Journal Website

Article history:
Received 06 July 2025
Revised 14 October 2025
Accepted 02 November 2025
Published online 01 January 2026

Journal of Psychosociological Research in Family and Culture

Open Peer-Review Report

E-ISSN: 3041-8550

Machine Learning Identification of Cultural Value Systems Using Collectivism, Power Distance, and Uncertainty Avoidance

Martín Echeverría[1], Chloe. Bennett[2*], Marco Quispe[3]

[1] Department of Psychology, University of the Republic, Montevideo, Uruguay
[2] Department of Counseling Psychology, University of Ottawa, Ottawa, Canada
[3] Department of Psychology, Universidad Mayor de San Andrés, La Paz, Bolivia

*** Corresponding author email address**: chloe.bennett@uottawa.ca

Editor	Reviewers
Sergii Boltivets Chief Researcher of the Department of Scientific Support of Social Formation of Youth. Mykhailo Drahomanov University, Ukraine sboltivets@ukr.net	**Reviewer 1:** Ali Khodaei Department of Psychology, Faculty of Educational Sciences and Psychology, Payam Noor University, Tehran, Iran. Email: alikhodaei@pnu.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:

In the introduction, the statement “traditional analytical approaches have often struggled to capture the complex, nonlinear, and interactive nature of cultural value systems” is conceptually appropriate but insufficiently substantiated, and it would benefit from a more rigorous methodological critique (e.g., limitations of SEM, regression assumptions, or multicollinearity issues) supported by empirical citations to strengthen the theoretical justification for machine learning adoption.

In the paragraph beginning “Collectivism represents the degree to which individuals prioritize group cohesion…,” the construct definition is adequate, but the operationalization is not sufficiently aligned with the later measurement section, and the authors should explicitly bridge theoretical definitions with measurement indicators to ensure construct validity and avoid conceptual–operational gaps.

The sentence “Despite the growing body of research on cultural values, several methodological limitations persist” introduces an important critique, yet it remains too general; I recommend specifying concrete limitations (e.g., measurement invariance across cultures, response bias, common method variance) and linking them directly to how machine learning addresses each limitation.

In the discussion, the claim "The descriptive results indicated moderate variability across all three cultural dimensions within the Canadian sample" would benefit from statistical support (e.g., variance measures, distribution plots), as "moderate variability" is currently an interpretive statement without quantitative backing.

The sentence "ensemble-based models… achieved superior performance" is accurate but incomplete, and the authors should discuss potential overfitting risks associated with ensemble methods and how cross-validation mitigated these risks, including reporting variance across folds.

The paragraph stating "These findings have important implications for understanding how cultural values influence behavior across domains" is overly broad, and I recommend narrowing the scope by linking specific model outputs (e.g., cluster membership probabilities) to concrete applied contexts such as policy segmentation or targeted interventions.

Response: Revised and uploaded the new document.

1.2. Reviewer 2

Reviewer:

In the methods section, the phrase "The sample size was determined based on both statistical power considerations for multivariate modeling and the requirements of machine learning algorithms" is insufficiently precise, and the authors should provide a formal justification (e.g., power analysis parameters, minimum sample size heuristics for ML models such as events-per-variable or learning curves).

Within the "Measures" subsection, the statement "All instruments have demonstrated robust psychometric properties in prior cross-cultural research" lacks specificity, and it is necessary to report actual reliability coefficients (Cronbach's alpha for this sample) and confirmatory factor analysis indices to demonstrate internal consistency and construct validity in the present dataset.

In the data analysis section, the sentence "handling of missing values באמצעות multiple imputation" contains a language inconsistency (the term "באמצעות"), which is inappropriate in an English manuscript and should be corrected to maintain linguistic coherence and academic professionalism.

The statement "feature selection techniques, including recursive feature elimination and mutual information criteria, were applied" requires further elaboration, as the authors do not report selection thresholds, ranking criteria, or whether feature selection was nested within cross-validation, raising concerns about potential data leakage.

In Table 1 interpretation, the claim "The strongest correlation was observed between power distance and uncertainty avoidance ($r = 0.48$, $p < 0.01$)" is accurate but underinterpreted, and the authors should discuss potential theoretical implications of this relatively strong association, including whether it suggests latent construct overlap or cultural co-dependence.

The sentence "Logistic regression showed comparatively lower performance, suggesting the presence of nonlinear relationships among variables" is logically plausible but methodologically incomplete, and the authors should statistically verify nonlinearity (e.g., interaction terms, residual diagnostics, or partial dependence plots) rather than inferring it solely from model performance differences.

In Table 2, although multiple performance metrics are reported, there is no mention of class imbalance handling, and the authors should clarify whether techniques such as SMOTE, class weighting, or stratification were applied, as these directly influence accuracy and F1-scores.

In the clustering results, the sentence "Cluster 2 exhibits the highest levels of collectivism, power distance, and uncertainty avoidance" is descriptive but lacks statistical validation, and the authors should report clustering validation metrics (e.g., silhouette score, Davies–Bouldin index) alongside justification for selecting $k = 3$.

The description "Cluster 3 reflects moderate levels across all dimensions, indicating a hybrid cultural orientation" is conceptually appealing but requires stronger empirical grounding, and the authors should consider conducting post hoc statistical comparisons (e.g., ANOVA across clusters) to validate these interpretations.

JPRFC
E-ISSN: 3041-8550

In the feature importance section, the statement “power distance was the most influential predictor” should be complemented with quantitative SHAP summary statistics (e.g., mean absolute SHAP values) to provide a more rigorous and transparent interpretation of variable importance.

The sentence “Uncertainty avoidance showed nonlinear contributions, with both very high and very low values exerting distinct predictive effects” is theoretically significant but insufficiently evidenced, and the authors should include visualization (e.g., SHAP dependence plots) to substantiate this claim.

Response: Revised and uploaded the new document.

2. Revised

Editor’s decision: Accepted.
Editor in Chief’s decision: Accepted.

JPRFC
E-ISSN: 3041-8550