Journal Website

Article history:
Received 03 August 2026
Revised 04 November 2026
Accepted 12 November 2026
Initial Publication 16 November 2026
Final Publication 01 September 2027

International Journal of Innovation Management and Organizational Behavior

Open Peer-Review Report

E-ISSN: 3041-8992

Validation of an Artificial Intelligence-Based Educational Leadership Model with a Foresight Approach in Schools of Razavi Khorasan Province

Ehsan. Taghaviyan[1], Roya. Afrasiabi[1*], Behrang. Esmaeilishad[1]

[1] Department of Educational Sciences, Boj.C., Islamic Azad University, Bojnord, Iran

*** Corresponding author email address**: afrassiabi@iau.ac.ir

Editor	Reviewers
Davood Feiz Affiliation: Professor of Management Department, Semnan University, Iran feiz1353@semnan.ac.ir	**Reviewer 1:** Marco Yamba-Yugsi Unidad Académica de Posgrado, Universidad Católica de Cuenca, Azuay 010101, Ecuador Email: marco.yamba@ucacue.edu.ec **Reviewer 2:** Manijeh Haghighinasab Assistant Professor, Department of Management, Alzahra University, Tehran, Iran Email: haghighinasab@srbiau.ac.ir

1. Round 1

1.1. Reviewer 1

Reviewer:

The manuscript should correct the conceptual treatment of AVE in the statement that “reliability of the instrument [was confirmed] based on Cronbach’s alpha, composite reliability, and average variance extracted (AVE).” Average variance extracted is an indicator of convergent validity rather than reliability. Cronbach’s alpha and composite reliability assess internal consistency, whereas AVE is used to evaluate the proportion of indicator variance captured by a latent construct. The manuscript should separate the assessment of internal consistency reliability, convergent validity, and discriminant validity throughout the Methods and Results sections and use standard PLS-SEM terminology consistently.

The authors report that face validity, content validity “(CVR and CVI),” and construct validity were confirmed, but the corresponding numerical evidence is not presented. For a researcher-developed instrument, the manuscript should report the number and expertise of the panel members, the decision rule applied to the content validity ratio, the calculation method and acceptable threshold used for the content validity index, and the actual CVR/CVI results. Without these details, readers cannot evaluate whether the questionnaire underwent an adequate content-validation process before factor analysis.

The manuscript reports both exploratory and confirmatory factor analysis but does not make clear whether these analyses were conducted using the same sample of 461 respondents. If the same observations were used first to discover the factor structure and then to confirm it, there is a risk of capitalization on chance and inflated evidence for construct validity. Ideally, the sample should be divided into independent EFA and CFA/PLS measurement-model subsamples, or the factor structure should be confirmed in a separate sample. If sample splitting was not feasible, this limitation should be acknowledged and the authors should explain why both EFA and confirmatory assessment were considered necessary.

The measurement-model evaluation is incomplete in the Results section. The abstract refers to "AVE above 0.59," "HTMT below 0.74," and "Cronbach's alpha above 0.89," yet no complete table of indicator loadings, Cronbach's alpha, rho_A, composite reliability, AVE, and HTMT values is provided in the translated Results. Before interpreting the structural paths, the manuscript should present comprehensive evidence that the reflective measurement models satisfy reliability, convergent validity, and discriminant validity requirements. In particular, HTMT confidence intervals should preferably be reported rather than relying only on point estimates.

The structural diagram and hypothesis structure need to be reconciled. Figure 2 displays PLD with "$R^2 = 0.000$" and appears to include a dashed connection toward professional development, while Table 4 contains no hypothesis predicting PLD and lists only PLD → AIL. At the same time, the figure description states that "five paths" originate from technological competencies, although the reported hypotheses directly connect TEC to ETH, DEC, FUT, and CHM, with no TEC → PLD hypothesis. The authors should revise the figure and narrative so that every displayed path corresponds exactly to a tested hypothesis and every endogenous R^2 value is meaningful.

The bootstrapping procedure in Table 4 needs methodological specification. The heading reports "N = 5,000," but 5,000 appears to refer to the number of bootstrap subsamples rather than the empirical sample size, which is 461. The table caption should therefore state "5,000 bootstrap resamples" rather than "N = 5,000." The authors should also report whether bias-corrected and accelerated or percentile confidence intervals were used, whether bootstrapping was performed with sign changes, and whether hypothesis tests were one-tailed or two-tailed.

Authors revised the manuscript and uploaded the new document.

1.2. Reviewer 2

Reviewer:

The paragraph beginning "Before model validation was conducted, the statistical assumptions were examined" should be reconsidered in relation to PLS-SEM. The manuscript places considerable emphasis on univariate normality based on Kolmogorov–Smirnov tests, although PLS-SEM does not assume multivariate normality in the same manner as covariance-based SEM. More importantly, the sentence "the use of bootstrapping in PLS-SEM adequately addresses concerns associated with non-normal data distributions" is too categorical. Bootstrapping provides nonparametric inference for standard errors and confidence intervals; it does not "correct" the underlying distribution. The authors should moderate this claim and preferably report skewness and kurtosis values if they are being used to support distributional adequacy.

Table 1 and its interpretation require statistical clarification. The manuscript classifies digital ethical leadership ($p = .048$) and strategic foresight ($p = .038$) as "approximately normal," although the Kolmogorov–Smirnov test formally rejects the null hypothesis of normality at $\alpha = .05$ for both variables. If the authors retain this table, they should state that the tests were statistically significant but that the observed departures were not considered problematic for the selected PLS-SEM estimator. The phrase "approximately normal" should not be presented as the direct outcome of a statistically significant normality test without supporting evidence from skewness, kurtosis, Q-Q plots, or distributional diagnostics.

There appears to be a numerical inconsistency in Table 2 that should be checked carefully. The manuscript states that exploratory factor analysis was conducted on "50 questionnaire items," yet Bartlett's test is reported with 660 degrees of freedom. For a correlation matrix containing 50 observed items, Bartlett's test would ordinarily have $50 \times 49 / 2 = 1{,}225$ degrees of freedom. The reported df = 660 therefore does not appear compatible with a 50-item EFA. The authors should verify whether

IJIMOB
E-ISSN: 3041-8992

the analysis was actually conducted on a smaller number of items, whether items had already been removed, or whether the reported degrees of freedom are a transcription error.

The exploratory factor analysis procedure requires fuller reporting beyond KMO and Bartlett's test. The manuscript states that "Table 3 presents the complete factor-loading matrix… after Varimax rotation," but it does not specify the factor extraction method, the criterion used to determine the number of factors, eigenvalues, scree-plot results, parallel analysis, or whether communalities were evaluated. Because retaining exactly seven factors may reflect the expected conceptual model, the authors should demonstrate that the empirical factor-retention evidence independently supported seven factors rather than imposing the theoretically expected structure.

The choice of Varimax rotation should be theoretically justified. Constructs such as technological competencies, ethical leadership, strategic foresight, change management, professional development, and AI-based educational leadership would reasonably be expected to correlate. Varimax assumes orthogonal factors and may therefore be less appropriate than an oblique rotation such as Promax or Direct Oblimin. The authors should either justify why factor independence was assumed or repeat the EFA using an oblique rotation and report the factor-correlation matrix. This is particularly important because the subsequent structural model explicitly assumes substantial relationships among these constructs.

There is an internal inconsistency between the narrative description of Table 3 and the displayed factor loadings. The text states that "the highest factor loading (0.852) belonged to item ETH2" and "the lowest loading (0.724) belonged to TEC5," whereas Table 3 reports the values only as 0.85 and 0.72. The authors should use a consistent number of decimal places throughout the text and tables. More importantly, if exact loadings were 0.852 and 0.724, these precise values should appear in the table rather than rounded values if they are subsequently interpreted as evidence that all loadings exceed a specific threshold.

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