



Character Scale for Assessing Discipline and Learning Responsibilities Character in Elementary School

Maria Sumunaringtyas*, Tiara Faniska D², Edi Istiyono³, Farida Agus Setiawati⁴ 

^{1,2,3,4} Research and Evaluation in Education Department, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia

ARTICLE INFO

Article history:

Received October 13, 2023

Accepted April 10, 2024

Available online May 25, 2024

Kata Kunci:

Skala Karakter, Penilaian Kedisiplinan, Tanggung Jawab Belajar, Karakter, Sekolah Dasar

Keywords:

Character Scale, Assessing discipline, Learning Responsibilities, Character, Elementary School



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Copyright © 2024 by Author. Published by Universitas Pendidikan Ganesha.

ABSTRAK

Penilaian merupakan proses penting dalam rangkaian pembelajaran yang perlu dilakukan sebaik mungkin menggunakan instrumen yang jelas dan berkualitas. Penilaian karakter disiplin dan tanggung jawab di salah satu sekolah dasar hanya dilakukan dengan berdasarkan pengamatan guru. Hal ini secara tidak langsung dapat menjadikan penilaian menjadi subjektif dari sudut pandang guru saja. Penelitian ini bertujuan untuk menghasilkan seperangkat instrumen penilaian karakter disiplin dan tanggung jawab berbentuk skala diri yang berkualitas. Jenis penelitian ini merupakan penelitian pengembangan yang dilakukan menggunakan tahap pengembangan instrumen. Jumlah butir yang dikembangkan dalam penelitian ini sebanyak 20 butir pernyataan. Setelah dianalisis menggunakan CFA, didapatkan total 14 butir pernyataan yang dapat digunakan dalam mengukur karakter disiplin dan tanggung jawab. Teknik pengumpulan data yang dilakukan dalam penelitian ini menggunakan lembar validasi dan skala diri dengan Likert 5 kategori. Validitas isi dihitung dengan formula Aikens dan menghasilkan skor yang tinggi. Sedangkan estimasi reliabilitas instrumen ini dihitung menggunakan Composite Reliability menghasilkan angka $>0,7$ yang artinya instrumen memiliki reliabilitas tinggi. Pembuktian validitas dan reliabilitas telah terpenuhi sehingga instrumen dapat digunakan dalam penilaian karakter disiplin dan tanggung jawab. Melalui instrumen yang valid dan reliabel, maka akan didapatkan data pengukuran yang lebih akurat dan objektifitas dalam penilaian karakter lebih terjamin, Implikasi penelitian terdapat pada skala diri yang dapat digunakan sebagai penyeimbang dalam penilaian karakter yang sifatnya laten.

ABSTRACT

Assessment is an important process in the learning series that needs to be carried out as well as possible using clear and high-quality instruments. Assessment of the character of discipline and responsibility in one elementary school was only carried out based on teacher observations. This can indirectly make the assessment subjective from the teacher's perspective only. This research aims to produce a collection of instruments for assessing the character of discipline and responsibility in the form of a quality self-scale. This type of research is development research carried out using the development instrument stage. The number of items developed in this research was 20 statements. After being analyzed using CFA, a total of 14 statements were obtained which can be used to measure discipline and responsibility character. The data collection technique used in this research used a validation sheet and a self-scale with a 5-category Likert. Validity is calculated using the Aikens formula and produces a high score. Meanwhile, the estimated reliability of this instrument was calculated using Composite Reliability, producing a number >0.7 , which means the instrument has high reliability. Proof of validity and reliability has been fulfilled so that the instrument can be used in assessing discipline and responsibility character. Through a valid and reliable instrument, more accurate measurement data will be obtained, and objectivity in character assessment will be guaranteed. The research implications are in the self-scale which can be used as a balance in latent character assessment.

1. INTRODUCTION

Character of responsibility and discipline is a character that comes from self-control. Self-control in this case is the ability to control, regulate, and direct oneself to behavior to achieve goals in the learning process (Dwi Marsela & Supriatna, 2019; Nugraha & Suyadi, 2019; Putri & Yarni, 2022). Students certainly

need a continuous process to have a good self-control to support their educational process. Based on research results, the character of learning responsibility and discipline has an influence on student learning outcomes (Prasasty, 2017; Susanti et al., 2019; Yanti et al., 2022). Seeing this both characters have a significant influence on student learning outcomes, these two characters should not be ignored in education. Character education on discipline and responsibility should be instilled from an early age. The character of discipline is obedience to respect and carry out a system that requires everyone to submit to applicable decisions, orders, or regulations (Akmaluddin & Haqqi, 2020; Elly, 2016; Jamaludin et al., 2023). In other words, discipline is compliance with the rules and regulations that have been set. It means, if there is a violation of these rules, they will get sanctions. Meanwhile, responsibility if viewed from the student's perspective, it can be said to be responsible if he carries out his obligations (Kamaluddin & Al-Asror, 2021; Safitri et al., 2020), as a student, such as doing assignments based on his own work, submitting assignments on time, carrying out class pickets, and working on group assignments together (Musfirah, 2019; Purwanto et al., 2020; Wulandari et al., 2023). Responsibility must be trained so that it becomes a habit and develops into character. Responsibility also shows whether a person has a good or bad character.

Responsibility refers to a person's attitude and behavior to carry out their duties and obligations as they should towards themselves, society, the environment and God. A person's duties and obligations in every environment make him able to understand positive and negative things and take a stand on both things (Blašková, 2014; Siburian, 2012). This means that responsibility is also related to the ability to make the most appropriate decisions within the limits of social norms in society. This ability also refers to the ability to determine attitudes and take risks regarding what one will do and has done. Responsibility is the actualization of awareness of attitudes and actions that have been carried out whether intentionally or not, both personally and socially, up to a higher level, namely the devotion of a servant to his God (Murti et al., 2020; Syafitri, 2017). Human consciousness is closely related to a heart and mind that is open and willing to accept information, knowledge, and things that are true. The demands of a person's responsibilities vary based on the period and space. That is, the higher a person's age, the greater the responsibility. The character indicators of responsibility of learning are as follows: (1) use time effectively, (2) make preparations before learning, (3) carry out individual assignments received, (4) carry out the discussion process, (5) work on questions or problems carefully (Prihastutia & Santa, 2020; Syifa et al., 2022).

A student can be said as a discipline person if he fulfills several indicators, namely not being late for school, carrying out picket schedules, throwing trash in its place, not making noise in class when the teacher not around, and sitting neatly (Anggraeni, 2021; Lewis, 2001). This is in line with the meaning of disciplinary character as the attitude of a person who is willing to comply with all applicable rules and norms in carrying out their duties and responsibilities (Halim, 2023). Discipline can be realized through repeated practices and self-intention, which in the process involves the values of order and obedience. Discipline can make someone know and differentiate between things they should do and what they should not do. Character can be formed through repeated actions that are carried out regularly until become a habits and then develop into character (Dini, 2018; Khoirroni et al., 2023; Septiani & Mukhtar, 2022). The existence of character education is expected to be able for the nation's successors to become generations of intelligent and character so that a dignified civilization can be created according to the ideals of the nation's founders. Discipline and responsibility character education should be implemented from an early age in the family and school. This is because character cannot be formed in a short time. There needs to be habits so that these good things become habits and develop into character. Meanwhile, in implementing character education, evaluation also needs to be carried out. The measure of the success of character education that has been carried out so far. In addition, as previously stated these two characters have a considerable influence on student learning outcomes, So this measurement of character should also be carried out seriously.

It should be noted that the objects measured in psychology are humans. Therefore, the measurements carried out should use quality tools and be carried out at the right time and under the right conditions so that the measurement results are not biased. Measuring the character of discipline and responsibility is the first step in efforts to improve the quality of character education (Azwar, 2012; Febrina Tarigan et al., 2022; Hikmah & Muslimah, 2021). Valid and reliable instruments will produce accurate measurements, conclusions and decisions based on accurate and reliable data will be very helpful for any party interested in the test results. Accompanied by correct interpretation, valid and reliable measurement results will lead to more precise and useful diagnoses and decisions (Puspasari & Puspita, 2022; Sahetapy et al., 2019). This is so that the measurement results obtained can be used as a basis for providing treatment so that the character of discipline and responsibility can be embedded. If the student's character of discipline and responsibility is high, then it is likely that the learning outcomes will be high. If these two things are done successfully, then educational goals can be achieved more optimally.

The fact that researchers found when examining the implementation of discipline and responsibility character education in one of the state elementary schools was that there was no instrument to assess discipline and responsibility character. In fact, in general the character education steps from planning to implementation in these schools are quite good. Meanwhile, at the evaluation stage, the assessment used by the teacher as a basis for evaluation is one-sided observation of the teacher. Even though teachers are said to be professional in their experience, it cannot be denied that someone's judgment can be wrong or biased. Various studies have stated that it is best to carry out assessments objectively (Hafidhoh & Rifa'i, 2021; Idrus, 2019; Lestari, 2022). This is in line with the statement that the subjects of psychological measurements are humans whose future can depend on the results of the assessment (Rudini, 2020). Therefore, so that the assessment does not result in bad decisions, the assessment should be balanced with information obtained objectively. Therefore, instruments for assessing the character of discipline and responsibility that have been tested for quality are very necessary so that they can be used by teachers to assess the character of discipline and responsibility of students.

One of the main concerns in every research activity in the field of psychology is the problem of how to obtain accurate, thorough, and objective information data. Accuracy, thoroughness, and objectivity of data in the field of psychology that uses a quantitative approach can only be obtained through measurement procedures because, with the measurement process, the conceptual research variables (theoretical construct) will be able to be quantified correctly based on observations of the indicators. operational (observable). Unfortunately, accurate and objective data in research in the field of psychology is not always easy to obtain, especially because the concept of latent variables being measured is not as easy to operationalize as in measuring physical variables. Psychometric experts (the science of measurement and psychological tests) have established several important criteria for any psychological measuring tool to be declared a good measuring tool, namely being able to produce data and provide accurate information. The criteria referred to include valid, reliable, objective, standard, economical, and practical (Linden, 1997; Lyublinskaya & Kaplon-Schilis, 2022). Valid and reliable properties are shown by the high accuracy and precision of the measurement results. A measuring instrument or test is said to be invalid if it is unable to produce accurate information about the attributes or variables it measures, that is, the score does not reflect the actual situation. If inaccurate information is used as a basis for consideration in making a conclusion or decision, then the conclusion and decision must also be an incorrect conclusion or decision. Meanwhile, evidence of instrument quality with item response theory can be shown through goodness of fit, item parameters, information function, and standard error measurement (Joarder et al., 2020; Nitko & Brookhart, 2014).

Validity is used to prove the quality of the instrument. This instrument will be proven by using CFA (Confirmatory Factor Analysis) construct validity. This is because the development of this instrument is based on the study of several theories that have developed. Confirmatory factor analysis (confirmation of the measurement model) is a part of SEM (Structural Equation Modelling) which is useful for testing how variables can be measured well in describes the number of factors, in CFA factors can also be said to be constructed (Hair et al., 2019; Mardhotillah & Yulhendri, 2022). The aims of this study produce a collection of instruments for assessing the character of discipline and responsibility in the form of a quality self-scale.

2. METHOD

This research is a type of development research with a proprietary instrument development model (Retnawati, 2016). This development model consists of 9 stages, namely, (1) determining the purpose of preparing the instrument, (2) looking for relevant theories or material coverage, (3) compiling instrument item indicators, (4) compiling instrument items, (5) content validation, (6) revision based on validator input, (7) conducting trials, (8) conducting analysis, and (9) assembling the instrument. The procedure is adjusted to the real conditions in the field, especially in the validation section. The process of content revision and validation will continue to be carried out until a valid instrument is obtained in terms of quality as stated by expert judgment. After the instrument is declared valid, the next stage will be carried out, namely testing, analysis, and instrument assembly. Respondents in this instrument trial were 50 elementary school students who were not bound by region. The data collection technique was carried out using the Google form. The items produced in this research consisted of 20 statements with a balanced proportion of favorable and unfavorable statements. The construct of this instrument consists of 10 indicators, each consisting of 2 statement items. The scoring in this instrument uses a Likert scale from 1-5.

The preparation of the instrument begins with a literature study to look for supporting theories for research and concepts regarding discipline and responsibility as research objects. The literature study used to construct the instrument for assessing the character of discipline and responsibility (Abdurrahman, 2018; Yanti et al., 2022). The next step is to determine the operational definition of the discipline and

responsibility variables. After the indicators have been prepared, the next step is to determine the scale according to the Likert Scale. Determination of this scale is also related to response criteria, 5 (always), 4 (often), 3 (sometimes), 2 (rarely), and 1 (never) for favorable items. Meanwhile, the scores for unfavorable items are reversed to 5 (never), 4 (sometimes), 3 (rarely), 2 (often), and 1 (always).

After the scale and scoring criteria have been determined, the steps taken by the researcher are to arrange the items according to the indicators by paying attention to the predetermined Likert Scale. Each indicator in this study consisted of 2 items, namely 1 item favorable and 1 item unfavorable. The total item statements produced are 20 items. These points were then proven for their content validity by V Aikens. There were 10 raters who assessed the validity of the content of this instrument, consisting of instrument development experts and elementary school teachers. The scale used to prove content validity consists of 4 categories which can be seen in Table 1.

Table 1. Content Validity Scoring Criteria

	Criteria	Score
STR	: Items do not match indicators and aspects/dimensions	1
TR	: Items correspond to indicators but do not correspond to aspects/dimensions	2
R	: Items are following indicators, aspects/dimensions, or instrument variables using inappropriate or less operational language	3
SR	: Items are relevant to indicators, aspects/ dimensions, and instrument variables with appropriate and operational language	4

Data result from expert judgment with the above criteria are then processed by using V Aikens Validity formula. Valid item criteria refer to the specified minimum index by with many raters of 10 and a scale of 4, the percentage is 95% (Aiken, 1985), The item is said to be valid if it is in the coefficient range of 0.73 to 1.00. If there are still items whose validity coefficient is < 0.73, revisions and reassessment will be carried out until the item's validity coefficient meets the criteria. After all items meet the content validity criteria, the next step is to conduct a trial. The trial data will be used for the validation of the CFA construct assisted by the SmartPLS software. Good items are shown by loading factor between 0.4 - 0.7 or more than 0.7. Meanwhile, good indicators should have loading factors ≥ 0.5 .

3. RESULT AND DISCUSSION

Result

This research aims to produce an instrument that can be used to measure the disciplinary character and responsibility of students whose quality has been proven. This instrument has gone through several stages in proving its quality. The initial step taken by the researcher was to conduct a literature review regarding the character of discipline and responsibility. The preparation of indicators has been adapted to the conditions encountered by researchers in the field. These indicators were then developed into statement items with a Likert Scale. The construct of this instrument can be seen more clearly in Table 3.

Table 3. Instrument Grid for Assessing Discipline and Responsibility Character

Variable	Indicator	Item Number	Favorable	Unfavorable
Discipline	Obey the school rules	1,2	1	1
	Leave school on time	3,4	1	1
	Get to class on time	5,6	1	1
Responsibility	Complete assignments at your own pace	7,8	1	1
	Complete tasks according to applicable regulations	9,1	1	1
	Turn in assignments on time	11,12	1	1
	Admit mistake	13,14	1	1
	Apologize when you make a mistake	15,16	1	1
	Take lessons seriously	17,18	1	1
	Make material notes	19,2	1	1

After the grid and items were prepared, the next step was content validation by experts consisting of 10 experts in the field of instrument development and elementary school teachers. The validity results are presented below using validity calculations according to Aikens in Table 4.

Table 4. Content Validity

Item	Minimum Index	Aikens Validity Coefficient	Note	Item	Minimum Index	Aikens Validity Coefficient	Note
1		1	Valid	11		0.97	Valid
2		1	Valid	12		1	Valid
3		0.97	Valid	13		0.97	Valid
4		0.93	Valid	14		1	Valid
5	0.73	1	Valid	15	0.73	0.97	Valid
6		1	Valid	16		1	Valid
7		1	Valid	17		0.97	Valid
8		1	Valid	18		1	Valid
9		1	Valid	19		0.97	Valid
10		1	Valid	20		1	Valid

The content validity above was not obtained in one assessment by expert judgment but was revised several times until the sentence structure and relevance of the items to the indicators were appropriate. Based on the results of the content validity as show in Table 4, it can be concluded that 20 items have been proven to be content-valid. This valid item then disseminated for testing. There were 50 respondents in this trial. Due to the limited data obtained by researchers, the analysis that will be used in this research refers to classical test theory. The data that has been obtained is used to prove the validity of the construct. Construct validity can be done through EFA or CFA. Because this instrument was prepared starting with a literature review, the construct validity that will be used is CFA with the help of SmartPLS software. The result of loading factors is show in Table 5.

Table 5. Loading Factors

Items	Variable	Loading Factor	Follow-up	Items	Variable	Loading Factor	Follow-up
T1	Discipline	0.796	Used	BT1	Responsibility	0.978	Used
T2	Discipline	0.919	Used	BT2	Responsibility	-0.267	Not used
B1	Discipline	0.406	Not used	KS1	Responsibility	0.799	Used
B2	Discipline	0.970	Used	KS2	Responsibility	0.664	Not used
MS1	Discipline	0.743	Used	MF1	Responsibility	0.808	Used
MS2	Discipline	0.844	Used	MF2	Responsibility	0.678	Not used
ST1	Responsibility	0.944	Used	BL1	Responsibility	0.852	Used
ST2	Responsibility	0.588	Not used	BL2	Responsibility	0.874	Used
KT1	Responsibility	0.951	Used	CT1	Responsibility	0.773	Used
KT2	Responsibility	-0.370	Not used	CT2	Responsibility	0.801	Used

Loading factors are the relationship between items and indicators. Base on Table 5, the desired loading factor score is > 0.7. Items with loading factor scores < 0.7 should not be taken into final instruments, or the construction of these indicators on variables is too weak. The inner model result is show in Table 6.

Table 6. Inner Model

Indicator	Inner	Variable
Obey the school rules	0.767	Discipline
Leave school on time	0.780	
Get to class on time	0.755	
Complete assignments at your own pace	0.646	
Complete tasks according to applicable regulations	0.651	
Turn in assignments on time	0.550	Responsibility
Admit mistake	0.627	
Apologize when you make a mistake	0.748	
Take lessons seriously	0.824	
Make material notes	0.814	

Inner is a coefficient that shows the correlation between indicators and variables. The inner coefficient above is the score of the path coefficients known as the total effect. The minimum inner that must

be met by each indicator of the variable is 0.3. Based on Table 6 shown above, all sub-indicators have met the requirements > 0.3 so that these indicators are declared meaningful for measuring the variables of responsibility and discipline. If the path coefficient score gets closer to 1, then the construction of the indicator on the variable is getting stronger. Reliability of discipline variables is show in Table 7.

Table 7. Reliability of Discipline Variables

Indicator	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted (AVE)
Obey the school rules	0.659	0.733	0.849	0.739
Leave school on time	1	1	1	1
Get to class on time	0.422	0.439	0.773	0.631
Discipline	0.624	0.711	0.757	0.372

Table 7 shows the reliability of the instrument that measures discipline character after item B1 was eliminated because the loading factor did not meet the requirements. The reliability of discipline character instruments can use the Composite Reliability coefficient because it produces high reliability. Furthermore, the reliability of the responsibility variable will be presented in Table 8.

Table 8. Reliability of Responsibility Variables

Indicator	Cronbach's Alpha	rho_A	Composite Reliability	Average Variance Extracted (AVE)
Complete assignments at your own pace	1	1	1	1
Complete tasks according to applicable regulations	1	1	1	1
Turn in assignments on time	1	1	1	1
Admit mistake	1	1	1	1
Apologize when you make a mistake	1	1	1	1
Take lessons seriously	0.657	0.657	0.854	0.745
Make material notes	0.386	0.388	0.765	0.620
Responsibility	0.780	0.852	0.821	0.279

Base on Table 8, the reliability of instruments that measure the character of responsibility using Composite Reliability is relatively high. The reliability was estimated after items with loading factor values < 0.7 were deleted, such as items ST2, KT2, BT2, KS2, and MF2.

Discussion

Assessment is an important and inseparable part of the learning process that must be carried out appropriately (Kurniawan et al., 2022; Magdalena et al., 2023). Character is a latent variable, namely a variable that cannot be measured directly, such as body weight or height. Therefore, in assessing this variable, measuring instruments are needed, which can be in the form of observation sheets, self-scales, or journals. The problem that researchers found was that the assessment of disciplinary character and responsibility was carried out by teachers through observation without clear instruments (Andersson et al., 2023; Shrestha et al., 2023). It can be concluded that the assessment is only based on the teacher's assumptions whose subjectivity is very high, even though one of the principles of assessment is objective, namely based on standards and not influenced by the subjectivity of the assessor (Ramatni et al., 2023; Siregar, 2018). Assessments that are based on high subjectivity without clear instruments can result in biased assessments or errors in measurement and be detrimental to students.

Bias in measurements can occur due to measurement errors. Measurement errors can occur due to limited information possessed by teachers or assessors, inaccurate assessment processes, and instruments that are of poor quality. In an effort to minimize errors in measurements, methods and procedures for constructing measuring instruments are needed which are expected to be able to produce items of good quality and respond well. Such items are items that have fewer sources of error (at least when viewed from the quality of the measuring instrument), making it possible to achieve a high level of reliability and validity. Then, the procedures for implementing test administration also need to receive attention in order to reduce the influence of irrelevant factors on test results (Gümüş et al., 2021; Retnawati, 2016). Therefore, it is very important to ensure that the instruments used in the assessment have high validity and reliability. The quality of instruments can be proved by using content validity, construct validity, and reliability

(Bakhshandeh Bavarsad et al., 2021; Park, 2023). On the other hand, reliability and validity are not only the properties of instrument, but of the scores and their interpretations. Therefore, better assessment methods of the construct validity of Likert scales are needed, and alternative estimation methods are recommended to avoid incorrect parameter estimates, such as factor loading coefficients, standard errors, and model fit statistics (Artino et al., 2014; Park, 2023). Based from the previous statement, we can conclude that validity and reliability are the important properties of the quality of instruments and how accurate the result of measurement interpretation. When the instruments are valid and reliable, the measurement interpretation also accurate.

If the instrument factors have been met in quality, then the next stage is the use of the instrument. Assessment using a character scale is more suitable when used for high class students. This is due to differences in students' cognitive development. High class students certainly have a higher ability to understand sentences than low class students (Kesuma et al., 2022; Ritonga et al., 2023). As researchers, we must ensure that the respondents in our research are on target. Errors that occur in measurement can be related to various problems regarding the sentence or editorial of the question which may be unclear, may be understood incorrectly, may contain sensitive. There are two types of bias which might be caused by a lack of measurement invariance and meanings so that the answer desired from the individual does not reflect the answer that should be could go on to affect the findings of measurement in psychometric in this study. The first is response shift bias and the second is differential item functioning. Response shift bias are incomplete response because the respondent do not fully conceptualize something they have yet to experience (Auld et al., 2017; Cartwright & Atwood, 2014). It means, when the respondents are student of lower class in elementary schools, the response shift bias probably happen because the student do not have experience or do not understanding the main question from the instrument. However, it does not rule out the possibility of this instrument being used for lower grade students with help and assistance from the teacher.

The instrument developed in this research is one answer to the difficulties experienced by teachers in assessing latent variables in education such as character or attitude (Darna, 2016; Taufan, 2022). Therefore, the implications of this research are useful for teachers in carrying out assessments, especially the character of discipline and responsibility for high class students. This instrument can be used as an assessment tool that can balance teacher assessments carried out through observation, so that it can be ensured that the assessment is more objective. The limitation of this research is that the instrument currently being developed can only measure the character of discipline and learning responsibility, which is a small part of the 18 character values (Nabilah et al., 2021; Zuhera & Habibah, 2017). Another limitation of this instrument is that there are not yet sufficient respondents for analysis using IRT. Research into the development of character assessment instruments is something that needs to be considered to be carried out in order to support the assessment process in character education.

4. CONCLUSION

Assessment instrument: Discipline and responsibility character education is one answer to the problems facing teachers in carrying out attitude and character assessments. The results of the analysis of the instruments that have been developed also have quality criteria instruments so that they can be used in the evaluation process. The instrument developed in this research meets the criteria of a good quality of instruments. The results of this research, showed that from 20 items developed only 14 valid items were produced that construct the character construct of discipline and responsibility according to CFA (Confirmatory Factor Analysis). It means, these 14 items can be used in assessing of the disciplinary and responsibility character of elementary school students. For similar researchers, it is advisable to make enough items for each indicator. This is because if the items represent the burnt indicators there are still other items that can represent them and using IRT analysis.

5. REFERENCES

- Abdurrahman, M. (2018). *Pendidikan Karakter Bangsa*. Adnin Foundation Publisher.
- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1). <https://doi.org/10.1177/0013164485451012>.
- Akmaluddin, & Haqqi, B. (2020). Kedisiplinan Belajar Siswa di Sekolah Dasar (SD) Negeri Cot Keu Eung Kabupaten Aceh Besar (Studi Kasus). *Journal of Education Science*, 5(2). <https://doi.org/https://jurnal.uui.ac.id/index.php/jes/article/viewFile/467/204>.
- Andersson, A., Brink, E., Young, K. H., & Skyvell Nilsson, M. (2023). Development and validation of experienced work-integrated learning instrument (E-WIL) using a sample of newly graduated

- registered nurses – A confirmatory factor analysis. *Nurse Education Today*, 128. <https://doi.org/10.1016/j.nedt.2023.105889>.
- Anggraeni, C. E. & M. S. (2021). Metode Pembiasaan Untuk Menanamkan Karakter Disiplin dan Tanggung jawab Di Ra Daarul Falaah Tasikmalaya. *Jurnal PAUD Agapedia*, 5(1), 100–109. <https://doi.org/10.17509/jpa.v5i1.39692>.
- Artino, A. R., Rochelle, J. S., Dezee, K. J., & Gehlbach, H. (2014). Developing questionnaires for educational research. *Medical Teacher*, 36(6), 463–474. <https://doi.org/10.3109/0142159X.2014.889814>
- Auld, G., Baker, S., McGirr, K., Osborn, K. S., & Skaff, P. (2017). Confirming the Reliability and Validity of Others' Evaluation Tools Before Adopting for Your Programs. *Journal of Nutrition Education and Behavior*, 49(5), 441–450. <https://doi.org/10.1016/j.jneb.2017.02.006>.
- Azwar, S. (2012). *Sikap Manusia: Teori dan Pengukurannya* (3rd ed.). Pustaka Pelajar.
- Bakhshandeh Bavarsad, M., Foroughan, M., Zanjari, N., Jorjoran Shushtari, Z., & Ghaedamini Harouni, G. (2021). Psychometric properties of modified MOS social support survey 5-item (MSSS-5-item) among Iranian older adults. *BMC Geriatrics*, 21(1). <https://doi.org/10.1186/s12877-021-02353-0>
- Blašková, M. (2014). Influencing Academic Motivation, Responsibility and Creativity. *Procedia - Social and Behavioral Sciences*, 159, 415–425. <https://doi.org/10.1016/j.sbspro.2014.12.399>.
- Cartwright, T. J., & Atwood, J. (2014). Elementary Pre-Service Teachers' Response-Shift Bias: Self-efficacy and attitudes toward science. *International Journal of Science Education*, 36(14), 2421–2437. <https://doi.org/10.1080/09500693.2014.925152>.
- Darna. (2016). Kesulitan Guru dalam Mengembangkan Penilaian Sikap Peserta Didik di SDN 75 Locok Kab. *Enrekang [UIN Alauddin Makassar]*. <http://repositori.uin-alauddin.ac.id/4804/>.
- Dini, P. P. (2018). Pendidikan Karakter Pada Anak Sekolah Dasar di Era Digital. *Jurnal Pendidikan Dasar*, 2(1). <http://journal.staincurup.ac.id/index.php/JPD>.
- Dwi Marsela, R., & Supriatna, M. (2019). Kontrol Diri: Definisi dan Faktor. *Journal of Innovative Counseling : Theory, Practice & Research*, 3(2), 65–69. http://journal.umtas.ac.id/index.php/innovative_counseling.
- Elly, R. (2016). Hubungan kedisiplinan terhadap hasil belajar siswa kelas v di SD Negeri 10 Banda Aceh. Pendidikan Guru Sekolah Dasar (PGSD) JURNAL PESONA DASAR. *Universitas Syiah Kuala*, 3(4), 43–53. <https://doi.org/https://jurnal.usk.ac.id/PEAR/article/view/7540>.
- Febrina Tarigan, E., Nilmarito, S., Islamiyah, K., Darmana, A., Retno, D., Suyanti, D., & Abstrak, I. A. (2022). Analisis Instrumen Tes Menggunakan Rasch Model dan Software SPSS 22.0. *JIPK*, 16(2). <http://journal.unnes.ac.id/nju/index.php/JIPK>.
- Gümüş, S., Hallinger, P., Cansoy, R., & Bellibaş, M. Ş. (2021). Instructional leadership in a centralized and competitive educational system: a qualitative meta-synthesis of research from Turkey. *Journal of Educational Administration*, 59(6), 702–720. <https://doi.org/10.1108/JEA-04-2021-0073>.
- Hafidhoh, N., & Rifa'i, M. R. (2021). Karakteristik Penilaian Pembelajaran pada Kurikulum 2013 di MI. In *Awwaliyah: Jurnal PGMI* (Vol. 4). <https://doi.org/https://ejournal.iai-tabah.ac.id/index.php/awaliyah/article/view/673>.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Pearson Education Inc.
- Halim, A. (2023). Pengaruh Gaya Kepemimpinan dan Disiplin Kerja terhadap Kinerja Pegawai pada Kantor Kecamatan Kalukku Kabupaten Mamuju. *Forecasting: Jurnal Ilmiah Ilmu Manajemen*, 2(1). <https://doi.org/https://stiemmamuju.e-journal.id/FJIIM/article/download/145/86>.
- Hikmah, & Muslimah. (2021). Validitas dan Reliabilitas Tes dalam Menunjang Hasil Belajar PAI. *Proceedings*, 1, 345–356. <https://jurnal.ar-raniry.ac.id/index.php/mudarrisuna/article/view/7839>.
- Idrus. (2019). Evaluasi Dalam Proses Pembelajaran. *Evaluasi Dalam Proses Pembelajaran*, 2, 920–935. <https://garuda.kemdikbud.go.id/documents/detail/1655265>.
- Jamaludin, U., Pribadi, R. A., & Yunus, F. A. M. (2023). Upaya dalam Meningkatkan Kedisiplinan Peserta Didik SDN Cibungur 1 melalui Program Late Vest and Goodness Project. *Didaktik: Jurnal Ilmiah PGSD FKIP Universitas Mandiri*, 9(2). <https://doi.org/10.36989/didaktik.v9i2.1208>.
- Joarder, M. H. R., Ashraf, M. A., & Ratan, S. R. A. (2020). Quality education from faculty management perspectives in private higher education: Does faculty commitment behaviour mediate? *International Journal of Education and Practice*. <https://doi.org/10.18488/journal.61.2020.81.190.206>.
- Kamaluddin, & Al-Asror, M. K. (2021). Pengaruh Kebijakan Terhadap Disiplin Kerja Karyawan BMT Baskara Cabang Purbolinggo Kecamatan Purbolinggo Kabupaten Lampung Timur Tahun 2020. *Jurnal Riset Ekonomi Islam*, 1(1). <https://doi.org/https://journal.steidarulquran.ac.id/index.php/jrei/article/view/5>.
- Kesuma, D. T., Yuliantini, N., & Bengkulu, U. (2022). Hubungan Antara Kemampuan Membaca Pemahaman

- dengan Hasil Belajar Siswa Kelas IV SDN 71 Kota Bengkulu Irfan Supriatna. *Juridikdas Jurnal Riset Pendidikan Dasar*, 5(1), 54–60. <https://doi.org/10.33369/juridikdas.4.2.172-178>.
- Khoirroni, I. A., Patinasarani, R., Hermayanti, N. I., & Santoso, G. (2023). *Pendidikan Karakter: Tingkat Anak Sekolah Dasar di Era Digital* (Vol. 2, Issue 2). <https://doi.org/10.9000/jpt.v2i2.372>.
- Kurniawan, M. F., Botifar, M., & Wanto, D. (2022). Evaluasi dalam Pembelajaran Pendidikan Islam di Sekolah Dasar Islam Terpadu An-Nida Kota Lubuk Linggau. *Inspirasi Dunia: Jurnal Riset Pendidikan Dan Bahasa*, 1(4), 23–35. <https://doi.org/https://journal.unimar-amni.ac.id/index.php/insdun/article/view/189>.
- Lestari, A. (2022). *Implementasi Nilai Karakter Disiplin dan Tanggung Jawab Pada Siswa Kelas III SD Negeri 280 Ongkoe Kabupaten Wajo [Universitas Bosowa]*. <https://repository.unibos.ac.id/xmlui/handle/123456789/2307>.
- Lewis, R. (2001). Classroom discipline and student responsibility: : the students' view. *Teaching and Teacher Education*, 17(3), 307–319. [https://doi.org/10.1016/S0742-051X\(00\)00059-7](https://doi.org/10.1016/S0742-051X(00)00059-7).
- Linden, W. J. (1997). Review test equating: Methods and practices. *Psychometrika*, 62(2), 287–290. <https://research.utwente.nl/files/6635362/Linden97book.pdf>.
- Lyublinskaya, I., & Kaplon-Schilis, A. (2022). Analysis of Differences in the Levels of TPACK: Unpacking Performance Indicators in the TPACK Levels Rubric. *Education Sciences*, 12(2). <https://doi.org/10.3390/educsci12020079>.
- Magdalena, I., Ramadhan, G., Wahyuni, H. D., & Safitri, N. D. (2023). Pentingnya Proses Evaluasi Dalam Pembelajaran Di Sekolah Dasar. *Ta'rim: Jurnal Pendidikan Dan Anak Usia Dini*, 4(3), 167–176. <https://doi.org/10.59059/tarim.v4i3.220>.
- Mardhotillah, N. I., & Yulhendri. (2022). Confirmatory factor analysis (cfa) interaksi dosen-mahasiswa dalam pembelajaran daring. *Jurnal Salingka Nagari*, 1(1), 93–100. <https://doi.org/10.24036/jsn.v1i1.13>.
- Murti, D. K., Gunarhadi, & Winarno. (2020). Development of Educational Comic with Local Wisdom to Foster Morality of Elementary School Students: A Need Analysis. *International Journal of Educational Methodology*, 6(2), 337–343. <https://doi.org/10.12973/ijem.6.2.337>.
- Musfirah. (2019). *Upaya Meningkatkan Kedisiplinan Peserta Didik dalam Menaati Tata Tertib di MTS DDI Kalumpang Kab.Pinrang*. <http://repository.iainpare.ac.id/1218/1/14.1100.085.pdf>.
- Nabilah, N., Karma, I. N., & Husniati, H. (2021). Identifikasi Kesulitan Guru dalam Melaksanakan Penilaian Autentik pada Kurikulum 2013 di SDN 50 Cakranegara. *Jurnal Ilmiah Profesi Pendidikan*, 6(4), 617–622. <https://doi.org/10.29303/jipp.v6i4.298>.
- Nitko, A. J., & Brookhart, S. M. (2014). *Educational assessment of students* (6th ed.). Pearson Education Limited.
- Nugraha, R., & Suyadi. (2019). Regulasi Diri dalam Pembelajaran. *Tarbiyah Al-Awlad: Jurnal Kependidikan Islam Tingkat Dasar*, 9(2). <https://doi.org/10.15548/alawlad.v9i2.1917>.
- Park, C. G. (2023). Implementing alternative estimation methods to test the construct validity of Likert-scale instruments. *Korean Journal of Women Health Nursing*, 29(2), 85–90. <https://doi.org/10.4069/kjwhn.2023.06.14.2>.
- Prasasty, A. T. (2017). Pengaruh disiplin dan motivasi belajar terhadap prestasi belajar matematika siswa kelas xsmk bina karya insan tangerang selatan. *UTILITY: Jurnal Ilmiah Pendidikan Dan Ekonomi*, 1(1). <https://doi.org/10.30599/utility.v1i1.64>.
- Prihastutia, M. W. A., & Santa. (2020). Analisis karakter tanggung jawab pada siswa kelas iv. *Jurnal Pendidikan Dan Pengajaran Guru Sekolah Dasar (JPPGuseda)*, 03, 128–132. <http://journal.unpak.ac.id/index.php/jppguseda>.
- Purwanto, D. I., Duwi Saputri, E., & Khoirotunnisa, A. U. (2020). Peran Guru PPKn dalam Meningkatkan Kedisiplinan pada Siswa Kelas VIII di MTS Hidayatul Athafal Kalirejo Bojonegoro. *Jurnal Pendidikan Edutama*, 1(1). [http://repository.ikipgribojonegoro.ac.id/1032/1/JURNAL SIAP DODI.pdf](http://repository.ikipgribojonegoro.ac.id/1032/1/JURNAL%20SIAP%20DODI.pdf).
- Puspasari, H., & Puspita, W. (2022). Uji Validitas dan Reliabilitas Instrumen Penelitian Tingkat Pengetahuan dan Sikap Mahasiswa terhadap Pemilihan Suplemen Kesehatan dalam Menghadapi Covid-19. *Jurnal Kesehatan*, 13(1). <http://ejurnal.poltekkes-tjk.ac.id/index.php/JK>.
- Putri, A. A., & Yarni, L. (2022). Pengendalian diri siswa di SMP Muhammadiyah Padang Panjang. *Orien: Cakrawala Ilmiah Mahasiswa*, 1(3), 273–278. <https://doi.org/10.30998/ocim.v1i3.6217>.
- Ramatni, A., Anjely, F., Cahyono, D., Rambe, S., & Shobri, M. (2023). Proses Pembelajaran dan Asesmen yang Efektif. *Journal on Education*, 05(04). <https://doi.org/10.31004/joe.v5i4.2687>.
- Retnawati, H. (2016). *Validitas reliabilitas & karakteristik butir*. Parama Publishing.
- Ritonga, A. A., Zahrani Purba, A., Nasution, F. H., Adriyani, F., & Azhari, Y. (2023). *Keterampilan Membaca pada Pembelajaran Kelas Tinggi di Tingkat SD/MI*. 2(3), 102–113. <https://doi.org/10.58192/insdun.v2i3.988>.

- Rudini, M. (2020). Efektifitas Analisis Butir Soal Mata Pelajaran Matematika pada Siswa Kelas IV dalam Meningkatkan Kualitas Guru di SDN Sabang. *Tolis Ilmiah: Jurnal Penelitian*, 2(1). <https://doi.org/10.56630/jti.v2i1.90>.
- Safitri, I., Sa'odah, & Magdalena, I. (2020). Analisis Perilaku Tanggung Jawab dalam Pembelajaran PKn Siswa Kelas IV SD Negeri Cipondoh 2 Kota Tangerang. *Trapsila: Jurnal Pendidikan Dasar*, 2(2), 36–49. <https://doi.org/10.30742/tpd.v2i2.1179>.
- Sahetapy, S. Y., Noach, R. M., & Samadara, P. D. (2019). Penyusunan Instrumen Pengukuran Ekspektasi Konsumen Retailer di Kota Kupang. *Jurnal STIEKEN*. <https://doi.org/https://journal.stieken.ac.id/index.php/penataran/article/view/390>.
- Septiani, T., & Mukhtar, T. (2022). Penerapan Pendidikan Karakter melalui Budaya di Sekolah Dasar. *COLLASE: Journal of Elementary Education*, 05, 3. <https://doi.org/10.22460/collase.v5i3.10838>.
- Shrestha, R., Kadel, R., & Mishra, B. K. (2023). A two-phase confirmatory factor analysis and structural equation modelling for customer-based brand equity framework in the smartphone industry. *Decision Analytics Journal*, 8. <https://doi.org/10.1016/j.dajour.2023.100306>.
- Siburian, P. (2012). Penanaman dan implementasi nilai karakter tanggung jawab. *Jurnal Generasi*, 5, 85–102. <https://doi.org/https://jurnal.unimed.ac.id/2012/index.php/gk/article/view/7215>.
- Siregar, L. A. (2018). *Penilaian Otentik dalam Kurikulum 2013*. *Jurnal STAIR Sibuhuan*. <https://ejournal.stai-br.ac.id/index.php/alrazi/article/download/26/21>.
- Susanti, E., Purwaningsih, E., & Rosyid, R. (2019). Pengaruh karakter disiplin siswa terhadap hasil belajar siswa madrasah aliyah al-mustaqim arang limbung. <https://doi.org/10.26418/jppk.v8i6.33577>.
- Syafitri, R. (2017). Meningkatkan Tanggung Jawab Belajar Melalui Strategi Giving Questions And Getting Answers Pada Siswa. *Jurnal Penelitian Dan Pengembangan Pendidikan*, 1(2), 57–63. <https://doi.org/10.23887/jppp.v1i2.12623>.
- Syifa, U. Z., Ardianti, S. D., & Masfuah, S. (2022). Analisis Nilai Karakter Tanggung Jawab Anak Dalam Pembelajaran Daring. *Jurnal Educatio FKIP UNMA*, 8(2), 568–577. <https://doi.org/10.31949/educatio.v8i2.2071>.
- Taufan, M. Y. (2022). Professional Development of Teachers, Competencies, Educational Facilities and Infrastructure on Teacher Performance and Learning Achievement of High School Students in Makassar City. *Golden Ratio of Social Science and Education*, 2(1), 24–38. <https://doi.org/10.52970/grsse.v2i1.168>.
- Wulandari, C. E., Haq, A. Z., & Milenia, B. F. (2023). Upaya Sekolah dalam Pembentukan Sikap Tanggung Jawab melalui Penugasan di MA Darul Islah, Lampung. In *The Strengthening of Pesantren Education Outcome Through The Synergy of Multidisciplinary Knowledge*. <https://doi.org/https://ejournal.unida.gontor.ac.id/index.php/shibghoh/article/view/10330>.
- Yanti, M. O. D., Sulistyarini, & Witarsa. (2022). Pengaruh karakter tanggung jawab terhadap hasil belajar perbankan dasar smk mandiri pontianak. *Jurnal Pendidikan Dan Pembelajaran*, 11(3). <https://doi.org/https://jurnal.untan.ac.id/index.php/jpdpb/article/view/53147/75676592469>.
- Zuhera, Y., & Habibah, S. (2017). Kendala Guru dalam Memberikan Penilaian Terhadap Sikap Siswa dalam Proses Pembelajaran Berdasarkan Kurikulum 2013 di SD Negeri 14 Banda Aceh. *Jurnal Ilmiah Pendidikan Guru Sekolah Dasar FKIP Unsyiah*, 2(1). <https://media.neliti.com/media/publications/187406-ID-kendala-guru-dalam-memberikan-penilaian.pdf>.