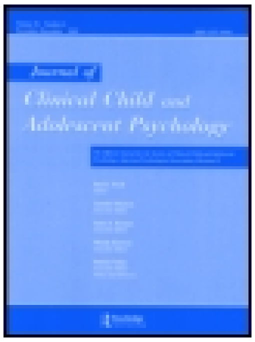

Journal of Clinical Child & Adolescent Psychology

ISSN: 1537-4416 (Print) 1537-4424 (Online) Journal homepage: https://www.tandfonline.com/loi/hcap20

Statistical Approaches for Highly Skewed Data: Evaluating Relations between Maltreatment and Young Adults' Non-Suicidal Self-injury

Ana Gonzalez-Blanks, Jessie M. Bridgewater & Tuppett M. Yates

To cite this article: Ana Gonzalez-Blanks, Jessie M. Bridgewater & Tuppett M. Yates (2020): Statistical Approaches for Highly Skewed Data: Evaluating Relations between Maltreatment and Young Adults' Non-Suicidal Self-injury, Journal of Clinical Child & Adolescent Psychology, DOI: 10.1080/15374416.2020.1724543

To link to this article: https://doi.org/10.1080/15374416.2020.1724543

Published online: 20 Feb 2020.

Journal of Clinical Child & Adolescent Psychology, 49(02), 1–15, 2020

ISSN: 1537-4416 print/1537-4424 online
DOI: https://doi.org/10.1080/15374416.2020.1724543

Statistical Approaches for Highly Skewed Data: Evaluating Relations between Maltreatment and Young Adults' Non-Suicidal Self-injury

Ana Gonzalez-Blanks, Jessie M. Bridgewater, and Tuppett M. Yates
Department of Psychology, University of California, Riverside

Objective: Clinical phenomena often feature skewed distributions with an overabundance of zeros. Unfortunately, empirical methods for dealing with this violation of distributional assumptions underlying regression are typically discussed in statistical journals with limited translation to applied researchers. Therefore, this investigation compared statistical approaches for addressing highly skewed data as applied to the evaluation of relations between child maltreatment and non-suicidal self-injury (NSSI). *Method*: College students ($N = 2{,}651$; 64.2% female; 85.2% nonwhite) completed the Child Abuse and Trauma Scale and the Functional Assessment of Self-Mutilation. Statistical models were applied to cross-sectional data to provide illustrative comparisons across predictions to a) raw, highly skewed NSSI outcomes, b) natural log, square-root, and inverse NSSI transformations to reduce skew, c) zero-inflated Poisson (ZIP) and negative-binomial zero-inflated (NBZI) regression models to account for both disproportionate zeros and skewness in the NSSI data, and d) the skew-*t* distribution to model NSSI skewness. *Results*: Child maltreatment was significantly and positively related to NSSI frequency in the raw, transformation, and zero-inflated models, but this relation was negative in the skew-*t* model. *Conclusions*: These findings highlight the importance of using zero-inflated models rather than transformation approaches to address data skew. Moreover, whereas the skew-*t* distribution has been used to model skewed *non-clinical* data, this study suggests that the skew-*t* approach may not be well-suited to address skewed *clinical* data.

Statistical regression is a useful and widely used tool to evaluate relations between a dependent variable and one or more explanatory variables given a set of controls (Cohen, Cohen, West, & Aiken, 2003). However, the accuracy of the obtained parameters rests on the degree to which the data adhere to specific assumptions relevant to the regression model itself and to the model errors (i.e., observed values minus values predicted from the *true* regression model for the population; Osborne & Waters, 2002; Williams, Grajales, & Kurkiewicz, 2013). At the model level, the dependent variable should be a linear function of the regression parameters, such that a graph of residuals (i.e., observed values minus values predicted by the *estimated* regression model; Weisberg, 2005) against predicted values of the dependent variable appears linear in form. At the level of errors, following estimation of the regression model, the residuals (as the best available proxy for errors) should feature a normal distribution and equivalent variance across all levels of the predictors (i.e., homoscedasticity; Cohen et al., 2003). Of the myriad response outcomes to which regression models have been applied across varied domains of psychological science (e.g., extraversion, achievement), clinical response outcomes often pose a special challenge because they tend to feature lower-base rates with an overabundance of zero values. Thus, the current paper drew on the clinical case of non-suicidal self-injury (NSSI; i.e., direct and deliberate damage of one's body tissue without suicidal intent; Nock, 2010) to illustrate the application of varied statistical approaches for addressing the kinds of highly skewed data distributions that characterize many clinical phenomena.

Correspondence should be addressed to Tuppett M. Yates, Department of Psychology, University of California, Riverside, CA 92521, USA. E-mail: Tuppett@ucr.edu

The Phenomenology of NSSI

Over the past 20 years, NSSI has emerged as a prominent public health concern, particularly among young people ages 15–25 (Whitlock, Eckenrode, & Silverman, 2006). A recent review across 128 prevalence estimates in the extant literature confirms significant rates of NSSI among adolescents (17.2%), young adults (13.4%), and adults (5.5%; Swannell, Martin, Page, Hasking, & St John, 2014). Despite elevated rates of NSSI within some sub-populations (e.g., psychiatric inpatient adolescents; Nock & Prinstein, 2004), pooled prevalence estimates indicate that NSSI has a relatively low base rate, such that the number of individuals who do *not* endorse NSSI far exceeds the number that does (i.e., an overabundance of zeros). Moreover, a relatively small subset of individuals who engage in NSSI do so at a high frequency (i.e., the distribution is positively skewed; Benjet et al., 2017; Gillies et al., 2018).

In light of the unique distributional features of NSSI, researchers typically use robust methods (e.g., Thomassin, Shaffer, Madden, & Londino, 2016) or transformations to render their data more amenable to statistical regression in the service of identifying predictors of NSSI. Commonly used transformation methods in the NSSI literature include natural log (e.g., Midkiff, Lindsey, & Meadows, 2018; You, Zheng, Lin, & Leung, 2016), square-root (e.g., Boone & Brausch, 2016; Robertson, Miskey, Mitchell, & Nelson-Gray, 2013), and inverse (e.g., Andover & Gibb, 2010; Buser, Pertuit, & Muller, 2019) approaches, with researchers typically adding a value of one to all scores to remove zero-values that cannot be transformed using natural log or inverse transformation approaches. Despite these efforts, NSSI data often remain highly skewed after transformation approaches.

In light of the prevalence of NSSI, as well as its clinical relevance as a marker of significant psychological distress and impairment (Benjet et al., 2017; Klonsky & Olino, 2008), research that uses appropriate methods to address the highly skewed nature of NSSI data is needed to evaluate NSSI predictors that have been suggested in prior studies. The etiologic contribution of child maltreatment experiences to later NSSI has received consistent empirical support in previous studies (e.g., Kaess et al., 2013; Lang & Sharma-Patel, 2011; Liu, Scopelliti, Pittman, & Zamora, 2018; Serafini et al., 2017). However, as researchers continue to struggle with best analytic practices to adequately address the distributional skewness of NSSI, this investigation compared approaches to address the expected skew in NSSI outcomes by conducting a series of regression analyses to evaluate previously-documented associations between child maltreatment and NSSI.

Contemporary Approaches to Address Distributional Violations

Data skewness refers to a lack of symmetry in the data distribution (i.e., one tail being longer than the other), which biases the estimation of parameters in regression models (e.g., Abbott & Gutgesell, 1994; Godfrey & Orme, 1991). Skewness estimates capture the degree of this asymmetry with positive skewness reflecting an over-abundance of low scores and a waning tail in the positive direction, and negative skewness reflecting an excess of high scores and a waxing tail from the negative direction (MacGillivray, 1986). As a result of their large data spreads (i.e., overdispersion), skewed distributions tend to feature larger variances than mean values. In turn, traditional regression models, which are fit by minimizing the sum of squared residuals, will lead to inaccurate parameter estimates and confidence intervals. In a series of simulated comparisons, Curran, West, and Finch (1996) found that skewness beyond a maximum level of 2 yielded unreliable parameter estimates. However, recent researchers have suggested even more conservative limits (e.g., skewness <1; Driscoll et al., 2019; Morse, Klingman, Jacob, & Kodali, 2019) and encouraged the application of statistical approaches to mitigate the negative impact of data skewness on parameter estimation.

Data skewness approaches share the common goal of supporting the reliable estimation of statistical parameters when observed data violate distributional assumptions of normality. Transformation approaches attempt to reduce the level of skewness, while other approaches, such as Poisson regression and the skew-*t* distribution, directly model skewness in the data when estimating regression parameters.

Variable Transformation Approaches

Transforming observed data is the most common approach to address skewness (Hall, 1992), and has been widely used in prior research on NSSI (e.g., Boone & Brausch, 2016; Buser et al., 2019; Midkiff et al., 2018). There are multiple forms of nonlinear and linear transformation approaches, such as *natural log, square-roo*t, and *inverse* transformations. Moreover, multiple transformations can be used to obtain skewness values nearest to zero with the fewest outliers. Natural log, square-root, and inverse transformations are used when the observed variable is characterized by moderate, substantial, and severe skewness, respectively (Tabachnick & Fidell, 2001). Although transformation approaches can complicate the interpretation of parameter estimates because they change the scaling of the variable, reversing the transformation before interpretation can ameliorate this problem. Transformation approaches are also valuable because they support standardized model fit indices (e.g., RMSEA, SRMR) and effect size estimates (e.g., Cohen's f^2). Thus, Tabachnick and Fidell (2001) recommend considering transformation in all skewness situations, unless there is a specific reason not to, such as when using non-parametric analyses.

Zero-Inflated Approaches

Zero-inflated models can be used to analyze data when there is an overabundance of zeros, which is the case for many clinical phenomena, and have been used in prior studies of NSSI (e.g., Allen, Fox, Schatten, & Hooley, 2019; Esposito, Bacchini, & Affuso, 2019; Kranzler et al., 2018; Vergara, Stewart, Cosby, Lincoln, & Auerbach, 2019). These models adopt a dual-process conceptualization of psychopathology wherein one process (or set of processes) is thought to predict disorder onset and a second process (or set of processes) is thought to account for disorder maintenance and/or severity (Lewinsohn, Allen, Seeley, & Gotlib, 1999; Yates, 2009). Thus, zero-inflated models yield two parameters, one that estimates the probability of whether or not an individual endorses symptomatology (e.g., 0 for non-injurers and 1 for injurers) using a logistic regression, and a second that estimates the degree to which the person expresses few or multiple symptoms (e.g., how many times a person has repeated NSSI) using either a Poisson or a negative binomial distribution.

Both zero-inflated Poisson (ZIP; Lambert, 1992) and negative binomial zero-inflated (NBZI; Minami, Lennert-Cody, Gao, & Román-Verdesoto, 2007) regression models evaluate covariates separately, yet simultaneously, to predict a) whether or not an individual expresses a symptom or engages in a behavior, and b) how many symptoms or how often behaviors are expressed. A ZIP model uses the Poisson distribution to account for the excess of zeros (Lambert, 1992), and to address some of the variation resulting from overdispersion in the outcome variable (i.e., when the variance of the variable is greater than its mean value). In contrast, the NBZI model pairs a Poisson approach with a multiplicative random effect to model overdispersion in the outcome variable (Gschlößl & Czado, 2008). The advantage of the NBZI model is that it can account for both the overabundance of zeros *and* for a greater level of overdispersion in the outcome than the ZIP model (Minami et al., 2007). Thus, when the outcome is overdispersed, the confidence intervals from the NBZI model are likely to be narrower than those from a ZIP model (Long, 1997). Despite their strengths, zero-inflated approaches require higher statistical power to detect effects than linear regressions (Dufour & Zung, 2005), and standardized model fit indices and estimates of effect sizes developed for linear regression analysis are not available (Muthén & Muthén, 1998–2017). Finally, the Monte Carlo integration required to estimate missing data in a Poisson-distributed outcome variable (Muthén & Muthén, 1998–2017), limits comparisons with approaches that use other missing data estimation methods (e.g., full-information maximum-likelihood).

Skew-t Distribution Approach

Although Azzalini (1985) introduced novel approaches to address moderately skewed data (e.g., the skew-normal distribution) and highly skewed data (e.g., the skew-*t* distribution), analytic software has only recently advanced to the point that these techniques can be carried out in a reasonable time frame. Thus, the current investigation is the first to our knowledge to apply this approach in the modeling of NSSI.

The skew-*t* distribution accounts for skewness (and kurtosis) by including a skew scale or shaping parameter with each variable and a degree of freedom parameter for the entire multivariate distribution when calculating the regression estimates (Lee & McLachlan, 2013). Theoretically, the multivariate skew-*t* distribution will equal the *t* distribution when there is no skewness in the data, and the degrees of freedom approach infinity (Asparouhov & Muthén, 2015). The most common method of parameter estimation using this distribution is maximum likelihood (Asparouhov & Muthén, 2015). A significant advantage of the skew-*t* approach is that skewness is modeled using predictors rather than removed by transforming variables (Muthén & Muthén, 1998–2017).

Summary

Individual approaches for dealing with data skewness have been presented in statistical journals, which has limited their uptake among applied researchers (e.g., Adarabioyo & Ipinyomi, 2019; Lee, Karrison, Nocon, & Huang, 2018; Tüzen & Erbaş, 2018). Thus, the current investigation sought to illustrate and compare these approaches as applied to a contemporary clinical issue, namely NSSI. As reviewed earlier, studies demonstrate that NSSI data are highly skewed such that most people do not engage in NSSI, a sizable minority of individuals have on at least one occasion, and a smaller number have done so on many occasions (Benjet et al., 2017; Gillies et al., 2018). Researchers studying NSSI have employed different statistical techniques to deal with the skewness of NSSI, including transformation approaches (e.g., Boone & Brausch, 2016; Buser et al., 2019; Midkiff et al., 2018) and zero-inflated approaches that address the overabundance of non-injurers in most samples using either Poisson (e.g., Esposito et al., 2019; Fox et al., 2018; Kranzler et al., 2018; Yates, Tracy, et al., 2008; You & Leung, 2012) or negative binomial distributions (e.g., Allen et al., 2019; Glenn, Kleiman, Cha, Nock, & Prinstein, 2016; Schoenleber, Berenbaum, & Motl, 2014; Vergara et al., 2019). However, clinical researchers generally, and those focused on NSSI in particular, have not compared different approaches for dealing with distributional violations of the normality assumption. Thus, the overarching goal of this investigation was to provide a novel comparison of multiple statistical approaches for addressing the distributional features of NSSI as applied to the evaluation of theoretically- and empirically- supported relations between child maltreatment and NSSI.

The Current Study

This investigation evaluated raw, transformation, zero-inflated, and skew-*t* distribution methods for dealing with skewed data when examining the relation between college students' self-reported histories of child maltreatment and NSSI. We compared multiple approaches for addressing the skewed distributional features of NSSI, including a) a natural log transformation of NSSI, b) a square-root transformation of NSSI, c) an inverse transformation of NSSI, d) a ZIP regression model, e) a NBZI regression model, and f) a skew-*t* distribution. Each approach for addressing the skewness of NSSI was evaluated with reference to the regression model results using observed (i.e., raw) NSSI data. Given the aforementioned variability in missing data estimation constraints across these varied approaches (e.g., Monte Carlo integration vs. full-information maximum-likelihood), predicted values were correlated with the raw NSSI variable to assess which approach most appropriately captured the expression of NSSI in this large and diverse college-student sample. Additional comparisons across effect size estimates (when available) and residual value distributions were also considered in evaluating these approaches.

All analyses controlled for youth gender, age, and ethnicity-race. Relative fit indices were evaluated to select the best-fitting model within each approach. This model was used to estimate predicted values of NSSI, which were then correlated with the raw NSSI scores and used to calculate residual values for further comparison. All predicted NSSI outcomes from these approaches were expected to correlate more strongly with the observed NSSI variable than with estimates using the raw data. However, we expected that zero-inflated models would yield the most accurate NSSI estimates because they can address the abundance of zeros that are expected in NSSI data.

METHOD

Participants

The present sample included 2,651 undergraduate students (M_{age} = 19.005, *SD* = 1.196) at a large University in the Western United States. The sample was predominantly female (64.2%), and nearly half (46.8%) the respondents were of Asian descent, with the remaining students identifying as Hispanic (29.0%), White (14.8%), Black (6.5%), or Multiracial/Other (3%). Unfortunately, data regarding specific countries of origin were not available.

Procedures

Participants were recruited from introductory psychology courses and received class credit in exchange for their completion of a 2-hour computerized survey. Informed consent was administered verbally and in writing, including specific reference to the potentially triggering nature of the questions posed, as well as confidentiality limitations pertinent to mandated reporting concerns. Participants completed the survey in small groups of up to 14 students using private cubicles in a laboratory setting under the supervision of a trained research assistant. Password-protected surveys were administered through a computerized survey management company. Responses were carefully monitored for completeness and the information was encrypted (until download) and identified by a code number to further ensure data security. Participants were informed that the purpose of the study was to examine relations between adaptation in young adulthood and various experiences in childhood and adolescence, including difficult life events. They were assured their participation was anonymous and no personally identifying information was collected. Participants were required to stay for the full 2-hour survey period to minimize the incentive to rush through the questionnaires. All participants received printed information and resource referrals for both campus- and community-based health and mental health support services. These procedures were reviewed and approved by the Human Research Review Board of the participating university.

Measures

Non-Suicidal Self-Injury (NSSI)

Participants completed the Functional Assessment of Self-Mutilation (FASM; Lloyd, Kelley, & Hope, 1997) to assess rates and methods of NSSI during the 12 months preceding the time of data collection. The reliability and validity of the FASM has been established across several studies (e.g., Guertin, Lloyd-Richardson, Spirito, Donaldson, & Boergers, 2001; Nock & Prinstein, 2004, 2005). Participants reported whether and how often they engaged in each of 11 different forms of NSSI behavior, including cutting or carving skin, picking at a wound, self-hitting, scraping skin to bleed, self-biting, picking areas of the body to bleed, inserting objects under skin or nails, self-tattooing, burning skin, pulling out hair, or erasing skin to bleed in the past 12 months. Participants rated the frequency of each form of NSSI on a 7-point scale that ranged across 0 (*0 times*), 1 (*1 time*), 2 (*2 times*), 3 (*3–5 times*), 4 (*6–10 times*), 5 (*11–20 times*), and 6 (*>20 times*). For these analyses, NSSI was operationally defined as the sum of cutting and burning behaviors, which is consistent with prior studies and recent measures of NSSI (e.g., the Self-Injurious Thoughts and Behaviors Interview; Nock, Holmberg, Photos, & Michel, 2007) that emphasize cutting and burning because they are most clearly identified as NSSI, are most commonly endorsed, and involve direct tissue damage.

TABLE 1
Means, Standard Deviations, Skewness, and Kurtosis for Study Variables

	N	M	SD	*Skewness*	*Kurtosis*
NSSI (raw)	2524	1.782	2.145	3.228	11.109
NSSI (square-root transformed)	2523	0.469	0.110	2.894	−0.595
NSSI (natural log transformed)	2523	0.206	0.454	2.366	4.367
NSSI (inverse transformed)	2523	0.723	0.077	2.128	2.907
Child Maltreatment	2511	1.973	0.498	0.928	0.531
Age (raw)	2561	19.00	1.192	1.256	1.045
Age (square-root transformed)	2561	4.471	0.131	1.196	0.834

Note. NSSI = Non-Suicidal Self-Injury.

Child Maltreatment

The Child Abuse and Trauma Scale (CATS) includes 38 items that assess the frequency and extent of child maltreatment experiences (Sanders & Becker-Lausen, 1995; e.g., “Did your parents ever hit or beat you when you did not expect it?” “Did your relationship with your parents ever involve a sexual experience?”). Items were rated on a five-point scale from 0 (*never*) to 4 (*always*) and composited to yield a total maltreatment score (α = .916). The CATS evidences concurrent validity with objective measures of child maltreatment and acceptable reliability in college student populations (Sanders & Becker-Lausen, 1995).

Data Analytic Plan

Analysis Software

Data preparation was conducted utilizing IMB SPSS 26 to evaluate descriptive statistics, perform transformations (i.e., square-root, natural log, and inverse), and evaluate Little’s MCAR test (1988). All regression models were evaluated using MPlus 8.3 (Muthén & Muthén, 1998–2017).

Data Preparation

Prior to analyses, all predictors and continuous covariates were evaluated for violations of normality assumptions. Participant age was square-root transformed, since it was the only exogenous variable that surpassed conservative recommended levels of skewness (i.e., >1; Tabachnick & Fidell, 2001).[1] Table 1 depicts pre- and post-transformation estimates of participant’s age and NSSI, as well as skewness, means, and standard deviations for all study variables. Outliers were evaluated using a multiple regression to obtain Mahalanobis distance scores.[2] Six scores with $p < .001$, which were derived from leverage scores, emerged as multivariate outliers and were deleted from all analyses.[3]

Data Analysis and Missingness

Data were missing on age (3.2%), child maltreatment (5.1%), and NSSI (4.7%). Missing data were handled using full-information maximum likelihood (FIML) estimation as supported by Little’s (1988) test, χ^2 (27) = 27.409, $p = .442$, which indicated the data met criteria for missing completely at random (MCAR). A multivariate analysis of variance (MANOVA) evaluated mean differences in age, child maltreatment, and NSSI by participant gender, ethnicity-race, and their interaction.

Model Evaluation

All variables were centered to reduce multicollinearity, since saturated models include all predictors and the interactions among covariates. Bias-corrected bootstrapped confidence intervals (CIs) across 10,000 resamples evaluated relations between child maltreatment and NSSI while holding participant age (square-root transformed), gender, and ethnicity-race constant (MacKinnon, Lockwood, & Williams, 2004).

Model Fitting Approach 1. Model fit was assessed sequentially using the criteria set forth by Hu and Bentler (1999) as informed by model fit chi-squares, the comparative fit index (CFI > .90), the Tucker-Lewis Index (TLI > .90), the standardized root mean square residual (SRMR < .08), and the root mean square error of approximation (RMSEA < .08).

[1] All model approaches were also evaluated using the raw age variable. Across all approaches, the age-transformed models evidenced better fit. However, there were minimal differences across raw and transformed analyses with regard to predicted relations with the raw NSSI outcome scores.

[2] Alternatively, following Tabachnick and Fidell (2001), outliers could be re-assessed after transformations of age to evaluate which transformation recommends the fewest outlier removals. However, to maintain consistency across approaches and to remain conservative, all outliers were removed prior to transformation.

[3] Analyses including outliers revealed no major differences in findings. Probability-probability and residual plots for assessing outliers are available upon request.

Failure to meet criteria on one or more fit indices was indicative of poor fit.

Model Fitting Approach 2. Information measures assessed model fit when the above fit indices were unavailable. These included Akaike's information criterion (AIC), the Bayesian information criterion (BIC), which allows for comparison of nested and non-nested models (Long, 1997), and the sample size adjusted BIC (BIC_{SSA}). Table 2 summarizes the resampling technique, missing data estimation, and model fitting for all model approaches.

Model Comparisons

Traditional methods (e.g., chi-square differences) were not available to compare all methods; therefore, outputted predicted values were correlated against the raw NSSI values to assess the predictive strength of each approach, and to evaluate the distributional pattern of residual values.

Model Approaches

The regression model evaluated the effect of child maltreatment (i.e., CATS Scores) on NSSI, while holding age (square-root transformed), gender, and ethnicity-race constant. All models included interactions among covariates (not shown) and used the same general equation:

$$\widehat{NSSI} = \beta_0 + \beta_1 CATS + \beta_2\, SQRT\_AGE + \beta_3\, GENDER + \beta_4 RACE$$

Raw Regression Approach (A). The raw NSSI variable was regressed on the predictors to evaluate the correlation of predicted values of NSSI to actual NSSI.

Transformation Approaches (B). NSSI was transformed with the natural log (i.e., ln[NSSI + 1], Model B.1), square-root (i.e., SQRT[NSSI + 1], Model B.2), and inverse (i.e., 1/NSSI + 1, Model B.3) approaches. The inverse transformation was then reversed to preserve the direction of the original scale. See Table 1 for pre- and post-transformation values of NSSI.

TABLE 2
Resampling, Missing Data, and Model Fitting by Model Approaches

Model Approach	*Model Reference*	*Resamples*	*Missing Data*	*Model Fitting Approach*
Raw	A	Bootstraps	ML	1
Transformed	B	Bootstraps	ML	1
ZIP	C.1	Montecarlo	MLR	2
NBZI	C.2	Bootstraps	ML	1
SKEW-T	D	Bootstraps	ML	2

Note. All approaches used 10,000 resamples.

Zero-Inflated Regression Model Approaches (C). The ZIP (Model C.1) and NBZI (Model C.2) regression approaches used the same general equation to evaluate the likelihood of NSSI endorsement $(\widehat{NSSI_{\#1}})$ and the frequency of NSSI endorsed $(\widehat{NSSI_0})$. The ZIP model used a logit and Poisson distribution for NSSI, and the NBZI model used a negative binomial distribution for NSSI.

Skew-t Distribution Approach (D). The skew-*t* distribution approach adds a skew scaling weight when estimating regression parameters using the same general equation to evaluate NSSI.

RESULTS

Descriptive & Bivariate Analyses

Out of the 2,651 participants, 16.3% reported having engaged in one or more episodes of NSSI during the past 12 months. A multivariate analysis of variance (MANOVA) evaluated mean differences in age, child maltreatment, and NSSI as a function of participant gender, ethnicity-race, and their interaction (See Table 3). There were significant main effects for participant gender (Wilks' $\lambda = .985$, $p < .001$) and ethnicity-race (Wilks' $\lambda = .970$, $p < .001$), as well as for their interaction (Wilks' $\lambda = .987$, $p < .001$). Between-subjects effects revealed significant gender differences in participant age, child maltreatment, and NSSI, with males being older than females, and females endorsing higher rates of child maltreatment and NSSI than males. Likewise, there were significant differences in all study variables across ethnic-racial groups. Bonferonni-corrected post-hoc tests indicated that White participants were older than Hispanic respondents, Hispanic and multiracial participants endorsed higher rates of child maltreatment than both White and Black respondents, and Asian participants reported lower levels of NSSI than White, Hispanic, and multiracial respondents. There was a significant gender*ethnicity-race interaction for maltreatment, such that females endorsed higher rates of maltreatment in all ethnic-racial groups, except among Black respondents where males endorsed higher rates than females. There was also a significant gender*ethnicity-race interaction for NSSI, such that females endorsed higher rates of NSSI than males among all ethnic-racial groups, except among Whites where rates did not differ significantly by gender. At the bivariate level, NSSI was positively related to child maltreatment ($r = .239$, $p < .001$), but age was not significantly related to child maltreatment or NSSI.

Model Fitting

Tables 4 and 5 show successive model fitting indices for all models. Estimates for both approaches were presented when available. Across all models, age tended to evidence

TABLE 3
Descriptive Data for Study Variables by Participant Gender and Ethnicity-Race

	Participant Gender		*Participant Ethnicity-Race*				
	Males M (SD)	*Females* M (SD)	*White* M (SD)	*Black* M (SD)	*Hispanic* M (SD)	*Asian* M (SD)	*Multi* M (SD)
N	942	1,693	379	166	741	1,197	76
Age	19.140[1]	18.910	19.160[a]	18.840	18.890	19.030	18.860
	(1.286)	(1.140)	(1.311)	(1.095)	(1.193)	(1.174)	(1.261)
Maltreatment	1.927[1]	1.998	1.888	2.006	1.925	2.010[ab]	2.146[ab]
	(0.449)	(.527)	(.510)	(.517)	(.498)	(.488)	(.572)
NSSI	1.428[1]	1.973	2.141	1.574	1.895	1.578[abc]	2.369
	(1.674)	(2.394)	(2.844)	(1.535)	(2.432)	(1.741)	(2.770)

Note: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$. [1]Significiant difference from female. [a]Significant difference from Hispanic. [b]Significant difference from White. [b]Significant difference from Multi. *N*s do not sum to 2,651 due to selective non-response for gender (n = 16) and ethnicity-race (n = 91).

nonsignificant effects and covariances. Models ending in 0 represent saturated models (i.e., all predictors and correlations among covariates are included) and models ending in F represent final best fitting models. The best fitting model for each approach was used to output predicted NSSI values.

Regression Model Analyses

A regression model evaluated the effect of child maltreatment on NSSI while holding age (square-root transformed), gender, and ethnicity-race constant. Table 6 provides estimates, standardized estimates, and model fit indices for the final model (A) of the raw data approach. The raw data regression model evidenced good fit (χ^2 [3] = 3.693, p = .297; RMSEA = .009, 90% CI [.000–.035], p = .998; CFI = .997; TLI = .993; SRMR = .009), and child maltreatment was related to higher NSSI (b = 1.069, β = .247, p < .001).

Transformation Approaches (Models B.1-B.3)

Table 6 presents estimates, standardized estimates, and model fit indices for the final model of each transformation approach.

Natural Log Transformation (B.1). The natural log NSSI regression model evidenced good fit (χ^2 [3] = 3.562, p = .313; RMSEA = .009, 90% CI [.000–.035], p = .998; CFI = .998; TLI = .994; SRMR = .009), and child maltreatment was related to higher NSSI (b = .847, β = 1.865, p < .001).

Square-Root Transformation (B.2). The square-root NSSI regression model evidenced good fit (χ^2 [3] = 3.618, p = .306; RMSEA = .009, 90% CI [.000–.035], p = .998; CFI = .997; TLI = .994; SRMR = .009), and child maltreatment was related to higher NSSI (b = .461, β = 4.247, p < .001).

Inverse Transformation (B.3). The inverse NSSI regression model evidenced good fit (χ^2 [3] = 3.614, p = .306; RMSEA = .009, 90% CI [.000–.035], p = .998; CFI = .997; TLI = .994; SRMR = .009), and child maltreatment was related to higher NSSI (b = .717, β = 9.159, p < .001).

Zero-Inflated Approaches (Models C.1 and C.2)

Table 7 provides estimates, standardized estimates, and model fit indices for the final model of the zero-inflated approaches.

Zero-Inflated Poisson Regression (ZIP) Approach (C.1). The ZIP regression model indicated that child maltreatment was associated with a decreased probability of endorsing NSSI (b#1 = − .499, β#1 = − .082, p < .001) and with greater frequency of NSSI when it was endorsed (b = .542, β = .851, p < .001).

Negative Binomial Zero-Inflated (NBZI) Approach (C.2). The NBZI regression model indicated that child maltreatment was associated with an increased probability of endorsing NSSI (b#1 = 2.046, β#1 = .126, p < .001) and with greater frequency of NSSI when it was endorsed (b = .529, β = .846, p < .001).

Skew-t Distribution Approach (D)

The Skew-*t* regression model indicated that child maltreatment was related to lower NSSI (b = − .007, β = − .005, p < .001; see Table 7).

Model Comparisons

Effect size estimates for the raw and transformation regression models can be found in Table 6. Across

TABLE 4
Comparison of Model Fit Indices of Raw and Transformation Approaches with 10,000 Bootstrap/Montecarlo Resamples

Models		*FP*	*LL-Base*	*LL Model*	*−2LL*	p	*DF SAT*	*−2LL SAT*	p *SAT*	*AIC*	*BIC*	*Adjusted BIC*	χ2	χ2p	*RMSEA*	*CI (90%)*	*RMSEA* p < *.05*	*CFI*	*TLI*	*SRMR*
Raw Models																				
A.0	Raw_SAT_SAge	20	−9300.006	−10953.180						18640.011	18757.620	18694.074								
A.1	rm SAge w MALTX	19	−9300.006	−10953.487	0.656	0.418	8	0.656	1.000	18638.667	18750.395	18690.027	0.656	0.418	0.000	0.000 0.048	0.961	1.000	1.007	0.004
A.2	rm SAge w ASIAN	18	−9300.006	−10954.943	2.932	0.087	9	3.588	0.936	18639.600	18745.448	18688.256	3.589	0.166	0.017	0.000 0.046	0.973	0.992	0.984	0.009
A.3-F	rm NSSI on SAge	17	−9300.006	−10955.099	0.104	0.747	10	3.692	0.960	18637.704	18737.671	18683.657	3.693	0.297	0.009	0.000 0.035	0.998	0.997	0.993	0.009
A.4	add SAge w ASIAN	18	−9300.006	−10953.644	3.692	0.055	9	0.000	1.000	18636.773	18742.621	18685.430	0.762	0.683	0.000	0.000 0.029	0.999	1.000	1.020	0.004
Transformed Models																				
B.1.0	Nlog_SAT_SAge	20	−5364.031	−5364.031						10768.062	10885.671	10822.125								
B.1.1	rm SAge w MALTX	19	−5364.031	−5364.342	0.622	0.430	8	0.622	1.000	10766.684	10878.412	10818.043	0.621	0.431	0.000	0.000 0.047	0.963	1.000	1.007	0.004
B.1.2	rm SAge w ASIAN	18	−5364.031	−5365.810	2.936	0.087	9	3.558	0.938	10767.620	10873.468	10816.277	3.558	0.169	0.017	0.000 0.046	0.974	0.993	0.985	0.009
B.1.3-F	rm NSSI on SAge	17	−5364.031	−5365.812	0.004	0.950	10	3.562	0.965	10765.625	10865.592	10811.578	3.562	0.313	0.008	0.000 0.035	0.998	0.999	0.994	0.009
B.1.4	add SAge w ASIAN	18	−5364.031	−5364.344	2.936	0.087	9	0.626	1.000	10764.688	10870.536	10813.344	0.626	0.731	0.000	0.000 0.027	0.999	1.000	1.021	0.004
B.2.0	SQRT_SAT_SAge	20	−1751.211	−1751.211						3542.422	3660.031	3596.485								
B.2.1	rm SAge w MALTX	19	−1751.211	−1751.518	0.614	0.433	8	0.614	1.000	3541.035	3652.763	3592.395	0.613	0.434	0.000	0.000 0.047	0.964	1.000	1.007	0.004
B.2.2	rm SAge w ASIAN	18	−1751.211	−1752.987	2.938	0.087	9	3.552	0.938	3541.974	3647.822	3590.630	3.552	0.169	0.017	0.000 0.046	0.974	0.993	0.985	0.009
B.2.3-F	rm NSSI on SAge	17	−1751.211	−1753.020	0.066	0.797	10	3.618	0.963	3540.040	3640.008	3585.993	3.618	0.306	0.009	0.000 0.035	0.998	0.997	0.994	0.009
B.2.4	add SAge w ASIAN	18	−1751.211	−1751.550	2.940	0.086	9	0.678	1.000	3539.101	3644.949	3587.757	0.678	0.712	0.000	0.000 0.028	0.999	1.000	1.020	0.004
B.3.0	INV2_SAT_SAge	20	−928.440	−928.440						1896.879	2014.488	1950.942								
B.3.1	rm SAge w MALTX	19	−928.440	−928.745	0.610	0.435	8	0.610	1.000	1895.491	2007.219	1946.850	0.611	0.434	0.000	0.000 0.047	0.966	1.000	1.008	0.004
B.3.2	rm SAge w ASIAN	18	−928.440	−930.215	2.940	0.086	9	3.550	0.938	1896.429	2002.277	1945.086	3.550	0.170	0.017	0.000 0.046	0.974	0.992	0.985	0.009
B.3.3-F	rm NSSI on SAge	17	−928.440	−930.246	0.062	0.803	10	3.612	0.963	1894.493	1994.460	1940.446	3.614	0.306	0.009	0.000 0.035	0.998	0.997	0.994	0.009
B.3.4	add SAge w ASIAN	18	−928.440	−928.777	2.938	0.087	9	0.674	1.000	1893.554	1999.401	1942.210	0.674	0.714	0.000	0.000 0.028	0.999	1.000	1.020	0.004

Note. F = Final Model; FP = Free Parameters; Adjusted BIC [(n* = (n + 2)/24)]; rm = Removed; SAT = Saturated Model; SAge = Square Root Age

TABLE 5
Comparison of Model Fit Indices of Zero-Inflated and Skew-*t* Distribution Approaches with 10,000 Bootstrap/Montecarlo Resamples

Models		*FP*	*LL-Base*	*LL Model*	*−2LL*	p	*DF SAT*	*−2LL SAT*	p *SAT*	*AIC*	*BIC*	*Adjusted BIC*
Zero-Inflated Models												
C.1.0	ZIP_SAT_SAge	24	−8378.829	−8378.829						16805.658	16946.788	16870.533
C.1.1	rm SAge w MALTX	23	−8378.829	−8379.143	0.628	0.428	9	0.628	1.000	16804.287	16939.536	16866.459
C.1.2	rm SAge w ASIAN	22	−8378.829	−8380.622	2.958	0.085	10	3.586	0.964	16805.245	16934.614	16864.714
C.1.3-F	rm NSSI on SAge	21	−8378.829	−8380.671	0.098	0.754	11	3.684	0.978	16803.341	16926.83	16,860.107
C.1.4	rm SQRTAge w SEX	20	−8378.829	−8392.169	22.996	0.000	12	26.68	0.009	16824.338	16941.946	16878.4
C.1.0	NBZI_SAT_SAge	25	−8110.36	−8110.36						16270.719	16417.73	16338.298
C.2.1	rm SAge w MALTX	24	−8110.36	−8110.673	0.626	0.429	9	0.626	1.000	16269.346	16210.476	16334.221
C.2.2	rm SAge w ASIAN	23	−8110.36	−8112.111	2.876	0.090	10	3.502	0.967	16270.221	16405.471	16332.393
C.2.3-F	rm NSSI on SAge	22	−8110.36	−8112.213	0.204	0.652	11	3.706	0.978	16268.427	16397.796	16327.896
C.2.4	rm SEX w MALTX	21	−8110.36	−8117.48	10.534	0.001	12	14.24	0.286	16276.96	16400.449	16333.726
Skew-*t* Distribution Models												
D.0	SKEWT_SAT_SAge	26	−7246.922	−7246.922						14545.845	14698.736	14698.739
D.1	rm SEX w MALTX	25	−7246.922	−7246.923	0.002	0.964	9	0.002	1.000	14543.846	14690.857	14611.424
D.2	rm SAge w CATS	24	−7246.922	−7248.611	3.376	0.066	10	3.378	0.971	1454.221	14686.351	14610.096
D.3-F	rm NSSI on ASIAN	23	−7246.922	−7248.645	0.068	0.794	11	3.446	0.983	14543.29	14678.54	14605.462
D.4*	rm SAge w ASIAN	22	−7246.922	−7254.187	11.084	0.001	12	14.53	0.268	14552.374	14681.743	14611.843

Note. F = Final Model; FP = Free Parameters; Adjusted BIC [(n* = (n + 2)/24)]; rm = Removed; SAT = Saturated Model; SAge = Square Root Age

these models, the relation between child maltreatment and NSSI was generally moderate in magnitude (Cohen, 1988) suggesting that, consistent with prior work (e.g., Kaess et al., 2013; Lang & Sharma-Patel, 2011; Liu et al., 2018; Serafini et al., 2017), child maltreatment emerged as a robust predictor of NSSI in this sample. The skewness and kurtosis of residual values for each of the models are reported in Tables 6 and 7. Although there was significant non-normality across residuals for all models, as confirmed by significant Shapiro-Wilk Normality Test estimates for all models, the zero-inflated models evidenced the lowest levels of skew and kurtosis, which suggests these approaches were best able to address the distributional features of NSSI.

TABLE 6
Estimates, Standardized Estimates, and Model Fit Indices for the Final Models of Raw and Transformation Approaches

	Raw.NSSI		*Nlog.NSSI*		*SQRT.NSSI*		*INV.NSSI*	
Estimate	*B*	*β*	*B*	*β*	*B*	*β*	*B*	*β*
NSSI Intercept	1.682***	0.780***	0.847***	1.865***	0.461***	4.247***	0.717***	9.159***
NSSI on SEX	0.429***	0.095***	0.098***	0.103***	0.024***	0.106***	0.017***	0.106***
NSSI on ASIAN	−0.368***	−0.085***	−0.069***	−0.075***	−0.015***	−0.070***	−0.011***	−0.070***
NSSI on MALTX	1.069***	0.247***	0.231***	0.254***	0.055***	0.253***	0.040***	0.252***
Chi-square: (DF) Value	(3) 3.693, p = .297		(3) 3.562, p = .313		(3) 3.618, p = .306		(3) 3.614, p = .306	
CFI	0.997		0.998		0.997		0.997	
TLI	0.993		0.994		0.994		0.994	
H0 Loglikelihood	−9301.852		−5365.812		−1753.02		−930.246	
H1 Loglikelihood	−9300.006		−5364.031		−1751.211		−928.44	
No of Free Parameters	17		17		17		17	
Akaike (AIC)	18637.704		10765.625		3540.04		1894.493	
Bayesian (BIC)	18737.671		10865.592		3640.008		1994.46	
Sample-Size Adjusted BIC	18683.657		10811.578		3585.993		1940.446	
RMSEA: Estimate	0.009		0.008		0.009		0.009	
RMSEA: 90% CI	0.00 0.035		0.00 0.035		0.00 0.035		0.00 0.035	
RMSEA: Probability	0.998		0.998		0.998		0.998	
SRMR	0.009		0.009		0.009		0.009	
R^2	0.080***		0.084***		0.083***		0.083***	
Cohen's f 2	0.087		0.092		0.091		0.091	
Residuals: Skewness	3.248		3.248		3.248		3.248	
Residuals: Kurtosis	11.362		11.363		11.363		11.363	

Note. ***p < .001

TABLE 7
Estimates, Standardized Estimates, and Model Fit Indices for the Final Models of the Zero-Inflated and Skew-*t* Distribution Approaches

	ZIP		*NBZI*			*SKEW*-t	
	B	*β*	*B*	*β*		*B*	*β*
NSSI#1 Intercept	−17.689***	−5.787***	−64.788	−7.990***			
NSSI intercept	0.457***	1.437***	0.454***	1.456***	NSSI intercept	0.959***	1.097***
NSSI on SEX	0.249***	0.375***	0.252***	0.388***			
NSSI on ASIAN	−0.198***	−0.310***	−0.189***	−0.302***			
NSSI on MALTX	0.542***	0.851***	0.529***	0.846***	NSSI on MALTX	−0.007***	−0.005***
NSSI#1 on SAge	−0.382	−0.016***	−2.755	−0.045***	NSSI on SAge	−0.035***	−0.008***
NSSI#1 on SEX	−5.121	−0.803***	−16.457	−0.972***	NSSI on SEX	0.004***	0.004***
NSSI#1 on ASIAN	−0.294	−0.048***	−5.692	−0.349***			
NSSI#1 on MALTX	−0.499	−0.082***	2.046	0.126***			
H0 Loglikelihood	−8380.671		−8112.213		H0 Loglikelihood	−7248.645	
H0 Scaling Correction Factor for MLR	0.953		0.795				
No. of Free Parameters	21		22		No. of Free Parameters	23	
Akaike (AIC)	16803.341		16268.427		Akaike (AIC)	14543.29	
Bayesian (BIC)	16926.83		16397.796		Bayesian (BIC)	14678.54	
Sample-Size Adjusted BIC	16860.107		16327.896		Sample-Size Adjusted BIC	14605.462	
Condition Number	0		0				
Residuals: Skewness	−2.238		−2.199			2.785	
Residuals: Kurtosis	4.561		3.516			9.671	

Note. ***$p < .001$; #1 denotes models predicting membership to the zero generating process.

Table 8 shows the correlations of predicted values with observed values of NSSI across all models. Differences among correlations were evaluated using the *Fisher* Z transformation. The correlations between NSSI and the predicted values of NSSI across all transformation models did not differ significantly from the raw data regression model, and the transformation approaches did not differ significantly from each other (all r values = .272, $p < .001$; all *Fisher* Z values = .984). The ZIP (r = .833, $p < .001$) and NBZI (r = .832, $p < .001$) models evidenced the strongest relations between reported NSSI and predicted values and these relations significantly differed from all transformation approaches (both *Fisher Z* values = − 32.285), but not from each other (*Fisher Z* = .100). The skew-*t* distribution showed the weakest relation between NSSI and the predicted values (r = .246, $p < .001$). This value was lower than those obtained from the zero-inflated approaches (*Fisher Z* values = 33.190$_{_ZIP}$ and 33.080$_{_NBZI}$), but did not differ from the transformation models (all *Fisher Z* values = 0.999).

DISCUSSION

The present investigation sought to illustrate the application of varied statistical techniques for addressing highly skewed clinical data as applied to the evaluation of theoretically-specified and empirically-validated associations between child maltreatment and NSSI. As in prior studies (Guertin et al., 2001; Lloyd et al., 1997, April; Nock & Prinstein, 2005; Yates, Tracy, et al., 2008), NSSI emerged as a highly-skewed variable with an overabundance of zeros. Researchers have employed various procedures to address this distributional pattern, such as transforming the

TABLE 8
Correlations among Actual NSSI and Predicted Values Across all Approaches

	1	*2*	*3*	*4*	*5*	*6*	*7*
1. NSSI							
2. Raw NSSI β	0.271554						
3. Nlog.NSSI β	0.272029	0.999842					
4. SQRT.NSSI β	0.272135	0.999643	0.999954				
5. INV.NSSI β	0.272111	0.999647	0.999956	0.999997			
6. ZIP.NSSI β	0.832722	0.241859	0.241027	0.240516	0.240524		
7. NBZI.NSSI β	0.831812	0.262764	0.261954	0.261448	0.261444	0.999357	
8. SKEWT.NSSI β	0.245870	0.914541	0.920753	0.923841	0.923784	0.188694	0.209159

Note. All correlations are significant at $p < .001$.

variable to reach acceptable levels (e.g., Afifi, Kotlerman, Ettner, & Cowan, 2007; Andover & Gibb, 2010; Robertson et al., 2013; Yates, Carlson, et al., 2008; You et al., 2016), and, less often, using zero-inflated approaches to model the overabundance of zeros as distinct from the skewed frequency of NSSI using either Poisson (Yates, Tracy, et al., 2008) or negative binomial models (e.g., Allen et al., 2019; Glenn et al., 2016; Schoenleber et al., 2014).

Across raw and transformation approaches, child maltreatment evidenced moderately strong relations with higher NSSI. Likewise, both zero-inflated models yielded significant and positive relations between child maltreatment and NSSI frequency among respondents who endorsed NSSI. However, the ZIP model yielded a negative, albeit small, relation between child maltreatment and the probability of NSSI endorsement, whereas the NBZI model supported a positive relation between child maltreatment and the likelihood of endorsing NSSI. Similarly, the skew-*t* model showed a small, but significant negative relation between child maltreatment and NSSI. The negative associations between child maltreatment and NSSI obtained from the ZIP and skew-*t* model were surprising. Relative to the NBZI correction, which can account for clinical phenomena where there is an increase in the distribution of frequencies at the tail end of the distribution, the ZIP approach may be less able to accommodate subsequent spikes yielding the appearance of a negative association between child maltreatment and the likelihood of endorsing NSSI. With regard to the skew-*t* distribution approach, additional research is needed to apply this method with both zero-inflated and non-zero-inflated skewed data to further understand if and how an excess of zeros might alter the performance of the skew-*t* distribution approach.

Across approaches, the predicted values of the ZIP and NBZI models evidenced the strongest relations with raw NSSI scores, whereas the skew-*t* distribution model predicted values with the weakest correlation to the observed data. This pattern was echoed in our evaluation of the residual values for each of the approaches. The superiority of the zero-inflated approaches may indicate that accounting for the overabundance of zeros is most important when distributions feature this type of skewed data. However, as noted earlier, there may be risks to approaches that fail to fully correct for overdispersion of outcome data in the ZIP model. Likewise, the nonnormality of the residual values from the zero-inflated approaches points to the need for ongoing efforts to develop and evaluate statistical approaches for dealing with highly skewed data.

Strengths & Limitations

The current study drew on a strong foundation of theoretical and empirical literature, as well as a large and diverse sample of participants who provided real-world NSSI data with unique skewed features (i.e., zero-inflated and overdispersed), to provide clinical researchers with a comparative evaluation of different methods for addressing highly skewed data in their statistical analyses. However, despite these strengths, the current study featured several limitations that qualify the interpretation of the data while highlighting valuable directions for future research.

First, although NSSI is most prevalent among late adolescents and young adults (Whitlock et al., 2006), college students from a public university in the western United states may not be representative of college-going youth at private institutions or in other regions and countries. Likewise, student samples may not be representative of clinical, incarcerated, or broader community populations (e.g., working young adults who are not attending school). Indeed, in a sample of adolescent inpatients, Lüdtke, In-Albon, Michel, and Schmid (2016) did not find a significant association between child maltreatment and NSSI. Given the potential for variability across samples, future studies must evaluate these associations in other adolescent and young adult populations.

Second, and related to the first concern, the generalizability of the obtained findings regarding the relative merits of statistical approaches for addressing skewness await replication in future studies using varied conceptualizations of NSSI, as well as in studies of other clinical phenomena with high public health impact despite relatively low base rates. The current study assessed NSSI outcomes using frequency bands, rather than count data. Although numerous studies have used similar frequency bands (e.g., Buser et al., 2019; Hasking, Boyes, Finlay-Jones, McEvoy, & Rees, 2019; Midkiff et al., 2018; Whitlock et al., 2013), this approach may result in reduced power or conflate potentially different degrees of NSSI. That said, extant measures that employ count report formats may inadvertently bias reports to favor particular frequency ranges by suggesting count metrics (e.g., 1, 5, 15 in Cerutti, Zuffianò, & Spensieri, 2018; 0, 10, 100, 500 in Nielsen, Sayal, & Townsend, 2017). Similarly, studies of NSSI vary with regard to duration, with some studies assessing NSSI events during the prior year as in the current study (e.g., Hasking et al., 2019; Muehlenkamp, Xhunga, & Brausch, 2019), in the past month (e.g., Auerbach et al., 2014; Fox et al., 2019), or over the course of the respondent's life time (e.g., Cerutti et al., 2018). It is not yet clear if and how these variations may influence NSSI distributions and, by extension, the relative effectiveness of statistical approaches for addressing data skewness. In addition to evaluating the applicability of these approaches to different conceptualizations of NSSI, future research should test whether or not zero-inflated models yield superior predictions across multiple forms of clinical symptom expression, or whether these approaches are particularly effective for modeling NSSI. Of note, in the absence of additional samples and metrics, Monte Carlo simulation studies may support the evaluation of skewness approaches by comparing "estimated" parameters to "true" parameters (Feinberg & Rubright, 2016).

Third, although the current study evaluated several prominent approaches for addressing data skewness, additional models should be evaluated in future studies. Many clinical phenomena, including NSSI, can be modeled using multinomial regression approaches (e.g., by recoding NSSI frequency into a three-point ordinal scale; Yates, Carlson, & Egeland, 2008). In addition, semi-parametric regression-type models may overcome some of the distributional limitations of generalized linear models (Akantziliotou, Rigby, & Stasinopoulos, 2002; Rigby & Stasinopoulos, 2001, 2005). For example, Generalized Additive Models for Location, Scale, and Shape (GAMLSS) address skewness and kurtosis issues using non-parametric smoothing functions, though their application to questions about NSSI and other clinical phenomena awaits further evaluation.

Fourth, the current study relied on retrospective, self-report data collected at a single time point using a survey platform, which raises concerns about the validity of the data due to shared method variance, and constrains our capacity to render directional interpretations of the obtained associations. Retrospective reports of child maltreatment may be affected by false reporting (Fergusson, Horwood, & Woodward, 2000), errors in recollection (DiLillo et al., 2006), and inconsistency in and/or inability to access traumatic memories (Widom & Morris, 1997). However, it is important to note that underreporting (i.e., false negatives) is far more common than overreporting (i.e., false positives; Hardt & Rutter, 2004). Moreover, empirical studies support the convergent validity of retrospective self-reports of maltreatment with child welfare records using both administrative data and prospective indices of adjustment (Dube, Williamson, Thompson, Felitti, & Anda, 2004; Shaffer, Huston, & Egeland, 2008). Although retrospective reports capture the totality of participant's lived experiences, rather than just those associated with reported and reportable events (Hardt & Rutter, 2004; Shaffer et al., 2008), the current findings would have been bolstered by concurrent administrative data reports, particularly given that both maltreatment and NSSI data were provided by a single informant in this study.

Conclusions

The current findings suggest that the unique skewness characteristics of NSSI may render specific analytic methods inappropriate or less accurate, including the use of skew-*t* distribution models, even though these approaches have been successful in modeling skewed non-clinical data (e.g., Asparouhov & Muthén, 2015). Apparent differences in predictive power across the approaches examined here highlight the importance of using zero-inflated models, rather than transformation or skew-*t* approaches. Conceptually, zero-inflated approaches offer a critical advantage by modeling symptom initiation (e.g., probability of occurrence) versus maintenance (e.g., frequency of occurrence). Indeed, by offering distinct parameter estimates, these models enable researchers to evaluate predictors of symptom onset as distinct from those accounting for symptom maintenance or exacerbation, which is consistent with broader tenets of developmental psychopathology emphasizing the possibility that distinct factors account for disorder onset versus course (Lewinsohn et al., 1999; Yates, Burt, & Troy, 2011). Efforts to elucidate distinct developmental pathways toward disorder onset versus recurrence will inform more targeted and efficacious prevention and intervention treatments, respectively. For example, prior work has suggested that social and peer factors are salient for NSSI initiation, whereas intrapersonal factors tend to be more strongly associated with NSSI recurrence (Brausch & Muehlenkamp, 2018; Muehlenkamp, Brausch, Quigley, & Whitlock, 2013). Moving forward, expanded efforts to evaluate putative mediators of the relation between child maltreatment and NSSI using zero-inflated modeling approaches will further refine extant treatment approaches.

ACKNOWLEDGMENTS

We gratefully acknowledge Dr. Chandra Reynolds who provided generous instruction and consultation throughout the execution of the analyses reported herein.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

REFERENCES

Abbott, R. D., & Gutgesell, H. P. (1994). Effects of heteroscedasticity and skewness on prediction in regression: Modeling growth of the human heart. In M. L. Johnson (Ed.), *Essential numerical computer methods* (pp. 67–81). Oxford, UK: Academic Press.

Adarabioyo, M. I., & Ipinyomi, R. A. (2019). Comparing zero-inflated poisson, zero-inflated negative binomial and zero-inflated geometric in count data with excess zero. *Asian Journal of Probability and Statistics*, 1–10. doi:10.9734/ajpas/2019/v4i230113

Afifi, A. A., Kotlerman, J. B., Ettner, S. L., & Cowan, M. (2007). Methods for improving regression analysis for skewed continuous or counted responses. *Annual Review Public Health*, *28*, 95–111. doi:10.1146/annurev.publhealth.28.082206.094100

Akantziliotou, K., Rigby, R. A., & Stasinopoulos, D. M. (2002). *The R implementation of generalized additive models for location, scale and shape*. Statistical modelling in Society: Proceedings of the 17th

International Workshop on statistical modelling (pp. 75–83). Chania, Greece: Statistical Modelling Society.

Allen, K. J., Fox, K. R., Schatten, H. T., & Hooley, J. M. (2019). Frequency of nonsuicidal self-injury is associated with impulsive decision-making during criticism. *Psychiatry Research*, *271*, 68–75. doi:10.1016/j.psychres.2018.11.022

Andover, M. S., & Gibb, B. E. (2010). Non-suicidal self-injury, attempted suicide, and suicidal intent among psychiatric inpatients. *Psychiatry Research*, *178*(1), 101–105. doi:10.1016/j.psychres.2010.03.019

Asparouhov, T., & Muthén, B. (2015). Structural equation models and mixture models with continuous nonnormal skewed distributions. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(1), 1–19. doi:10.1080/10705511.2014.947375

Auerbach, R. P., Kim, J. C., Chango, J. M., Spiro, W. J., Cha, C., Gold, J., … Nock, M. K. (2014). Adolescent nonsuicidal self-injury: Examining the role of child abuse, comorbidity, and disinhibition. *Psychiatry Research*, *220*(1–2), 579–584. doi:10.1016/j.psychres.2014.07.027

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, *12*, 171–178.

Benjet, C., González-Herrera, I., Castro-Silva, E., Méndez, E., Borges, G., Casanova, L., & Medina-Mora, M. E. (2017). Non-suicidal self-injury in Mexican young adults: Prevalence, associations with suicidal behavior and psychiatric disorders, and DSM-5 proposed diagnostic criteria. *Journal of Affective Disorders*, *215*, 1–8. doi:10.1016/j.jad.2017.03.025

Boone, S. D., & Brausch, A. M. (2016). Physical activity, exercise motivations, depression, and nonsuicidal self-injury in youth. *Suicide and Life-Threatening Behavior*, *46*(5), 625–633. doi:10.1111/sltb.2016.46.issue-5

Brausch, A. M., & Muehlenkamp, J. J. (2018). Perceived effectiveness of NSSI in achieving functions on severity and suicide risk. *Psychiatry Research*, *265*, 144–150. doi:10.1016/j.psychres.2018.04.038

Buser, T. J., Pertuit, T. L., & Muller, D. L. (2019). Nonsuicidal self-injury, stress, and self-differentiation. *Adultspan Journal*, *18*(1), 4–16. doi:10.1002/adsp.2019.18.issue-1

Cerutti, R., Zuffianò, A., & Spensieri, V. (2018). The role of difficulty in identifying and describing feelings in non-suicidal self-injury behavior (NSSI): Associations with perceived attachment quality, stressful life events, and suicidal ideation. *Frontiers in Psychology*, *9*, 318. doi:10.3389/fpsyg.2018.00318

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawerence Associates.

Cohen, J. E. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnomality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*(1), 16–29. doi:10.1.1.382.6856

DiLillo, D., Foriter, M. A., Hayes, S. A., Trask, E., Perry, A. R., Messman-Moore, T., … Nash, C. (2006). Retrospective assessment of childhood sexual and physical abuse. *Assessment*, *13*, 297–312. doi:10.1177/1073191106288391

Driscoll, C. F. B., Murray, C. B., Holbein, C. E., Stiles-Shields, C., Cuevas, G., & Holmbeck, G. N. (2019). Camp-based psychosocial intervention dosage and changes in independence in young people with spina bifida. *Developmental Medicine & Child Neurology*, *61*(12), 1392–1399. doi:10.1111/dmcn.v61.12

Dube, S. R., Williamson, D. F., Thompson, T., Felitti, V. J., & Anda, R. F. (2004). Assessing the reliability of retrospective reports of adverse childhood experiences among adult HMO members attending a primary care clinic. *Child Abuse & Neglect*, *28*, 729–737. doi:10.1016/j.chiabu.2003.08.009

Dufour, J. P., & Zung, N. T. (2005). *Poisson structures and their normal forms*. Basel, Switzerland: Birkhauser.

Esposito, C., Bacchini, D., & Affuso, G. (2019). Adolescent non-suicidal self-injury and its relationships with school bullying and peer rejection. *Psychiatry Research*, *274*, 1–6. doi:10.1016/j.psychres.2019.02.018

Feinberg, R. A., & Rubright, J. D. (2016). Conducting simulation studies in psychometrics. *Educational Measurement: Issues and Practice*, *35*(2), 36–49. doi:10.1111/emip.12111

Fergusson, D. M., Horwood, L. J., & Woodward, L. J. (2000). The stability of child abuse reports: A longitudinal study of the reporting behaviour of young adults. *Psychological Medicine*, *30*(3), 529–544. doi:10.1017/S0033291799002111

Fox, K. R., Huang, X., Linthicum, K. P., Wang, S. B., Franklin, J. C., & Ribeiro, J. D. (2019). Model complexity improves the prediction of nonsuicidal self-injury. *Journal of Consulting and Clinical Psychology*, *87*(8), 684. doi:10.1037/ccp0000421

Fox, K. R., Ribeiro, J. D., Kleiman, E. M., Hooley, J. M., Nock, M. K., & Franklin, J. C. (2018). Affect toward the self and self-injury stimuli as potential risk factors for nonsuicidal self-injury. *Psychiatry Research*, *260*, 279–285. doi:10.1016/j.psychres.2017.11.083

Gillies, D., Christou, M. A., Dixon, A. C., Featherston, O. J., Rapti, I., Garcia-Anguita, A., … Christou, P. A. (2018). Prevalence and characteristics of self-harm in adolescents: Meta-analyses of community-based studies 1990–2015. *Journal of the American Academy of Child & Adolescent Psychiatry*, *57*(10), 733–741. doi:10.1016/j.jaac.2018.06.018

Glenn, C. R., Kleiman, E. M., Cha, C. B., Nock, M. K., & Prinstein, M. J. (2016). Implicit cognition about self-injury predicts actual self-injurious behavior: results from a longitudinal study of adolescents. *Journal of Child Psychology and Psychiatry*, *57*(7), 805–813. doi:10.1111/jcpp.2016.57.issue-7

Godfrey, L. G., & Orme, C. D. (1991). Testing for skewness of regression disturbances. *Economic Letters*, *37*, 31–34. doi:10.1016/0165-1765(91)90238-G

Gschlößl, S., & Czado, C. (2008). Modeling count data with overdispersion and spatial effects. *Statistical Papers*, *49*, 531–552. doi:10.1007/s00362-006-0031-6

Guertin, T., Lloyd-Richardson, E., Spirito, A., Donaldson, D., & Boergers, J. (2001). Self-mutilative behavior in adolescents who attempt suicide by overdose. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(9), 1062–1069. doi:10.1097/00004583-200109000-00015

Hall, P. (1992). On the removal of skewness by transformation. *Journal of the Royal Statistical Society: Series B (Methodological)*, *54*(1), 221–228. doi:10.1111/rssb.1992.54.issue-1

Hardt, J., & Rutter, M. (2004). Validity of adult retrospective reports of adverse childhood experiences: Review of the evidence. *Journal of Child Psychology and Psychiatry*, *45*, 260–273. doi:10.1111/jcpp.2004.45.issue-2

Hasking, P., Boyes, M. E., Finlay-Jones, A., McEvoy, P. M., & Rees, C. S. (2019). Common pathways to NSSI and suicide ideation: The roles of rumination and self-compassion. *Archives of Suicide Research*, *23*(2), 247–260. doi:10.1080/13811118.2018.1468836

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. doi:10.1080/10705519909540118

Kaess, M., Parzer, P., Mattern, M., Plener, P. L., Bifulco, A., Resch, F., & Brunner, R. (2013). Adverse childhood experiences and their impact on frequency, severity, and the individual function of nonsuicidal self-injury in youth. *Psychiatry Research*, *206*(2–3), 265–272. doi:10.1016/j.psychres.2012.10.012

Klonsky, E. D., & Olino, T. M. (2008). Identifying clinically distinct subgroups of self-injurers among young adults: A latent class analysis. *Journal of Consulting and Clinical Psychology*, *76*(1), 22. doi:10.1037/0022-006X.76.1.22

Kranzler, A., Fehling, K. B., Lindqvist, J., Brillante, J., Yuan, F., Gao, X., … Selby, E. A. (2018). An ecological investigation of the emotional context surrounding nonsuicidal self-injurious thoughts and behaviors in adolescents and young adults. *Suicide and Life-Threatening Behavior*, *48*(2), 149–159. doi:10.1111/sltb.2018.48.issue-2

Lambert, D. (1992). Zero-inflated poisson regression, with an application to defects in manufacturing. *Techonometrics*, *34*(1), 1–14. doi:10.2307/1269547

Lang, C. M., & Sharma-Patel, K. (2011). The relation between childhood maltreatment and self-injury: A review of the literature on conceptualization and intervention. *Trauma, Violence, & Abuse*, *12*(1), 23–37. doi:10.1177/1524838010386975

Lee, S. M., Karrison, T., Nocon, R. S., & Huang, E. (2018). Weighted zero-inflated poisson mixed model with an application to medicaid utilization data. *Communications for Statistical Applications and Methods*, *25*(2), 173–184. doi:10.29220/CSAM.2018.25.2.173

Lee, S. X., & McLachlan, G. J. (2013). On mixtures of skew normal and skew *t*- distributions. *Advances in Data Analysis and Classification*, *7*(3), 241–266. doi:10.1007/s11634-013-0132-8

Lewinsohn, P. M., Allen, N. B., Seeley, J. R., & Gotlib, I. H. (1999). First onset versus recurrance of depression: Differential process of psychological risk. *Journal of Abnormal Psychology*, *108*(3), 483–489. doi:10.1037/0021-843X.108.3.483

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

Liu, R. T., Scopelliti, K. M., Pittman, S. K., & Zamora, A. S. (2018). Childhood maltreatment and non-suicidal self-injury: A systematic review and meta-analysis. *The Lancet Psychiatry*, *5*(1), 51–64. doi:10.1016/S2215-0366(17)30469-8

Lloyd, E. 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 Society for Behavioral Medicine, New Orleans, LA.

Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage Publications.

Lüdtke, J., In-Albon, T., Michel, C., & Schmid, M. (2016). Predictors for DSM-5 nonsuicidal self-injury in female adolescent inpatients: The role of childhood maltreatment, alexithymia, and dissociation. *Psychiatry Research*, *239*, 346–352. doi:10.1016/j.psychres.2016.02.026

MacGillivray, H. L. (1986). Skewness and asymmetry: Measures and orderings. *The Annals of Statistics*, *14*, 994–1011. doi:10.1214/aos/1176350046

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, *39*(1), 99–128. doi:10.1207/s15327906mbr3901_4

Midkiff, M. F., Lindsey, C. R., & Meadows, E. A. (2018). The role of coping self-efficacy in emotion regulation and frequency of NSSI in young adult college students. *Cogent Psychology*, *5*(1), 1520437. doi:10.1080/23311908.2018.1520437

Minami, M., Lennert-Cody, C. E., Gao, W., & Román-Verdesoto, M. (2007). Modeling shark bycatch: The zero-inflated negative binomial regression model with smoothing. *Fisheries Research*, *84*(2), 210–221. doi:10.1016/j.fishres.2006.10.019

Morse, C. D., Klingman, K. J., Jacob, B. L., & Kodali, L. (2019). Exercise and insomnia risk in middle-aged women. *The Journal for Nurse Practitioners*, *15*(3), 236–240. doi:10.1016/j.nurpra.2018.10.020

Muehlenkamp, J., Brausch, A., Quigley, K., & Whitlock, J. (2013). Interpersonal features and functions of nonsuicidal self-injury. *Suicide and Life-Threatening Behavior*, *43*(1), 67–80. doi:10.1111/sltb.2013.43.issue-1

Muehlenkamp, J. J., Xhunga, N., & Brausch, A. M. (2019). Self-injury age of onset: A risk factor for NSSI severity and suicidal behavior. *Archives of Suicide Research*, *23*(4), 551–563. doi:10.1080/13811118.2018.1486252

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Nielsen, E., Sayal, K., & Townsend, E. (2017). Functional coping dynamics and experiential avoidance in a community sample with no self-injury vs. non-suicidal self-injury only vs. those with both non-suicidal self-injury and suicidal behaviour. *International Journal of Environmental Research and Public Health*, *14*(6), 575. doi:10.3390/ijerph14060575

Nock, M. K. (2010). Self-injury. *Annual Review of Clinical Psychology*, *6*, 339–363. doi:10.1146/annurev.clinpsy.121208.131258

Nock, M. K., Holmberg, E. B., Photos, V. I., & Michel, B. D. (2007). Self-injurious thoughts and behaviors interview: Development, reliability, and validity in an adolescent sample. *Psychological Assessment*, *19*(3), 309–317.

Nock, M. K., & Prinstein, M. J. (2004). A functional approach to the assessment of self-mutilative behavior. *Journal of Counseling & Clinical Psychology*, *72*(5), 885–890. doi:10.1037/0022-006X.72.5.885

Nock, M. K., & Prinstein, M. J. (2005). Contextual features and behavioral functions of self-mutilation among adolescents. *Journal of Abnormal Psychology*, *114*(1), 140–146. doi:10.1037/0021-843X.114.1.140

Osborne, J. W., & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research, and Evaluation*, *8*(1), 2.

Rigby, R. A., & Stasinopoulos, D. M. (2001, July). *The GAMLSS project: A flexible approach to statistical modelling*. New trends in statistical modelling: Proceedings of the 16th international workshop on statistical modelling (Vol. 337, p. 345). Odense, Denmark: University of Southern Denmark.

Rigby, R. A., & Stasinopoulos, D. M. (2005, June). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *54*(3), 507–554. doi:10.1111/j.1467-9876.2005.00510.x

Robertson, C. D., Miskey, H., Mitchell, J., & Nelson-Gray, R. (2013). Variety of self-injury: Is the number of different methods of non-suicidal self-injury related to personality, psychopathology, or functions of self-injury? *Archives of Suicide Research*, *17*(1), 33–40. doi:10.1080/13811118.2013.748410

Sanders, B., & Becker-Lausen, E. (1995). The measurement of psychological maltreatment: Early data on the child abuse and trauma scale. *Child Abuse & Neglect*, *19*(3), 315–323. doi:10.1016/S0145-2134(94)00131-6

Schoenleber, M., Berenbaum, H., & Motl, R. (2014). Shame-related functions of and motivations for self-injurious behavior. *Personality Disorders: Theory, Research, and Treatment*, *5*(2), 204. doi:10.1037/per0000035

Serafini, G., Canepa, G., Adavastro, G., Nebbia, J., Belvederi Murri, M., Erbuto, D., … Amore, M. (2017). The relationship between childhood maltreatment and non-suicidal self-injury: A systematic review. *Frontiers in Psychiatry*, *8*, 149. doi:10.3389/fpsyt.2017.00149

Shaffer, A., Huston, L., & Egeland, B. (2008). Identification of child maltreatment using prospective and self-report methodologies: A comparison of maltreatment incidence and relation to later psychopathology. *Child Abuse & Neglect*, *32*, 682–692. doi:10.1016/j.chiabu.2007.09.010

Swannell, S. V., Martin, G. E., Page, A., Hasking, P., & St John, N. J. (2014). Prevalence of nonsuicidal self-injury in nonclinical samples: Systematic review, meta-analysis and meta-regression. *Suicide and Life-Threatening Behavior*, *44*(3), 273–303. doi:10.1111/sltb.2014.44.issue-3

Tabachnick, B. G., & Fidell, L. S. (2001). Using multivariate statistics. Needham Heights, MA: Allyn and Bacon.

Thomassin, K., Shaffer, A., Madden, A., & Londino, D. L. (2016). Specificity of childhood maltreatment and emotion deficit in nonsuicidal self-injury in an inpatient sample of youth. *Psychiatry Research*, *244*, 103–108. doi:10.1016/j.psychres.2016.07.050

Tüzen, M. F., & Erbaş, S. (2018). A comparison of count data models with an application to daily cigarette consumption of young persons. *Communications in Statistics-Theory and Methods*, *47*(23), 5825–5844. doi:10.1080/03610926.2017.1402050

Vergara, G. A., Stewart, J. G., Cosby, E. A., Lincoln, S. H., & Auerbach, R. P. (2019). Non-suicidal self-injury and suicide in depressed adolescents: Impact of peer victimization and bullying. *Journal of Affective Disorders*, *245*, 744–749. doi:10.1016/j.jad.2018.11.084

Weisberg, S. (2005). *Applied linear regression* (Vol. 528). Hoboken, NJ: John Wiley & Sons.

Whitlock, J., Muehlenkamp, J., Eckenrode, J., Purington, A., Abrams, G. B., Barreira, P., & Kress, V. (2013). Nonsuicidal self-injury as a gateway to suicide in young adults. *Journal of Adolescent Health*, *52*(4), 486–492. doi:10.1016/j.jadohealth.2012.09.010

Whitlock, J. L., Eckenrode, J., & Silverman, D. (2006). Self-injurious behaviors in a college population. *Pediatrics*, *117*(6), 1939–1948. doi:10.1542/peds.2005-2543

Widom, C. S., & Morris, S. (1997). Accuracy of adult recollection of childhood victimization: Part 2 childhood sexual abuse. *Psychological Assessment*, *9*, 34–46. doi:10.1037/1040-3590.9.1.34

Williams, M. N., Grajales, C. A. G., & Kurkiewicz, D. (2013). Assumptions of multiple regression: Correcting two misconceptions. *Practical Assessment, Research, and Evaluation*, *18*(1), 11.

Yates, T. M. (2009). Developmental pathways from child maltreatment to nonsuicidal self-injury. In M. Nock (Ed.), *Understanding non-suicidal self-injury: Current science and practice* (pp. 117–137). Washington, DC: American Psychological Association.

Yates, T. M., Burt, K. B., & Troy, M. F. (2011). A developmental approach to clinical research, classification & practice. In D. Cicchetti & G. I. Roisman (Eds.), *The Minnesota symposia on child psychology: The origins and organization of adaptation and maladaptation* (Vol. 36, pp. 231–282). New York, NY: Wiley.

Yates, T. M., Carlson, E., & Egeland, B. (2008). A prospective study of child maltreatment and self-injurious behavior in a community sample. *Development and Psychopathology*, *20*, 651–671. doi:10.1017/S0954579408000321

Yates, T. M., Tracy, A. J., & Luthar, S. S. (2008). Nonsuicidal self-injury among "privileged" youths: Longitudinal and cross-sectional approaches to developmental process. *Journal of Consulting and Clinical Psychology*, *76*(1), 52–62. doi:10.1037/0022-006X.76.1.52

You, J., & Leung, F. (2012). The role of depressive symptoms, family invalidation and behavioral impulsivity in the occurrence and repetition of non-suicidal self-injury in Chinese adolescents: A 2-year follow-up study. *Journal of Adolescence*, *35*(2), 389–395. doi:10.1016/j.adolescence.2011.07.020

You, J., Zheng, C., Lin, M. P., & Leung, F. (2016). Peer group impulsivity moderated the individual-level relationship between depressive symptoms and adolescent nonsuicidal self-injury. *Journal of Adolescence*, *47*, 90–99. doi:10.1016/j.adolescence.2015.12.008