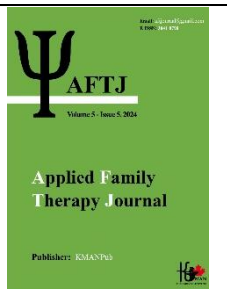
KMAN PUBLISHING INSTITUTE
Journal Website

Article history:
Received 12 May 2026
Revised 14 August 2026
Accepted 22 August 2026
Published online 01 September 2026

Applied Family Therapy Journal

OPEN PEER-REVIEW REPORT



E-ISSN: 3041-8798

Predicting Adolescent Emotional Problems From Parenting Styles, Family Functioning, and Parent–Child Communication Using LightGBM

Wei Jie. Lim[1], Julien. Dupont[2*], Aigerim. Sadykova[3]

[1] Department of Organizational Psychology, Universiti Putra Malaysia, Serdang, Malaysia
[2] Department of Neuropsychology, University of Bordeaux, Bordeaux, France
[3] Department of Applied Psychology, Al-Farabi Kazakh National University, Almaty, Kazakhstan

*** Corresponding author email address**: julien.dupont@u-bordeaux.fr

Editor	Reviewers
Monika Szczygieł Department of Psychology, Jagiellonian University, Krakow, Poland monika.szczygiel@uj.edu.pl	**Reviewer 1:** Masoud Asadi Assistant Professor, Department of Psychology and Counseling, Arak University, Arak, Iran. Email: m-asadi@araku.ac.ir **Reviewer 2:** Stephen C. L. Lau Professor (Assistant) at Washington University in St, Louis, United States. Email: lauc@wustl.edu

1. Round 1

1.1. Reviewer 1

Reviewer:

In the Introduction paragraph beginning “Traditional approaches to studying adolescent emotional problems have relied heavily on correlation, multiple linear regression, logistic regression, and structural equation modeling,” the rationale for selecting LightGBM is currently somewhat generic. The authors should provide a more technically specific justification for choosing LightGBM over other gradient-boosting methods, such as XGBoost or CatBoost. For example, the manuscript should discuss LightGBM’s leaf-wise tree growth, histogram-based feature binning, handling of nonlinearities and interactions, computational efficiency, and regularization options. Such detail is especially important because XGBoost is later included as a comparator, and readers need to understand why LightGBM was selected a priori as the principal model rather than simply because it produced the best result.

The Introduction repeatedly emphasizes that family processes may involve nonlinear and interactive effects, particularly in the paragraph stating that “the same level of authoritarian parenting, for example, may have different implications depending on whether family functioning is high or low.” This is theoretically compelling, but the manuscript does not clearly identify which interactions were expected to be most important. The authors should strengthen the theoretical model by specifying

plausible interactions before model estimation, such as family dysfunction × authoritarian parenting, open communication × authoritative parenting, or problematic communication × permissive parenting. Doing so would help distinguish theory-driven interpretation from post hoc interpretation of SHAP patterns.

For the McMaster Family Assessment Device, the manuscript correctly notes that "higher scores indicate poorer family functioning," but the findings and SHAP interpretation occasionally refer simply to "family functioning" rather than consistently to "family dysfunction." This terminology could create confusion regarding the direction of effects. The authors should standardize the variable label throughout the Methods, tables, Figure 1, Results, and Discussion. If higher FAD scores represent poorer functioning, the predictor should consistently be labeled "family dysfunction" or "poorer family functioning," particularly in figures where a reader may otherwise interpret larger values as healthier functioning.

The Data Analysis section states that "remaining item-level missing values represented less than 2% of observations and were handled through median imputation performed within the training data to prevent information leakage." This procedure requires further clarification because the predictors were scale totals based on multiple questionnaire items. The authors should explain whether median imputation was performed at the individual-item level before scale construction or at the scale-score level after calculating scores. If item-level median imputation was used, the manuscript should justify this choice because questionnaire scoring rules frequently recommend person-mean substitution or prorating when only a small number of items are missing. Moreover, any transformation learned from the training data should also have been applied unchanged to the test set, and this should be stated explicitly.

The manuscript reports an 80/20 split yielding "a training set comprising 80% of participants (n = 594) and an independent holdout test set comprising 20% (n = 148)." The authors should clarify whether this split was stratified according to outcome distribution, gender, age, school, or another factor. Because participants were clustered within schools, random allocation of individual adolescents to training and test sets may allow students from the same school to appear in both partitions, potentially inflating predictive performance through shared contextual characteristics. A more rigorous approach would be group-based splitting at the school level or nested resampling that preserves cluster independence. At minimum, the authors should discuss this issue as a potential source of optimistic performance estimation.

The hyperparameter-tuning paragraph provides final values such as "learning rate of .035, 420 estimators, 24 leaves, maximum depth of 6," but the search procedure is not sufficiently reproducible. The manuscript should specify whether grid search, randomized search, Bayesian optimization, Optuna, or another procedure was used; the full parameter ranges searched; the optimization metric; and the number of candidate configurations evaluated. Because the stated goal is predictive modeling, reproducibility of tuning is essential. The authors should also explain how early stopping was implemented in conjunction with five-fold cross-validation and whether the reported 420 estimators represent the median, mean, or final refitted value selected after tuning.

Table 2 reports five individual cross-validation folds and a mean R^2 of .623, followed by a holdout R^2 of .638. The manuscript interprets the similarity as evidence that "there was no indication of substantial model overfitting," but this claim should be made more cautiously. Similar cross-validation and holdout performance does not, by itself, establish absence of overfitting, particularly if hyperparameter tuning and model selection were conducted repeatedly on the same training data. The authors should report training performance, learning curves, or preferably nested cross-validation to quantify optimism more rigorously. Confidence intervals for R^2, RMSE, and MAE would also substantially improve interpretation of model stability.

Response: Revised and uploaded the manuscript.

1.2. Reviewer 2

Reviewer:

In the "Study Design and Participants" section, the authors report that participants were recruited from "Île-de-France, Auvergne-Rhône-Alpes, Nouvelle-Aquitaine, Occitanie, and Hauts-de-France," but no information is provided regarding the number of schools, number of classes, regional allocation of participants, or whether sampling probabilities differed across

AFTJ Applied Family Therapy Journal
E-ISSN: 3041-8798

regions. Because the study claims a multistage cluster-sampling procedure, the sampling frame should be described in considerably more detail. The authors should specify how schools were identified, how many schools agreed to participate, whether public and private schools were sampled proportionately, how classes were selected within schools, and whether clustering by school was considered analytically.

The same participant paragraph reports that "Of 789 adolescents initially invited, 759 returned consent materials and participated. Seventeen cases were excluded because more than 20% of questionnaire responses were missing, resulting in a final analytic sample of 742 adolescents." The flow of participants should be reported more transparently. The manuscript should explain whether the 30 nonparticipants differed systematically from participants, whether any adolescents were excluded for eligibility reasons other than missingness, and whether missingness was associated with age, gender, school type, or emotional-problem scores. A brief missing-data analysis would be useful because exclusion based on questionnaire completeness may introduce selection bias if adolescents with greater emotional difficulties were disproportionately represented among incomplete responders.

In the ethics portion of the participant section, the authors state that "The research procedures were conducted in accordance with the principles of the Declaration of Helsinki and applicable French regulations concerning research with minors." This statement is insufficient for publication in most journals. The manuscript should identify the approving ethics committee or institutional review board, provide the approval number and approval date, and clarify the exact consent/assent procedure. If parental consent was written and adolescent assent was obtained separately, this should be stated explicitly. The authors should also clarify whether any procedures were in place for responding to adolescents who reported severe emotional symptoms or potential self-harm risk.

In the description of the Strengths and Difficulties Questionnaire, the authors state that "the Emotional Symptoms subscale consists of five items" and that the scale was treated as a continuous outcome. The choice of a five-item subscale as the sole outcome requires additional justification. The authors should discuss whether the Emotional Symptoms subscale provides adequate reliability and dimensional precision for machine-learning prediction, especially because the outcome range is restricted to 0–10. It would also be useful to report whether the French adolescent self-report version has validated cut-points and whether the authors considered predicting the broader SDQ internalizing score, a clinically significant category, or a validated depression/anxiety measure.

The measurement section states that "The French-language items were administered using established translation procedures" for the Parental Authority Questionnaire, but this wording is too vague. The authors should specify whether an officially validated French version was used or whether the research team translated the scale specifically for the present study. If translation was undertaken, the manuscript should describe forward translation, backward translation, expert review, pilot testing, and any cultural adaptation. This is important because parenting-style constructs may have culturally specific interpretations, and measurement error introduced by unvalidated translations could substantially affect machine-learning feature importance.

Response: Revised and uploaded the manuscript.

2. Revised

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

AFTJ Applied Family Therapy Journal
E-ISSN: 3041-8798