

Classifying Panic-Symptom Risk among Taiwanese Youth Based on Anxiety Sensitivity, Interoceptive Awareness, Sleep Disturbance, and Perceived Stress Using a Support Vector Machine

Keisuke. Mori¹, Chen. Yu-Han^{2*}, Minh. Tran Quoc³

¹ Department of Behavioral Psychology, Kyoto University, Kyoto, Japan

² Department of Counseling Psychology, National Taiwan University, Taipei, Taiwan

³ Department of Psychology, Vietnam National University, Hanoi, Vietnam

* Corresponding author email address: yuhan.chen@ntu.edu.tw

E d i t o r	R e v i e w e r s
Muhammad Rizwan  Associate Professor, Department of Psychology, Haripur University, Islamabad, Pakistan muhammad.rizwan@uoh.edu.pk	Reviewer 1: Mehdi Rostami  Department of Psychology and Counseling, KMAN Research Institute, Richmond Hill, Ontario, Canada. Email: dr.mrostami@kmanresce.ca Reviewer 2: Azade Abooei  Department of Counseling, Faculty of Humanities, University of Science and Art, Yazd, Iran. Email: a.aboei@tea.sau.ac.ir

1. Round 1

1.1. Reviewer 1

Reviewer:

In the eligibility paragraph, the exclusion criterion referring to “an acute psychiatric crisis requiring immediate intervention” requires operational clarification. The manuscript should explain how such crises were identified in an online or questionnaire-based study, whether a screening item was used, who reviewed responses, and what referral protocol was activated. This is both a methodological and ethical issue, particularly because the sample included minors and participants reporting potentially severe anxiety symptoms.

The sentence “Traditional Chinese versions of all instruments were administered” should be expanded to document the source and validation status of each translated measure. The authors should specify whether validated Taiwanese versions of the Anxiety Sensitivity Index–3, Multidimensional Assessment of Interoceptive Awareness–2, Pittsburgh Sleep Quality Index, Perceived Stress Scale–10, and Panic Disorder Severity Scale–Self-Report were used. If any measure was newly translated, the manuscript should describe forward translation, back-translation, expert review, cognitive interviewing, and preliminary psychometric evaluation.

Table 1 reports very large standardized mean differences, including Cohen's $d = 2.00$ for anxiety sensitivity and $d = 1.71$ for perceived stress. These effects are unusually large for community psychological data and should be carefully verified. The authors should confirm the calculations, report group sample sizes directly in the table, provide confidence intervals for effect sizes, and examine whether the cutoff-based grouping mechanically inflated differences. Because the panic-risk classification was derived from a symptom measure correlated with these predictors, dichotomization may exaggerate between-group contrasts.

The sentence in the Findings section stating that “no predictor demonstrated variance inflation sufficient to warrant removal” is not adequately supported because no numerical diagnostics are reported. The authors should provide a correlation matrix or supplementary table containing variance inflation factors, tolerance statistics, and pairwise correlations. Although variance inflation factors are primarily associated with linear models, strong dependence among predictors can also destabilize permutation importance and complicate interpretation in nonlinear models.

Authors uploaded the revised manuscript.

1.2. Reviewer 2

Reviewer:

The operationalization of elevated panic-symptom risk requires stronger justification. The paragraph stating that “participants with a total score of 8 or higher were classified as having elevated panic-symptom risk” should cite validation evidence for this threshold and explain whether the cutoff has been established in Taiwanese adolescents and young adults. A threshold developed in adult clinical populations may not possess equivalent sensitivity and specificity in a community youth sample. The authors should conduct sensitivity analyses using alternative thresholds or model the panic-severity score continuously.

The authors should address the potential circularity created by predicting a questionnaire-defined panic-risk category from conceptually overlapping self-report constructs. The outcome measure contains anticipatory anxiety, avoidance, distress, and functional impairment, while anxiety sensitivity and interoceptive variables assess closely related fear and appraisal processes. In the paragraph beginning, “The binary outcome variable represented elevated versus low panic-symptom risk,” the authors should discuss construct overlap and demonstrate that the model is not merely reproducing shared wording or general negative affectivity.

The treatment of the Multidimensional Assessment of Interoceptive Awareness–2 requires additional psychometric justification. The sentence “the eight subscale scores were entered separately into the machine-learning model” is theoretically reasonable, but the manuscript should report correlations among subscales, variance inflation indicators, and evidence of subscale reliability in this specific youth sample. Several subscales contain relatively few items, and coefficients near .70 may not ensure stable individual-level prediction. The authors should also clarify whether reverse-scored dimensions were interpreted consistently.

The missing-data procedure needs greater transparency. The manuscript states that participants with more than 10% missing data were excluded and that remaining values were median-imputed within the training pipeline. The authors should report the number of participants affected, the percentage of missingness for each variable, whether missingness differed by panic-risk group, and whether the missing-at-random assumption was examined. A sensitivity analysis using multiple imputation or complete cases would help establish whether the chosen procedure materially influenced model performance.

In the Data Analysis section, the statement that hyperparameters were optimized using “a logarithmically spaced hyperparameter grid” is insufficiently reproducible. The manuscript should report the exact values tested for C and γ , the scoring metric used to select the optimal combination, the number of repeated cross-validation cycles, and the software package and version used. The final selected values, $C = 10$ and $\gamma = .05$, should be linked explicitly to the predefined search space rather than reported only in the Findings section.

The analytic pipeline may contain optimistic bias unless the threshold selection, probability calibration, and hyperparameter tuning were fully nested. The paragraph describing repeated stratified 10-fold cross-validation should clarify whether an inner

cross-validation loop was used for tuning and calibration while an outer loop estimated performance. If the same cross-validated predictions were used both to optimize hyperparameters and select the .41 decision threshold, the reported internal performance may be overly optimistic. A fully nested cross-validation design would provide a more defensible estimate.

The choice of the Youden index for selecting the final threshold requires clinical justification. The sentence “the threshold was determined by maximizing the Youden index while also considering sensitivity” is ambiguous because it does not specify the decision rule. The authors should state whether a minimum sensitivity was imposed, how ties were resolved, and why false negatives and false positives were treated as having approximately equal costs. In a youth screening context, the consequences of missing a high-risk participant may differ substantially from unnecessary referral.

Authors uploaded the revised manuscript.

2. Revised

Editor’s decision after revisions: Accepted.

Editor in Chief’s decision: Accepted.