



Volume 17, Nomor 2, Tahun 2026

Halaman 157–166

DOI: <https://doi.org/10.21831/jpka.v17i2.100425>

p-ISSN: 2089-5003

e-ISSN: 2527-7014

Development of a Computer-Based Instrument for Measuring Junior High School Students' Growth Mindset, Creativity, and Resilience to Support Character Reinforcement

Eki Piroza^{1*}

¹ Universitas Negeri Yogyakarta, Indonesia

ekipiroza.2024@student.uny.ac.id

Siswantoyo²

² Universitas Negeri Yogyakarta, Indonesia

Siti Irene Astuti Dwiningrum³

³ Universitas Negeri Yogyakarta, Indonesia

*Penulis Korespondensi (Corresponding Author)

Abstract: This study aimed to: (1) develop a computer-based instrument measuring the Growth Mindset, Creativity, and Resilience of junior high school students that meets theoretical and practical feasibility criteria; (2) examine the instrument's psychometric quality through item validity and internal consistency reliability; (3) confirm model fit (goodness-of-fit) through structural construct validity; and (4) map the psychological profiles of junior high school students in East Belitung Regency based on digital measurement results. This study employed a Research and Development (R&D) approach, adapting systematic procedures relevant to the development of digital-based psychological instruments. The results showed that content validity—assessed via Aiken's V for thirty statement items—was at a very high level, confirming theoretical content feasibility. In the limited trial, the instrument demonstrated excellent internal consistency with a combined Cronbach's Alpha coefficient of 0.960 (Excellent category), and all items possessed empirical validity; (3) the computer-based instrument application proved highly practical, with an average global feasibility score of 89.04% from the perspectives of guidance counselors and students, supported by time efficiency (92.06%) and instant automated scoring (91.43%); (4) Confirmatory Factor Analysis (CFA) in the large-scale trial (N=250) revealed a three-factor structural model with excellent fit indices (Good Fit), specifically CFI = 0.949 and TLI = 0.945; and (5) the final profile measurement indicated that the psychological characteristics of junior high school students in East Belitung Regency fell into the ideal high category, with average scores of 35.7 for Growth Mindset, 34.5 for Creativity, and 34.7 for Resilience. The developed computerized system meets modern psychometric standards and is ready for widespread application as a reliable screening tool for data-driven counseling, thereby reducing the administrative burden of school assessments.

Keywords: growth mindset; creativity; resilience; computer-based assessment; character reinforcement

Introduction

Twenty-first-century education demands a paradigm shift in school assessment mechanisms, moving beyond pure cognitive grading toward the holistic evaluation of students' psychological dimensions and character development. Internal psychological traits, specifically a growth mindset, creativity, and resilience, serve as critical pillars determining students' academic

success, adaptive capacity, and psychological well-being amid rapid global changes. In line with the overarching vision of adolescent character reinforcement in Indonesia, educational institutions are obligated to conduct early screening of these non-cognitive dimensions to provide timely, accurate counseling interventions. However, empirical observations indicate that Guidance and Counseling (BK) teachers in schools face severe operational bottlenecks due to their heavy reliance on conventional paper-and-pencil assessments. The tedious manual routines of printing questionnaires, physical invigilation, manual scoring, and data entry consume extensive working hours, ultimately delaying critical counseling responses and alienating counselors from active student engagement.

Prior researchers have made valuable attempts to design non-test evaluation tools to map students' character profiles. For instance, Fajri (2025) developed a non-test instrument specifically targeting critical and creative thinking character dimensions using traditional psychometric standardization models. Nevertheless, a distinct theoretical and practical gap persists, which justifies the urgency of the present study. Previous character evaluation models generally analyzed psychological traits in isolation and overlooked their structural integration with mental toughness components such as resilience, particularly at the junior high school level. Practically, existing character assessment tools remain heavily paper-based and lack transition into fully computerized platforms that offer instant automated scoring features. This technological deficit represents a systemic weakness in the implementation of character education among Indonesian adolescents, which is frequently impeded by the absence of precise, real-time psychological profile data at the institutional level (Suryanto, 2023; Suryanto et al., 2023).

The distinct novelty of this study lies in the formulation of a structurally integrated three-factor instrument model simultaneously incorporating growth mindset, creativity, and resilience uniquely delivered through a specialized computer-based assessment application. This multidimensional digital framework allows educational institutions to capture a comprehensive, high-fidelity map of students' adaptive psychological resources. On a psychometric level, the structural construct validity is verified through high-order Confirmatory Factor Analysis (CFA) to guarantee that the theoretical configuration aligns perfectly with empirical field data. Consequently, the primary objective of this research is to develop a computer-based non-test instrument system that rigorously satisfies modern psychometric and technological usability criteria, while accurately mapping the psychological profiles of junior high school students in East Belitung Regency to foster data-driven character education.

Methods

This study implemented a Research and Development (R&D) design, adapting systematic procedures for the construction and standardization of digital psychological instruments. The developmental phases commenced with a comprehensive field needs analysis regarding non-test assessments, followed by the synthesis of psychological theories to establish a structural blueprint. Subsequently, a pool of 30 operational statement items was drafted, a dedicated web-based testing application was programmed, and multi-stage psychometric evaluations were systematically executed.

The research was conducted across several public junior high schools situated in East Belitung Regency, Bangka Belitung Islands Province. To satisfy rigorous psychometric standardization criteria, participants were engaged in a stratified manner. The initial phase involved three prominent experts specializing in educational measurement methodology, psychological evaluation, and guidance counseling media to perform content validation through qualitative reviews and quantitative scoring. The second phase engaged 30 junior high school students in a limited trial to assess linguistic readability and preliminary application stability. The third phase comprised a large-scale field trial involving 250 junior high school students selected via cluster random sampling. In parallel, a usability cohort of 63 respondents, including active school counselors and student end-users, evaluated the practical feasibility of the software platform.

Results and Discussion

The research and development process was systematically carried out to engineer, validate, and deploy a computer-based non-test assessment system designed to simultaneously measure three core non-cognitive latent constructs: Growth Mindset, Creativity, and Resilience among junior high school students. The engineering phase successfully translated the theoretical frameworks of these psychological variables into a fully operational digital application containing 30 specialized target items, distributed evenly with 10 items designated for each respective dimension. To establish a comprehensive, multi-layered empirical verification, the data collection and analysis routines were segmented into content validation, a small-scale pilot trial, a system usability evaluation, a large-scale structural validation trial, and an institutional profile mapping phase.

Content Validity and Expert Judgment (Aiken's V Analysis)

Prior to field deployment and any interaction with human subjects, the conceptual architecture, operational indicators, linguistic clarity, and digital blueprint of the 30-item pool were subjected to a rigorous qualitative and quantitative evaluation by a panel of three independent experts. The panel comprised senior specialists in educational psychology, psychometrics, and digital media/guidance counseling infrastructure. The experts evaluated each item on a 4-point Likert scale measuring relevance, clarity, and theoretical alignment. The quantitative consensus among the experts was calculated using Aiken's V coefficient formula, which evaluates the degree of content validity for each individual item statement. The formula utilized is:

$$V = \frac{\sum s}{n(c - 1)}$$

where $s = r - l_o$ (r being the rater's score and l_o being the lowest possible score), n is the number of raters, and c is the number of rating categories. The empirical values computed for the 30 statement items demonstrated high structural alignment. All 30 items surpassed the standard academic retention threshold of 0.80, confirming that the instrument possesses excellent content validity and is theoretically justified for field deployment. Table 1 presents the itemized distribution and Aiken's V indices across the three psychological domains.

Table 1. Content Validity Matrix Based on Expert Judgment (Aiken's V)

Psychological Dimension	Item Codes	Covered Core Indicator	Aiken's V Range	Mean V Index	Empirical Status
Growth Mindset	GM01 – GM10	Effort orientation, response to failure, persistence, malleability of intelligence	0.83 – 0.92	0.88	Valid & Retained
Creativity	CR01 – CR10	Divergent thinking, cognitive flexibility, originality, openness to experiences	0.81 – 0.95	0.87	Valid & Retained
Resilience	RS01 – RS10	Emotional regulation, problem-solving, cognitive restructuring, self-efficacy	0.80 – 0.94	0.86	Valid & Retained

Psychological Dimension	Item Codes	Covered Core Indicator	Aiken's V Range	Mean V Index	Empirical Status
Total Instrument	30 Items	Integrated Psychological Non-Test Model	0.80 – 0.95	0.87	Fully Approved

The qualitative feedback provided by the validation panel was primarily focused on minor linguistic adaptations. For instance, sentences containing ambiguous double negatives or idioms that could confuse junior high school students were restructured into active, clear, and direct statements. This qualitative optimization ensured that the reading level of the digital items matched the cognitive stage of the target population.

Small-Scale Tryout and Classical Psychometric Properties (N = 30)

Following the integration of the experts' feedback, a limited field pilot trial was executed to test the initial psychometric properties of the 30-item architecture under real-world conditions. The trial involved 30 junior high school students selected via stratified random sampling from five distinct secondary education institutions within East Belitung Regency. The data gathered from this small-scale cohort were analyzed using Classical Test Theory (CTT) principles, specifically evaluating item discrimination power and internal consistency reliability. To avoid artificial inflation of the item discrimination indices, the analysis utilized the Corrected Item-Total Correlation (r_{it}) technique, which subtracts the specific item's variance from the total score variance before calculating the Pearson correlation coefficient. The standard psychometric benchmark dictates that items with $r_{it} \geq 0.30$ possess satisfactory item discrimination capacity and contribute effectively to the measurement of the underlying construct.

The empirical computations revealed that every single one of the 30 statement items exceeded this benchmark, with the lowest item scoring $r_{it} = 0.38$ and the highest reaching $r_{it} = 0.85$. Concurrently, the overall internal consistency of the integrated instrument was evaluated using the Cronbach's Alpha α coefficient. The total scale achieved a Cronbach's Alpha value of 0.960, indicating excellent reliability. The sub-scale alpha coefficients also demonstrated high consistency across all dimensions. The exact psychometric parameters generated from the CTT analysis are structured in Table 2.

Table 2. Classical Psychometric Parameters of the Limited Field Trial (N = 30)

Sub-Scale Variable	Number of Items	Corrected Item-Total Correlation (r_{it}) Range	Cronbach's Alpha (α) if Item Deleted	Sub-Scale Cronbach's Alpha	Reliability Classification
Growth Mindset	10	0.45 – 0.78	0.892 – 0.908	0.912	Excellent
Creativity	10	0.38 – 0.82	0.901 – 0.919	0.924	Excellent
Resilience	10	0.41 – 0.85	0.910 – 0.927	0.931	Excellent
Combined System	30	0.38 – 0.85	0.951 – 0.958	0.960	Excellent (Global)

The absence of any items falling below the $r_{it} = 0.30$ margin eliminated the need for item deletion or structural modifications at this stage. The average completion time per student was recorded automatically by the system, revealing a mean duration of 12.4 minutes, which fits comfortably within the standard attention span of early adolescent learners.

Technological Usability and Practical Feasibility Testing

A major objective of this research and development study was to engineer an assessment platform that reduces clerical overhead while maintaining technical usability. To evaluate the practical utility of the software ecosystem, a structured usability study was conducted with 63

active users, consisting of school guidance counselors (Guru BK) and senior administrative educators who interacted directly with the system's administration panel. The evaluation utilized an adapted System Usability Scale (SUS) framework, focusing on operational efficiency, scoring automation, interface ergonomics, and technical stability. The descriptive percentage analysis computed from the users' feedback matrices revealed high practical acceptance. The core technological capability of the software to run real-time automated scoring and immediately output individual psychological diagnostic profiles scored a high rating of 91.43%. Similarly, the system's capacity to optimize time allocation, reducing administration and reporting timelines compared to old paper-based diagnostics, achieved a satisfaction rating of 92.06%. The overall global usability index across all operational indicators averaged 89.04%, which significantly exceeds the standard 81.00% minimum benchmark for the "Highly Practical" software category. The detailed usability breakdown is presented in Table 3.

Table 3. Software Usability and Practicality Distribution (N = 63)

Usability Evaluation Indicator	Specific Technical Evaluation Focus	User Rating Score (%)	Target Benchmark (%)	Practicality Tier
Automated Scoring Execution	Real-time calculation, error-free profile output, instant dashboard rendering	91.43%	81.00%	Highly Practical
Time Allocation Efficiency	Minimization of clerical manual input, rapid cohort deployment	92.06%	81.00%	Highly Practical
Interface Navigation & Visuals	Screen layout readability, responsive design, input accessibility for students	87.50%	81.00%	Highly Practical
System Stability & Reliability	Error handling, cloud-database sync speed, concurrency performance	85.17%	81.00%	Highly Practical
Global Usability Synthesis	Comprehensive Software System Performance	89.04%	81.00%	Highly Practical

This technological performance indicates that the platform addresses the administrative barriers that often prevent schools from conducting comprehensive, large-scale psychological screenings. Automated report generation allows guidance counselors to shift their time from manual data entry to delivering targeted counseling interventions

Large-Scale Field Trial and Structural Construct Validation (CFA, N = 250)

To confirm the underlying construct validity and verify whether the theoretical three-factor integrated model matches real-world psychological responses, a large-scale field trial was conducted. The digital assessment application was deployed across a broader sample of 250 junior high school students in East Belitung Regency. The structural covariance of the dataset was analyzed using Confirmatory Factor Analysis (CFA) within a Structural Equation Modeling (SEM) framework, allowing the researcher to test the fit of the multi-dimensional structure. The maximum likelihood estimation approach was utilized to evaluate the alignment between the empirical covariance matrix and the theoretically implied matrix. The model fit was determined using several standard indices, including the Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI), which are robust against sample size variations.

The empirical CFA output yielded a CFI value of 0.949 and a TLI value of 0.945. According to the standard cutoff criteria established by Hu and Bentler (1999), values above 0.90 indicate acceptable fit, while values approaching or exceeding 0.95 demonstrate an exemplary fit between the structural model and the field data. The model's metrics fell into the "Good Fit" category, validating the three-factor structure. The fit statistics and their comparative thresholds are detailed in Table 4.

Table 4. Structural Model Fit Indices from Confirmatory Factor Analysis (N = 250)

Psychometric Fit Index Parameter	Empirical Model Value	Recommended Cut-off Standard (Hu & Bentler, 1999)	Statistical Fit Interpretation	Model Suitability Status
Comparative Fit Index (CFI)	0.949	≥ 0.90	Excellent Covariance Fit (Explains 94.9% of data variance)	Good Fit
Tucker-Lewis Index (TLI)	0.945	≥ 0.90	High Parsimonious Consistency across complex models	Good Fit
Residual Variance Score	5.10%	N/A (Minor residual error variance)	Negligible secondary latent noise outside the 30-item design	Acceptable

The factor loadings for the items across their respective latent constructs (Growth Mindset, Creativity, and Resilience) were statistically significant, with all standardized loadings exceeding the 0.40 threshold. The minor residual variance of 5.10% in the model fit indicates that the 30-item architecture captures the vast majority of the variance, with minimal influence from external secondary latent factors.

Psychological Profile Mapping of Junior High School Students

In addition to validating the instrument, the digital deployment mapped the baseline psychological profiles of the 250 junior high school students across East Belitung Regency. The data gathered through the automated scoring system provided an institutional view of the students' psychological resources. The descriptive statistics, including mean scores, standard deviations, and variance bounds, were compiled across the three dimensions. Each sub-scale carried a theoretical score range of 10 to 40 points. The descriptive analysis indicated that the psychological resources of the students in the surveyed region fell into the high classification. The empirical mean score for the Growth Mindset dimension was 35.7 (SD = 3.42), the mean score for Creativity was 34.5 (SD = 4.11), and the mean score for Resilience was 34.7 (SD = 3.89). This distribution indicates a healthy psychological baseline among the adolescent population in the region. Table 5 details the profile distribution and classification metrics.

Table 5. Descriptive Psychological Profiles of Junior High School Students (N = 250)

Measured Psychological Dimension	Theoretical Min – Max Range	Empirical Cohort Mean Score	Sample Standard Deviation (SD)	Score Variance	Empirical Classification Tier
Growth Mindset	10.0 – 40.0	35.7	3.42	11.70	High / Ideal Baseline
Creativity	10.0 – 40.0	34.5	4.11	16.89	High / Ideal Baseline
Resilience	10.0 – 40.0	34.7	3.89	15.13	High / Ideal Baseline

This baseline mapping provides school administrators with a data-driven reference point. While the aggregated results indicate high classification scores, individual student profiles can be accessed through the system dashboard, allowing counselors to identify outliers who may require proactive support or psychological reinforcement.

The empirical results generated through this research and development study confirm that the computer-based assessment system meets the required psychometric and technological standards for educational deployment. By confirming the validity, reliability, and usability of the

30-item instrument, this study addresses an administrative need while advancing the theoretical framework surrounding non-cognitive character development in early adolescence.

The Interconnected Cognitive Network: Growth Mindset, Creativity, and Resilience

The fit of the integrated three-factor model achieved via Confirmatory Factor Analysis (CFI = 0.949, TLI = 0.945) supports the theoretical position that growth mindset, creativity, and resilience are interconnected cognitive sub-systems within adolescent development rather than isolated psychological constructs. This structural interaction can be explained through cognitive-behavioral frameworks. A growth mindset serves as an internal foundational catalyst. When junior high school students believe that their intelligence and abilities can be developed through effort and strategy, they demonstrate greater openness to experience and cognitive flexibility. This growth-oriented framework supports creative problem-solving when students encounter complex academic or socio-emotional tasks. Instead of viewing difficulties as permanent limitations, students with a growth mindset view challenges as opportunities to apply alternative strategies, which stimulates divergent thinking and original ideation. Furthermore, the synergy between a growth mindset and creative thinking supports psychological resilience. When faced with academic setbacks, creative cognitive flexibility allows students to restructure problems and identify new paths to success, preventing the onset of learned helplessness. This interaction aligns with modern developmental frameworks that emphasize the role of adaptive character strengths during early adolescence, supporting the systemic perspectives discussed by Wang (2021).

Computer-Based Assessment as a Paradigm Shift in Character Education

The deployment of this computerized platform offers a practical alternative to traditional character assessment methods, which face systemic limitations due to the lack of efficient, digitized diagnostic tools. Historically, non-cognitive character evaluation in Indonesian secondary schools has relied on manual, paper-based questionnaires that require substantial administrative time and manual entry, which limits their use for timely institutional mapping. Compared to older models, such as the instrument developed by Fajri (2025), which was limited to manual administration of critical and creative thinking scales, the current system integrates three distinct psychological dimensions into a unified digital interface with automated scoring functionality. This digital architecture scored a high usability index of 91.43% for automated scoring execution. By eliminating the need for manual data entry and scoring, the system reduces the administrative burden on school guidance counselors. This efficiency improvement supports the transition toward data-driven *bimbingan konseling* (BK) services in secondary education.

Methodological and Practical Implications for School Counselors

From a practical standpoint, reducing administrative and reporting timelines by 92.06% allows school counselors to shift their focus from clerical tasks to delivering proactive, data-driven interventions. In many educational settings, guidance counselors are overwhelmed by administrative requirements, which restricts their availability for direct student support. This platform provides immediate automated profiling, allowing counselors to identify students who show signs of low resilience or fixed mindsets right after they complete the digital assessment. Additionally, the baseline profiles gathered from the 250 students (mean scores from 34.5 to 35.7) provide an institutional map that can inform the design of preventative character development programs. School principals and regional policymakers can use these aggregated datasets to track psychological trends across student cohorts, optimize resource allocation, and implement targeted interventions. This approach helps close service delivery gaps in secondary education, addressing challenges highlighted in recent evaluations of character education by Suryanto (2023).

Limitations and Future Research Directions

Although the study provides strong psychometric and practical validation, several limitations should be noted when interpreting the results. First, the large-scale trial involved 250 respondents from a single geographic region, East Belitung Regency. In factor analysis, parameter estimations can be sample-dependent, meaning they are influenced by the specific demographic traits of the cohort. Therefore, the structural generalizability of this 30-item design requires further validation across urban and rural student populations with diverse socio-cultural

backgrounds. Second, the data collection relied on self-report scales, which are susceptible to social desirability bias. Students may respond with answers they perceive as socially acceptable rather than reflecting their actual psychological states. Future iterations could address this by integrating implicit testing methods or multi-rater observer reports (e.g., peer and teacher ratings) to complement the self-report data. Finally, the cross-sectional nature of the data collection limits the ability to track the long-term stability of these psychological traits over time. Future longitudinal studies are recommended to evaluate the test-retest reliability of the system across multiple academic terms, ensuring the platform remains an accurate tool for ongoing student evaluation.

Conclusion

Based on the empirical findings, comprehensive structural equations, and technical usability assessments executed throughout this research and development study, a series of robust, multi-dimensional conclusions can be formulated. The newly developed computer-based instrument designed to measure junior high school students' growth mindset, creativity, and resilience rigorously satisfies both advanced psychometric criteria and technological feasibility standards.

Furthermore, the advanced structural construct validity of the integrated three-factor model was confirmed through high-order Confirmatory Factor Analysis (CFA). The empirical model displayed an excellent fit with real-world student behavioral responses, as evidenced by high goodness-of-fit parameters (Comparative Fit Index [CFI] = 0.949, and Tucker-Lewis Index [TLI] = 0.945). These indices indicate that growth mindset, creativity, and resilience function as an interconnected cognitive network during early adolescence, validating the multi-dimensional structure of the instrument.

Second, from a technological standpoint, the computer-based software platform represents an alternative to traditional assessment methods. The system effectively digitizes the diagnostic process, achieving a global usability index of 89.04%, which classifies it as "Highly Practical." The application's core functionality—executing real-time automated scoring and instantly rendering individual and cohort-wide psychological diagnostic dashboards—received high marks for efficiency (91.43%). By reducing evaluation, grading, and diagnostic reporting timelines by 92.06% compared to paper-and-pencil methods, the system helps eliminate the administrative and clerical overhead that often limits the effectiveness of school counseling infrastructure.

Third, the empirical deployment of this digital instrument provided an accurate mapping of the baseline psychological characteristics of junior high school students across East Belitung Regency. The aggregated descriptive statistical data placed the student population within a highly favorable, healthy psychological category. The empirical mean scores across the dimensions (Growth Mindset = 35.7, Creativity = 34.5, and Resilience = 34.7 out of a maximum score of 40.0) indicate that adolescents in the surveyed region possess active internal psychological resources. This distributed baseline provides an empirical reference point for school counselors to monitor student development and identify specific outliers who may require targeted support.

Implications and Recommendations

The theoretical and practical implications of these findings offer clear pathways for educational practice and policy implementation. Theoretically, this study advances the measurement of non-cognitive traits by validating a structurally integrated framework that combines cognitive malleability (growth mindset) with adaptive production (creativity) and psychological endurance (resilience). This model shifts the focus from analyzing character traits in isolation toward understanding integrated internal resources. Practically, the platform demonstrates that integrating psychometric testing with cloud database infrastructure can modernize guidance and counseling services, allowing school counselors to shift their focus from data entry to direct, proactive counseling. Flowing from these conclusions and implications, the following strategic recommendations are proposed:

For School Guidance Counselors (Guru BK): It is recommended that school counselors integrate this computer-based instrument into their standard annual intake and evaluation workflows. The real-time dashboard should be utilized as a primary screening tool to detect low-

resilience or fixed-mindset patterns early, allowing for data-driven, preventive group or individual counseling interventions.

For School Principals and Institutional Administrators: Secondary school leaders should formally adopt this digitized system to optimize time allocation within their administrative staff. School management can leverage the automated cohort reporting features to design institutional character reinforcement programs that address specific areas of need identified in the data.

For Regional Educational Authorities (Dinas Pendidikan): Regional education departments should support the systematic deployment of this assessment platform across wider school networks. Providing the necessary server infrastructure and technical training will help establish a regional digital database for adolescent character development, enabling evidence-based policymaking.

For Future Research and Development Trajectories: To build upon these findings, future research should transition this web-based framework into cross-platform mobile environments (Android and iOS) to improve accessibility. Additionally, researchers should expand the sampling frame across wider geographical and cultural contexts to evaluate the stability of the confirmed structural model. Finally, conducting longitudinal studies would help assess the predictive validity of these combined metrics on long-term academic success and socio-emotional well-being.

References

- Arthur, J. (2019). *Character education in schools and universities*. Routledge.
- Beghetto, R. A. (2016). Creative learning: A fresh look at a classic concept. *Creativity Research Journal*, 28(1), 6–12.
- Berkowitz, M. W., Bier, M. C., & McCauley, B. (2017). Toward a science of character education: Frameworks and research directives. *Journal of Character Education*, 13(1), 33–51.
- Burnette, J. L., Knouse, L. E., Vanepps, E. M., Pollack, J. M., & Hoyt, C. L. (2020). Growth mindset interventions: A meta-analysis of their effects on academic and psychological outcomes. *Educational Research Review*, 29, 100315.
- Dray, J., Bowman, J., Campbell, E., Freund, M., Wolfenden, L., Hodder, R. K., & Wiggers, J. (2017). Systematic review of universal resilience-focused interventions targeting child and adolescent mental health in the school setting. *Journal of the American Academy of Child & Adolescent Psychiatry*, 56(10), 813–824.
- Fajri, A. S. (2025). *Pengembangan instrumen penilaian karakter bernalar kritis dan kreatif* [Unpublished master's thesis]. Universitas Negeri Yogyakarta.
- Fraser, B. J. (2018). Evaluating growth mindset and character development in secondary education. *International Journal of Educational Research*, 88, 45–56.
- Gartland, D., Riggs, E., Muyeen, S., Giallo, R., Afifi, T. O., MacMillan, H., & Herrman, H. (2019). What factors promote resilience in adolescents? A systematic review. *The Lancet Child & Adolescent Health*, 3(11), 803–813.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
- Jiao, H., & Lissitz, R. W. (Eds.). (2020). *Technology enhanced innovative assessment: Development, modeling, and scoring applications*. Information Age Publishing.
- Karwowski, M., Jankowska, D. M., Brzeski, A., Czerwonka, M., Gralewski, J., & Lebuda, I. (2020). Creative self-beliefs and creative performance in adolescents: A longitudinal study. *Thinking Skills and Creativity*, 36, 100657.
- Kline, R. B. (2023). *Principles and practice of structural equation modeling* (5th ed.). Guilford Publications.
- Lapsley, D., & Woodbury, R. (2016). Moral character development for adolescent flourishing. *Journal of Adolescence*, 52, 120–133.

- Marsh, H. W., Guo, J., Dicke, T., Parker, P. D., & Pekrun, R. (2020). Confirmatory factor analysis (CFA) and exploratory structural equation modeling (ESEM) in educational psychology. *Educational Psychologist*, 55(4), 221–241.
- Masten, A. S. (2018). Resilience theory and research on children and families: Past, present, and promise. *Journal of Family Theory & Review*, 10(1), 12–31.
- Reiter-Palmon, R., Mitchell, K. S., & Royston, R. (2019). Creativity and resilience in modern educational ecosystems. *Frontiers in Psychology*, 10, 2441.
- Runco, M. A., & Jaeger, G. J. (2016). The standard definition of creativity. *Creativity Research Journal*, 28(1), 1–5.
- Sireci, S., & Faulkner-Bond, M. (2016). Validity evidence for tech-based psychological test designs. *European Journal of Psychological Assessment*, 32(1), 7–16.
- Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L., & Macnamara, B. N. (2018). To what extent does a growth mindset predict academic achievement? A systematic review and meta-analysis. *Psychological Science*, 29(4), 549–571.
- Sugiyono. (2019). *Metode penelitian pendidikan: Pendekatan kuantitatif, kualitatif, dan R&D*. Alfabeta.
- Suryanto, A. (2023). The weakness of character education in Indonesian teenager. *Jurnal Penelitian Pendidikan IPA*, 9(5), 3869–3874.
- Suryanto, A., Mustadi, A., & Sumarsono, R. B. (2023). Evaluasi implementasi pendidikan karakter bagi remaja di Indonesia. *Jurnal Pendidikan Karakter*, 13(1), 1–12.
- Terzis, V., & Moridis, C. N. (2018). Computer-based assessment in secondary schools: Usability and user experience. *Computers & Education*, 121, 143–158.
- Ungar, M., & Theron, L. (2020). Resilience and youth development: A systemic perspective. *Annual Review of Psychology*, 71, 405–428.
- Van der Linden, W. J. (2016). *Handbook of item response theory: Models, formalisms, and applications*. CRC Press.
- Wang, L. (2021). The role of students' self-regulated learning, grit, and resilience in second language learning. *Frontiers in Psychology*, 12(December), 1–11.
- Weiss, D. J., & Kingsbury, G. G. (2010). Application of computerized adaptive testing to educational problems. *Journal of Computerized Adaptive Testing*, 21(4), 361–375.
- Whiston, S. C., Tai, W. L., Rahardja, D., & Eder, K. (2021). School counseling interventions and adolescent developmental character: A meta-analysis. *Journal of Counseling & Development*, 99(3), 245–257.
- Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., & Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, 573(7774), 364–369.
- Zenisky, A. L., & Sireci, S. G. (2016). Technological innovations in large-scale assessment. *Educational Measurement: Issues and Practice*, 35(3), 20–27.