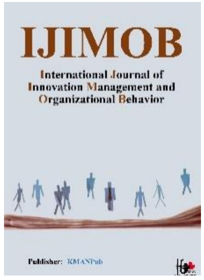
PUBLISHING INSTITUTE
Journal Website
Article history:
Received 20 September 2025
Revised 10 December 2025
Accepted 12 December 2025
Published online 01 April 2026

International Journal of Innovation Management and Organizational Behavior

Open Peer-Review Report




E-ISSN: 3041-8992

# Using Random Forests to Predict Team Creativity from Psychological Diversity and Emotional Intelligence

Mustafa. Al-Hadithi[1], Dimitrios. Papadakis[2*]

[1] Department of Business Administration, University of Mosul, Mosul, Iraq
[2] Department of Business Administration, Athens University of Economics and Business, Athens, Greece

*** Corresponding author email address**: dpapadakis@aueb.gr

| Editor | Reviewers |
|---|---|
| Arpana Rai<br>Assistant Professor, Indian Institute of Management, Udaipur, Mumbai, India<br>arpana.rai@iimu.ac.in | **Reviewer 1:** Mohammad Esmaeil Fadaeinejad<br>Associate Prof., Department of Financial Management and Insurance, Shahid Beheshti University, Tehran, Iran.<br>Email: m-fadaei@sbu.ac.ir<br>**Reviewer 2:** Rezvan Hosseingholizadeh<br>Associate Professor, Department of Educational Management and Human Resource Development, Ferdowsi University of Mashhad, Mashhad, Iran. Email: rhgholizadeh@um.ac.ir |

## 1. Round 1

### *1.1. Reviewer 1*

Reviewer:

The manuscript repeatedly discusses employee innovation behavior (pp.2–3) and then pivots to supervisor-rated team creativity (p.5) without a clean definitional bridge. Because innovation behavior typically includes idea implementation, while creativity focuses on novelty/usefulness generation, you should explicitly define the outcome construct, distinguish it from innovation behavior, and justify why cited studies about innovation behavior (pp.2–3) are relevant to predicting *team creativity* as measured here (p.5).

Emotional intelligence is measured via WLEIS at the individual level and "aggregated to team level given sufficient within-team agreement" (p.5), while psychological diversity is computed as within-team variance/SD of Big Five traits (p.5). However, the paper does not report standard aggregation statistics (e.g., $r_{wg}$, ICC(1), ICC(2)) or the exact decision thresholds used; without these, the legitimacy of team-level EI and the stability of diversity indices across 65 teams is difficult to evaluate (p.5).

The Methods specify RMSE, MAE, and $R^2$(p.5) and claim the model explains "a substantial portion of variance" and outperforms linear models (p.9), but the numerical results, baseline comparisons, and (critically) the specific linear benchmark(s) are not presented in the available results narrative (pp.8–9). The paper should report exact metrics on test data, include a transparent comparison model (e.g., OLS/PLS with interactions), and provide a table of hyperparameters and performance (p.5; p.9).

The discussion argues that variance in conscientiousness harms cohesion and creativity (p.9), but this could depend on task type, team workflow interdependence, and whether strong process leadership exists. Since the context is Greek corporate teams across multiple industries (p.5), consider moderating explanations (e.g., task routineness, deadline pressure, creative vs. compliance-heavy work) and avoid treating conscientiousness diversity as universally harmful without testing interactions (p.5; p.9–p.10).

Authors revised the manuscript and uploaded the new document.

### *1.2. Reviewer 2*

Reviewer:

The modeling unit appears to be teams (65 teams) with an 80/20 split (p.5), implying ~52 training and ~13 test observations—extremely small for tuning Random Forest hyperparameters via "grid search with cross-validation" (p.5). This design can yield unstable performance estimates and variable importance rankings; you should consider repeated cross-validation, nested CV, or leave-one-out/bootstrapped validation at the team level, and report variability/CI of metrics (p.5, pp.8–9).

The paper uses k-NN imputation and standardization (z-scores) (p.5). Random Forests generally do not require feature scaling, and k-NN imputation can distort variance-based diversity indices (especially if imputation is done before computing within-team SDs). Please specify the pipeline order (impute → compute diversity indices → aggregate EI → model?) and justify why these preprocessing steps are appropriate given the constructs (p.5).

The manuscript reports using "mean decrease in Gini impurity" to rank predictors (p.5). Gini impurity is standard for classification; for regression forests, impurity is typically measured via MSE/variance reduction, and importance can be biased in correlated predictors. I recommend reporting permutation importance (and/or SHAP) and clearly stating the exact implementation used, because the substantive claims hinge on EI being the "single most critical determinant" (pp.8–9; p.5).

Several passages read causally (e.g., EI "functions as a regulatory mechanism" that "unlocks" diversity and "drives" creativity) (p.4; p.9), but the design is cross-sectional and predictive, and the authors themselves note causal ambiguity (p.10). Please calibrate language to prediction/association, or add a causal identification strategy (longitudinal/experimental) if causal claims are intended (p.10).

Authors revised the manuscript and uploaded the new document.

## 2. Revised

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



IJIMOB
E-ISSN: 3041-8992