

## Machine Learning Detection of Online Social Exclusion and Its Association with Adolescent Loneliness

Hamish. Rutherford<sup>1</sup>, Fernanda. Ortega<sup>2\*</sup>, Anouk. Vermeulen<sup>3</sup>

<sup>1</sup> Department of Psychology, University of Otago, Dunedin, New Zealand

<sup>2</sup> Department of Social Psychology, National Autonomous University of Mexico (UNAM), Mexico City, Mexico

<sup>3</sup> Department of Developmental Psychology, Utrecht University, Utrecht, Netherlands

\* Corresponding author email address: fernanda.ortega@unam.mx

---

### E d i t o r

### R e v i e w e r s

Gholamreza Rajabi  
Professor of Counseling  
Department, Shahid Chamran  
University, Ahvaz, Iran  
rajabireza@scu.ac.ir

**Reviewer 1:** Karim Afshariniya   
Assistant Professor, Department of Psychology, Kermanshah Branch, Islamic Azad  
University, Kermanshah, Iran Email: k.afsharineya@iauksh.ac.ir  
**Reviewer 2:** Keivan Kakabreae   
Assistant Professor, Department of Psychology, Kermanshah Branch, Islamic Azad  
University, Kermanshah, Iran. Email: keivan@iauksh.ac.ir

---

### 1. Round 1

#### 1.1. Reviewer 1

Reviewer:

The sentence “including cyber-ostracism, silent ignoring, and algorithmically mediated invisibility” introduces a novel term without operational clarification. Please define “algorithmically mediated invisibility” more precisely and explain whether it was empirically measured in this study or referenced conceptually.

The incremental  $\Delta R^2 = 0.06$  is statistically significant. However, please contextualize its practical significance. Is a 6% increase substantively meaningful in adolescent mental health prediction models? Compare effect magnitude with prior studies.

Authors uploaded the revised manuscript.

#### 1.2. Reviewer 2

Reviewer:

In the paragraph discussing cyberbullying literature, the manuscript states that exclusion and aggression may operate through distinct behavioral patterns. However, the operational distinction in your ML model is unclear. Please clarify whether exclusionary linguistic markers were explicitly differentiated from aggressive or hostile language during manual coding.

In Table 1, you report ML-detected exclusion probability ( $M = 0.37$ ,  $SD = 0.18$ ). The interpretation that “approximately one-third of analyzed digital interactions were classified as containing exclusionary markers” appears speculative. Clarify whether this reflects interaction-level probability or participant-level aggregated score.

Given that self-reported exclusion and ML-detected exclusion correlate at  $r = 0.54$  (Table 2), please report exact VIF and tolerance values to confirm absence of multicollinearity concerns in Step 3 of the regression model (Table 3).

Authors uploaded the revised manuscript.

## 2. Revised

Editor's decision after revisions: Accepted.

Editor in Chief's decision: Accepted.