

Cross-Lagged Structural Equation Modeling of Loneliness, Problematic Social Media Use, Self-Esteem, and Depressive Symptoms Among Adolescents

Shantel. Rowe¹, Nika. Kovač^{2*}, Marco. Quispe³

¹ Department of Psychology, University of the West Indies, Kingston, Jamaica
² Department of Social Psychology, University of Ljubljana, Ljubljana, Slovenia
³ Department of Psychology, Universidad Mayor de San Andrés, La Paz, Bolivia

* Corresponding author email address: nika.kovac@ff.uni-lj.si

E d i t o r	R e v i e w e r s
John S. Carlson ¹  Distinguished Professor of the Department of Educational Psychology, Michigan State University, East Lansing, MI, United carlosoj@msu.edu	Reviewer 1: Parvaneh Mohammadkhai ¹  Professor, Department of Clinical Psychology, University of Rehabilitation Sciences and Social Health, Tehran, Iran. Email: Pa.mohammadkhani@uswr.ac.ir Reviewer 2: Mahdi Khanjan ¹  Associate Professor, Department of Psychology, Allameh Tabataba'i University, Tehran, Iran. Email: khanjani_m@atu.ac.ir

1. Round 1

1.1. Reviewer 1

Reviewer:

The manuscript needs considerably more detail on the linguistic and cultural validity of all instruments in the Slovenian context. The Data Collection Tools section repeatedly notes that the measures have shown validity “across different cultural contexts,” but it does not state whether validated Slovenian versions of the UCLA Loneliness Scale, Social Media Disorder Scale, Rosenberg Self-Esteem Scale, and PHQ-9 were used. The authors should identify the specific Slovenian translations, describe translation and back-translation procedures if new translations were prepared, and provide evidence of psychometric validity in Slovenian adolescents. This is especially important because longitudinal comparisons assume equivalent interpretation of items across time and cultural appropriateness of the measures.

The treatment of the PHQ-9 raises an ethical and procedural issue that should be addressed explicitly. The instrument includes an item assessing thoughts related to death or self-harm, yet the manuscript does not describe a risk-management protocol for participants endorsing this item. The Ethics subsection should therefore indicate whether responses were reviewed in real time, whether predefined thresholds triggered referral, whether school psychologists or clinical professionals were

available, and how confidentiality was balanced against safeguarding obligations for minors. This information is necessary in research involving adolescents and depressive symptom assessment and should not be omitted from the procedural description.

The paragraph reporting attrition states that “comparisons between participants who completed all three assessments and those who did not complete the final assessment revealed no statistically significant baseline differences,” but no inferential statistics are provided. The authors should report the actual tests, test statistics, degrees of freedom, effect sizes, and p values for attrition comparisons rather than relying on a general statement. In addition, it would be preferable to model attrition more comprehensively, for example through logistic regression predicting Time 3 participation from baseline demographic and psychological variables. Failure to find significant differences in several separate tests does not necessarily establish that missingness is ignorable or fully support the missing-at-random assumption.

The equality-constrained analysis is potentially valuable but is insufficiently documented. The Findings section states that “constraining corresponding longitudinal paths to equality resulted in only a negligible change in global model fit, $\Delta CFI = -.003$ and $\Delta RMSEA = .001$,” yet the manuscript does not report the chi-square difference test, degrees-of-freedom change, or which specific paths were constrained simultaneously versus individually. The authors should provide the complete constrained-model fit statistics and clearly distinguish between omnibus temporal-invariance testing and tests of individual paths. This would permit readers to determine whether the conclusion that cross-lagged effects were stable across intervals is adequately supported.

The treatment of age and gender covariates should be expanded. The Findings section states that “gender was significantly associated with depressive symptoms, with girls reporting higher symptom levels than boys,” while age showed small associations with problematic social media use and depressive symptoms, but exact coefficients, confidence intervals, and p values are not reported. Because adolescence from ages 13 to 18 encompasses substantial developmental variation, age may also moderate rather than merely confound the structural relationships. The authors should report the full covariate estimates and consider testing whether major cross-lagged paths differ by gender or developmental stage, provided the sample size permits statistically credible multi-group or interaction analyses.

Authors uploaded the revised manuscript.

1.2. Reviewer 2

Reviewer:

Table 1 is informative but currently presents a very large 12-variable correlation matrix without confidence intervals, while the manuscript interprets numerous differences in correlation magnitude descriptively. The authors should consider providing 95% confidence intervals for the principal longitudinal correlations or, at minimum, avoid implying meaningful differences between coefficients unless those differences have been formally tested. In addition, because repeated measures of the same constructs are naturally highly correlated, the authors should discuss whether multicollinearity was examined before SEM estimation and whether the correlations created any estimation instability.

The longitudinal measurement invariance analysis requires much greater methodological transparency. The paragraph beginning “The results of the longitudinal measurement analyses are presented in Table 2” reports configural, metric, scalar, and residual invariance, but the Methods section does not explain precisely how the measurement model was specified. The authors should state whether all four constructs were modeled as latent variables using individual items, parcels, or total scores; whether residuals of identical indicators were correlated across waves; which identification method was used; and whether categorical items from the Social Media Disorder Scale were modeled differently from continuous indicators. Without these details, the reported invariance tests cannot be adequately evaluated or reproduced.

There is a potentially important conceptual and statistical inconsistency between the measurement-invariance analysis and the subsequent description of the cross-lagged model. The manuscript reports longitudinal CFA and invariance testing, which implies latent-variable modeling, but Tables 3 and 4 present paths among constructs in a manner that could also reflect observed composite scores. The authors should clearly state whether the final cross-lagged model was a latent-variable CLPM or an observed-score path model. If latent factors were used, factor loadings, residual structures, and latent variances should be

described; if composite scores were used, the rationale for conducting longitudinal measurement invariance prior to an observed-score model should be clarified.

The choice of a traditional cross-lagged panel model should be reconsidered or, at minimum, explicitly defended. The manuscript states that the model “made it possible to determine whether an earlier level of one variable predicted subsequent changes in another variable,” but conventional CLPMs conflate stable between-person differences with within-person temporal dynamics. Consequently, the reported cross-lagged coefficients cannot automatically be interpreted as within-person effects. The authors should either estimate a random-intercept cross-lagged panel model as a robustness analysis or substantially temper statements about individual-level developmental change. Comparison of CLPM and RI-CLPM results would considerably strengthen the article, particularly because all four constructs exhibit strong temporal stability.

The interpretation of loneliness as the “most prominent prospective predictor” should be made more statistically rigorous. The Discussion states that “the strongest prospective cross-domain effects consistently originated from loneliness,” but differences between the standardized coefficients have not been formally tested. For example, the loneliness-to-depression coefficients of $\beta = .21$ and $.22$ are numerically larger than several other paths, but this alone does not demonstrate that they are statistically stronger. The authors should conduct Wald tests or parameter-constraint comparisons if they wish to claim superiority of particular pathways; otherwise, the wording should be revised to indicate that loneliness displayed among the larger and most consistently significant cross-domain associations.

Authors uploaded the revised manuscript.

2. Revised

Editor’s decision after revisions: Accepted.

Editor in Chief’s decision: Accepted.