



LightGBM Classification of Academic Procrastination Risk Among Adolescents Based on Executive Function, Motivation, Smartphone Use, and Learning Strategy Variables

Jean-Paul. Nkurunziza¹, Lars. Becker^{2*}

¹ Department of Clinical Psychology, University of Rwanda, Kigali, Rwanda



² Department of Cognitive Psychology, Ludwig Maximilian University of Munich, Munich, Germany

* Corresponding author email address: lars.becker@psy.lmu.de

Editor

John S. Carlson¹
Distinguished Professor of the
Department of Educational
Psychology, Michigan State
University, East Lansing, MI,
United
carlsoj@msu.edu

Reviewers

Reviewer 1: Mohammad Salehi¹
Associate Professor, Department of Educational Management, Sari Branch, Islamic
Azad University, Sari, Iran. Email: drsalehi@iausari.ac.ir
Reviewer 2: Mahdi Khanjani¹
Associate Professor, Department of Psychology, Allameh Tabataba'i University,
Tehran, Iran.
Email: khanjani_m@atu.ac.ir

1. Round 1

1.1. Reviewer 1

Reviewer:

In the paragraph beginning “A growing body of evidence identifies executive function deficits as core cognitive correlates of procrastination,” the authors discuss working memory, inhibitory control, and cognitive flexibility as key predictors. However, the manuscript does not sufficiently justify why these specific executive-function dimensions were selected over other relevant constructs such as attentional control, cognitive persistence, or emotional regulation. A stronger rationale linking the BRIEF-SR subdomains to the target classification outcome is needed.

In the Data Analysis section, the authors indicate that “academic procrastination scores were transformed into categorical risk groups based on percentile-based classification thresholds.” This methodological decision requires stronger justification. Categorizing a continuous variable can reduce statistical information and potentially affect model performance. The authors should explain why classification was preferred over regression and provide evidence supporting the selected percentile cutoffs.

The sentence “Cases with substantial missing data were removed, while minor missing values were imputed using median-based procedures” lacks important methodological detail. The authors should define what constituted “substantial”

missingness, report the exact number of excluded cases, and explain why median imputation was chosen instead of more sophisticated methods such as multiple imputation or k-nearest-neighbor imputation.

Authors uploaded the revised manuscript.

1.2. Reviewer 2

Reviewer:

The paragraph beginning “In parallel, the pervasive use of smartphones and digital media has introduced novel behavioral risk factors” presents smartphone use as a major contributor to procrastination. Nevertheless, the discussion remains largely descriptive. The authors should provide a more detailed explanation of the mechanisms through which smartphone dependence affects procrastination, including attentional fragmentation, reward sensitivity, digital multitasking, and interference with executive control processes.

The statement “Traditional correlational analyses provide valuable insights into the relationships among executive function, motivation, smartphone use, and learning strategies; however, they often fail to capture complex, nonlinear interactions” requires additional methodological justification. The authors should elaborate on why LightGBM was specifically chosen over alternative machine-learning approaches such as Random Forest, XGBoost, CatBoost, Support Vector Machines, or Neural Networks, and discuss the comparative advantages of LightGBM for this educational classification problem.

In the final paragraph of the Introduction, the authors state that “this study aims to develop a predictive LightGBM model to classify the risk of academic procrastination.” However, no explicit research questions or hypotheses are presented. Although machine-learning studies often emphasize prediction rather than hypothesis testing, the inclusion of clearly defined research objectives and expected predictor importance rankings would strengthen the conceptual coherence of the study.

In the Study Design and Participants section, the authors report that participants were recruited from four German federal states using a multistage cluster sampling procedure. However, the manuscript does not provide sufficient information regarding the sampling stages, the number of schools approached, school participation rates, classroom selection procedures, or reasons for nonparticipation. Additional details are necessary to evaluate sample representativeness and potential sampling bias.

The sentence “Students with diagnosed intellectual disabilities or severe psychiatric conditions that could interfere with questionnaire completion were excluded” requires clarification. The authors should specify how these conditions were identified, whether exclusions were based on school records, parental reports, self-reports, or clinical diagnoses, and how many participants were excluded based on these criteria.

In the Measures section, the manuscript provides descriptions of the PASS, BRIEF-SR, AMS, SAS-SV, and MSLQ instruments; however, no reliability coefficients are reported for the current sample. For each scale and subscale used in model training, the authors should report internal consistency statistics (e.g., Cronbach’s alpha, McDonald’s omega) derived from the present dataset to demonstrate measurement reliability.

The paragraph describing the Motivated Strategies for Learning Questionnaire indicates that “selected subscales relevant to learning strategy utilization were employed.” This statement is insufficiently precise. The authors should explicitly identify which subscales were included, the number of items retained from each subscale, and whether any adaptations or item exclusions were performed before analysis.

Authors uploaded the revised manuscript.

2. Revised

Editor’s decision after revisions: Accepted.

Editor in Chief’s decision: Accepted.