KMAN PUBLISHING INSTITUTE
Journal Website

Article history:
Received 08 January 2026
Revised 25 February 2026
Accepted 30 March 2026
Published Online 10 June 2026

Journal of Adolescent and Youth Psychological Studies

Open peer-review report

E-ISSN: 2981-2526

Random Forest Classification of Non-Suicidal Self-Injury Risk among Adolescents Using Trauma History, Shame, Impulsivity, and Peer Victimization Variables

Rúnar. Guðmundsson[1], Valentina. Paredes[2*], Ilze. Kalniņa[3]

[1] Department of Psychology, University of Iceland, Reykjavík, Iceland
[2] Department of Social Psychology, Pontificia Universidad Católica de Chile, Santiago, Chile
[3] Department of Clinical Psychology, University of Latvia, Riga, Latvia

*** Corresponding author email address**: valentina.paredes@uc.cl

Editor	Reviewers
Seyed Ali Darbani Assistant Professor, Department of Psychology and Counseling, South Tehran Branch, Islamic Azad University, Tehran, Iran Ali.darbani@iau.ac.ir	**Reviewer 1:** Mohammad Salehi Associate Professor, Department of Educational Management, Sari Branch, Islamic Azad University, Sari, Iran. Email: drsalehi@iausari.ac.ir **Reviewer 2:** Kamdin Parsakia Department of Psychology and Counseling, KMAN Research Institute, Richmond Hill, Ontario, Canada. Email: kamdinarsakia@kmanresce.ca

1. Round 1

1.1. Reviewer 1

Reviewer:

The paragraph beginning with “Machine learning approaches offer significant advantages for psychological risk assessment” (p. 4) presents Random Forest as a preferred analytical method but does not justify why this algorithm was selected over other commonly used classifiers such as XGBoost, LightGBM, Support Vector Machines, Logistic Regression, or Neural Networks. The authors should provide a methodological rationale and ideally compare Random Forest performance against benchmark models to demonstrate its relative suitability for NSSI risk classification.

In the Study Design and Participants section, the authors report that “A multistage cluster sampling strategy was employed” (p. 4). However, critical information regarding the number of schools selected, school participation rates, classroom selection procedures, and refusal rates is absent. These details are necessary for evaluating sample representativeness and potential selection bias. The sampling procedure should therefore be described more comprehensively.

The statement that “Participants were categorized into low-risk and high-risk groups based on established clinical cutoff scores on the self-injury assessment measure” (p. 4) is insufficiently detailed. The manuscript does not specify the exact cutoff score, the empirical basis for its selection, or whether it has been validated in Chilean adolescent populations. Since the target

variable determines the entire machine-learning classification process, greater transparency regarding classification criteria is essential.

In the Measures section, the authors repeatedly state that previous studies confirmed the reliability and validity of the Spanish versions of the instruments. However, the manuscript does not report internal consistency coefficients (e.g., Cronbach's alpha, McDonald's omega) for the current sample. For example, in the paragraph describing the Experience of Shame Scale, no reliability statistics are presented. Psychometric properties should be reported for all measures within the current dataset rather than relying solely on previous validation studies.

A major concern involves Figure 1. The figure title reads "SHAP Summary Plot Illustrating the Relative Contributions of Executive Function, Motivation, Smartphone Use, and Learning Strategy Variables to Academic Procrastination Risk Classification" (p. 9), whereas the current study investigates NSSI risk using trauma history, shame, impulsivity, and peer victimization. This discrepancy strongly suggests that the figure was copied from another manuscript and not properly revised. The figure title, caption, and content must be corrected immediately.

Related to Figure 1, the text states that "Figure 1 illustrates the receiver operating characteristic curve for the final Random Forest classifier" (p. 9). However, the caption references a SHAP summary plot rather than an ROC curve. This inconsistency raises concerns regarding manuscript preparation quality and figure validity. The authors should carefully review all figures, captions, and references to ensure alignment with the study objectives and analyses.

Authors uploaded the revised manuscript.

1.2. Reviewer 2

Reviewer:

The paragraph describing the Functional Assessment of Self-Mutilation (FASM) (pp. 4–5) lacks important information regarding scoring procedures. The authors should clarify whether total scores, frequency counts, or dichotomous classifications were used, how multiple self-injury behaviors were aggregated, and whether severity weighting was applied. Such details are critical for replicability.

In the Data Analysis section, the authors state that "Missing data representing less than 5% of observations were handled using multiple imputation procedures" (p. 5). However, the specific imputation algorithm, number of imputations performed, predictor matrix specifications, and convergence diagnostics are not reported. Since imputation can substantially affect model performance, these methodological details should be included.

The authors indicate that "Continuous variables were standardized prior to model development" (p. 5). This decision requires clarification because Random Forest algorithms are generally insensitive to variable scaling. The authors should explain the rationale for standardization and discuss whether preprocessing was applied within cross-validation folds to prevent data leakage.

The paragraph stating that "To address potential class imbalance between high-risk and low-risk NSSI groups, the Synthetic Minority Oversampling Technique (SMOTE) was applied exclusively to the training data" (p. 5) lacks sufficient detail. The authors should report the imbalance ratio before and after oversampling, the SMOTE parameters used, and whether sensitivity analyses were conducted to evaluate the impact of oversampling on predictive performance.

In the sentence "Hyperparameter optimization was conducted through grid search combined with five-fold cross-validation" (p. 6), the manuscript fails to specify the hyperparameter search space. The ranges tested for number of trees, maximum depth, minimum samples split, minimum samples leaf, and maximum features should be reported, preferably in a dedicated table, to ensure methodological reproducibility.

The manuscript reports impressive predictive metrics, including "accuracy of 89.2%, precision of 86.1%, recall of 84.5%, F1-score of 85.3%, and AUC of 0.938" (p. 7). However, no confidence intervals, bootstrap estimates, or variability measures are provided. Because machine-learning performance estimates can fluctuate substantially across data partitions, confidence intervals should be reported to demonstrate model stability and reliability.

JAYPS Adolescent and Youth Psychological Studies
E-ISSN: 2981-2526

Table 1 presents correlations among study variables (p. 6), and the authors conclude that "variance inflation factor analyses indicated no problematic multicollinearity." However, VIF values are not reported. Since trauma history, shame, and peer victimization show moderately strong intercorrelations, the actual VIF and tolerance statistics should be presented to substantiate this conclusion.

In Table 2, group comparisons between high-risk and low-risk adolescents are reported solely through t-tests (p. 6). Given the large sample size, statistical significance is expected even for relatively modest differences. The authors should additionally report effect size measures such as Cohen's d and confidence intervals to facilitate interpretation of the practical significance of these findings.

The discussion of Table 4 states that "Shame emerged as the most influential predictor, accounting for 32.1% of the model's predictive contribution" (p. 7). The authors should clarify whether these importance values represent Gini importance, mean decrease in impurity, permutation importance, or another metric. Since traditional Random Forest importance measures are known to exhibit bias under certain conditions, the choice of importance metric requires explicit justification.

Authors uploaded the revised manuscript.

2. Revised

Editor's decision after revisions: Accepted.
Editor in Chief's decision: Accepted.

JAYPS Adolescent and Youth Psychological Studies
E-ISSN: 2981-2526