2016-08

Who Are We Missing? Nondisclosure In Online Suicide Risk Screening Questionnaires

Matthew C. Podlogar, Megan L. Rogers, Bruno Chiurliza, Melanie A. Hom, Mirela Tzoneva and Thomas Joiner

The publisher's version of record is available at https://doi.org/10.1037/pas0000242

Who Are We Missing? Non-Disclosure in Online Suicide Risk Screening Questionnaires

Matthew C. Podlogar, B.A. [1]

Megan L. Rogers, B.A. [1]

Bruno Chiurliza, B.S. [1]

Melanie A. Hom, B.A. [1]

Mirela Tzoneva, B.S. [1]

Thomas Joiner, Ph. D. [1]

Author Note

[1] Department of Psychology, Florida State University, Tallahassee, FL.

This research was supported in part by a grant from the Military Suicide Research Consortium, an effort supported by the Office of the Assistant Secretary of Defense for Health Affairs under Award No. (W81XWH-10-2-0181). Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the Military Suicide Research Consortium or the Department of Defense.

Correspondence concerning this article should be addressed to Matthew Podlogar, Department of Psychology, Florida State University, Tallahassee, FL. Contact: podlogar@psy.fsu.edu

Abstract

The use of self-report surveys for suicide risk screening is a key first step in identifying currently suicidal individuals and connecting them with appropriate follow-up assessment and care. Online methods for suicide risk screening are becoming more common, yet they present a number of complexities compared to traditional methods. This study aimed to assess whether forcing item responses may unintentionally hide or misrepresent otherwise useful missing suicide risk data. We investigated in a secondary analyses of three independent samples of undergraduates (*n*s = 1306, 694, 172) whether participants who chose not to respond specifically to current suicide risk screening items (i.e., Non-Disclosers) scored significantly different from other risk response groups (i.e., Deniers, Lower-Risk Endorses, and Higher-Risk Endorsers) on auxiliary measures related to suicidality. Multivariate Analysis of Variance (MANOVA) tests for each sample revealed that Non-Disclosers were rare (*n*s = 7, 6, 7) and scored significantly higher than Deniers and similarly to Endorsers on suicide risk related measures. In one sample, Non-Disclosers tended to score higher than all groups on suicide risk related measures. These findings suggest that non-disclosure for suicide risk screening questions is a preferred option for a distinct group of respondents who are likely at elevated suicide risk. Allowing for and flagging Non-Disclosers for follow-up suicide risk assessment may be an ethical and feasible way to enhance the sensitivity of online suicide risk screenings for weary respondents, who if forced, may choose to under-report their suicide risk and misrepresent data.

Keywords: Web-based Survey; Suicidality; Missing Data; Risk Assessment; Multiple Imputation

Who Are We Missing? Non-Disclosure in Online Suicide Risk Screening Questionnaires

With over 800,000 individuals estimated to die by suicide globally each year (World Health Organization, 2014), it is imperative that efforts be made to improve the identification of individuals at elevated risk for suicide. Researchers and clinicians have traditionally relied on paper-and-pencil self-report surveys and/or face-to-face interactions to screen an individual's level of current suicide risk (Bryan & Rudd, 2006; Joiner, Walker, Rudd, & Jobes, 1999). However, the use of technology in the mental health care field, including in suicide risk assessment, has burgeoned in recent years (Kazdin & Blase, 2011), and suicidal individuals who are less likely to seek help from traditional sources are more likely to seek help online (Harris, McLean, & Sheffield, 2009a, 2009b). As of recently, web-based suicide screening methods are being utilized in a wide array of settings, including university counseling centers, guided and self-help web/mobile mental health interventions, and within large public social media websites (see Christensen, Batterham, & O'Dea, 2014). Additionally, given that suicide is the second leading cause of death among college-aged individuals (i.e., age 19 to 24; Centers for Disease Control and Prevention, 2015), and colleges are under immense pressure to identify and help students who are at risk (Scelfo, 2015; Schwartz, 2010), web-based suicide intervention programs are being developed for college students specifically (Haas et al., 2008; King et al., 2015). Yet despite the growing use of web-based suicide risk assessment and intervention, little is known about the novel complexities associated with online suicide risk screening methods and how they should be managed.

Complexities of Web-Based Self-Report Suicide Risk Screening

There are numerous advantages to using web-based self-report surveys as a suicide risk screening method. First, web-based surveys are very efficient. They allow researchers and

clinicians to reach a greater number of individuals in a fraction of the time compared to traditional survey methods. Second, web-based screenings allow for instantaneous calculations of a suicide symptom severity score. By using an established cutoff score as a guide, survey administrators can be immediately and automatically alerted by survey software to take further action if an individual is identified to be at elevated risk for suicide (see Michaels, Chu, Silva, Schulman, & Joiner, 2014 for recommended practices for web-based screening follow-up). Third, problems such as inaccurate data entry, errors in measure completion, and illegible survey responses are significantly reduced or eliminated through the use of computerized measures. Finally, past studies have indicated that computer-based self-report surveys can result in greater disclosure of sensitive personal information, including mental health problems (Tourangeau & Yan, 2007; Turner et al., 1998). It follows then, that web-based self-report screening surveys may increase individuals' likelihood of reporting suicidal thoughts or plans.

Despite the many advantages of using web-based self-report measures for assessing suicide risk, there are still a number of issues that challenge their utility. Some of these problems are inherent to self-report measures of suicidal symptoms themselves. Contrary to the goals of screening, mounting evidence suggests that individuals with higher levels of suicidal ideation are actually *less* likely to self-disclose thoughts of suicide or seek help for suicidal ideation (Apter, Horesh, Gothelf, Graffi, & Lepkifker, 2001; Carlton & Deane, 2000; Deane, Wilson, & Ciarrochi, 2001; Horesh, Zalsman, & Apter, 2004). One reason for this might be that self-report surveys rely on higher levels of insight and assume that individuals will be able to accurately identify and classify their own symptoms. Past and recent events, current mood states, and openness to help-seeking can all bias individuals' self-report responses and affect their willingness to disclose this information (Brener, Billy, & Grady, 2003). Individuals may also

have strongly ambivalent suicidal ideation (i.e., feel both the wish to live and wish to die), and despite generally experiencing thoughts of suicide, may not be able or willing to endorse or deny any of the presented suicide risk item responses at a given moment. Further, none of the available item responses may "fit" with their current self-conception. One study of over 1,000 participants who completed an online survey found that 94% of highly suicidal participants indicated they engaged in these ambivalence debates, and over one-third of those participants reported doing so frequently (Harris, McLean, Sheffield, & Jobes, 2010). An additional limitation of using self-report measures for suicidal risk screening is that individuals may not understand the consequences of reporting suicidal ideation and may fear involuntary hospitalization or other adverse results if they disclose thoughts of suicide (Cigularov, Chen, Thurber, & Stallones, 2008). As a result, they may deny or minimize their suicidal symptoms to protect against unknown consequences, compromising their ability to receive information about other perhaps wanted help, as well compromising the accuracy of the measure's results.

Many of these challenges are compounded when self-report suicide risk screens are completed online. For one, web-based surveys do not allow for two-way questioning during the informed consent process in the same way that in-person assessments do. The online screening format reduces opportunities for clinicians to ask participants follow-up questions, and likewise, reduces opportunities for participants to clarify the meaning of certain questions or terms. Suicide risk assessment terms such as "imminent risk" or "safety" are commonly used within consent forms; however, they may be unclear to a layperson, and they may additionally be interpreted differently by different researchers (e.g., what constitutes "imminent risk"). Web-based surveys also do not afford an opportunity for clinicians or researchers to build rapport with individuals. During in-person assessments, clinicians are able to slowly increase the intensity of

the interview by asking general questions first (e.g., "How have things been going for you recently?") to lead up to questions about suicidal ideation (e.g., "Have you ever thought about killing yourself?"; see Bryan & Rudd, 2006). For web-based surveys, however, it is much more difficult to build trust and create an environment in which the participant feels comfortable disclosing sensitive information. Similarly, without two-way communication it is difficult for the researcher to have confidence in the integrity of the participant's response data, especially with such sensitive and important information as indication of suicide risk.

The Question of Forcing Responses

The complexities with online suicide risk assessment culminate in an important issue, namely, of whether to "force responses," or to alternatively allow for missing data. Computerized surveys can be designed to require that participants provide a response to a question before moving on to the next item (i.e., "forcing responses"). Intuitively, forcing responses may seem necessary to ensure the effectiveness of automated online suicide risk screening. For example, measure cutoff scores are only interpretable if all items are answered. However, considering the aforementioned limitations with online self-report suicide risk screening, forcing responses on suicide risk screening items may actually increase the rate of data misrepresentation and potentially lead to "forcing" false-negatives and consequently underestimating suicide risk. Even more concerning, forcing responses would make the data necessarily complete, and would limit researchers' abilities to retrospectively distinguish if risk underreporting occurred at all through missing data techniques (See Enders, 2010; Rubin, 1976).

On the other hand, allowing for non-disclosure (e.g., allowing participants to skip items, select statements of non-disclosure, or provide free response answers) may be a legal, ethical, and logistical liability for researchers. Allowing for non-disclosure would require establishing

screening alert rules that override empirically derived measure cut-scores. For legal and ethical reasons, researchers may arguably feel the need to follow up on all participants who choose not to disclose risk, which may not be logistically feasible as online studies bring the capability of very rapid and large scale participant screening. In agreement with these concerns, we believe that the most balanced online suicide risk assessment should take such ethical and logistical issues into consideration, in addition to addressing the need for maximizing screening sensitivity and response data accuracy.

The Present Study

This study examined data from three past online surveys where participants were allowed to skip items. We aimed to test in these samples: (a) if those who chose not to respond specifically to suicide risk screening items (i.e., Non-Disclosers) composed a unified group of participants distinct from other risk response groups (i.e., Deniers, Lower-Risk Endorses, and Higher-Risk Endorsers) on measures related to suicidality; and (b) if real-time assessment of Non-Disclosers' suicide risk could be ethical and feasible in web-based studies. Using Rubin's (1976) terminology, we predicted that univariate missing suicide risk data (i.e., non-disclosure) would not be missing completely at random (i.e., not MCAR), but that missingness would instead be associated with non-missing suicide-related auxiliary measure scores. If MCAR was indeed rejected among univariate suicide risk non-disclosers, we further hypothesized: (a) that Non-Disclosers' scores on suicide related measures would be significantly and reliably higher than the scores of Deniers, and not significantly different from Lower- or Higher-Risk Endorsers; and (b) that Non-Disclosers would represent a small group of respondents that could be ethically and feasibly flagged for follow-up assessment if needed.

Method

Participants

Archived screening data from three independent samples, each using web-based surveying, were analyzed: (1) a Fall 2011-Spring 2012 research screening survey of undergraduate of students (n = 1319), (2) a Fall 2012 research screening survey of undergraduate students (n = 707), and (3) a Summer 2014 online study of undergraduate student participants (n = 184). As shown in Table 1, participants in complete samples showed multiple missing data patterns on suicide risk screening scores (i.e., DSI-SS), including missing data in none, some, or all of the auxiliary measures. However, participants missing only DSI-SS data (ns = 7, 6, 7) made up 67% of all participants missing DSI-SS data in general (ns = 11, 11, 8). In line with our aim to study individuals who did not disclose current suicide risk specifically, only these participants with univariate DSI-SS missing data patterns were classified as "Non-Disclosers" and included in further analyses.

Tables 2 through 4 present age, gender, and race distributions, as well as mean suicide risk screening scores, for the remaining selected participants in each sample (ns = 1306, 694, 172). Sample statistics were further divided into suicide screening risk groups according to recommended DSI-SS cut scores (Joiner, Pfaff, & Acres, 2002). Mean suicide risk level was significantly different across all risk groups in each sample, supporting the validity of the chosen cut points.[1] Only 5% of each sample exceeded the second suicide risk group cut score, suggesting relatively lower suicide risk among participants. No significant differences of age, gender, or race were found across the suicide risk groups within each sample, nor across samples. Across samples, participants reported a mean age of 18.47 (SD = .77, Range = 17 to 24), were 69.8% female, 69.9% Caucasian/White, 17.4% Hispanic, 8.6% African-American/Black, 3.4% Asian/Asian-American, and 0.7% Other race.

Procedure

All procedures in the current analyses, as well as in the original data collection, were reviewed and approved by the University Institutional Review Board. Samples' 1 and 2 suicide risk related data were extracted from a broad departmental online survey administered to introductory Psychology classes for the purpose of screening for eligibility to participate in certain research projects. The online survey was voluntary, though completed by the vast majority of eligible students, and was accessible online from campus computers or students' personal computers for a period of 7 to 10 days at the beginning of each semester. Students were compensated with academic research credit for participating. Screening participants were informed that the survey assessed "behaviors, characteristics, attitudes, and beliefs across a variety of domains including personality, habits, mental health, and social interactions," and all participants were provided with contact information for mental health resources if they were to experience distress.[2] Items were not randomized across administrations, and suicide risk related items were clustered together in approximately the middle of the survey.

Sample 3 suicide risk related data were extracted from an online survey study administered to student research participants for the purpose of investigating the relationship between religiosity and suicidal behavior. The survey was accessible online throughout the school semester from campus computers or students' personal computers, and questionnaires were presented in random order to control for potential order effects. Students were compensated with academic research credit for participating. In this study, data were collected with active suicide risk monitoring, and participants provided consent to be individually contacted for safety follow-up if their scores indicated they were currently at risk for suicide.

In all samples, participants were informed during the consent process that they may skip any items or questionnaires they choose.

Measures

Depressive Symptom Inventory-Suicidality Subscale (DSI-SS; Joiner, Pfaff, & Acres, 2002). The DSI-SS consists of four self-report item groups focusing on current frequency and intensity of suicidal thoughts and desires. The four item groups are: suicidal thoughts, suicidal plans, control over suicidal thoughts, and urges to kill one-self. Each item group is scored by a 4-point ordinal set of statements ranging from 0 to 3. Total scores can range from 0 to 12, with higher scores representing increased severity of current suicide risk. The DSI-SS has demonstrated good reliability and validity characteristics among young people age 15-24 years (Joiner, Pfaff, et al., 2002). In Sample 2, a single modified DSI-SS item (i.e., past suicidal plans) was also used as an independent measure of past suicidality; specifically, the question was asked, "in the past…" as opposed to "in the last two weeks…" Cronbach's alpha for the DSI-SS was good ($\alpha = 0.84$).

Suicidal Behaviors Questionnaire-Revised (SBQ-R; Osman et al., 2001). The SBQ-R is a brief 4-item self-report measure of past suicidal behavior, including frequency and severity of past ideation and attempts, and self-prediction of future behavior. Responses for each of the four items are coded on ordinal scales, for a total score range of 3 to 18. Osman et al. (2001) found in a community sample that a total score of seven effectively discriminated those at elevated suicide risk. In Sample 2, only the first item of the SBQ-R was used. In Sample 3, the full scale was used and was shown to have good internal consistency ($\alpha = 0.79$).

Interpersonal Needs Questionnaire (INQ; Van Orden, Cukrowicz, Witte, & Joiner, 2012). The INQ is a 15-item self-report measure designed to assess the degree to which an

individual experiences the psychological states of perceived burdensomeness (INQ-PB; 6 items; e.g., "These days I feel like a burden on the people in my life") and thwarted belongingness (INQ-TB; 9 items; e.g., "These days, I feel disconnected from other people"), as described by the interpersonal theory of suicide (Van Orden et al., 2010). Responses are rated on Likert-type scales with values ranging from 1 ("not at all true for me") to 7 ("very true for me"). Both subscales have been shown to be predictive of suicidal ideation (Van Orden, Witte, Gordon, Bender, & Joiner, 2008). Cronbach's alpha for the INQ-TB subscale was good ($\alpha = 0.88$), and for the INQ-PB subscale was excellent ($\alpha = 0.95$).

Sample 1 Belonging and Burden Proxy Items. In Sample 1, *Thwarted Belonging* was measured by two items, "How often do you feel that people are around you but not with you?" and "How often do you feel that there are people you can turn to?" (reverse-scored). Each item was scored on a Likert-type scale from 1 ("never") to 4 ("always"). Cronbach's alpha was poor for the item pair ($\alpha = 0.47$), though both are valid indicators of the full INQ-TB (Van Orden et al., 2012). *Perceived Burden* was measured by a single item from the INQ-PB, "These days I feel like a burden on the people in my life," and was scored on the same 7-point scale as the INQ.

Acquired Capability for Suicide Scale – Fearlessness About Death (FAD; Ribeiro et al., 2013). The FAD is a 7-item revised measure of the original Acquired Capability for Suicide Scale (Bender, Gordon, Bresin, & Joiner, 2011; Van Orden et al., 2008). The FAD is intended to measure fearlessness about death, which has been shown to be related to pain tolerance, the experience of painful and provocative events, and risk for suicidal behavior. Items include, "The fact that I am going to die does not affect me," and "I am not at all afraid to die." Responses are rated on a Likert scale with values ranging from 0 ("not at all like me") to 4 ("very much like me"), with a total score ranging from 0 to 28. The FAD has demonstrated convergent and

discriminant validity, and is associated with suicidal behavior specifically, as opposed to other indicators of suicide risk such as current depression or suicidal ideation (Ribeiro et al., 2013). Cronbach's alpha for the FAD was good ($\alpha = 0.84$).

Beck Hopelessness Scale (BHS; Beck, Weissman, Lester, & Trexler, 1974). The BHS is a 20-item, true-false measure that assesses an individual's negative expectations for the future (e.g., "I might as well give up because I can't make things better for myself"). Each item is scored as 0 or 1, with a total score ranging from 0 to 20. The BHS has been shown to be a highly sensitive measure for predicting future suicidal behavior (Beck et al., 1985; McMillan et al., 2007). Cronbach's alpha for the BHS was good ($\alpha = 0.85$).

Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996). The BDI-II is a 21-item measure used to assess the presence and severity of depressive symptoms within the past two weeks. Severity is assessed on a 4-point Likert-type scale, with responses ranging from 0 to 3 for each item. Higher scores are indicative of more severe depressive symptoms, a known risk factor for suicidal ideation (Nock, Hwang, Sampson, & Kessler, 2010). The BDI-II has demonstrated convergent and discriminant validity in college samples (Steer & Clark, 1997). Cronbach's alpha for the BDI-II was excellent ($\alpha = 0.92$).

Beck Anxiety Inventory (BAI; Beck & Steer, 1990). The BAI is a 21-item measure of current anxiety symptoms assessed on a 4-point Likert-type scale, with responses ranging from 0 to 3 for each item. Higher aggregate scores represent more severe levels of anxiety, a known risk factor for suicidal behavior (Nock et al., 2010). The BAI has demonstrated convergent and discriminant validity in college samples (Creamer, Foran, & Bell, 1995). Cronbach's alpha for the BAI was excellent ($\alpha = 0.91$).

Levenson Locus of Control (LOC) Scales – Short Form (Levenson, 1973). Ten items from the original Multidimensional Locus of Control (LOC) Scales were selected in concordance with those used in the National Comorbidity Survey ("National Comorbidity Survey," 2005) as psychometrically distinct and reliable measures of internal LOC (e.g., "My life is determined by my own actions"), external chance LOC (e.g., "When I get what I want, it is usually because I am lucky"), and external power LOC (e.g., "My life is chiefly controlled by powerful others"). Each item was rated on a Likert-type scale from 0 ("not true at all") to 3 ("very true"), with total scores summed for each of the three subscales. Locus of control was included in this study as an exploratory predictor of suicide risk given evidence that external locus of control is associated with depression and adolescent suicidality (Benassi, Sweeney, & Dufour, 1988; Pearce & Martin, 1993). Internal reliabilities were in the poor to acceptable range for the internal LOC and external chance LOC subscales (αs = 0.59, 0.61; respectively), and in the acceptable to good range for the external power LOC subscale (α = 0.79).

Data Analysis

SPSS software (IBM Corp., 2011) was used for all analyses. Groups of current suicide risk were determined according to DSI-SS scores, with two cut points set at: (1) scores greater than zero (i.e., "Lower-Risk Endorsers"), and (2) scores greater than or equal to three (i.e., "Higher-Risk Endorsers"), as suggested by Joiner et al. (2002). Participants who scored zero on the DSI-SS were categorized as "Deniers." Participants who did not answer some or all of the DSI-SS items, despite completing all other items in the survey (i.e., participants with a univariate DSI-SS pattern of missing data), were categorized as "Non-Disclosers." Although discriminating Endorsers into "Higher-Risk" and "Lower-Risk" groups was not necessary for missing data analyses, and in practice the cutoff scores could be moved to alter mean difference contrast tests,

Joiner and colleagues (2002) make a strong argument for why the cutoff scores of zero and three are points of interest for many researchers and clinicians. We therefore decided for descriptive reasons to include the above defined groups in our results, recognizing the limitations of categorical suicide risk classification (see Footnote 1).

Group comparison analyses were limited to univariate patterns of missingness on the DSI-SS only, such that subjects with other patterns of missing data were excluded list-wise. Because data were isolated to a univariate pattern of missing data for DSI-SS only, Little's (1988) MCAR test was able to be used to test if such univariate DSI-SS missingness was generally MCAR, or was alternatively associated with auxiliary measure scores. Multivariate Analysis of Variance (MANOVA) tests were also used to compare across groups mean distributions on individual auxiliary measures and further characterize associations with missingness (Rubin, 1976). Auxiliary measures in each analysis were limited to other suicide risk-related indicators available in the sample (e.g., past suicidality, hopelessness, perceived burdensomeness, thwarted belongingess, etc.).

After group comparisons, three sets of multiple imputation were used to predict missing DSI-SS scores, each using different assumptions. First, all available data from the original datasets, such that there were general patterns of missingness, were used to compute 10 cycles of multivariate imputation by Markov chain Monte Carlo (MCMC) fully conditional specification multiple imputation, with 100 iterations between cycles. Predicted DSI-SS score distributions of Non-Disclosers were then reported from the fully imputed datasets. Exploratory tests showed no problems with convergence or proportion of missing data (Bodner, 2008; Enders, 2010). The other two sets of multiple imputations were univariate, such that the available data were limited to univariate patterns of DSI-SS missingness, as was done in the mean comparisons, and

predicted DSI-SS scores were computed using 10 cycles of linear regression estimation from complete data of the auxiliary measures (Marini, Olsen, & Rubin, 1980). The univariate imputations were computed either: (a) utilizing all compete auxiliary data, or (b) utilizing only complete auxiliary data from participants with non-zero suicide risk scores (i.e., excluding Deniers). The last condition was included to allow for the possibility that Non-Disclosers may be more similar to Endorsers than Deniers. Given this assumption, multiple imputation excluding the Deniers' data may be less biased because the Deniers were very large groups (i.e., 6.7 to 7.2 times the size of the Endorsers groups), and their auxiliary data would be highly influential of imputed predictions.

Results

Measure score distributions and pair-wise inter-correlations, across all samples, are presented in Table 5. Reported scores on the suicide risk measure (i.e., DSI-SS) were significantly related to nearly all of the auxiliary measures, with the exceptions of internal and external-chance locus of control. Similarly, nearly all of the auxiliary measures were significantly inter-correlated. Fearlessness about death scores were a notable exception, in that they were associated with measures of perceived burdensomeness and suicidality (e.g., SBQ-R, DSI-SS), but not with measures of depression, anxiety, hopelessness, or locus of control. This finding is consistent with the conceptualization of fearlessness about death as an independent predictor of suicide risk, and not necessarily associated with other mood- and anxiety-related risk factors (Ribeiro et al., 2013).

For Samples 1 and 2, Little's (1988) MCAR test of the univariate missing DSI-SS pattern was non-significant ($\chi^2_{\text{Sample 1}} = 7.337(3)$, $p = .062$; $\chi^2_{\text{Sample 2}} = 4.810(3)$, $p = .186$), and MCAR could not be ruled out, though Sample 1 data trended away from MCAR. For Sample 3, Little's

MCAR test was highly significant ($\chi^2_{\text{Sample 3}} = 42.032(10)$, $p < .001$), suggesting that univariate missing DSI-SS data was not MCAR, but was instead associated with some factor(s) reliably distinguishing Non-Disclosers from other respondents.

MANOVA results for the auxiliary measure mean comparisons are presented in Table 6. Multivariate omnibus tests (i.e., Wilks' lambda) were significant for all samples; however, heterogeneity of group variances and group sizes warrant caution for interpretation. Specifically, the Deniers group generally had both the largest number of subjects and the smallest variance, making the *F* statistics more liberal (i.e., artificially low *p* values). To correct for inequality of variances and group sizes, follow-up tests were conducted excluding the Deniers group, such that only Non-Disclosers and Endorsers were compared. Homogeneity of variance held in these comparisons and all significance trends between Non-Disclosers and Endorsers remained the same as in the full model. Follow-up contrast tests for pair-wise group comparisons (i.e., at Bonferroni corrected α levels of 0.017 for three planned comparisons) were conducted in all significant one-way ANOVAs.

In Sample 1, Non-Disclosers were found to report significantly higher levels of the proxy measures *Thwarted Belonging* ($t = 2.79$, $p = .005$, $d = 1.05$) and *Perceived Burden* ($t = 2.47$, $p = .014$, $d = 0.94$) than Deniers. In Sample 2, Non-Disclosers were found to report significantly higher levels of past suicidal behavior (i.e., SBQ-R and the DSI-SS single item variations) than Deniers (ts = 3.02, 2.35; ps = .003, .019, ds = 1.24, 0.96; respectively), but significantly lower levels than Higher-Risk Endorsers (ts = -2.89, -3.18; ps = .004, .002, ds = 1.28, 1.41; respectively). However, the significant difference between Non-Disclosers' and Deniers' mean past DSI-SS single item variation did not survive Bonferroni correction. In Sample 3, Non-Disclosers were found to report the highest mean levels across all groups in many of the

auxiliary measures, including past suicidal behavior, thwarted belongingness, perceived burdensomeness, hopelessness, measures of depression and anxiety, and external power locus of control. Mean levels of Non-Disclosers were significantly higher than Deniers in all of the above cases (*t*s = 3.37 to 7.71; *p*s = .001 to < .001; *d*s = 1.31 to 2.98), and also significantly higher than Lower-Risk Endorsers (*t*s = 2.90 to 3.50; *p*s = .004 to .001; *d*s = 1.36 to 1.64) for past suicidal behavior, perceived burdensomeness, and measures of depression and anxiety. Mean levels on auxiliary measures for Non-Disclosers in Sample 3 were generally not significantly different from those of Higher-Risk Endorsers.

Predicted suicide risk scores of Non-Disclosers from the three different multiple imputations are presented in Table 7. Predicted DSI-SS score means of Non-Disclosers ranged from 0.52 (*SE* = 0.34) to 4.39 (*SE* = 1.70), depending on the sample used and whether all data were included in imputation or only complete data from Endorsers. Predicted scores were generally higher when auxiliary data were limited to univariate non-responders, and again when limited to Endorsers' data only. Of note, Sample 3 which showed the highest rate of non-disclosure (i.e., 4.1%), also predicted the highest levels of suicide risk among the Non-Disclosers.

Discussion

Using online self-report data, this study aimed to test: (a) if individuals who chose not to respond to specifically suicide risk screening items (i.e., Non-Disclosers) composed a unified group of participants distinct from other risk response groups (i.e., Deniers, Lower-Risk Endorsers, and Higher-Risk Endorsers) on measures related to suicidality, and (b) if real-time assessment of Non-Disclosers' suicide risk could be ethical and feasible.

Regarding the first aim, despite skipping current suicide risk items, Non-Disclosers across all three samples were found to report significantly higher scores than those of Deniers on nearly all of the other suicide-related variables, suggesting that Non-Disclosers of current suicide risk are a distinct response group at greater than zero suicide risk. The only exceptions to this trend were fearlessness about death and the exploratory measures of internal and external-chance locus of control (discussed below). Further, effect sizes for the mean differences were large to very large (i.e., all had Cohen's *d*s greater than 0.8). For example on the BDI-II, Non-Disclosers (M = 25.29, SE = 2.49) scored on average nearly 20 points higher than Deniers (M = 6.22, SE = 0.55). Likewise, SBQ-R scores of Non-Disclosers in Sample 3 (M = 10.29, SE = 0.76) were not only on average six points higher than those of Deniers (M = 4.31, SE = 0.17), but clearly exceeded the SBQ-R discriminant risk score (i.e., total score greater than seven; Osman et al., 2001). This evidence supports the idea that individuals who choose not to disclose current suicide risk information are at significantly higher risk for suicide than individuals who explicitly deny suicide risk. Despite being limited by group size and variance differences favoring significance tests, the effect sizes found for these differences provide a strong argument for the validity of these results.

The exceptions of fearlessness about death and internal and external-chance locus of control are notable. Fearlessness about death is often regarded as only a moderator of other suicide risk indicators, and although associated with suicidal behavior, it is not generally associated with suicidal ideation *per se* (Ribeiro et al., 2013). Given that the other suicide risk variables in this study associated with non-disclosure are also known to be associated with suicidal ideation, and fearlessness about death is not associated with either, these findings suggest that non-disclosure may be more related to levels of suicidal ideation than potential for

suicidal behavior. The exploratory finding that Non-Disclosers reported significantly elevated levels of specifically external-power locus of control is also consistent with this hypothesis, given that higher reported levels of external locus of control have been associated with increased depression and suicidal ideation (Benassi et al., 1988; Pearce & Martin, 1993). On the other hand, this finding may alternatively suggest that Non-Disclosers are particularly sensitive to perceived power dynamics, and that a unique contributor to non-disclosure may be a perceived lack of control or understanding of how suicide risk response data will be handled. In contrast to the otherwise consistent difference trends found in this study between suicide risk Deniers and Non-Disclosers, both of the above exceptions present avenues in need of further research to better characterize unique contributing factors to suicide risk non-disclosure.

Results were less clear from analyses aiming to distinguish Non-Disclosers from Higher- versus Lower-Risk Endorsers. In Sample 1, Non-Disclosers were not significantly different from either group of Endorsers in the proxy measures of *Thwarted Belonging* or *Perceived Burden*. In Sample 2, Non-Disclosers responded more similarly to Lower-Risk Endorsers, at significantly lower levels than Higher-Risk Endorsers, on both measures of past suicidal behavior. In Sample 3, however, Non-Disclosers reported the highest levels of all response groups on nearly every measure associated with suicide risk. They responded more similarly to Higher-Risk Endorsers, and significantly higher than Lower-Risk Endorsers, on measures of past suicidal behavior, depression, anxiety, hopelessness, and perceived burdensomeness. These differences may be attributable to a sample effect, such that more higher-risk participants were Non-Disclosers in Sample 3 than were in Samples 1 and 2. No significant differences were found among the Non-Disclosers across samples with respect to age, gender, or race—though cell sizes were small. However, the testing procedure in Sample 3 was notably different than in Samples 1 and 2, such

that only in Sample 3 were participants actively screened for suicide risk. Participants in Sample 3 were informed that some questions would be about suicide, and they consented to be contacted by the study investigator if their responses indicated that they "may benefit from further resources." Though there may be other sample differences, this finding suggests that conducting active suicide risk screening in online surveys may promote non-disclosure from higher-risk participants.

Regarding the second aim to investigate whether allowing for non-disclosure may be ethical and feasible, the above findings suggest that Non-Disclosers should be categorized as at least above zero suicide risk, and in cases of active suicide risk monitoring, closer to higher suicide risk. Allowing for non-disclosure of suicide risk screening items could be ethical if follow-up assessment or other appropriate suicide risk intervention is therefore conducted for all participants who do not disclose current suicide risk—but would that be feasible? In Sample 3, with active risk monitoring, the amount of non-disclosure was high relative to the amount of higher-risk endorsement, at nearly one-to-one. However, the overall frequency of non-disclosure was still small (less than 5% of the total sample), and we argue of a manageable proportion to conduct follow-up assessment. Given the finding that Non-Disclosers in Sample 3 were predicted by multiple imputation to possibly be at higher-risk for suicide, flagging them for follow-up would have nearly doubled the amount of (predicted) higher-risk participants identified. Although these findings are unable to be compared to counterfactual data of forced responding, we argue that given the reporting limitations of self-report online suicide screening, the mere potential for such greatly increased screening sensitivity (i.e., from if one were to rely on self-report risk endorsement only) outweighs the potential cost in specificity (i.e., from following-up with potential false-positive Non-Disclosers). The present data support the idea that

non-disclosure is itself an indicator of suicide risk, one that although rarely chosen, may in fact be unethical to omit or ignore.

Limitations and Future Directions

Given that the samples observed in the present study contained relatively healthy student populations and cross-sectional self-report data, the generalizability of these findings calls for some caution. Similarly, the Non-Discloser groups were small and not demographically diverse, and again should not be generalized to all populations. Furthermore, this study did not directly test whether forcing suicide risk screening responses indeed increases underreporting and masks the Non-Disclosers group as proposed, or if that concern is in fact invalid. Although this paper argues that the benefits of allowing for non-disclosure outweigh the potential risks of forcing responses, future research is needed to directly test the validity of such risks. Additionally, more qualitative and quantitative research is needed on suicide risk non-disclosure in general, including further study among more diverse participant samples, under various conditions (e.g., with or without being actively screened, with or without being presented an explicit "non-disclosure" option), by comparing other measurement modalities (e.g., utilizing clinician-reported or behaviorally measured suicide risk), and with assessment of suicide risk at multiple time points.

Conclusion

The findings from this study strongly suggest that missing data specifically on suicide risk screening items do not occur completely at random. Instead, selective non-disclosure can be intentional and defines a group of respondents that are likely at some level of elevated suicide risk, based on the information that they do provide. In every sample of this study, participants were given the option to deny suicide risk and deny referral for services. Likewise, participants

were also given the option to invite further services and endorse if they felt they were at risk for suicide. Considering its limitations, this study found that a small and distinct group of participants opted not to select either option, but instead chose to remain unselected or undecided. We argue that only when we allow for those participants to make a selection of non-disclosure will their true response be recorded, and it is then the ethical obligation of the survey administrator(s) to follow up with them and address their concerns that caused non-disclosure, whatever those may be. If screening surveys, on the other hand, force responses and do not allow for non-disclosure, the investigating researchers and clinicians may unintentionally "miss" otherwise critical missing data.

References

Apter, A., Horesh, N., Gothelf, D., Graffi, H., & Lepkifker, E. (2001). Relationship between self-disclosure and serious suicidal behavior. *Comprehensive Psychiatry*, *42*(1), 70–75. doi:10.1053/comp.2001.19748

Beck, A. T., & Steer, R. A. (1990). *Manual for the Beck Anxiety Inventory*. New York: Psychological Corporation.

Beck, A. T., Steer, R. A., Beck, J. S., & Newman, C. F. (1993). Hopelessness, depression, suicidal ideation, and clinical diagnosis of depression. *Suicide and Life-Threatening Behavior*, *23*(2), 139–45.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the BDI-II*. San Antonio, TX: Psychological Corporation.

Beck, A. T., Steer, R. A., Kovacs, M., & Garrison, B. (1985). Hopelessness and eventual suicide: A 10-year prospective study of patients hospitalized with suicidal ideation. *American Journal of Psychiatry*, *142*(5), 559–563.

Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: The Hopelessness Scale. *Journal of Consulting and Clinical Psychology*, *42*(6), 861–5.

Benassi, V. A., Sweeney, P. D., & Dufour, C. L. (1988). Is there a relation between locus of control orientation and depression? *Journal of Abnormal Psychology*, *97*(3), 357–67.

Bender, T. W., Gordon, K. H., Bresin, K., & Joiner, T. (2011). Impulsivity and suicidality: The mediating role of painful and provocative experiences. *Journal of Affective Disorders*, *129*(1-3), 301–7. doi:10.1016/j.jad.2010.07.023

Bodner, T. E. (2008). What improves with increased missing data imputations? *Structural Equation Modeling*, *15*(4), 651–675. doi:10.1080/10705510802339072

Brener, N., Billy, J., & Grady, W. (2003). Assessment of factors affecting the validity of self-reported health-risk behavior among adolescents: Evidence from the scientific literature. *Journal of Adolescent Health*, *33*(6), 436–57.

Bryan, C. J., & Rudd, M. (2006). Advances in the assessment of suicide risk. *Journal of Clinical Psychology*, *62*(2), 185–200.

Carlton, P. A., & Deane, F. P. (2000). Impact of attitudes and suicidal ideation on adolescents' intentions to seek professional psychological help. *Journal of Adolescence*, *23*(1), 35–45. doi:10.1006/jado.1999.0299

Centers for Disease Control and Prevention. (2015). Web-based Injury Statistics Query and Reporting System (WISQARS). Retrieved from http://www.cdc.gov/injury/wisqars

Christensen, H., Batterham, P. J., & O'Dea, B. (2014). E-Health Interventions for Suicide Prevention. *International Journal of Environmental Research and Public Health*, *11*(8), 8193–8212. doi:10.3390/ijerph110808193

Cigularov, K., Chen, P., Thurber, B., & Stallones, L. (2008). What prevents adolescents from seeking help after a suicide education program? *Suicide and Life-Threatening Behavior*, *38*(1), 74–86.

Creamer, M., Foran, J., & Bell, R. (1995). The Beck Anxiety Inventory in a non-clinical sample. *Behaviour Research and Therapy*, *33*(4), 477–485. doi:10.1016/0005-7967(94)00082-U

Deane, F. P., Wilson, C. J., & Ciarrochi, J. (2001). Suicidal ideation and help-negation: Not just hopelessness or prior help. *Journal of Clinical Psychology*, *57*(7), 901–914. doi:10.1002/jclp.1058

Enders, C. K. (2010). *Applied Missing Data Analysis*. New York: Guilford Press.

Haas, A., Koestner, B., Rosenberg, J., Moore, D., Garlow, S., Sedway, J., … Nemeroff, C. (2008). An interactive web-based method of outreach to college students at risk for suicide. *Journal of American College Health*, *57*(1), 15–22. doi: 10.3200/JACH.57.1.15-22

Harris, K. M., McLean, J. P., & Sheffield, J. (2009a). Examining suicide-risk individuals who go online for suicide-related purposes. *Archives of Suicide Research*, *13*(3), 264–276. doi:10.1080/13811110903044419

Harris, K. M., McLean, J. P., & Sheffield, J. (2009b). Solving suicidal problems online: Who turns to the Internet for help? *Advances in Mental Health*, *8*(1), 28–36. doi:10.5172/jamh.8.1.28

Harris, K. M., McLean, J. P., Sheffield, J., & Jobes, D. A. (2010). The internal suicide debate hypothesis: Exploring the life versus death struggle. *Suicide and Life-Threatening Behavior*, *40*(2), 181–92. doi:10.1521/suli.2010.40.2.181

Horesh, N., Zalsman, G., & Apter, A. (2004). Suicidal behavior and self-disclosure in adolescent psychiatric inpatients. *The Journal of Nervous and Mental Disease*, *192*(12), 837–842. doi:10.1097/01.nmd.0000146738.78222.e5

IBM Corp. (2011). IBM SPSS Statistics for Windows. Armonk, NY: IBM Corp.

Joiner, T., Pfaff, J. J., & Acres, J. G. (2002). A brief screening tool for suicidal symptoms in adolescents and young adults in general health settings: Reliability and validity data from the Australian National General Practice Youth Suicide Prevention Project. *Behaviour Research and Therapy*, *40*(4), 471–81.

Joiner, T., Walker, R. L., Rudd, M. D., & Jobes, D. A. (1999). Scientizing and routinizing the assessment of suicidality in outpatient practice. *Professional Psychology: Research and Practice*, *30*(5), 447–453. doi:10.1037//0735-7028.30.5.447

Kazdin, A. E., & Blase, S. L. (2011). Rebooting psychotherapy research and practice to reduce the burden of mental illness. *Perspectives on Psychological Science*, *6*(1), 21–37. doi:10.1177/1745691610393527

King, C. A., Eisenberg, D., Zheng, K., Czyz, E., Kramer, A., Horwitz, A., & Chermack, S. (2015). Online suicide risk screening and intervention with college students: A pilot randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *83*(3), 630–636. doi:10.1037/a0038805

Levenson, H. (1973). Multidimensional locus of control in psychiatric patients. *Journal of Consulting and Clinical Psychology*, *41*(3), 397–404.

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198–1202. doi:10.1080/01621459.1988.10478722

Marini, M. M., Olsen, A. R., & Rubin, D. B. (1980). Maximum-likelihood estimation in panel studies with missing data. *Sociological Methodology*, *11*, 314–357.

McMillan, D., Gilbody, S., Beresford, E., & Neilly, L. (2007). Can we predict suicide and non-fatal self-harm with the Beck Hopelessness Scale? A meta-analysis. *Psychological Medicine*, *37*(6), 769–78. doi:10.1017/S0033291706009664

Michaels, M. S., Chu, C., Silva, C., Schulman, B. E., & Joiner, T. (2014). Considerations regarding online methods for suicide-related research and suicide risk assessment. *Suicide and Life-Threatening Behavior*, 1–8. doi:10.1111/sltb.12105

National Comorbidity Survey. (2005). Retrieved April 6, 2015, from http://www.hcp.med.harvard.edu/ncs/

Nock, M. K., Hwang, I., Sampson, N. A., & Kessler, R. C. (2010). Mental disorders, comorbidity and suicidal behavior: Results from the National Comorbidity Survey Replication. *Molecular Psychiatry*, *15*(8), 868–76. doi:10.1038/mp.2009.29

Osman, A., Bagge, C. L., Gutierrez, P. M., Konick, L. C., Kopper, B. A., & Barrios, F. X. (2001). The Suicidal Behaviors Questionnaire-Revised (SBQ-R): Validation with clinical and nonclinical samples. *Assessment*, *8*(4), 443–454. doi:10.1177/107319110100800409

Pearce, C. M., & Martin, G. (1993). Locus of control as an indicator of risk for suicidal behaviour among adolescents. *Acta Psychiatrica Scandinavica*, *88*(6), 409–414.

Ribeiro, J. D., Witte, T. K., Van Orden, K. A., Selby, E. A., Gordon, K. H., Bender, T. W., & Joiner, T. (2013). Fearlessness about death: The psychometric properties and construct validity of the revision to the Acquired Capability for Suicide Scale. *Psychological Assessment*. doi:10.1037/a0034858

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*(3), 581. doi:10.2307/2335739

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.

Scelfo, J. (2015, July 27). Suicide on campus and the pressure of perfection. *The New York Times*. Retrieved from http://nyti.ms/1VIuROq

Schwartz, A. J. (2010). College student suicide in the United States: 1990-1991 through 2003-2004. *Journal of American College Health*, *54*(6), 341–352. doi:10.3200/JACH.54.6.341-352

Steer, R. A., & Clark, D. A. (1997). Psychometric characteristics of the Beck Depression Inventory-II with college students. *Measurement & Evaluation in Counseling & Development*, *30*(3), 128.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, *133*(5), 859–83.

Turner, C. F., Ku, L., Rogers, S. M., Lindberg, L. D., Pleck, J. H., & Sonenstein, F. L. (1998). Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. *Science*, *280*(5365), 867–873. doi:10.1126/science.280.5365.867

Van Orden, K. A., Cukrowicz, K. C., Witte, T. K., & Joiner, T. (2012). Thwarted belongingness and perceived burdensomeness: Construct validity and psychometric properties of the Interpersonal Needs Questionnaire. *Psychological Assessment*, *24*(1), 197–215. doi:10.1037/a0025358

Van Orden, K. A., Witte, T. K., Cukrowicz, K. C., Braithwaite, S. R., Selby, E. A., & Joiner, T. (2010). The interpersonal theory of suicide. *Psychological Review*, *117*(2), 575–600. doi:10.1037/a0018697

Van Orden, K. A., Witte, T. K., Gordon, K. H., Bender, T. W., & Joiner, T. (2008). Suicidal desire and the capability for suicide: Tests of the interpersonal-psychological theory of suicidal behavior among adults. *Journal of Consulting and Clinical Psychology*, *76*(1), 72–83. doi:10.1037/0022-006X.76.1.72

World Health Organization (WHO). (2014). *Preventing suicide: A global imperative*. Geneva.

Footnotes

[1] Suicide risk cutoff groups were used for largely descriptive purposes. As with many, if not all, mental health constructs, suicide risk may be better conceptualized on a continuum. Suicide risk is a complex product of risk factors and warning signs and must be determined with clinical judgment. Because DSI-SS score cutoffs do not indicate absolute risk groups, we therefore use "Higher-Risk" and "Lower-Risk" group labels to emphasize the relativity of our screening classifications.

[2] The procedures of these research screening surveys have since changed to include active suicide risk monitoring and follow-up assessment of individuals who report elevated suicide risk.

Table 1

Suicide Risk Screen (DSI-SS) Missing Data Patterns

			Measures (X = Missing Data)								
	Pattern	*n*	DSI-SS	FAD	Thwarted Belonging	Perceived Burden†					
Sample 1	1*	1299									
	2*	7	X								
	3	2	X	X							
	4	1	X		X	X					
	5	1	X	X		X					
			DSI-SS	FAD	SBQ-R†	DSI-SS†					
Sample 2	1*	688									
	2*	6	X								
	3	2	X		X	X					
	4	1	X	X							
	5	1	X		X						
	6	1	X	X	X	X					
			DSI-SS	FAD	BAI	INQ-TB	INQ-PB	BHS	BDI-II	SBQ-R	LOC-I/ ExCh/ ExP
Sample 3	1*	165									
	2*	7	X								
	3	1	X		X	X	X	X			

Note. Missing data patterns with complete DSI-SS scores not shown (*n*s = 9, 8, 11). *DSI-SS* = Depression Symptom Inventory - Suicidality Subscale. *FAD* = Acquired Capability for Suicide Scale - Fearlessness About Death. *SBQ-R* = Suicidal Behavior Questionnaire - Revised. *INQ-TB/ PB* = Interpersonal Needs Questionnaire - Thwarted Belongingness subscale/ Perceived Burdensomeness subscale. *BHS* = Beck Hopelessness Scale. *BDI-II* = Beck Depression Inventory-II. *BAI* = Beck Anxiety Inventory. *LOC-I/ ExCh/ ExP* = Locus of Control Short Form - Internal subscale/ External-Chance subscale/ External-Power subscale.

† Single-item variation.

* Selected for univariate missing data analyses

Table 2

Sample 1 Participant Characteristics for Univariate Analyses

	Total *n* = 1306	Deniers *n* = 1125 (86.1%)	Lower-Risk Endorsers *n* = 108 (8.3%)	Higher-Risk Endorsers *n* = 66 (5.1%)	Non-Disclosers *n* = 7 (0.5%)
DSI-SS[a]	0.30 ± 0.96	0.00 ± 0.00	1.21 ± 0.41	3.94 ± 1.25	- -
Age[b]	18.48 ± 0.68	18.48 ± 0.68	18.57 ± 0.74	18.35 ± 0.69	18.57 ± 0.79
Gender[c]					
Female	899 (69%)	781 (69%)	76 (70%)	40 (61%)	2 (29%)
Male	407 (31%)	344 (31%)	32 (30%)	26 (39%)	5 (71%)
Race[c]					
White	923 (71%)	803 (71%)	71 (66%)	42 (64%)	7 (100%)
Black	111 (8%)	100 (9%)	8 (7%)	3 (5%)	0 (0%)
Hispanic	228 (17%)	188 (17%)	24 (22%)	16 (24%)	0 (0%)
Asian	38 (3%)	28 (2%)	5 (5%)	5 (8%)	0 (0%)
Other Race	6 (<1%)	6 (1%)	0 (0%)	0 (0%)	0 (0%)

Note. All values represent frequencies with percentages for each category by suicide screening risk level, except for age and DSI-SS where the Mean and Standard Deviation are listed. *DSI-SS* = Depression Symptom Inventory - Suicidality Subscale. *Deniers* = score of zero on the DSI-SS. *Lower-Risk Endorsers* = score greater than zero but less than three on the DSI-SS. *Higher-Risk Endorsers* = score greater than or equal to three on the DSI-SS. *Non-Disclosers* = some or all of the DSI-SS scores missing.

[a] $p < .001$; One sample t-tests for non-zero groups. $p < .001$; Independent samples Welch t-test between non-zero groups.

[b] $p > .05$; One-way ANOVA.

[c] $p > .05$; Pearson chi-squared.

Table 3

Sample 2 Participant Characteristics for Univariate Analyses

	Total *n* = 694	Deniers *n* = 609 (87.8%)	Lower-Risk Endorsers *n* = 45 (6.5%)	Higher-Risk Endorsers *n* = 34 (4.9%)	Non-Disclosers *n* = 6 (0.9%)
DSI-SS[a]	0.31 ± 1.03	0.00 ± 0.00	1.40 ± 0.50	4.32 ± 1.36	- -
Age[b]	18.35 ± 0.65	18.35 ± 0.65	18.36 ± 0.68	18.35 ± 0.69	18.17 ± 0.41
Gender[c]					
Female	489 (70%)	426 (70%)	33 (73%)	26 (76%)	4 (67%)
Male	205 (30%)	183 (30%)	12 (27%)	8 (24%)	2 (33%)
Race[c]					
White	485 (70%)	429 (70%)	32 (71%)	21 (62%)	3 (50%)
Black	58 (8%)	51 (8%)	0 (0%)	5 (15%)	2 (33%)
Hispanic	120 (17%)	100 (16%)	11 (24%)	8 (24%)	1 (17%)
Asian	25 (4%)	23 (4%)	2 (4%)	0 (0%)	0 (0%)
Other Race	6 (1%)	6 (1%)	0 (0%)	0 (0%)	0 (0%)

Note. All values represent frequencies with percentages for each category by suicide screening risk level, except for age and DSI-SS where the mean ± standard deviation are listed. *DSI-SS* = Depression Symptom Inventory - Suicidality Subscale. *Deniers* = score of zero on the DSI-SS. *Lower-Risk Endorsers* = score greater than zero but less than three on the DSI-SS. *Higher-Risk Endorsers* = score greater than or equal to three on the DSI-SS. *Non-Disclosers* = some or all of the DSI-SS scores missing.

[a] $p < .001$; One sample t-tests for non-zero groups. $p < .001$; Independent samples Welch t-test between non-zero groups.

[b] $p > .05$; One-way ANOVA.

[c] $p > .05$; Pearson chi-squared.

Table 4

Sample 3 Participant Characteristics for Univariate Analyses

	Total *n* = 172	Deniers *n* = 144 (83.7%)	Lower-Risk Endorsers *n* = 13 (7.6%)	Higher-Risk Endorsers *n* = 8 (4.7%)	Non-Disclosers *n* = 7 (4.1%)
DSI-SS[a]	0.33 ± 1.11	0.00 ± 0.00	1.38 ± 0.51	4.50 ± 2.07	- -
Age[b]	18.85 ± 1.39	18.83 ± 1.42	19.31 ± 1.44	18.63 ± 1.06	18.71 ± 1.25
Gender[c]					
Female	129 (75%)	108 (75%)	9 (69%)	6 (75%)	6 (86%)
Male	43 (25%)	36 (25%)	4 (31%)	2 (25%)	1 (14%)
Race[c]					
White	111 (65%)	89 (62%)	10 (77%)	6 (75%)	6 (86%)
Black	17 (10%)	15 (10%)	1 (8%)	1 (13%)	0 (0%)
Hispanic	30 (17%)	28 (19%)	0 (0%)	1 (13%)	1 (14%)
Asian	10 (6%)	8 (6%)	2 (15%)	0 (0%)	0 (0%)
Other Race	4 (2%)	4 (3%)	0 (0%)	0 (0%)	0 (0%)

Note. All values represent frequencies with percentages for each category by suicide screening risk level, except for age and DSI-SS where the mean ± standard deviation are listed. *DSI-SS* = Depression Symptom Inventory - Suicidality Subscale. *Deniers* = score of zero on the DSI-SS. *Lower-Risk Endorsers* = score greater than zero but less than three on the DSI-SS. *Higher-Risk Endorsers* = score greater than or equal to three on the DSI-SS. *Non-Disclosers* = some or all of the DSI-SS scores missing.

[a] $p < .001$; One sample t-tests for non-zero groups. $p < .01$; Independent samples Welch t-test between non-zero groups.

[b] $p > .05$; One-way ANOVA.

[c] $p > .05$; Pearson chi-squared.

Table 5

Descriptive Statistics and Inter-Correlations of Variables Used in Analyses

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13	14.	15.
All Samples															
1. DSI-SS	0.84														
2. FAD	.069**	0.84													
Sample 1															
3. Thwarted Belonging	.250***	-.062*	0.47												
4. Perceived Burden†	.295***	-.121***	.491***	-											
Sample 2															
5. SBQ-R†	.502***	.080*	-	-	-										
6. DSI-SS†	.499***	.090*	-	-	.871***	-									
Sample 3															
7. SBQ-R	.565***	.182*	-	-	-	-	0.79								
8. INQ-TB	.321***	.132	-	-	-	-	.513***	0.88							
9. INQ-PB	.396***	.162*	-	-	-	-	.369***	.579***	0.95						
10. BHS	.481***	0.103	-	-	-	-	.452***	.613***	.589***	0.85					
11. BDI-II	.529***	0.063	-	-	-	-	.651***	.613***	.495***	.708***	0.92				
12. BAI	.367***	-0.070	-	-	-	-	.500***	.413***	.302***	.517***	.662***	0.91			
13. LOC-I	-.045	-.139	-	-	-	-	-.153*	-.251***	-.240**	-.333***	-.321***	-.204**	0.59		
14. LOC-ExCh	.086	.056	-	-	-	-	.070	.295***	.181*	.271***	.223**	.253***	-.113	0.61	
15. LOC-ExP	.222**	-.001	-	-	-	-	.134	.341***	.326***	.362***	.302***	.229**	-.150*	.360***	0.79
n	2180	2191	1317	1316	702	700	184	180	180	176	184	182	184	184	184
Mean	0.30	14.42	4.15	2.39	1.62	1.50	4.90	20.76	8.14	2.76	8.16	11.66	9.48	4.05	1.11
Standard Dev.	0.99	6.67	1.23	1.54	0.83	0.79	2.52	9.92	4.84	3.33	8.20	9.85	1.63	2.22	1.34
Skewness	4.18	< .01	0.45	0.99	1.40	1.51	1.40	0.64	3.34	2.19	1.64	1.51	-0.30	0.20	1.36

Note. Cronbach's alphas are listed along the diagonal. See Table 1 for full measure names.

† Single-item variation.

*$p < .05$, **$p < .01$, ***$p < .001$

Table 6

Auxiliary Measure Mean Comparisons (MANOVAs) across DSI-SS Risk Response Categories

			Non-Disclosers		[A] Deniers (DSI-SS = 0)					[B] Lower-Risk Endorsers (DSI-SS < 3)					[C] Higher-Risk Endorsers (DSI-SS ≥ 3)				
	Wilks' λ	*F*	*M*	*SE*	*M*	*SE*	*t*	*p*	*d [95% CI]*	*M*	*SE*	*t*	*p*	*d [95% CI]*	*M*	*SE*	*t*	*p*	*d [95% CI]*
Sample 1	0.871	20.51***																	
Thwarted Belonging [A]		31.23***	5.29	(0.45)	4.03	(0.04)	-2.79	.005	1.05 [0.31, 1.80]	4.68	(0.11)	-1.32	.188		5.23	(0.15)	-0.12	.901	
Perceived Burden† [A]		48.28***	3.57	(0.55)	2.20	(0.04)	-2.47	.014	0.94 [0.19, 1.68]	3.25	(0.14)	-0.56	.573		4.05	(0.18)	0.82	.415	
FAD		4.12	12.71	(2.50)	14.83	(0.20)	0.85	.398		14.22	(0.64)	0.58	.559		17.52	(0.81)	1.83	.068	
Sample 2	0.689	30.79***																	
SBQ-R† [A, C]		92.30***	2.33	(0.29)	1.46	(0.03)	-3.02	.003	1.24 [0.43, 2.04]	2.44	(0.11)	0.36	.717		3.24	(0.12)	2.89	.004	1.28 [0.36, 2.19]
DSI-SS† [(A), C]		90.33***	2.00	(0.27)	1.36	(0.03)	-2.35	.019	0.96 [0.16, 1.77]	2.40	(0.10)	1.38	.169		2.94	(0.11)	3.18	.002	1.41 [0.48, 2.32]
FAD		1.49	14.00	(2.75)	13.57	(0.27)	-0.16	.875		15.18	(1.00)	0.40	.688		15.38	(1.16)	0.46	.643	
Sample 3	0.382	6.05***																	
SBQ-R [A, B]		38.24***	10.29	(0.76)	4.31	(0.17)	-7.71	< .001	2.98 [2.15, 3.81]	7.00	(0.56)	-3.50	.001	1.64 [0.56, 2.69]	9.25	(0.71)	-1.00	.319	
INQ-TB [A]		17.00***	36.00	(3.30)	18.74	(0.73)	-5.11	< .001	1.98 [1.18, 2.77]	30.77	(2.42)	-1.28	.202		27.50	(3.08)	-1.88	.061	
INQ-PB [A, B]		12.17***	15.57	(1.65)	7.35	(0.36)	-4.87	< .001	1.88 [1.09, 2.67]	8.46	(1.21)	-3.48	.001	1.63 [0.55, 2.67]	13.38	(1.54)	-0.97	.332	
FAD		2.49	16.57	(2.33)	12.88	(0.51)	-1.55	.124		15.15	(1.71)	-0.49	.624		17.50	(2.18)	0.29	.771	
BHS [A, (B)]		17.20***	7.86	(1.12)	2.11	(0.25)	-5.00	< .001	1.94 [1.14, 2.72]	4.54	(0.82)	-2.39	.018	1.12 [0.12, 2.09]	7.38	(1.05)	-0.31	.754	
BDI-II [A, B]		35.56***	25.29	(2.49)	6.22	(0.55)	-7.48	< .001	2.90 [2.07, 3.72]	15.31	(1.83)	-3.23	.001	1.52 [0.46, 2.54]	21.25	(2.33)	-1.18	.238	
BAI [A, B]		12.64***	25.86	(3.45)	10.38	(0.76)	-4.38	< .001	1.70 [0.91, 2.48]	13.46	(2.53)	-2.90	.004	1.36 [0.33, 2.36]	25.25	(3.22)	-0.13	.898	
LOC-I		1.50	8.43	(0.60)	9.56	(0.13)	1.83	.070		9.77	(0.44)	1.79	.075		9.00	(0.56)	0.69	.490	
LOC-ExCh		0.78	4.29	(0.83)	3.95	(0.18)	-0.39	.696		4.23	(0.61)	-0.05	.958		5.13	(0.78)	0.73	.464	
LOC-ExP [A]		7.57***	2.57	(0.48)	0.92	(0.11)	-3.37	.001	1.31 [0.53, 2.08]	1.62	(0.35)	-1.61	.109		2.38	(0.45)	-0.30	.765	

Note. Limited to univariate missing data patterns on the DSI-SS. See Table 1 for full measure names.

[A, B, C] Significant one-way ANOVA contrasts (i.e., *Non-Disclosers* versus *Group A/B/C*). Parentheses indicate a loss of significance after Bonferroni correction (i.e., $\alpha = .017$).

† Single-item variation.

***$p < .001$

Table 7

Multiple Imputation Predictions of Missing DSI-SS Scores

	Multivariate Using All Data		Univariate Using Complete Data		Univariate Using Endorsers' Data Only	
	M	*SE*	*M*	*SE*	*M*	*SE*
Sample 1	0.52	(0.34)	0.67	(0.51)	2.30	(0.76)
Sample 2	0.61	(0.48)	1.15	(0.65)	3.33	(1.30)
Sample 3	1.95	(0.52)	2.47	(0.58)	4.39	(1.70)
Average	**1.03**	**(0.45)**	**1.43**	**(0.58)**	**3.34**	**(1.25)**

Note. Means and standard errors computed from 10 repetitions of imputation. Multivariate imputation computed by Markov chain Monte Carlo (MCMC) fully conditional specification multiple imputation, with 100 iterations between cycles. Univariate imputation computed by linear regression estimation from complete auxiliary data. Standard errors computed according to Rubin's (1987) rules. Auxiliary predictors were the same as those used in MANOVA analyses.