Predicting Cyberbullying Perpetration via Random Forest Modeling of Moral Disengagement and Empathy Deficits

Tomas. Jankauskas[1], Kabelo. Radebe[2*], Amira. Chennoufi[3]

[1] Department of Social Psychology, Vilnius University, Vilnius, Lithuania
[2] Department of Health Psychology, North-West University, Potchefstroom, South Africa
[3] Department of Clinical Psychology, University of Sfax, Sfax, Tunisia

* Corresponding author email address: kabelo.radebe@nwu.ac.za

Editor	Reviewers
Muhammad Rizwan Associate Professor, Department of Psychology, Haripur University, Islamabad, Pakistan muhammad.rizwan@uoh.edu.pk	**Reviewer 1:** Mohammad Hassan Ghanifar Assistant Professor, Department of Psychology, Birjand Branch, Islamic Azad University, Birjand, Iran. Email: ghanifar@iaubir.ac.ir **Reviewer 2:** Abotaleb Saadati Shamir Assistant Professor, Department of Educational Sciences, University of Science and Research, Tehran, Iran. Email: psychology@iau.ac.ir

1. Round 1

1.1. Reviewer 1

Reviewer:

You state that "Moral disengagement refers to cognitive mechanisms that allow individuals to disengage internal moral standards from harmful conduct." While Bandura's framework is referenced conceptually, the manuscript would benefit from a clearer articulation of how each specific mechanism (e.g., dehumanization vs. diffusion of responsibility) theoretically maps onto online affordances such as anonymity or algorithmic amplification. Please expand the theoretical integration.

You discuss reciprocal amplification cycles between moral disengagement and aggression (e.g., longitudinal findings), yet your design is cross-sectional. This creates a conceptual tension. Please explicitly acknowledge this theoretical–methodological discrepancy earlier in the Introduction and clarify that your predictive modeling does not test reciprocal causality.

Was there imbalance between high vs. low cyberbullying categories after dichotomization? If so, were techniques such as class weighting or SMOTE applied? Without addressing imbalance, accuracy may overestimate performance.

Gender and grade differences are reported with t and F values but no effect sizes (e.g., Cohen's d, η^2). Given the sample size (N = 742), statistically significant differences may be small in magnitude. Please include effect sizes and interpret practical significance.

Authors uploaded the revised manuscript.

1.2.   Reviewer 2

Reviewer:

The manuscript states that multi-stage cluster sampling ensured "geographical and socio-demographic representativeness." However, no socio-economic indicators or school-type characteristics are reported. Please provide demographic stratification information (urban/rural distribution, SES proxies, school size) to substantiate representativeness claims.

You note that instruments were "adapted for the South African context" and piloted with 30 students. Please clarify whether formal cultural validation procedures (e.g., CFA, measurement invariance testing, back-translation if applicable) were conducted. A pilot test alone does not ensure construct validity across cultural groups.

You state that items were reverse-coded so that higher scores reflected "lower empathic responsiveness." However, it is unclear whether empathy deficits were treated as a unidimensional construct or whether cognitive and affective empathy were modeled separately. Given theoretical distinctions, please justify collapsing them (if done).

The manuscript indicates that missing data (<3%) were handled via multiple imputation using chained equations. Please report the number of imputations, convergence diagnostics, and whether imputation was performed prior to train/test split (which may risk data leakage) or within training folds only.

You mention that cyberbullying perpetration scores were dichotomized using a "theoretically informed percentile cutoff." Please specify the exact cutoff percentile and provide justification. Dichotomization may reduce statistical power and distort distributional information; a sensitivity analysis using continuous modeling would strengthen robustness.

The manuscript notes hyperparameter optimization via five-fold cross-validation but does not report final tuned values (e.g., optimal mtry, tree depth, minimum node size). Please provide these parameters for reproducibility.

Authors uploaded the revised manuscript.

2.   Revised

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