



Journal of Education, Teaching, and Learning is licensed under
A [Creative Commons Attribution-Non Commercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

Multiple Choice Question Type Character Self-Assessment Instrument for Upper Grade Elementary School: Validity & Reliability

Dita Hardiani¹⁾, Farida Agus Setiawati²⁾, Asyraf Suryadin³⁾ ✉

¹⁾ State university of Yogyakarta, Indonesia
E-mail: dita0178pasca.2021@student.uny.ac.id

²⁾ State university of Yogyakarta, Indonesia
E-mail: faridaagus@uny.ac.id

✉³⁾ Muhammadiyah University of Bangka Belitung, Indonesia
E-mail: asyraf.suryadin@unmuhbabel.ac.id

✉ Correspondence Author

Article Information:

Received August 20th, 2026

Revised September 4th, 2026

Accepted September 29th, 2026

Keywords: Instrument Development;
Character Self-Assessment; Validity;
Reliability

© **Copyright:** 2023. Authors retain copyright and grant the JETL (Journal of Education, Teaching and Learning) right of first publication with the work simultaneously licensed under a [Creative Commons Attribution License](https://creativecommons.org/licenses/by-nc/4.0/)

Abstract

This research builds upon previous research in the field. The objective of this study is to ascertain the outcomes of the development of character assessment instruments that focus on information regarding the quality of 21 independent character instrument items that have been developed. The character self-assessment instrument employs a multiple-choice question format with three answer choices. The instrument's quality is evaluated through the assessment of its validity and reliability. In terms of validity, the instrument is reviewed from the content and construct aspects. In terms of content, the instrument was analysed using Aikens' V with the assistance of five experts from the fields of measurement, basic education and psychology. This resulted in an Aikens' V coefficient in the valid category with an average of 0.933 for all items. In terms of the construct aspect, the instrument was analysed using confirmatory factor analysis, which resulted in 15 valid items with factor loading above 0.3 and fit to the fit of the construct model used theory. While the reliability of the instrument was analysed using McDonald's omega reliability, which produced a reliability that is categorised as very high with a value of 0.837. Based on the quality obtained during the development stage, the instrument can be declared valid and reliable for the measurement of independent character in students.

How to cite: Hardiani, D., Setiawati, F.A. & Suryadin, A. (2025). Multiple Choice Question Type Character Self-Assessment Instrument for Upper Grade Elementary School: Validity & Reliability. *JETL (Journal of Education, Teaching and Learning)*, 10(2), 530-540. doi: <http://dx.doi.org/10.26737/jetl.v10i2.9696>

INTRODUCTION

The urgency of character as a foundation in learners is very important. This departs from the magnitude of impact of character on the survival of student that can no longer be ignored. One of the many examples of a negative culture that has taken root and continues to grow in educational

circles even at the elementary school level is bullying (Rasmadi, 2024). The passing of time seems to realize many people about the urgency of character. The government, especially the Ministry of education and culture, always integrates various forms of character education in every curriculum that is launched. More details are exemplified in the last curricula that were passed to strengthen character education (Kemendikbud, 2025) and the project to strengthen the student profile of Pancasila (Zuchron, 2021).

The urgency of character that continues to be left unchecked will present degradation which is a description of categorization that is at a level of less or even very less. Especially in the independent character of students. Independent character makes a person not dependent on others or free to make decisions and have confidence. (Haiq et al., 2020). Independence provides self-regulation that can control the response within the self (Jeona et al., 2021). So that it can have a positive impact on students (Volz & Masicampo, 2021). Independent character will foster children's morals and personality, especially in carrying out tasks and activities according to their needs and age stages (Wulandari et al., 2018). The cultivation of independent character will present humans who can read opportunities in the future (Sumahamijaya et al., 2018) so that students can determine careers after completing their education.

Information about learners' independent character can be obtained through assessment activities of each target element. So it requires the help of many assessment treatments to obtain categorization information (Tim Pusat Penilaian Pendidikan. (2019)). The use of independent character instruments that are procedurally appropriate and of good quality will be able to produce a portrait of independent character. The use of procedurally appropriate and good quality independent character instruments will be able to produce a genuine portrait of the independent character of students and then will be the right basis for determining policies to improve the character degradation that occurs.

In the assessment of independent character, the response format should also consider the developmental characteristics of elementary school students. The multiple-choice format was selected because it provides clear and structured response alternatives that can facilitate students in identifying responses that most closely represent their behavior. Previous studies have shown that social and emotional characteristics of elementary school students can be assessed using alternative response formats, including forced-choice and situational judgment items, which allow students to select responses representing their behavioral tendencies (Murano et al., 2021). Moreover, self-report assessment for children should use items and response formats that are appropriate to their cognitive and developmental characteristics to support the validity of the resulting data (Abrahams et al., 2019). Therefore, the multiple-choice format in this study is not intended to distinguish correct from incorrect answers, as in cognitive tests, but to represent different behavioral tendencies related to students' independent character.

The diversity of the target elements of the assessment will require many types of instruments so the development of instruments that are specific to the target characteristics is needed. Instrument development is the process of developing a measuring instrument that is by the will of the researcher because in researching the researcher must carefully select the instrument used. (Paz et al., 2023).

Many stages of instrument development have been proposed such as by (Azwar, 2022; Istiyono, 2020; MacKenzie et al., 2011; Mardapi, 2017). Instrument development generally begins with determining the theoretical construct and then the specification of the instrument to the quality of the instrument developed. Instrument development includes measuring perceptions and behaviors that represent certain constructs, which are confirmed through proving validity, and reliability. (Kristina Mikkonen et al., 2022).. Instruments with good quality proof will be able to produce data. So the validity and reliability of the instrument are very important because they reflect the quality of the instrument.

Validity is the degree of accuracy and accuracy of an instrument in research that is supported by empirical facts and theoretical reasons. (Istiyono, 2020). Validity is also defined similarly by Mardapi (2017) namely as evidence and theoretical support for the interpretation of test scores for the purpose of test development. In the development of validity instruments, it is a must to obtain valid/correct measurement results. (Arikunto, 2018). Validity is divided into 4 types, namely: a) Content validity, which relates to the accuracy of the instrument in measuring the content that should be measured (Sudjana, 2026). b) Predictive validity, relating to the ability of the instrument to show the level of accuracy of the score with future test performance (Rakhmat & Solehuddin, 2006). c) Concurrent validity relates to a measure of how well a particular test correlates with a previous measure (Sekaran, 2006), and d) construct validity, which relates to the extent to which empirical evidence and theoretical foundations support the accuracy of the interpretation and use of measurement scores in accordance with the research objectives and context (Zumbo, 2026).

After validity, reliability is known. Reliability is the degree of consistency between two measurement scores on the same object despite using different measuring instruments and different scales (Reynolds et al., 2010). Reliability in short can be interpreted as a coefficient of constancy or stability of measurement results. Reliability in short can be interpreted as the coefficient of sharpness or stability of measurement results (Retnawati, 2017).

Although previous studies have examined self-report and alternative response formats for assessing social and emotional characteristics of elementary school students, limited attention has been given to the development and psychometric evaluation of a self-assessment instrument specifically designed to measure independent character in upper-grade elementary students. In particular, the literature discussed above has not sufficiently addressed the use of a multiple-choice response format in which the alternatives represent different behavioral tendencies rather than correct or incorrect answers. Moreover, evidence regarding the content validity, construct validity, and reliability of such an instrument remains necessary to ensure that the resulting scores can appropriately represent students' independent character. Therefore, this study addresses this gap by evaluating the psychometric quality of a multiple-choice self-assessment instrument for independent character among upper-grade elementary school students.

Accordingly, this study aims to evaluate the psychometric quality of a multiple-choice self-assessment instrument developed to measure independent character among upper-grade elementary school students. Specifically, the study examines the instrument's content validity, construct validity, and reliability. By providing empirical evidence for these three psychometric properties, this study is expected to contribute a developmentally appropriate assessment instrument that can support the measurement of independent character in elementary education.

METHODS

This research uses descriptive analysis that focuses on the quality of the instrument developed. This study applies the procedure for preparing psychological scales according to Azwar (2022) which consists of 10 stages, namely: a) Determining the measuring domain / selecting theoretical constructs; b) Developing a grid (blue-print), determining test specifications and item specifications; c) Validating behavioral indicators; d) Writing items; e) Reviewing and improving items; f) Item content validation; g) Field test; h) Validating constructs; i) Assembling tests and estimating reliability; j) Establishing empirical validity. Assemble the test and estimate reliability the quality of the instrument consists of content validity, construct validity, and instrument reliability.

The population of this study were all high-grade students at the elementary school level in Yogyakarta City. The research sample was determined by purposive sampling technique which

determined as many as 550 students in grades V and VI from 13 elementary schools in Yogyakarta City. The subjects in the research sample were in the age range of 10-12 years..

This study uses data collection techniques in the form of distributing questionnaires in the form of content review sheet questionnaires given to experts and readability questionnaires and instrument questionnaires that have been developed to test respondent students. In proving the content validity of the instrument, Aiken's validation index was used. Aiken's formula is formulated to calculate the content-validity coefficient based on the results of the assessment of the expert panel. The expert panel involved in the content analysis were teaching lecturers at the University of Yogyakarta and Muhammadiyah University of Bangka Belitung with expertise in measurement, assessment, educational evaluation; basic education; and psychology. The Aiken validation index is formulated as follows (Istiyono, 2020):

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

The standard reference for the validity of items is reviewed according to Aiken's minimum standard table. For a maximum level of relevance in a review format of 4 with 5 experts, the instrument will be valid if it obtains an Aiken's v value of at least 0.86.

Furthermore, construct validity was assessed for items that had been declared valid at the content validity stage. Construct validity was examined using confirmatory factor analysis (CFA) to determine whether the observed indicators adequately represented the hypothesized construct. The standardized factor loading of each item and the overall model fit were evaluated. In this study, a factor loading of 0.30 was used as the minimum retention criterion. Although loadings of 0.50 or higher are generally considered more practically meaningful, a loading of 0.30 has been recognized as the minimum interpretable level in factor analysis (Hair et al., 2019). Hair et al. (2019) also indicate that a loading of 0.30 can be statistically meaningful when the sample size exceeds 350 observations. Given the relatively large sample in this study (N = 550), the 0.30 threshold was used as a minimum screening criterion rather than as evidence of strong indicator quality. Item retention was therefore considered together with the previously established content validity, overall CFA model fit, and instrument reliability. The criteria for model fit are presented in the following table:

Table 1. Criteria of Goodness of Fit

No	Goodness Of Fit Index	Score Limit	Criteria
1	DF	≥ 0	Over identified
2	Chi-square	$\leq \alpha, df$	
3	Probability	$\geq 0,05$	Fit
4	CMIN/DF	≤ 2	Fit
5	GFI	$\geq 0,900$	Fit
6	AGFI	$\geq 0,900$	Fit
7	CFI	$\geq 0,900$	Fit
8	TLI/NNFI	$\geq 0,900$	Fit
9	NFI	$\geq 0,900$	Fit
10	IFI	$\geq 0,900$	Fit
11	RMSEA	$\leq 0,08$	Fit
12	RMR	$\leq 0,05$	Fit
13	SRMR	$\leq 0,08$	Fit

Model fit assessment or model suitability in CFA analysis testing can be declared suitable if at least one of the model feasibility test methods is met or the results meet at least 4-5 different model fit indices (Haryono, 2017)

After proving the validity, the reliability coefficient of the instrument that has been declared valid at the construct validation stage will be estimated. In this study, reliability testing will be carried out using construct reliability, specifically McDonald's coefficient omega (ω). McDonald's coefficient omega is done using only the loading factor. The formula for estimating ω is as follows (Katama et al in Retnawati, 2017).

$$\omega = \left(\frac{\sum_{i=1}^n \lambda_i^2}{\sum_{i=1}^n \lambda_i^2 + \sum_{i=1}^n (1 - \lambda_i^2)} \right) \quad (2)$$

Description :

ω = McDonald's coefficient omega

λ_i = loading factor

The minimum acceptable reliability value is more than 0.6 in order to get a high category, while if it is > 0.8 it will be categorized as very high Guilford (Istiyono, 2020).

RESULT AND DISCUSSION

1. Content Validity of The Instrumen

Content validity of independent character multiple choice question type at the elementary school level was conducted with the help of 5 experts. Analysis was carried out on 21 instrument items that were successfully developed with the help of the Excel application with the following results:

Table 2. Content Validation Results

Item	Aiken V Coefficient	Category	Item	Aiken V Coefficient	Category
Item 1	0,94	Valid	Item 11	0,87	Valid
Item 2	0,94	Valid	Item 12	0,94	Valid
Item 3	0,94	Valid	Item 13	0,94	Valid
Item 4	1	Valid	Item 14	1	Valid
Item 5	0,94	Valid	Item 15	1	Valid
Item 6	0,87	Valid	Item 16	1	Valid
Item 7	0,87	Valid	Item 17	1	Valid
Item 8	1	Valid	Item 18	0,94	Valid
Item 9	0,94	Valid	Item 19	0,94	Valid

Item 0,87	Valid	Item 0,87	Valid
10		20	
		Item 0,87	Valid
		21	

Based on the table above, it is known that the Aiken coefficient value obtained by each item is > 0.86 and can be declared valid. Based on the average Aiken coefficient for each item, a value of $0.933 > 0.86$ is also obtained and valid. Aiken's coefficient both per item and on average is in very high criteria so that overall the average coefficient obtained by 21 instrument items is declared valid and can be used for the next stage.

2. Construct Validity

The construct validity of the instrument was analyzed by confirmatory factor analysis (CFA), precisely with the first-order confirmatory factor analysis assisted by the R program. The test was conducted on 550 data obtained from the research sample. The use of confirmatory type factor analysis aims to confirm the representation of each item against the theorized model. Proof of construct validity through CFA is seen from the loading factor/factor loading value of each item. Testing is carried out in 2 stages, namely before and after modifying the model to get the items as a whole with a factor loading > 0.3 . And also reviewed the suitability of the model based on existing criteria. The following is a picture of the factor analysis output from the R program.

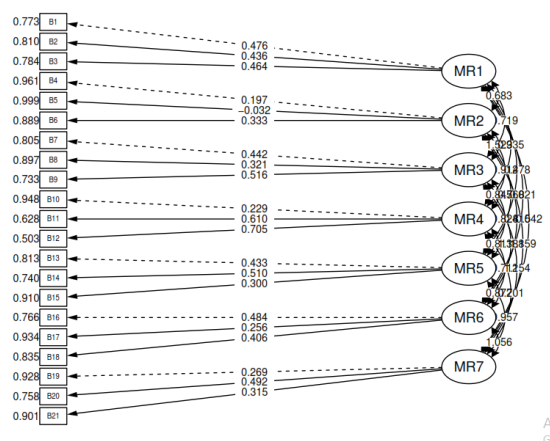


Fig. 1 First-Order Confirmatory Factor Analysis Before Modification

Based on Figure 1, of the 21 items analyzed, there are 5 items that have a factor load < 0.3 , namely items B4, B5, B10, B17, and B19 so they must be dropped, because they do not have a sufficient factor load value to represent the independent character construct. The analysis continued with the following results .

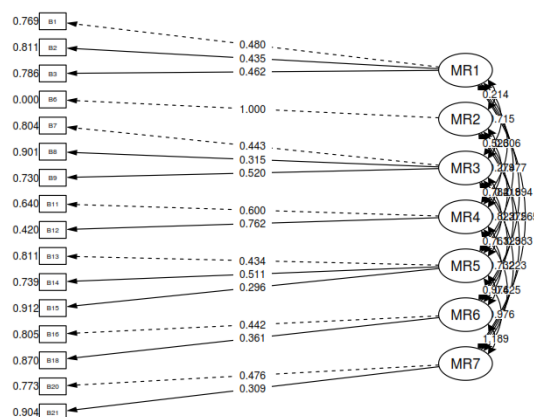


Fig. 2 First-Order Confirmatory Factor Analysis After Modification 1

Based on Figure 2 of the 16 items analyzed there is still 1 item with a factor load < 0.3 , namely item 15, and must be dropped. The analysis continued with the remaining items to obtain the following results.

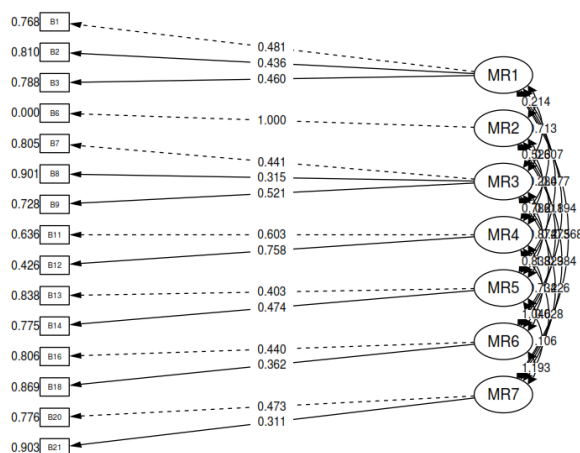


Fig. 3 First-Order Confirmatory Factor Analysis After Modification 2

Based on Figure 3, the remaining 15 items had factor loadings above 0.30, indicating that they adequately represented their respective dimensions of independent character. Therefore, all 15 retained items were considered acceptable in terms of construct validity. Overall, six items (B4, B5, B10, B15, B17, and B19) were removed because their factor loadings were below the minimum criterion of 0.30, indicating weak empirical relationships with their intended dimensions.

From a content perspective, the discarded items may require further review regarding wording clarity, developmental appropriateness, and ease of interpretation by upper-grade elementary students. Although the removal of these items did not eliminate any of the seven indicators, Indicator 2 was represented by only one retained item, resulting in more limited content coverage than the other indicators. Future studies should consider revising or developing additional items for this indicator. The construct validity analysis was then continued by evaluating the overall fit of the hypothesized model, as presented in Table 3.

Table 3. Out model fit results of First-Order Confirmatory Factor Analysis

No	Goodness	Value	Value	Criteria
----	----------	-------	-------	----------

Of Fit Index			Limit	
1	DF	70	> 0	Over identified
2	Chi-square	0.000	$< 0,35$	Fit
3	Probability	0.000	$> 0,05$	Unfit
4	DF/CMIN	2,278	> 2	Unfit
5	CFI	0.913	$\geq 0,900$	Fit
6	TLI	0.870	$\geq 0,900$	Marginal fit
7	RMSEA	0.048	$\leq 0,08$	Fit
8	SRMR	0.041	$< 0,08$	Fit

Based on Table 3, 6 criteria are obtained which are said to be good/fit for the model, namely SRMR, RMSEA, CFI, Chi-square, DF with over-identified criteria which means that the analyzed data can be observed, TLI with marginal fit criteria. With the 6 criteria for model fit obtained, the theoretical construct tested can be declared fit to the model.

3. Reliability

Reliability estimation is evidenced by the value of the construct reliability coefficient precisely omega McDonald's. The analysis was carried out using the factor load value of each item using the Excel application with the following results.

Table 4. Out model fit results of First-Order Confirmatory Factor Analysis

Butir	Λ	δ	λ_2	$1-\lambda_2$
B1	0,481	0,732	0,231	0,769
B2	0,436	0,803	0,190096	0,809904
B3	0,460	0,82	0,2116	0,7884
B6	1,000	0	1	0
B7	0,441	0,815	0,194481	0,805519
B8	0,315	0,9	0,099225	0,900775
B9	0,521	0,717	0,271441	0,728559
B11	0,603	0,634	0,363609	0,636391
B12	0,758	0,429	0,574564	0,425436
B13	0,403	0,818	0,162409	0,837591
B14	0,474	0,751	0,224676	0,775324
B16	0,440	0,894	0,1936	0,8064
B18	0,362	0,807	0,131044	0,868956
B20	0,473	0,868	0,223729	0,776271
B21	0,311	0	0,096721	0,903279
JUMLAH	7,478	9,988	4,168556	10,83144
Ω	0,8377359			

The analysis results state that the McDonald's omega coefficient value of 0.837 is in the very high category. So it can be concluded that the developed upper grade multiple choice question type independent character assessment instrument is very reliable in measuring the independent character of students and is able to provide stable results in repeated measurements.

4. Discussion

Based on the feasibility of the independent character assessment instrument developed, it can be stated that it has good feasibility. From the content aspect, the reviewed items received an Aiken v coefficient of 0.933 in line with research conducted by Anculle-Arauco et al. (2024; and Curcio et al., (2022). In terms of constructs, 15 valid items were produced with a standard loading factor of more than 0.3 in line with research conducted by Asbupel & Supahar (2021). The elimination of six items from the original 21-item instrument indicates that content validity alone does not necessarily ensure that all items will adequately represent the intended construct empirically. Although these items had previously been considered relevant by experts, their low factor loadings suggest weaker relationships with the latent construct of independent character. This may indicate that some behavioral indicators were interpreted differently by students, were less clearly understood, or were less representative of independent character in the context of upper-grade elementary students. However, because this study did not specifically investigate students' interpretations of the eliminated items, these possible explanations should be treated cautiously. The findings emphasize that character assessment for elementary students requires both expert judgment and empirical validation to ensure that the items are developmentally appropriate and consistently represent the intended construct. The 15 items consist of 3 items in Indicator 1, 1 item in Indicator 2, 3 indicators in Indicator 3, 2 items in Indicator 4, 2 items in Indicator 5, 2 items in Indicator 6, and 2 items in Indicator 7. The model fit obtained in the development stage is in line with the research conducted Safitri (2020) and Hayono, (2017). In terms of instrument reliability, it has an omega McDonald's value of 0.8672527 with a very high-reliability category in line with research conducted by Görlich (2023).

CONCLUSIONS

Based on the results of this study, the developed independent character assessment instrument demonstrated satisfactory evidence of content validity, construct validity, and reliability. The average Aiken's V coefficient was 0.933, indicating adequate content validity based on expert judgment. Construct validity analysis showed that 15 of the initial 21 items were retained with factor loadings above 0.30 and an acceptable overall model fit. In addition, McDonald's omega coefficient was 0.837, indicating high internal consistency.

These findings support the use of the instrument for assessing independent character among upper-grade elementary school students within the context and population examined in this study. However, the findings should not be interpreted as evidence of universal applicability. Further studies are recommended to examine the instrument across different populations and educational contexts, as well as to evaluate its practical effectiveness before broader implementation.

ACKNOWLEDGMENT

Thanks to Farida Agus Setiawati as the supervisor in the implementation of this research and to Asyraf Suryadin as one of the instrument validators in this study.

CONFLICTS OF INTEREST STATEMENT

There is no conflict of interest in the writing of the manuscript.

AUTHOR CONTRIBUTIONS

Dita Hardiani conducted the research on the subject matter, while Farida Agus Setiawati served as the supervisor and assisted in the completion of the study. Meanwhile, Asyraf Suryadin contributed as an instrument validator and ensured alignment with the latest template.

REFERENCES

- Abrahams, L., Pancorbo, G., Primi, R., Santos, D., Kyllonen, P. C., John, O. P., & De Fruyt, F. (2019). Social-emotional skill assessment in children and adolescents: Advances and challenges in personality, clinical, and educational contexts. *Psychological Assessment*, 31(4), 460–473. <https://doi.org/10.1037/pas0000591>
- Anculle-Arauco, V., Krüger-Malpartida, H., Arevalo-Flores, M., Correa-Cedeño, L., Mass, R., Hoppe, W., & Pedraz-Petrozzi, B. (2024). Content validation using Aiken methodology through expert judgment of the first Spanish version of the Eppendorf Schizophrenia Inventory (ESI) in Peru: A brief qualitative report. *Spanish Journal of Psychiatry and Mental Health*, 17(2), 110–113. <https://doi.org/10.1016/j.rpsm.2022.11.004>
- Arikunto, S. (2018). *Dasar-dasar Pendidikan* (3rd ed.). Bumi Aksara.
- Asbupel, F., & Supahar. (2021). *Pengembangan Instrumen Penilaian Sikap Pada Pembelajaran Kimia Berbentuk Self-Assessment dan Peer-Assessment* [Universitas Negeri Yogyakarta]. <http://eprints.uny.ac.id/id/eprint/73000>
- Azwar, S. (2022). *Penyusunan Skala Psikologi Ed 2*. Pustaka Belajar. <https://books.google.co.id/books?id=H4P1DwAAQBAJ&lpg=PP1&hl=id&pg=PR4#v=onepage&q&f=false>
- Curcio, F., Vaquero-Abellán, M., Zicchi, M., Ez zinabi, O., & Romero-Saldaña, M. (2022). Translation and cross-cultural adaptation of the Neonatal Skin Risk Assessment Scale (NSRAS) to Italian. *Journal of Tissue Viability*, 31(4), 693–698. <https://doi.org/10.1016/j.jtv.2022.10.001>
- Görlich, Y. (2023). Development of Creative Process Assessment Scale (CPAS). *Journal of Creativity*, 33(1), 100042. <https://doi.org/10.1016/j.yjoc.2023.100042>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning.
- Haiq, U., Martati, B., & Setiawan, F. (2020). An Analysis Of Independent Character Of Primary School Students In Learning From Home During The Covid-19 Pandemic (A Case Study at SD Muhammadiyah 8 Surabaya). *Widyagodik : Jurnal Pendidikan Dan Pembelajaran Sekolah Dasar*, 8(1), 55–65. <https://doi.org/10.21107/widyagodik.v8i1.8798>
- Hayono, S. (2017). *Metode SEM Untuk Penelitian Manajemen Amos Lisrel PLS*. Luxima.
- Istiyono, E. (2020). Pengembangan instrumen penilaian dan analisis hasil belajar Fisika dengan teori tes klasik dan modern. *Urnal Manajemen Pendidikan Islam*, 1(9), 478–492.
- Jeona, Y.-H., Milne, N., Kaizik, C., & Resnick, B. (2021). Improving functional independence: Dementia rehabilitation programs. In *Dementia Rehabilitation* (pp. 227–261). Elsevier. <https://doi.org/10.1016/B978-0-12-818685-5.00013-1>
- Kemdikbud. (2025). *Surat Edaran Bersama Menteri Pendidikan Dasar Dan Menengah Republik Indonesia Tentang Penguatan Pendidikan Karakter Melalui Pembiasaan Di Satuan Pendidikan*.
- Kristina Mikkonen, Tomietto, M., & Watson, R. (2022). Instrument development and psychometric testing in nursing education research. *Nurse Education Today Journal*, 119, 1–6. <https://doi.org/10.1016/j.nedt.2022.105603>
- Lawrence. (1994). *Question to ask when evaluating test*.

<http://www.ericfacility.net20ericdigest/ed.385607.Html>

- MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques. *Management Information Systems Research Center, University of Minnesota*, 35(2). <https://doi.org/doi.org/10.2307/23044045>
- Mardapi, D. (2017). *Pengukuran, penilaian dan evaluasi pendidikan* (2nd ed.). Parama Publishing.
- Murano, D., Lipnevich, A. A., Walton, K. E., Burrus, J., Way, J. D., & Anguiano-Carrasco, C. (2021). Measuring social and emotional skills in elementary students: Development of self-report Likert, situational judgment test, and forced choice items. *Personality and Individual Differences*, 169, 110012. <https://doi.org/10.1016/j.paid.2020.110012>
- Paz, V., Dashti, H. S., Burgess, S., & Garfield, V. (2023). Selection of genetic instruments in Mendelian randomisation studies of sleep traits. *Sleep Medicine*, 112(May), 342–351. <https://doi.org/10.1016/j.sleep.2023.10.036>
- Rakhmat, C., & Solehuddin, M. (2006). *Pengukuran dan Penilaian Hasil Belajar*. Andira.
- Rasmadi, S. (2024). Viral Siswa SD Indramayu Di-bully-Ditelanjangi, Polisi Turun Tangan. <https://news.detik.com/berita/d-7228615/viral-siswa-sd-indramayu-di-bully-ditelanjangi-polisi-turun-tangan>
- Retnawati, H. (2017). *Validitas Reliabilitas & Karakteristik Butir* (Cet.2 (ed.)).
- Reynolds, C. R., Livingston, R. B., & Willson, V. (2010). *Measurement and assessment in education*. NY: Pearson.
- Safitri, A. I. D. (2020). *Analisis Faktor yang Mempengaruhi Penggunaan Konten Digital pada Media Sosial Menggunakan Structural Equation Modeling dengan Variabel Moderasi Digital Content Type*. 1–162.
- Sekaran, U. (2006). *Metodologi Penelitian untuk Bisnis (IV)*. Salemba Empat.
- Sudjana, N. (2026). *Penilaian Hasil Proses Belajar Mengajar*. PT Remaja Rosdakarya.
- Sumahamijaya, S., Yasben, D., & Dana, D. A. (2018). *Pendidikan Karakter Mandiri dan Kewiraswastaan: Suatu Upaya bagi Keberhasilan Pendidikan Berbasis Luas/Broad Based Education dan Life Skills*. Angkasa.
- Tim Pusat Penilaian Pendidikan. (2019). *Model Penilaian Karakter* (Asrijanty & D. Hadiana (eds.)). Pusat Penilaian Pendidikan Badan Penelitian dan Pengembangan kementerian Pendidikan dan Kebudayaan.
- Volz, S., & Masicampo, E. J. (2021). Self-regulatory processes and personality. In *The Handbook of Personality Dynamics and Processes* (pp. 345–363). Elsevier. <https://doi.org/10.1016/B978-0-12-813995-0.00014-5>
- Wulandari, D. A., Saefuddin, S., & Muzakki, J. A. (2018). Implementasi Pendekatan Metode Montessori Dalam Membentuk Karakter Mandiri Pada Anak Usia Dini. *Awlady : Jurnal Pendidikan Anak*, 4(2), 1. <https://doi.org/10.24235/awlady.v4i2.3216>
- Zuchron, D. (2021). *Tuna Pancasila* (A. L. SP, D. Gunawan, E. R. Widodo, & E. Purnawanarni, Eds.). Direktorat Sekolah Dasar Direktorat PAUD, Dikdas, dan DIkmen Kementrian Pendidikan, Kebudayaan, Riset dan Teknologi.
- Zumbo, B. D. (2026). *Measurement Validity: Essential in Principle, Subordinated in Practice*. Quality of Life Research, 35(10). DOI: 10.1007/s11136-026-04355-0.