boards. In order to ensure that we obtained a sufficiently large number of currently self-injurious adolescents, one of the announcements specifically requested participants with a recent history of SITB. Inclusion criteria were age 12–19 years and provision of written informed consent to participate in the research, with parental consent also required for those less than 18 years old. The study included adolescents with no history of any form of self-injury, to serve as a comparison group, and potential participants were excluded only if they demonstrated an impaired ability to comprehend and effectively participate in the study (because of factors such as an inability to speak or write English fluently or the presence of gross cognitive impairment due to psychosis, mental retardation, dementia, intoxication, or the like). No one who responded to the advertisements and met inclusion criteria was excluded from the study because of these factors. In all cases in which the participant was less than 18 years old ($n = 47$), a parent or guardian accompanied the participant to the laboratory and also provided data for this study. These were biological mothers (76.5%), biological fathers (7.4%), other biological relatives (5.9%), adoptive/foster mothers (2.9%), and other nonbiologically related guardians (7.4%). More detailed participant demographic information is presented in Table 1.

Assessment

SITBI. Participants were administered the SITBI, a structured interview with 169 items in five modules that assesses the presence, frequency, and characteristics of five types of SITB: (a) suicidal ideation ("Have you ever had thoughts of killing yourself?"), (b) suicide plans ("Have you ever actually made a plan to kill yourself?"), (c) suicide gestures ("Have you ever done something to lead others to believe you wanted to kill yourself when you really had no intention of doing so?"), (d) suicide attempts ("Have you ever made an actual attempt to kill yourself in which you had at least some intent to die?"), and (e) nonsuicidal self-injury ("Have you ever done something to purposely hurt yourself without intending to die?").[2] Items were created and worded so as to be consistent with the commonly accepted definitions of each type of SITB, with a special emphasis on clarity and specificity in assessing each self-injury-related construct (O'Carroll et al., 1996). For instance, we used the item "Have you ever had thoughts of killing yourself?" rather than "Have you ever had thoughts of suicide?" because the latter may be interpreted more loosely, to include mere consideration of the concept or philosophy of suicide, without necessarily implying a contemplation of engaging in the act. When assessing suicide attempts, we included the phrase "in which you had some intent to die" because of prior evidence that many individuals respond affirmatively to questions about having made a "suicide attempt" even when they lacked intent to die,

[1] Our definition of *suicide gesture* falls under the classification of *instrumental suicide-related behaviors* used by O'Carroll and colleagues (1996). However, we use the former term in order to distinguish this behavior from nonsuicidal self-injury, which would also fall under O'Carroll and colleagues' classification of instrumental suicide-related behaviors.

[2] A copy of the SITBI is available from the first author.

Table 1
Demographic and Diagnostic Characteristics of Adolescent Participants

Variable	%	*M*	*SD*
Age in years		17.1	1.9
Sex (% female)	77.7		
Race/ethnicity			
European American	73.4		
African American	3.2		
Hispanic	6.4		
Asian	5.3		
Biracial	10.6		
Other	1.1		
Annual household income			
$0–$20,000	10.0		
$21,000–$40,000	18.8		
$41,000–$60,000	17.5		
$61,000–$80,000	16.3		
$81,000–$100,000	15.0		
>$100,000	22.5		
DSM–IV Diagnosis from K–SADS			
Any mood disorder[a]	33.3		
Any anxiety disorder[b]	46.8		
Any impulse-control disorder[c]	11.7		
Any eating disorder[d]	6.4		
Any substance use disorder[e]	13.8		
Any *DSM–IV* disorder	61.7		
Number of *DSM–IV* disorders		1.6	1.8

Note. K–SADS = Schedule for Affective Disorders and Schizophrenia for School Aged Children.
[a] Major depressive and bipolar disorders. [b] Panic disorder, separation anxiety, phobias, generalized anxiety, and obsessive-compulsive disorder. [c] Oppositional defiant, conduct, and attention-deficit/hyperactivity disorders. [d] Bulimia and anorexia nervosa. [e] Alcohol and drug abuse and/or dependence.

thereby obscuring important differences related to the presence or absence of intent to die in self-injurious behavior (Nock & Kessler, 2006).

The SITBI is comprised of five modules that correspond to the five types of SITB. Each module begins with a screening question that asks about the lifetime presence of that thought or behavior. If the initial screening question is endorsed, then the module is included in the interview. If the initial screening item is denied, then the questions from that module are skipped. For example, if a respondent denies ever having suicidal ideation, this participant is not asked additional questions about various aspects of suicidal ideation, and the interviewer proceeds to the screening question for the next module. However, if the suicidal ideation screening question is endorsed, the interviewer administers the entire module corresponding to suicidal ideation. Beyond lifetime presence, the SITBI assesses the frequency of each type of thought or behavior in the respondent's lifetime, past year, and past month, as well as the age of onset of each thought or behavior endorsed. The SITBI also assesses the severity of each thought or behavior endorsed, on average and at the worst point, as well as providing an open-ended question about the methods of self-injury used. For instance, assessment of the severity of suicide ideation is assessed using questions such as "At the worst point how intense were your thoughts of killing yourself?" rated on a 0 ("*low/little*") to 4 ("*very much/severe*") scale. For self-injurious behaviors, participants indicate whether they received medical treatment as a result of the self-injury.

Next, the SITBI assesses the self-reported function of each type of SITB via an open-ended question, followed by four questions about the extent (on the 0 to 4 scale) to which the respondent has engaged in each thought or behavior for the purpose of emotion regulation (i.e., to escape aversive feelings or generate feelings) or communication with others (i.e., to get attention from others or escape from others). Our inclusion of these questions was based on prior research demonstrating that these are among the most common functions of such behaviors (Boergers, Spirito, & Donaldson, 1998; Hawton, Cole, O'Grady, & Osborn, 1982; Nock & Prinstein, 2004, 2005). The SITBI also assesses the extent to which respondents believe different factors may have contributed to their behavior (on the 0 to 4 scale), including "family," "friends," "relationships," "peers," "work/school," and "mental state."

Respondents are also asked about other characteristics of their SITB, including the extent to which they have experienced physical pain (measured on the 0 to 4 scale and administered only for the modules inquiring about self-injurious *behavior*); the percentage of SITB episodes during which they have used alcohol or drugs; and the length of time they typically spend thinking about the behavior before engaging in it (i.e., impulsiveness). The SITBI also assesses how many of the respondents' peers have engaged in each thought or behavior, as well as to what extent (on the 0 to 4 scale) their friends' experiences have influenced their engagement in the SITB. For each of these peer-related items, the respondent is asked to provide separate responses for the period of time before the onset of that SITB and for the period since the first time the respondent engaged in that SITB. The purpose of this distinction is to allow for the examination of peer influence on the initiation versus maintenance of each type of SITB. Finally, the SITBI assesses respondents' self-reported likelihood (on the 0 to 4 scale) that they will engage in each SITB at some time in the future. Administration of the SITBI takes approximately 3–15 minutes, depending on the number of modules administered.

As described, most of the SITBI items ask for quantitative information from the respondent, but the interview also obtains some qualitative and open-ended responses. The wording of the SITBI is appropriate for both adolescents and adults. It can also be used to interview the parent of an adolescent, in cases where it is desirable to have multiple informants. In instances in which parents or guardians are available, the parent or guardian is interviewed separately by the same interviewer and answers each question about the child. The optional parent interview takes an additional 3–15 minutes, depending on the number of modules administered. Although prior work has shown poor parent–child agreement on the presence of suicide-related constructs (Prinstein & Nock, 2003; Prinstein, Nock, Spirito, & Grapentine, 2001), we included this option in the belief that parent reporting may help clinicians identify additional cases and may help researchers better understand the reasons for nonagreement (De Los Reyes & Kazdin, 2005).

The SITBI is intended to be administered by master's- and doctoral-level clinicians and researchers, as well as closely supervised bachelor's-level research assistants. The SITBI is intended to be administered exactly as worded. However, the interviewers should be knowledgeable about categories of SITB and may use follow-up questioning to clarify responses. It is particularly im-

portant for interviewers to be clear on the definitions of suicidal plans and gestures, as well as the boundaries between ideation, plans, and gestures, as these categories are not as familiar to respondents as suicide attempts or acts of self-injury. Follow-up probing should be done sparingly, and many interviews will not require any probing.

Other self-injury related measures. All respondents also participated in two additional interviews and completed one rating scale that assessed some of the SITB-related constructs measured by the SITBI. First, we administered the Schedule for Affective Disorders and Schizophrenia for School Aged Children—Present and Lifetime Version (K–SADS–PL; Kaufman, Birmaher, Brent, Rao, & Ryan, 1997). The K–SADS–PL is a widely used semi-structured diagnostic interview designed to assess current and past episodes of 33 different mental disorders according to the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM–IV*; American Psychiatric Association, 1994). We were particularly interested in examining the correspondence between responses on the SITBI and the three items in the major depressive disorder module focused on suicidal ideation, suicide attempts, and NSSI. In the current study, the presence of each SITB was denoted by a score of 2 or 3 using the K–SADS–PL standard 0 to 3 scoring procedures. These items are commonly used as measures of self-injury-related constructs in research studies (e.g., Nock & Kazdin, 2002). The K–SADS–PL was administered by the first author and four trained and supervised research assistants. Independent rating of the K–SADS–PL was completed for 20 videotaped interviews and demonstrated good interrater reliability across all diagnoses (average $\kappa = .93$) and for the three self-injury items mentioned above ($\kappa = .90$, .83, and .71, respectively). Summary diagnostic characteristics of the current sample are presented in Table 1.

Second, we administered the Functional Assessment of Self-Mutilation (FASM; Lloyd, Kelley, & Hope, 1997). The FASM is a structured interview that evaluates the characteristics of NSSI, such as the frequency, functions, and age of onset of these behaviors. Previous research has described the factor structure and has supported the reliability and validity of the FASM among adolescent psychiatric inpatients (Nock & Prinstein, 2004, 2005), as well as among adolescents recruited from the community (Lloyd et al., 1997). Confirmatory factor analysis has supported a 4-function model of NSSI assessed by the FASM's 21 function items: automatic negative reinforcement (2 items, $\alpha = .62$; e.g., "To stop bad feelings"); automatic positive reinforcement (3 items, $\alpha = .69$; e.g., "To feel something, even if it was pain"); social negative reinforcement (4 items, $\alpha = .76$; e.g., "To avoid doing something unpleasant you don't want to do"); and social positive reinforcement (12 items, $\alpha = .85$; e.g., "To get other people to act differently or change"; Nock & Prinstein, 2004). In the current study, we examined the correspondence between the 4 function items included in the SITBI and the 4 corresponding subscales of the FASM, measured by 21 items. The demonstration of strong correspondence between the 4 SITBI items and the 21 FASM items would support the construct validity of the SITBI and would offer a more efficient method of assessing the behavioral functions of NSSI.

Third, all respondents completed the Beck Scale for Suicide Ideation (BSI; A. T. Beck, Steer, & Ranieri, 1988), a self-report instrument consisting of 21 items, rated on a 0 to 2 scale, that assesses the presence and severity of current suicidal ideation. The BSI is a widely used measure of suicidal ideation that has been shown to have strong internal consistency reliability, convergent validity (via significant relations with self-report measures of depressed mood and hopelessness), and divergent validity (via a nonsignificant relation with a self-report measure of anxiety) among clinical samples (A. T. Beck & Steer, 1991). The BSI showed adequate internal consistency reliability in the current sample, with a Cronbach's alpha of .85.

Procedure

Data collection. All data were collected as part of a behavioral laboratory study of SITB, which was approved by the Harvard University Institutional Review Board. Participants completed all of the interviews and rating scales described above during one baseline laboratory visit, for which the participant was paid $100. Participants and their parents were administered consent and debriefing procedures together but completed all portions of the assessment separately. Risk assessment interviews were conducted during the debriefing sessions, and safety planning was conducted as needed, which in some cases involved informing parents of elevated risk of self-injury and/or making referrals to outpatient treatment services. Participants were contacted 6 months after the laboratory visit, at which time the lifetime presence and frequency items from the SITBI were readministered via telephone, in order to evaluate the test–retest reliability of those components of the SITBI. Seventy-six (80.9%) participants provided data in follow-up interviews. Eighteen were not included in follow-up because they either were unable to be located at the time of follow-up ($n = 6$), did not respond to repeated requests to schedule an interview ($n = 9$), or refused to participate at the time of the interview ($n = 3$). Participants in the follow-up interviews did not differ significantly from those not included on age, gender, ethnicity, or clinical severity, or on the presence at the baseline interview of suicidal ideation, suicide plans, suicide gestures, suicide attempts, or NSSI.

Interviewers. Interviews were conducted by the first author, as well as two clinical psychology graduate students and two post-baccalaureate research assistants. Interviewers (one male, four female) were assigned to participants randomly, and, once assigned, an interviewer met with both the adolescent and parent from the same family. Prior to data collection, all interviewers participated in several training sessions on the administration of the SITBI, led by the first author. Training included review of the SITBI items and practice administering the interview. Interviewers received ongoing supervision, and all interviews were audio- and videotaped in their entirety for the purposes of supervision and reliability analyses. A random sample of these tapes was used to test the interrater reliability of the interviews.

Data Analyses

As described, rather than assessing a single construct (e.g., suicidal ideation) using numerous items, the SITBI was designed to efficiently examine a fairly broad range of constructs using a minimal number of items. Therefore, factor analyses and internal consistency reliability analyses were not conducted, as they would not be theoretically or empirically meaningful for this measure. Instead, we first examined the descriptive statistics (*M*, *SD*, or rate

of endorsement) for each of the quantitatively measured SITBI items. Second, we examined the interrater reliability of each quantitatively measured SITBI item. Third, we examined the test–retest reliability of the presence and frequency items from baseline to 6-month follow-up. We focused specifically on assessing the stability of the presence and frequency items because these are the central items of the measure and because we wanted to limit the length of follow-up interviews, given that participants were not provided additional monetary compensation for these interviews. Fourth, we examined interinformant agreement between adolescents and parents on the presence of each type of SITB. Fifth and finally, we examined the construct validity of the SITBI by testing the correspondence between SITBI and related measures on the presence and functions of several different types of SITB. Although we provide initial data on the reliability of all of the quantitative SITBI items and the validity of items assessing the presence and frequency of SITB, we did not report on the validity of some other constructs assessed by the SITBI, such as the experience of pain, use of alcohol and drugs, and influences on the performance of SITB. Also, given the breadth of the SITBI items, we were not able to include data on each of the items for each outcome in this initial article, but instead we focused on the variables likely to be of most interest and use to researchers and clinicians.

Results

Descriptive Statistics for the SITBI

The frequencies, means, and standard deviations of participants' responses to the SITBI are presented in Table 2. There were no consistent differences as a function of sex or age (i.e., younger vs. older than 18 years). As shown in Table 2, although this was a community/outpatient sample, there was a fairly high rate of SITB, and each form of SITB had an average age of onset of 13–14 years. The most strongly endorsed function for each form of SITB was automatic negative reinforcement, followed by automatic positive reinforcement, which highlights the importance of these functions and suggests that different forms of SITB may serve similar functions. The only deviation from this pattern was for suicide gestures, which adolescents reported were performed for the purposes of social reinforcement, consistent with the definition of

Table 2
Frequencies, Means, and Standard Deviations of Responses on the Self-Injurious Thoughts and Behaviors (SITBs) Interview

	Suicidal ideation		Suicide plan		Suicide gesture		Suicide attempt		NSSI	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%
Presence										
Lifetime	66	70.2	35	37.2	21	22.3	27	28.7	64	68.1
Past year	52	55.3	23	24.5	12	12.8	14	14.9	56	59.6
Past month	32	44.0	12	12.8	2	2.1	2	2.1	45	47.9
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Frequency										
Lifetime	82.00	158.95	8.83	28.44	4.76	31.50	1.17	3.31	709.29	3911.06
Past year	31.91	89.57	4.20	21.77	3.44	30.94	0.43	1.59	42.63	111.66
Past month	9.62	55.42	0.50	1.96	0.12	1.03	0.02	0.15	14.03	72.52
Average age of onset	13.32	2.56	13.86	2.70	13.35	3.76	14.11	2.14	13.52	2.68
Severity (worst point; 0–4 scale)	3.32	0.81	3.55	0.75	—	—	2.82	1.29	2.27	1.71
Severity (average; 0–4 scale)	2.17	0.88	2.75	1.00	—	—	2.41	1.14	1.71	0.71
Reported function (0–4 scale)										
Automatic negative reinforcement	2.39	1.23	2.59	1.36	1.26	1.24	2.96	1.01	3.06	1.02
Automatic positive reinforcement	1.34	1.40	1.28	1.30	1.00	1.15	1.44	1.28	2.08	1.48
Social negative reinforcement	—	—	—	—	1.42	1.50	1.23	1.42	0.45	0.89
Social positive reinforcement	—	—	—	—	2.35	1.69	1.22	1.45	0.92	1.24
Precipitants (0–4 scale)										
Family	2.25	1.30	2.18	1.42	2.19	1.38	2.04	1.30	1.95	1.37
Friends	1.79	1.39	1.86	1.30	1.88	0.96	1.74	1.42	1.50	1.25
Relationship	1.75	1.49	1.79	1.52	1.50	1.55	1.65	1.43	1.91	1.50
Peers	1.59	1.33	1.71	1.38	1.63	1.31	1.65	1.43	1.34	1.18
Work/school	1.59	1.27	1.68	1.25	1.94	1.34	1.17	1.15	1.43	1.20
Mental state	3.18	0.97	3.21	1.10	3.25	0.93	3.17	1.11	3.38	0.91
Characteristics of SITBs										
Physical pain experienced (0–4 scale)	—	—	—	—	—	—	1.89	1.34	1.88	1.05
Alcohol/drug use (% of time)	5.95	14.63	19.44	22.14	5.50	22.35	8.33	23.00	5.75	15.2
No. peers with behavior before 1st time	0.96	1.34	0.59	1.10	0.86	1.46	0.67	1.14	1.09	1.47
No. peers with behavior after 1st time	3.88	4.45	2.23	3.36	3.93	7.82	3.67	4.44	5.53	6.68
Peer influence before 1st time (0–4 scale)	0.73	1.15	0.22	0.74	0.58	1.07	0.26	0.81	0.83	1.23
Peer influence after 1st time (0–4 scale)	0.57	0.89	0.42	0.99	0.50	0.92	0.52	0.99	0.62	1.00
Future likelihood of this behavior (0–4 scale)	2.09	1.30	1.39	1.34	0.85	1.26	1.08	1.14	2.37	1.45

Note. NSSI = nonsuicidal self-injury.

such behaviors. Across each form of SITB, adolescents consistently reported the top three precipitants as their mental state at the time, family factors, and problems with friends.

In reporting on the characteristics of SITB, adolescents reported experiencing only a moderate amount of physical pain during suicide attempts and NSSI. In addition, they reported using alcohol during only a small percentage of the time in which they experienced suicidal ideation, suicide gestures, suicide attempts, or NSSI (M = 5.50% to 8.33%), but more often during times when they made suicide plans (M = 19.44%), suggesting that alcohol use may play a larger role in making suicide plans than in other forms of SITB. Adolescents in this study reported not having many friends who had engaged in each form of SITB before they themselves had ever done so (M = 0.59–1.09), but they reported having a much higher number of self-injurious friends after they had done so (M = 2.23–5.53). However, participants reported that the behavior of their friends did not have much influence on their own SITB at either time. The likelihood reported by adolescents that they would engage in each form of SITB in the future was lowest for suicide gestures, followed by suicide attempt, suicide plan, suicide ideation, and NSSI. This order is consistent with estimates of the prevalence of each form of SITB, providing some support for the validity of these responses.

Interrater Reliability

The demonstration that different interviewers score responses on the SITBI in a reliable fashion is especially important in the assessment of SITB, given the inconsistencies in the definitions used by different researchers and clinicians. We tested the interrater reliability of the SITBI by using the kappa statistic to examine the extent to which two independent raters of randomly selected interviews (n = 21; 22.3%) agreed on the presence versus absence of each outcome, and we examined the ratings of frequency, severity, and other continuous items using correlation coefficients. *Kappa* (κ) is a chance corrected statistic varying from -1 to $+1$, with zero representing chance agreement between raters. Values greater than .75 represent excellent agreement beyond chance; values from .40 to .75 represent fair to good agreement; and values below .40 represent poor agreement beyond chance (Fleiss, Levin, & Paik, 2003).

Our examination revealed perfect agreement between raters for the lifetime presence of suicidal ideation (Item 1), suicide gesture (Item 58), suicide attempt (Item 84), and NSSI (Item 143; all κs = 1.0), and excellent agreement for suicide plan (Item 30; κ = .90), supporting the interrater reliability of the SITBI classifications. Interrater reliability was also perfect (κ = 1.0) for the presence of each outcome in the past year (Items 5, 34, 62, 89, and 147) and past month (Items 6, 35, 63, 90, and 148), as well as for all of the other items assessed quantitatively in the SITBI (all κs and rs were 1.0). The responses to items assessed qualitatively (e.g., self-injury methods used, open-ended reasons given for SITB) were recorded in each case, but we did not assess the reliability of these responses here, given the nature of the data.

Test–Retest Reliability

We examined the test–retest reliability of the SITBI by evaluating the correspondence between the reported lifetime presence (κ) and frequency (one-way random effects intraclass correlation coefficient; ICC; Shrout & Fleiss, 1979) of each type of SITB reported at the baseline interview and the presence and frequency reported during the 6-month follow-up (i.e., the new lifetime frequency minus the frequency of that outcome during the 6 months since the last interview). Test–retest reliability for the presence versus absence of each lifetime outcome reported at baseline and 6-month follow-up (same items noted above) was strong for suicidal ideation (κ = .70), suicide plan (κ = .71), suicide attempt (κ = .80), and NSSI (κ = 1.0). However, agreement was poor for suicide gesture (κ = .25). Closer examination revealed that the reason for this poor agreement was a lower rate of endorsement of lifetime suicide gestures at the follow-up interview. More specifically, of 18 participants who reported a lifetime history of suicide gesture during the baseline interview and were available for the 6-month follow-up interview, only 5 reported a lifetime history of suicide gesture in the follow-up interview. The SITBI items also demonstrated strong test–retest reliability for the lifetime frequency of suicidal ideation (ICC = .74, $p < .001$), suicide attempt (ICC = .50, $p < .001$), and NSSI (ICC = .71, $p < .001$); slightly weaker reliability for suicide plans (ICC = .23, $p < .01$); and poor reliability for suicide gestures (ICC = .01, *ns*).

Interinformant Agreement

We also examined the agreement between adolescents and their parents on the presence versus absence of each outcome at the baseline interview. Agreement was strong for suicidal ideation (κ = .75), suicide attempt (κ = .67), and NSSI (κ = .91). Parent–adolescent agreement was fair for the presence of suicide plan (κ = .44) and poor for suicide gesture (κ = .21). Lack of agreement was due specifically to parents underreporting adolescent-reported SITB, with the exception that lack of agreement for suicide gesture was due both to parents reporting this behavior when adolescents did not and adolescents reporting it when parents did not.

Construct Validity

In addition to establishing the SITBI as a reliable assessment instrument, it is important to demonstrate that it is a valid measure of SITB-related constructs (Cronbach & Meehl, 1955). We examined the construct validity of the SITBI by testing the correspondence of responses to SITBI items assessing the presence and frequency of suicidal ideation, suicide attempts, and NSSI to responses elicited by similar items from the K–SADS–PL, SSI, and FASM. We also tested the correspondence between the SITBI and FASM items assessing the behavioral functions of NSSI.

We first examined the agreement between the SITBI and K–SADS–PL on the presence versus absence of suicidal ideation, suicide attempts, and NSSI. Both the SITBI and K–SADS–PL use an interview format; however, the K–SADS–PL items ask about the presence of behaviors during the past 6 months, whereas the SITBI asks about the past year and past month. We chose to examine the SITBI items inquiring about the past year in order to have a longer period of overlapping time between measures (i.e., 6 months rather than 1 month). Despite the different time frames used, there was good agreement between these measures on the presence of suicide attempt (κ = .65) and NSSI (κ = .74), but

slightly lower agreement on the presence of suicidal ideation (κ = .48).

Next, we examined the correspondence between the SITBI and the BSI on the presence of suicidal ideation. The BSI inquires about suicidal ideation during the past week, whereas the closest time frame from the SITBI is the item inquiring about suicidal ideation during the past month. Nevertheless, interpreting endorsement of a 1 or 2 on the BSI items assessing the presence of active (Item 4) and passive (Item 5) suicidal ideation as denoting the presence of suicidal ideation—as is recommended by the BSI Manual (A. T. Beck & Steer, 1991)—agreement on the presence of suicidal ideation between the SITBI and BSI was good (κ = .59).

Finally, we examined the agreement between the SITBI and the FASM on the presence, frequency, and functions of NSSI. There was perfect agreement between these two measures on the presence of NSSI (κ = 1.0), and excellent agreement on the lifetime frequency of NSSI (r = .99). We assessed the correspondence between the two measures on the assessment of the behavioral functions of NSSI by examining the correlations of the 4 functional items from the SITBI and the corresponding 4 functional subscales from the FASM (composed of 21 items). There were large and statistically significant correlations between the SITBI and the FASM in the assessment of automatic positive reinforcement (r = .71), automatic negative reinforcement (r = .72), social positive reinforcement (r = .73), and social negative reinforcement (r = .64) functions of NSSI (all ps $<$ .001),[3] supporting the construct validity of the SITBI assessment of the functions of NSSI.

Discussion

This study reports on the development of the SITBI, a new comprehensive measure of suicidal ideation, suicide plans, suicide gestures, suicide attempts, and NSSI. This preliminary examination provided descriptive information about the SITBI as well as initial evidence for the interrater reliability, test–retest reliability, interinformant agreement, and construct validity of this new measure. Although many measures are available to assess specific self-injury-related constructs (e.g., continuous measure of suicidal ideation, dichotomous measures of suicide attempt), few measure the important distinctions among different types of self-injurious thoughts and behaviors (Joiner et al., 2005; Nock et al., in press). In addition, many existing measures provide researchers and clinicians with scores that are somewhat arbitrary and thus not readily interpretable in many settings (Blanton & Jaccard, 2006; Kazdin, 2006). The SITBI addresses many limitations of prior work by measuring a range of self-injury-related constructs and by doing so using metrics (e.g., presence and frequency of SITB) that are intended to be useful to both researchers and clinicians. Several aspects of the initial findings warrant additional comment.

The SITBI assesses a broader range of SITB than any measures previously reported in the literature (Brown, 2000; Goldston, 2000; Nock et al., in press; Range & Knott, 1997; Reynolds, 1990), and it does so in a manner that is both comprehensive and consistent across each behavior. This is significant because previous measures have limited the ability of researchers to compare the characteristics of different forms of SITB, insofar as they used vague or inconsistent definitions, targeted one type of SITB (e.g., suicidal ideation), or assessed different aspects of each SITB. The development of a measure that assesses the presence, frequency, and characteristics of suicidal ideation, suicide plans, suicide gestures, suicide attempts, and NSSI, using consistent methods and items, allows researchers to make valid comparisons of different SITB both within and across studies.

The excellent interrater reliability observed for each of the SITB examined suggests that use of the SITBI yields strong agreement between raters on the presence of each of these behaviors. This level of agreement is likely the result of using clear and specific definitions for each type of SITB examined. Although strong interrater reliability is a necessity for any useful assessment method, the current data are especially meaningful given inconsistencies in the definitions of SITB that have characterized this area to date (Linehan, 1997; O'Carroll et al., 1996).

Our analyses also revealed strong test–retest reliability for the presence and frequency of each of the SITB assessed, with the exception of suicide gestures. The strong test–retest reliability coefficients for most of these variables indicate that the SITBI yields data that is consistent from one administration to the next. The reliability of the SITBI compares favorably with that of other SITB-related assessment instruments (Brown, 2000; Goldston, 2000). The relatively poor reliability of the suicide gesture item resulted primarily from fewer adolescents endorsing the suicide gesture item during the follow-up than in the baseline interview. The reason that some adolescents changed their reports for this behavior but not the others is not completely clear. One possible explanation is that the consequences of a suicidal gesture are typically less severe than the consequences of the other categories of suicidal ideation or behaviors and are therefore more likely to be forgotten over a period of 6 months. Another possibility is that the poor reliability observed resulted from a lack of clarity in the wording of this item of the SITBI, or with the concept of a suicide gesture more generally. In addition, a suicide gesture may be construed as being manipulative of another person's feelings and behaviors, and can imply a lack of authenticity that may be socially undesirable for adolescents to admit repeatedly. Nevertheless, although suicide gestures are not intended to result in death, prior findings indicate that they occur in approximately 2% of the general population and are associated with elevated levels of psychopathology (Nock & Kessler, 2006), highlighting their importance as a focus of future study. The apparent prevalence of suicide gestures, coupled with the inconsistency with which they were reported over time in this study, suggest that more attention needs to be given to the measurement and study of these behaviors, and significant improvements made.

Our analyses revealed strong parent–adolescent agreement on the presence of suicidal ideation, suicide plans, suicide attempts, and NSSI (κs from .44 to .91) Prior work has shown poor parent–adolescent agreement on such constructs using structured diagnostic interviews and rating scales for suicidal ideation (κ = .21) and suicide attempts (κ = .23; Prinstein et al., 2001). The stronger agreement between parents and adolescents obtained in the current

[3] In the current study, we classified Item 2 of the FASM ("To relieve feeling numb or empty") on the Automatic Positive Reinforcement (APR) subscale rather than the Automatic Negative Reinforcement subscale given its theoretical similarity to that behavioral function. There was also empirical support for this adjustment, as it increased the internal consistency reliability of the APR function from α = .35 to α = .61.

study was most likely due to the use of common items and methods across informants in the SITBI. This elevated rate of agreement highlights the benefit of using common measurement methods and items across informants. It is possible, though, that parent–adolescent agreement was inflated slightly because this was a laboratory based study in which parents were required to know that we were examining SITB, and parents in such studies may know more about their children's SITB than parents in epidemiologic or treatment samples. This remains an important question for future work using the SITBI and other measures of SITB.

We also demonstrated that responses on the SITBI corresponded closely to responses on other measures of SITB, providing support for the construct validity of this new measure. This correspondence was obtained despite the use of different time frames across measures, as well as different measurement scales across instruments. For instance, the SITBI inquires about the presence versus absence of SITB in one's lifetime, over the past year, and over the past month, whereas the K–SADS–PL assesses such behaviors over the past 6 months using a *trichotomous scale* (i.e., absent, subthreshold, threshold). Although the use of a broader measurement scale in the K–SADS–PL and BSI undoubtedly captures greater variability in responses, the scores obtained can be difficult to interpret and use in clinical settings, and their use can lead to greater disagreement between raters and clinicians. Ultimately, the researcher or clinician's choice of assessment instrument should be guided by his or her goals in a given situation. For instance, where the goal of assessment is to measure change over the course of treatment, the K–SADS–PL items and BSI are more likely to be sensitive to such change; whereas, if the goal of assessment is to determine the presence versus absence of SITB, we recommend using an instrument such as the SITBI.

Our finding demonstrating strong correspondence between the 4 items measuring behavioral functions of NSSI and the 4 respective subscales of the FASM suggests that use of these specific items captures much of the variance in the 21-item FASM and that the SITBI may thus provide a more efficient method for assessing the functions of SITB. Indeed, using the 4 items from the SITBI rather than the 21 from the FASM allows one to examine the behavioral functions of suicidal ideation, suicide plans, suicide gestures, suicide attempts, and NSSI in the time it would otherwise take to examine the functions of only one such behavior.

On balance, the results of this study should be viewed in the context of several important limitations. First, this initial study included a relatively small sample of paid, self-referred adolescents and parents, and therefore data on the scores, reliability, and validity of this new measure should be considered preliminary. Also, although the SITBI is designed and worded to be administered to adults as well as adolescents, this sample included only adolescents and their parents. Additional studies are needed to test the generality of these results.

Second, although the SITBI assesses a wide range of SITB and characteristics of these behaviors, it is limited in its scope. For instance, the SITBI does not ask about potentially self-destructive behaviors that are not directly self-injurious (e.g., alcohol and substance use) and does not attempt to assess psychological states likely to be associated with SITB (e.g., depressed mood, hopelessness, desperation), as some other existing measures do. Our intention was to create a measure that focuses specifically on SITB themselves; however, learning how SITB interact with other self-destructive behaviors and psychological states is integral to understanding the phenomena, as well as to providing the most specific treatments. Therefore, it may be useful to develop additional modules of the SITBI to correspond to integrative research or clinical interests. For example, a module on eating disorders could be constructed by adapting the questions from the five original modules on frequency, duration, peer-affiliation, pain, and more to focus on disordered eating rather than suicidal ideation. This method would allow direct comparison of the characteristics of SITB and disordered eating, among many possible topics. Possible avenues of research also include childhood abuse, perfectionistic thinking, compulsive behavior, risk-taking behavior, disturbed sleeping patterns, and emotional overinvolvement of parents.

Finally, it is important to note that structured interviews such as the SITBI represent only one of multiple assessment methods that should be used in the evidence-based assessment of SITB (Nock et al., in press; Prinstein et al., 2001). The SITBI may be best conceptualized and administered as an initial, broad screening measure that provides basic data about each type of SITB; those interested in obtaining more detailed information about SITB endorsed on the SITBI could follow up by administering more focused, in-depth measures such as the BSI, Suicide Intent Scale (R. W. Beck, Morris, & Beck, 1974), or FASM.

Several important research directions follow from this work. It will be important to administer and examine the reliability and validity of the SITBI among an adult population. In addition, modifications to the version of the SITBI examined here may be needed in order to increase its usefulness for research and clinical purposes. Toward this end we have added questions assessing *thoughts* of NSSI, in order to better understand the relations between thoughts and behaviors for NSSI, as has been so useful in the study of suicidal ideation and suicide attempts for many years. In addition, we have created a 72-item short form for use in situations where less time is available for assessment. This short form retains questions about presence, frequency, severity, duration, and probability of future SITB but excludes items related to the functions of the behavior, the experience of pain, and the influence of peers. This has reduced administration time by almost half. As mentioned above, there is a great need for assessment instruments that measure a wide range of SITB, and that do so consistently and in a way that is useful to researchers and clinicians. We are hopeful that the development and continued evaluation of such instruments will help to advance and improve the understanding, assessment, and treatment of SITB.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Beck, A. T., & Steer, R. A. (1991). *Manual for the Beck Scale for Suicide Ideation.* San Antonio, TX: The Psychological Corporation.

Beck, A. T., Steer, R. A., & Ranieri, W. F. (1988). Scale for Suicide Ideation: Psychometric properties of a self-report version. *Journal of Clinical Psychology, 44,* 499–505.

Beck, R. W., Morris, J. B., & Beck, A. T. (1974). Cross-validation of the Suicidal Intent Scale. *Psychological Reports, 34,* 445–446.

Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *American Psychologist, 61,* 27–41.

Boergers, J., Spirito, A., & Donaldson, D. (1998). Reasons for adolescent

suicide attempts: Associations with psychological functioning. *Journal of the American Academy of Child and Adolescent Psychiatry, 37,* 1287–1293.

Brown, G. K. (2000). *A review of suicide assessment measures for intervention research with adults and older adults* (Technical report submitted to NIMH under Contract No. 263-MH914950). Bethesda, MD: National Institute of Mental Health.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302.

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin, 131,* 483–509.

Fleiss, J. L., Levin, B., & Paik, M. C. (2003). *Statistical methods for rates and proportions* (3rd ed.). Hoboken, NJ: Wiley.

Goldston, D. B. (2000). *Assessment of suicidal behaviors and risk among children and adolescents* (Technical report submitted to NIMH under Contract No. 263-MD-909995). Bethesda, MD: National Institute of Mental Health.

Hawton, K., Cole, D., O'Grady, J., & Osborn, M. (1982). Motivational aspects of deliberate self-poisoning in adolescents. *British Journal of Psychiatry, 141,* 286–291.

Hawton, K., & van Heeringen, K. (Eds.). (2000). *The international handbook of suicide and attempted suicide.* Chichester, West Sussex, England: Wiley.

Hirschfeld, R. M., & Blumenthal, S. J. (1986). Personality, life events, and other psychosocial factors in adolescent depression and suicide. In G. L. Klerman (Ed.), *Suicide and depression among adolescents and young adults* (pp. 215–253). Washington, DC: American Psychiatric Press.

Jacobs, D. (Ed.). (1999). *The Harvard Medical School guide to suicide assessment and intervention.* San Francisco, CA: Jossey-Bass.

Joiner, T. E., Jr., Walker, R. L., Pettit, J. W., Perez, M., & Cukrowicz, K. C. (2005). Evidence-based assessment of depression in adults. *Psychological Assessment, 17,* 267–277.

Kaufman, J., Birmaher, B., Brent, D. A., Rao, U., & Ryan, N. D. (1997). Schedule for Affective Disorders and Schizophrenia for School Age Children, Present and Lifetime Version (K–SADS–PL): Initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 980–988.

Kazdin, A. E. (2006). Arbitrary metrics: Implications for identifying evidence-based treatments. *American Psychologist, 61,* 42–49.

Kessler, R. C., Berglund, P., Borges, G., Nock, M. K., & Wang, P. S. (2005). Trends in suicide ideation, plans, gestures, and attempts in the United States, 1990–1992 to 2001–2003. *Journal of the American Medical Association, 293,* 2487–2495.

Kessler, R. C., Borges, G., & Walters, E. E. (1999). Prevalence of and risk factors for lifetime suicide attempts in the National Comorbidity Survey. *Archives of General Psychiatry, 56,* 617–626.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., et al. (1994). Lifetime and 12-month prevalence of *DSM–III–R* psychiatric disorders in the United States: Results from the National Comorbidity Survey. *Archives of General Psychiatry, 51,* 8–19.

Linehan, M. M. (1997). Behavioral treatments of suicidal behaviors: Definitional obfuscation and treatment outcomes. *Annals of the New York Academy of Sciences, 836,* 302–328.

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

Maris, R. W., Berman, A. L., & Silverman, M. M. (2000). *Comprehensive textbook of suicidology.* New York: Guilford Press.

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

Nock, M. K., & Kazdin, A. E. (2002). Examination of affective, cognitive, and behavioral factors and suicide-related outcomes in children and young adolescents. *Journal of Clinical Child and Adolescent Psychology, 31,* 48–58.

Nock, M. K., & Kessler, R. C. (2006). Prevalence of and risk factors for suicide attempts versus suicide gestures: Analysis of the National Comorbidity Survey. *Journal of Abnormal Psychology, 115,* 616–623.

Nock, M. K., & Prinstein, M. J. (2004). A functional approach to the assessment of self-mutilative behavior. *Journal of Consulting and Clinical Psychology, 72,* 885–890.

Nock, M. K., & Prinstein, M. J. (2005). Clinical features and behavioral functions of adolescent self-mutilation. *Journal of Abnormal Psychology, 114,* 140–146.

Nock, M. K., Wedig, M. M., Janis, I. B., & Deliberto, T. L. (in press). Self-injurious thoughts and behaviors. In J. Hunsely & E. Mash (Eds.), *A guide to assessments that work.* New York: Oxford University Press.

O'Carroll, P. W., Berman, A. L., Maris, R., & Moscicki, E. (1996). Beyond the tower of Babel: A nomenclature for suicidology. *Suicide & Life-Threatening Behavior, 26,* 237–252.

Prinstein, M. J., & Nock, M. K. (2003). Parents under-report children's suicide ideation and attempts. *Evidence Based Mental Health, 6,* 12.

Prinstein, M. J., Nock, M. K., Spirito, A., & Grapentine, W. L. (2001). Multimethod assessment of suicidality in adolescent psychiatric inpatients: Preliminary results. *Journal of the American Academy of Child and Adolescent Psychiatry, 40,* 1053–1061.

Range, L. M., & Knott, E. C. (1997). Twenty suicide assessment instruments: Evaluation and recommendations. *Death Studies, 21,* 25–58.

Reynolds, W. M. (1990). Development of a semistructured clinical interview for suicidal behaviors in adolescents. *Psychological Assessment,* 2, 382–390.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428.

Received February 13, 2006
Revision received March 23, 2007
Accepted April 9, 2007 ■