Proceedings Academic Universitas Jambi
Volume 2, Issue 3, September, 2026

Developing a HEXACO-PI-R-Based Leadership Style Assessment Scale: An Item Response Theory and Generalizability Theory Approach

Mohamad Muspawi[1*], Mulyadi[2], Masbirorotni[3], Yantoro[4], Ajuzar Fiqhi[5]

[1 2345]Universitas Jambi, Indonesia

[*]Coressponding Authors: mohamad.muspawi@unja.ac.id

Abstract

Background: Educational leadership is crucial in determining strategic direction and school quality. However, current leadership assessment instruments heavily rely on the Big Five model and Classical Test Theory (CTT), frequently failing to capture moral integrity accurately and remaining prone to rater bias. **Objective**: This study aims to develop and validate a robust Leadership Style Assessment (LSA) scale model based on the HEXACO Personality Inventory-Revised (HEXACO-PI-R) utilizing Item Response Theory (IRT) and Generalizability Theory (GT). **Methods**: Employing a Research and Development (R&D) design, this conceptual study targets educational leaders across 302 public schools in Batang Hari Regency, Indonesia. The systematic 7-step procedure includes construct determination, blueprinting, expert validation, field trials, and psychometric finalization. IRT is utilized to calibrate item characteristics for measurement invariance, while GT is applied to evaluate measurement consistency across multiple sources of error variance. **Results**: As a study protocol, the anticipated outcomes include a validated multidimensional scale achieving an Aiken's V > 0.80 for content validity, and high measurement dependability indicated by IRT information functions and GT coefficients exceeding 0.80. **Conclusion**: The proposed LSA scale will provide a highly accurate, bias-free diagnostic profile of ethical leadership, serving as a vital decision-support system for educational talent management.

Keywords: Leadership Assessment; HEXACO-PI-R; Item Response Theory; Generalizability Theory; Educational Management; Moral Integrity.

Cite This Article

Muspawi M., Mulyadi M., Masbirorotni M., Yantoro Y., Fiqhi, A. (2026). Developing a HEXACO-PI-R-Based Leadership Style Asessment Scale: An Item Response Theory and Generalizability Theory Approach. Proceedings Academic Universitas Jambi, 2(3): e12345. DOI: 10.22437/proca.v2i3.60029

Editor

Dr. Azolla Degita Azis, MBA., M.Ak.

Article info

Received: August 05, 2026 . Revised: August 19, 2026 . Accepted: September 02, 2026

© 2026 The Author(s). Published with license by LPPM Universitas Jambi. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0 International). This license enables reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator.

INTRODUCTION

The introduction should explain the problems, scale, chronology, and state of the study. Please note that the last paragraph of the introduction should include the aims or purpose of the study. Educational leadership is globally recognized as a fundamental catalyst in organizational management, driving strategic direction, shaping school culture, and ensuring sustainable continuous quality improvement in education. Despite its critical role, the valid and reliable measurement of leadership styles remains a persistent empirical challenge. A major problem in current assessment practices is the overwhelming reliance on conventional tools that predominantly focus on task-oriented and relationship-oriented behaviors. These instruments frequently fail to capture the critical dimensions of moral integrity and ethics, which are currently at the forefront of educational governance and accountability.

Chronologically, the evaluation of leadership traits over the past two decades has been heavily dominated by the Five-Factor Model (FFM) or Big Five personality traits. While the Big Five has provided a useful foundational framework, recent psychometric literature consistently highlights its limitations in predicting counterproductive work behaviors and ethical leadership. In response to this gap, the state of the study has shifted towards the HEXACO personality model. The HEXACO Personality Inventory-Revised (HEXACO-PI-R) explicitly integrates a sixth dimension, Honesty-Humility, which has demonstrated superior predictive validity regarding authentic, ethical, and non-exploitative leadership behaviors compared to the traditional FFM.

Beyond theoretical limitations, a significant methodological problem persists in the scale of current leadership measurements. Most existing instruments in Indonesia rely heavily on Classical Test Theory (CTT). Instruments developed under CTT are inherently sample-dependent, meaning item difficulty and discrimination parameters fluctuate based on the test-takers' characteristics. Furthermore, CTT assumes a uniform standard error of measurement across all individuals, rendering the resulting leadership profiles susceptible to rater bias and imprecision.

Addressing both the theoretical and methodological gaps, the purpose of this study is to develop and validate the Leadership Style Assessment (LSA) scale model based on the HEXACO-PI-R framework, utilizing the advanced dual psychometric approach of Item Response Theory (IRT) and Generalizability Theory (GT).

METHODS

Study design and setting

This study employs a Research and Development (R&D) design focused on the construction, calibration, and standardization of a psychological measurement instrument. The study is situated within the context of the public education sector, targeting educational leadership evaluation.

Population, samples and sampling

The target population for this study comprises the executive leaders—specifically principals—of 302 public primary and secondary schools in Batang Hari Regency, Jambi Province, Indonesia. To fulfill the strict assumptions required for advanced psychometric modeling, the sample size is mathematically determined using the item-to-respondent ratio formula: $n = k \times i$, where n is the total sample size, i is the number of instrument items, and k represents the constant ratio. This ensures the stability of item parameter estimations during IRT analysis.

Instruments and criteria

The primary instrument developed is the Leadership Style Assessment (LSA) scale. The scale is theoretically anchored in the six domains of the HEXACO-PI-R model: Honesty-Humility, Emotionality, Extraversion, Agreeableness, Conscientiousness, and Openness to Experience. Each domain is operationalized into behavioral indicators specific to school management contexts. The instrument utilizes a standardized Likert- type scale to capture attitude and behavioral tendencies. Validity is initially ensured through expert judgment (quantified via Aiken's V), while reliability is established through advanced probabilistic models rather than traditional internal consistency metrics.

Procedure and data collection

The research procedure follows a systematic 7-step R&D framework: (1) determining the HEXACO-PI-R- based leadership construct, (2) formulating the instrument blueprint, (3) drafting the attitude scale items, (4) conducting expert judgment validation, (5) administering the initial field trial to the target principals, (6) executing psychometric analysis and revising the instrument based on empirical data, and (7) finalizing the standardized instrument.

Statistical analysis

Data analysis utilizes a multidimensional psychometric approach. Item Response Theory (IRT) is applied to calibrate item characteristics, specifically discrimination and difficulty parameters, ensuring measurement invariance across different respondent traits. Generalizability Theory (GT) is employed to decompose and evaluate measurement consistency by isolating various sources of error variance simultaneously (e.g., rater, item, and time variance). Additionally, Confirmatory Factor Analysis (CFA) is conducted to ensure the empirical factor structure strictly aligns with the theoretical HEXACO model. Statistical software such as R (utilizing specific packages for IRT and GT) will facilitate the analysis.

Ethical concideration

The study is conducted in accordance with the Declaration of Helsinki. The research protocol, encompassing the conceptual design and intended data collection procedures, was reviewed and approved by the Ethics Committee of Universitas Jambi, ensuring that all participants' rights, confidentiality, and professional integrities are strictly protected.

RESULTS

As this manuscript serves as a conceptual paper and study protocol, empirical data collection is currently in the preparatory phase. Consequently, the results presented herein reflect the rigorous hypothesized psychometric targets established for the finalized scale. The anticipated outcomes focus on achieving a robust, bias-free measurement model capable of precision profiling. Table 1 outlines the targeted psychometric parameters that the LSA scale must satisfy to be deemed standardized for widespread implementation.

Table 1. Anticipated Psychometric Targets for the LSA Scale

Psychometric Domain	Analytical Approach	Targeted Threshold / Criteria
Content Validity	Expert Judgment (Aiken's V)	V > 0.80 indicating strong expert consensus
Construct Validity	Confirmatory Factor Analysis (CFA)	RMSEA ≤ 0.08; CFI ≥ 0.90; Factor loadings (λ) ≥ 0.50
Item Calibration	Item Response Theory (IRT)	Invariant difficulty and discrimination parameters
MeasurementReliability	IRT Information Function & SEM	Test Information > 0.80; SEM < 0.5 across trait levels
ScoreDependability	Generalizability Theory (GT)	G−Coefficient ≥ 0.80 across error variance sources

DISCUSSION

The conceptualized development of the HEXACO-based Leadership Style Assessment scale represents a significant departure from traditional evaluative paradigms in educational management. By anchoring the instrument in the HEXACO-PI-R framework, this study proactively addresses the pressing need to evaluate moral integrity—specifically through the Honesty-Humility trait. In contemporary educational governance, identifying leaders who are not only administratively competent but also inherently ethical is paramount. The anticipated findings align with recent macro-theories suggesting that moral dimensions provide a more accurate predictor of long-term organizational health than basic managerial competencies.

Methodologically, the transition from CTT to a combined IRT and GT framework provides a revolutionary advantage. While CTT evaluates reliability globally, the proposed IRT approach allows for local reliability estimation, ensuring that the scale measures both highly effective and highly ineffective leadership behaviors with equal precision. Concurrently, the application of GT addresses the multifaceted nature of educational assessments, where rater bias and contextual variances frequently distort results. By isolating these error sources, the finalized scale promises unparalleled dependability.

Despite these strengths, the study acknowledges certain limitations. As an attitude scale, the instrument is inherently subject to social desirability bias, especially among executive leaders. While the Honesty- Humility dimension of the HEXACO model naturally mitigates some of this bias by detecting deceptive tendencies, future empirical iterations may benefit from integrating multi-rater 360-degree feedback to cross-validate self-reported profiles.

CONCLUSIONS

This study outlines a robust methodological protocol for developing a highly accurate, bias-free Leadership Style Assessment scale anchored in the HEXACO-PI-R framework. By integrating Item Response Theory and Generalizability Theory, the

research aims to overcome the traditional limitations of Classical Test Theory and the Big Five model, specifically by bringing ethical and moral leadership traits to the forefront of assessment. Practically, this standardized instrument is projected to serve as a vital decision-support system for educational policymakers. It is highly recommended that future research translates this validated scale into a Computerized Adaptive Testing (CAT) platform, further optimizing talent management and leadership selection in the education sector.

CONFLICT OF INTEREST

"The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest."

FUNDING

"This research was financially supported by the Institute for Research and Community Service (LPPM) of Universitas Jambi, under the 2026 Fiscal Year Budget (DIPA), Grant Number: SP DIPA-139.03.2.693426/2026, dated December 1, 2025."

DECLARATION OF ARTIFICIAL INTELLIGENCE USE

"This study used artificial intelligence (AI) tools and methodologies in the following capacities: **Manuscript writing support**: AI-based language model, Gemini was employed to do language refinement (improving the grammar, sentence structure, and readability of the manuscript).

PLUS

"We confirm that all AI-assisted processes were critically reviewed by the authors to ensure the integrity and reliability of the results. The final decisions and interpretations presented in this article were solely made by the authors."

REFERENCES

[1] M. Muspawi, K. Anwar, dan K. Imron Rosyadi, "The Influence of Headmaster Leadership and Teacher Pedagogical Competence on Teacher Performance," 2020.

[2] S. Suryadi, M. Muspawi, dan K. A. Rahman, "The Influence of Leadership Style on Organizational Culture at SMA Negeri 3 Muaro Jambi," *JMKSP (Jurnal Manajemen, Kepemimpinan, dan Supervisi Pendidikan)*, vol. 9, no. 1, 2024, doi: 10.31851/jmksp.v9i1.13492.

[3] N. L. A. N. I. Sari, M. Muspawi, dan L. Pratama, "Transformational Leadership of School Principals in the Implementation of the Independent Curriculum," *Thawalib J. Kependidikan Islam*, vol. 6, no. 1, 2025, doi: 10.54150/thawalib.v6i1.585.

[4] M. Masbirorotni, R. Pratama, N. Fajaryani, dan M. Abrar, "The Transformational Leadership and School Culture on The Learning Organization Capacity in Vocational Schools," *Edukasi J. Pendidik. dan Pengajaran*, vol. 9, no. 2, 1970, doi: 10.19109/ejpp.v9i2.13248.

[5] S. Summiyani, F. Chaniago, H. Hilmi, A. Fiqhi, dan I. N. Hazairin, "Peran Kepala Sekolah dalam Pembinaan Akhlak Siswa melalui Pengelolaan Kantin Kejujuran di Sekolah Menengah Atas 5 Jambi," *EDUKATIF J. ILMU Pendidik.*, vol. 4, no. 2, 2022, doi: 10.31004/edukatif.v4i2.2482.

[6] G. S. et Al, "Penerapan gaya kepemimpinan dan perilaku kepemimpinan yang efektif untuk meningkatkan kinerja guru," *J. Media Inform.*, vol. 6, no. 2, 2025.

[7] K. Leithwood, A. Harris, dan D. Hopkins, "Seven strong claims about successful

school leadership revisited," *Sch. Leadersh. Manag.*, vol. 40, no. 1, 2020, doi: 10.1080/13632434.2019.1596077.
[8] D. M. Netolicky, *Transformational professional learning: Making a difference in schools*. 2019. doi: 10.4324/9780429324277.
[9] R. Y. Deluma, Z. MS, dan M. Asmawi, "The Effect of Principal Leadership Style, Work Motivation, and Professional Competence of Primary School Teacher Performance in Kendari City," *JETL (Journal Educ. Teach. Learn.*, vol. 5, no. 1, 2020, doi: 10.26737/jetl.v5i1.1134.
[10] F. Achmadan, E. S. Jaya, dan S. R. Asih, "Properti psikometris Big Five Inventory–2 versi Indonesia," *J. Psikol. Sos.*, vol. 20, no. 2, 2022, doi: 10.7454/jps.2022.15.
[11] S. Huda dan E. Sulistyaningrum, "Pengaruh Kepribadian terhadap Upah Pekerja di Indonesia: Analisis Big Five Personality Traits," *J. Ekon. Indones.*, vol. 12, no. 3, 2024, doi: 10.52813/jei.v12i3.364.
[12] J. L. Pletzer, M. Bentvelzen, J. K. Oostrom, dan R. E. de Vries, "A meta-analysis of the relations between personality and workplace deviance: Big Five versus HEXACO," *J. Vocat. Behav.*, vol. 112, 2019, doi: 10.1016/j.jvb.2019.04.004.
[13] I. Zettler, I. Thielmann, B. E. Hilbig, dan M. Moshagen, "The Nomological Net of the HEXACO Model of Personality: A Large-Scale Meta-Analytic Investigation," *Perspect. Psychol. Sci.*, vol. 15, no. 3, 2020, doi: 10.1177/1745691619895036.
[14] A. Z. Irta, W. M. Angguna, dan Z. Z. Irawan, "Adaptasi dan Properti Psikometrik Alat Ukur Kepribadian HEXACO Versi Indonesia," *J. Pendidik. dan Konseling*, vol. 5, no. 2, hal. 2577–2588, 2023.
[15] M. Chajewski, "Classical test theory," in *International Encyclopedia of Education: Fourth Edition*, 2022. doi: 10.1016/B978-0-12-818630-5.10008-9.
[16] D. McNeish dan D. Dumas, "Reliability representativeness: How well does coefficient alpha summarize reliability across the score distribution?," *Behav. Res. Methods*, vol. 57, no. 3, 2025, doi: 10.3758/s13428-025-02611-8.
[17] S. E. Embretson dan S. P. Reise, *Item response theory for psychologists*. 2013. doi: 10.4324/9781410605269.
[18] G. J. Bean dan N. K. Bowen, "Item Response Theory and Confirmatory Factor Analysis: Complementary Approaches for Scale Development," *J. Evidence-Based Soc. Work (United States)*, vol. 18, no. 6, hal. 597–618, 2021, doi: 10.1080/26408066.2021.1906813.
[19] A. Khodi, "The affectability of writing assessment scores: a G-theory analysis of rater, task, and scoring method contribution," *Lang. Test. Asia*, vol. 11, no. 1, 2021, doi: 10.1186/s40468-021-00134-5.
[20] A. Pundt, "Leadership style assessment (LSA)," *Zeitschrift fur Arbeits- und Organ.*, vol. 61, no. 3, 2017, doi: 10.1026/0932-4089/a000245.
[21] K. Shams, T. R. Steffensmeier, dan A. J. Wefald, "Bridging the Assessment Gaps: The Engaged Validation of Adaptive Leadership Behavior and Self-Awareness Scales," *J. Leadersh. Organ. Stud.*, vol. 32, no. 1, 2025, doi: 10.1177/15480518241296665.
[22] S. Hassan, M. F. Malik, S. Raza, M. Suhardi, dan W. Merdiani, "Personality and Humbleness: The Role of the HEXACO Model of Personality in Development of Humble Leaders," *SAGE Open*, vol. 13, no. 4, hal. 1–13, 2023, doi: 10.1177/21582440231216172.
[23] M. Tarigan dan F. Fadillah, "Uji Validitas Eksternal Tes Kepribadian HEXACO-60," *Intuisi J. Psikol. Ilm.*, vol. 13, no. 1, hal. 38–47, 2021, doi:

10.15294/intuisi.v13i1.28408.
[24] A. McCulloch, “Measurement Models for Psychological Attributes,” *J. R. Stat. Soc. Ser. A Stat. Soc.*, vol. 187, no. 2, 2024, doi: 10.1093/jrsssa/qnad107.
[25] M. Amirrudin, K. Nasution, dan S. Supahar, “Effect of Variability on Cronbach Alpha Reliability in Research Practice,” *J. Mat. Stat. dan Komputasi*, vol. 17, no. 2, 2020, doi: 10.20956/jmsk.v17i2.11655.
[26] S. Ntumi dan T. Bulala, “Multilevel Psychometric Analysis of Clinician Burnout Using Bayesian G-Theory and IRT Fusion Models,” *Inquiry*, vol. 62, 2025, doi: 10.1177/00469580251371389.
[27] W. P. Vispoel, “Using Generalizability Theory to Estimate the Dependability of Individual Cut Scores within Biomedical and Other Assessments,” *Biomed. J. Sci. Tech. Res.*, vol. 63, no. 1, 2025, doi: 10.26717/bjstr.2025.63.009827.
[28] M. Muspawi, M. Sobri, A. Muh, dan Z. Mustofa, “Development of an Islamic Religious Mentoring System (SIMA ’ I) based on a Mobile Application System Using the ADDIE Model,” *Tadris J. Pendidik. Islam*, vol. 19, no. 1, 2024.