Research Article https://doi.org/10.12973/eu-jer.11.2.935

European Journal of Educational Research
Volume 11, Issue 2, 935 - 947.
ISSN: 2165-8714
https://www.eu-jer.com/

Construction of the Character Assessment Instrument for 21st Century Students in High Schools

Wiwin Mistiani
Universitas Negeri Yogyakarta, INDONESIA

Edi Istiyono*
Universitas Negeri Yogyakarta, INDONESIA

Amir Syamsudin
Universitas Negeri Yogyakarta, INDONESIA

Received: September 4, 2021 ▪ Revised: December 8, 2021 ▪ Accepted: February 17, 2022

Abstract: The study of character becomes a very important discussion in the 21st century. So that the integration of character values is very important both in the process and in educational assessment. The purpose of this study was to test the validity and reliability of the character assessment instrument for 21st-century high school students. The research approach used was quantitative with a sample of 200 high school students. Data analysis carried out includes validity and reliability tests. The test results of the instrument showed that the construct of the student character assessment instrument was declared valid and reliable. The content validity test shows the value of Aiken's > .80 in the high category. In the construct validation test with EFA, all variables have a loading factor > .5. In the CFA test, the model is declared fit with the estimated standard loading value of .40 and the t-count value > 1.96. Meanwhile, while in testing the reliability of the instrument obtained composite > .70 Cronbach's Alpha reliability > .70 which means reliable. So that this instrument is declared valid and reliable to measure the character of students in high school.

Keywords: *Assessment construct, character, validity, reliability.*

To cite this article: Mistiani, W., Istiyono, E., & Syamsudin, A. (2022). Construction of the character assessment instrument for 21st century students in high schools. *European Journal of Educational Research, 11*(2), 935-947. https://doi.org/10.12973/eu-jer.11.2.935

Introduction

The study of character has become a very urgent and serious subject of attention among psychologists and educational practitioners, especially in facing the 21st century. The 21st Century Education framework in the World Economic Forum lists three main skills that must be possessed by students, including basic literacy, character quality competition (Klaus, 2016). Likewise, the Centre for Curriculum Redesign mentions four skills that must be possessed by students including knowledge, skills, character, and metacognition (Bialik & Fadel, 2015). The two institutions agreed to establish the character as a very important skill to be taught to students in facing the 21st century. Strengthening character in the world of education has a very big influence in preparing a strong nation generation, both in intellectual, emotional, and spiritual aspects. Previous studies such as research by Harun et al. (2021) with good character, students can face the challenges of the 21st century.

Likewise in research by Suyitno et al. (2019), explained that collaboration between strong character and high competence can increase the nation's competitiveness to answer the challenges of the 21st century. Likewise, research by Nurhasanah and Nida (2016) confirms that with good character students can improve achievement academically. Therefore, countries in the world including Indonesia pay full attention to character development, especially to face the 21st century realizing a society with noble character, morals, ethics, culture, and civility based on the philosophy of Pancasila. The philosophy of Pancasila is a life that reflects the five precepts, namely belief in one God, just and civilized humanity, Indonesian unity, democracy led by wisdom that is present in deliberation/representation, and social justice for all Indonesian people (Kaelan, 2005).

Likewise, in 2010 the Indonesian government launched and implemented the policy of the National Movement for National Action Plan for National Character Education. Unfortunately, this policy has not been carried out evenly in schools throughout Indonesia. The next strategic step taken by the Ministry of Education and Culture to continue, expand, strengthen, optimize and deepen character education in schools by bringing up the Strengthening Character Education (PPK) movement. In 2020 The Minister of Education issues a policy on integrating character in assessment and is a requirement for student graduation. The policy is called the character survey. This step is taken as an effort to maximize

* **Corresponding author:**
Edi Istiyono, Universitas Negeri Yogyakarta, Indonesia. ✉ edi_istiyono@uny.ac.id

character strengthening through assessment. Aspects of the characters measured in the character survey are the six main characters that are very much needed in facing the 21st century. The six characters include:

Faithful character and fear of God Almighty is the character of students who understand religious teachings and beliefs and apply these understandings in everyday life. while Education Assessment Center Team (2019) explains that belief in God Almighty is a reflection of religious attitudes. The character of global diversity is the character of students who maintain their noble culture, locality, and identity, and keep an open mind in interacting with other cultures. The key elements of global diversity include the attitude of students who want to know and appreciate culture, have intercultural communication skills in interacting with others and are responsible for maintaining diversity by living in harmony and harmony with people of different races, religions, customs and groups (Wahyuningsih et al., 2021).

The character of global diversity is the character of students who maintain their noble culture, locality, and identity, and keep an open mind in interacting with other cultures. Indicators of the character of global diversity include the attitude of students who want to know and respect the culture, have intercultural communication skills in interacting with others and are responsible for maintaining diversity by living in harmony and harmony with people of different races, religions, customs and groups (Wahyuningsih et al., 2021)

The mutual corporation character relates to the character of Indonesian students who can carry out activities together voluntarily so that the activities carried out can run smoothly, easily and lightly. In line with that, the Education Assessment Center Team (2019) explains that mutual corporation is an attitude of students that reflects the action of working hand in hand and helping each other in solving problems/work. The elements of a mutual corporation are collaboration, caring, and sharing (Wahyuningsih et al., 2021)

Independent character, regarding the character of Pancasila students who do not depend on others and are responsible for their duties and obligations. Independent can be interpreted as an attitude that does not depend on others and uses energy, time of mind to achieve the desired goals (Education Assessment Center Team, 2019). Indicator of independent attitude is self-awareness and the situation at hand and self-regulation.

The character of critical reasoning is that students can objectively process information both qualitatively and quantitatively, build relationships between various information, analyze information, evaluate, and conclude. Critical reasoning is also an activity and assessment process that directs and regulates problem-solving and decision making (Facione, 2015; Facione & Gittens, 2015). Piirto (2011) views creative thinking skills as part of divergent thinking and divides creative thinking skills into five aspects consisting of fluency, flexibility, originality, elaboration, and transformation. Meanwhile, Hadar and Tirosh (2019) defines creative thinking as the attitude of students who can generate ideas and ideas in the problem-solving process. With the character of critical thinking, students can solve the problems they face. One cannot study well without thinking well. So critical thinking is related to career success, but also to success in higher education. Based on previous surveys, it was found that critical thinking and creativity are considered beneficial to improve students' ability in problem solving and decision making through collaboration and effective communication, especially in facing the 21st century (Selman & Jaedun, 2020).

Sixth, creativity is the character of students who can develop and produce something, both original and useful ideas/works, and have an impact on society and the nation. The key element of being creative consists of generating original ideas and producing original works and actions (Wahyuningsih et al., 2021). McGregor (2007) creative is defined as an ability to create or provide a unique idea from an alternative point of view. describes five aspects/dimensions of creative thinking, namely 1) fluency 2) flexibility 3) originality and 4) elaboration. Creativity skills are one of the important skills that must be possessed and are the key to effective learning in the 21st century. Creativity has been considered to play an important role throughout a person's life span and is directly related to the development of content knowledge and skills (Ritter & Mostert, 2017).

The measurement of these six aspects of character is very important to do in preparing students to face the 21st century which is full of challenges and global competition. To measure the six aspects of the character, a valid and reliable measuring instrument is needed. Valid relates to the extent to which the accuracy and accuracy of a measuring instrument in an assessment are supported by empirical facts and theoretical reasons (Azwar, 2010). An instrument is said to have high validity if the instrument performs its measuring function, or provides appropriate measuring results (Istiyono, 2020). The validity test used is content and construct validity. While reliable is related to the level of consistency or constancy of an instrument from two measurement results on the same object (Mehrens & Lehman, 1973). This reliability can be identified from the value of the correlation coefficient between the two observed scores obtained from the measurement results using parallel instruments (Allen & Yen, 1979). Instruments that have reliability will show the consistency of measurement, even though the measurement is carried out several times (Herwin & Mardapi, 2017). So, that by testing a valid and reliable assessment instrument can be obtained. So, the purpose of this study was to test the validity and reliability of the character assessment instrument for high school students.

Methodology

Research Design

The research approach used is a quantitative approach using validity and reliability analysis. Validity and reliability testing aims to determine the quality of the instruments used to measure the character of students in high school.

Sample and Data Collection

The sample in this study was 200 students from three high schools in Palu, Indonesia. According to Yadama and Pandey (1995) the adequacy of the sample greatly affects construct validity in determining the fit model in factor analysis. Samples allowed must be at least 100 samples (Kline, 2014; MacCallum et al., 1999; O'Rourke & Hatcher, 2013). From this opinion, the sample of this study was 200 students, so it was considered sufficient.

The variables studied in this study are 21st-century characters as measured by six aspects of character which include faith and fear of God Almighty, global diversity, independence, mutual cooperation, critical and creative reasoning. The measurement of these 6 main character aspects is limited to the indicators formulated by the Ministry of Education and Culture namely:

Table 1. The Distribution of Aspect, Indicators and Items

Aspect	Indicators	Measurement Items	Code
Faith and fear of God Almighty (IMTK)	Carrying out God's Commands	Students carry out worship according to their religion and beliefs	IMTK1
		Students actively participate in religious activities at school	IMTK2
		Students actively participate in religious activities (ta'lim assemblies) outside the school	IMTK3
	Avoid God's Prohibition	Students stay away from religious prohibitions (stealing/adultery/drinking / drugs)	IMTK4
Global Nationality (BKGL)	Appreciate culture	Students are proud to wear Indonesian traditional clothes	BKGL1
	Accept the difference	Students do not discriminate between friends in socializing	BKGL2
	live in peace	Students live in harmony with friends of different religions, races, customs and groups	BKGL3
Mutual cooperation (GTRG)	Cooperate	Students work together in doing group assignments	GTRG1
	Concern	Students help friends/other people who are affected by the disaster by raising funds	GTRG2
	share	Students are willing to share with friends who are in trouble by setting aside some of their pocket money.	GTRG3
Independent (MNDR)	self-regulation	Learners have targets in learning	MNDR1
		Have your strategy in learning	MNDR2
		Students can manage their study time well	MNDR3
	Have awareness of self and situation	Students know their strengths and weaknesses	MNDR4
Critical Reasoning (KRTS)	Questioning information	Students question something that they do not know.	KRTS1
	Analyze and evaluate reasoning	Students check the truth of the information before disseminating it	KRTS2
	Making decisions based on facts	Students make decisions/conclusions based on facts	KRTS3
Creative (KRTF)	Elaboration	Students are able to develop or enrich the ideas/works of others	KRTF1
	Generate original ideas/ideas	Students spark many ideas, suggestions in solving problems that other people don't think of	KRTF2
		Students make reports/works in detail and different from the theme	KRTF3

Analyzing of Data

The first validity test is content validity testing. According to Mardapi (2017), determination of content validity can be done through Focus group discussions (FGD) or determination through experts. The results of the assessment are then analyzed using Aiken's (1985) formula, which is as follows:

$v = \frac{\Sigma_s}{n(c-1)}$ ..... (1)

Information:

s = r-l0

V= index of expert agreement regarding item validity
r = Number given by expert
l_o = the lowest score of validity assessment (this case=1)
n = number of experts
c = The highest expert rating

If Aiken's validity index is less than or equal to .4 it is said to be less valid, if Aiken's validity index is between .4-.8 it is said to be moderately valid, and if Aiken's validity index is greater than .8 it is said to be very valid.

The next test is the construct validity test can be done by testing EFA (Exploratory Factor Analysis) and CFA (Confirmatory Factor Analysis). According to Child, (2006), the EFA test was carried out to determine the grouping design that refers to the data obtained in the field. Whereas CFA is used to confirm measurement design that has been carried out so that it can be seen the suitability and accuracy of grouping each item of the instrument (Yong & Pearce, 2013). Other than that, according to Brown and Moore (2012) and Harrington (2009) confirmatory factor analysis (CFA) is used to study the measurement design and can be used to identify the influence/relationship between dimensions/aspects, indicators, and observation items. Although the two analyzes have slightly different meanings and purposes, they can be used simultaneously and complement each other in a measurement. (Natalya & Purwanto, 2018). So that in this study the researchers used two types of analysis.

EFA testing is done using SPSS.16 Software. Provided that the instrument is said to have construct validity if the loading factor is greater than .5. The cumulative presentation shows whether or not a factor is used as an indicator with the number of presentations > 50% (Solimun, 2002). Factors that have an Eigen Value of more than 1 are factors that can be used as an indicator of traits/traits. Confirmatory Factor Analysis Testing was carried out using the LISREL 8.50 software. In the confirmatory factor analysis, there are two variables, namely the latent variable and the indicator variable. According to Nasir et al. (2015) latent variables can be interpreted as variables that cannot be modified/formed and constructed directly. The indicator variables are observed variables and can be measured directly. CFA analysis can be measured using the following formula:

Indicator = $x = \lambda\xi + \delta$.... (2)

As for x = factor of the indicator variable, λ = loading factor, ξ = exogenous latent variable and δ = error (Nasir et al., 2015).

According to Mulaik (2010) and Prudon (2015) evidence of construct validity in CFA can be seen from the t-value > 1.96 and the standardized factor load value of more than .4. Meanwhile, to test the fit of the model between the empirical data and the designed model, the goodness of fit model test was conducted. The model is said to be fit if the chi-square value is less than 2 df (Arbuckle, 1997), the p-value is more than .05, Root Mean Square Error of Approximation (RMSEA) is less than or equal to .08 and the Goodness of Fit Index (GFI) is greater than or equal to or close to .9 or close to one (Ferdinand, 2002; Sarwono, 2010)

The last test is reliability testing. The reliability test used in this study refers to Composite Reliability or reliability coefficients (ρc) (Fornell & Larcker, 1981; Ghozali & Fuad, 2014), which are presented in the following equation:

$$\rho c = \frac{(\Sigma \lambda i)^2}{(\Sigma \lambda i)^2 + \Sigma_i var(\varepsilon i)} = \rho c = \frac{(\Sigma \lambda i)^2}{(\Sigma \lambda i)^2 + \Sigma_i (1-\lambda_i^2)} \text{ ..... (3)}$$

Where is λ_i=the component loading or factor loading indicator i and var $(\varepsilon_i) = (1 - \lambda_i^2)$. The construct reliability coefficient has a value between 0 to 1. The closer to 1, the higher the reliability coefficient. The criteria are considered reliable if ρc > .70. Instrument reliability based on internal consistency is determined using the Alpha coefficient from Cronbach (Cronbach's Alpha). Alpha coefficient calculation is done with SPSS program. The criterion of a good reliability coefficient is at least .7 (Nunnally, 1978).

Findings / Results

Validity of Character Assessment Instruments for Students in High School

The first test is content validation testing. The content validation test tests the readability of the character assessment instrument which includes aspects of 1) the accuracy of indicators and instrument items, 2) the use of writing rules which include a) the accuracy of the use and writing of letters, c) the accuracy of word writing, d) the accuracy of the use of

punctuation marks, 3) the clarity of the sentences used. The assessment uses a scale of 5 with the lowest score being 1 and the highest being 5. Content validation is carried out by expert judgment as many as 7 experts consisting of 2 (two) measurement experts and 5 (five) character education practitioners. The results of the assessment by the expert were then analyzed using the Aiken formula to obtain the Aiken index. The results of content validity are shown as follows:

Table 2. The results of the validation of the content of the expert judgment

Aspect	Indicators	Item Code	V-Aiken
Faith and fear of God Almighty (IMTK)	Carrying out God's Commands	IMTK1	.857
		IMTK2	.821
		IMTK3	.893
	Staying Away from Religious Prohibition	IMTK4	.893
Global Nationality (BKGL)	Appreciate culture	BKGL1	.821
	Accept the difference	BKGL2	.929
	live in peace	BKGL3	.893
Mutual cooperation (GTRG)	Cooperate	GTRG1	.821
	Concern	GTRG2	.929
	Share	GTRG3	.893
Independent (MNDR)	Self-regulation	MNDR1	.821
		MNDR2	.893
		MNDR3	.857
	Have awareness of self and situation	MNDR4	.929
Critical Reasoning (KRTS)	Questioning information	KRTS1	.964
	Analyze and evaluate reasoning	KRTS2	.893
	Making decisions based on facts	KRTS3	.929
Creative (KRTF)	Elaboration	KRTF1	.893
	Generate original ideas/ideas	KRTF2	.929
		KRTF3	.929

Based on the table above, it is found that the Aiken index value is in the range of .821 to .964. V-Aiken values with a range of more than .80 are categorized as high. Thus, the developed instrument items meet the valid criteria and can be tested further.

The second validity test is constructed validity. This is done in two stages. The first stage is a unidimensional test which is carried out using the exploratory factor analysis method with the SPSS 16 program. This test is carried out to test whether the item only measures one ability (Retnawati, 2017). The main requirement before carrying out the unidimensional test is to look at the standard Kaiser-Meyer Olkin – MSA > .50 and the significance of the unidimensional Bartlett's test < .05 (Anderson et al., 2001). Bartlett's test was used to test the normality of the data. Meanwhile, the KMO -MSA Test evidence was used to determine the adequacy of the sample. Based on the results of the analysis obtained the value of KMO-MSA is .797 > .50. While the sig value of Bartlett's test is .000 < .05, which means it is significant. This indicates that the instrument is feasible for factor analysis. Furthermore, to obtain items that represent dimensions, the next step is the extraction process so that several factors are formed. Each factor formed has an eigenvalue, and a factor that has more than one eigencoefficient then that factor is applied or used (Santoso, 2012). The unidimensional assumption is considered fulfilled if the assessment instrument contains one dominant aspect that is used to see a person's ability (Hambleton & Swaminathan, 1985). Furthermore, if in the factor analysis one dominant factor is found, then that factor becomes the single dimension of an instrument (Naga, 1992). If the eigenvalue of the first factor has a value up to several times the eigenvalue, the second factor and so on are almost the same, then it is said that the unidimensional condition is fulfilled. The following table 4 shows the total cumulative variance.

Table 4. Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% Of Variance	Cumulative %	Total	% Of Variance	Cumulative %	Total	% Of Variance	Cumulative %
1	5.052	25.260	25.260	5.052	25.260	25.260	2.826	14.131	14.131
2	2.102	1.508	35.769	2.102	1.508	35.769	2.301	11.506	25.637
3	1.934	9.669	45.437	1.934	9.669	45.437	2.232	11.159	36.796
4	1.769	8.844	54.281	1.769	8.844	54.281	2.216	11.080	47.876
5	1.751	8.756	63.037	1.751	8.756	63.037	2.143	1.716	58.592
6	1.222	6.110	69.148	1.222	6.110	69.148	2.111	1.555	69.148

In table 4 above, it can be seen that there are 6 forming factors with eigenvalues > 1. Factor one has an eigenvalue of 5.052, factor two has an eigenvalue of 2.102 and factor three has an eigen value of 1.934, factor four has an eigenvalue of 1.769, factor five has an eigen value of 1.769 eigen value 1.751 and a factor of six has an eigenvalue of 1.222. Using this procedure, the number of components with eigenvalues > 1 will be calculated as the number of extracted factors which are then assigned to the model. Gorsuch proposed a scree plot on the criteria for Eigenvalues > 1 (Pett et al., 2003). However, this study relies on the second approach, namely the percentage of variance. The explained total variance table shows six components, each of which has a different explained variance value. Component one accounted for 25.26% of the variance, component two accounted for 1.51% of the variance, factor three explained 9.67% of the variance, factor four explained 8.84% of the variance, factor five explained variance of 8.76% and factor six explained variance of 6.11%. Of the six factors, the total variance is 69.15%. In terms of the variance described, although many researchers stop the factor extraction process when the total variance described reaches 50-80%, there are no definite guidelines for a specific threshold (Pett et al., 2003). The last factor criterion is not less than 5% of the variance described (Pett et al., 2003). Because the variance described is more than 5% or equal to 69.15%, it can be concluded that the assessment instrument is valid and able to explain the character of students. Unidimensional evidence can also be seen in the following scree plot image.

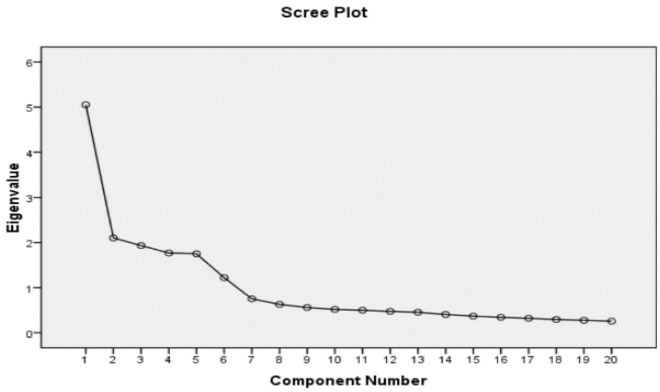

Figure 1. Unidimensional Test Scree plot

If seen in the figure, the scree plot shows that the distance between component one and component two is several times the distance between the other components. A steep scree plot indicates the presence of a dominant component. That is, the student character instrument only measures one dimension or one factor.

The next test is to test the fit of the model. The model fit test was carried out using 2nd order confirmatory factor analysis (CFA) using Lisrel.85. The CFA process is carried out to see the dimensionality of the developed model. The results of the 2nd order CFA analysis are as follows:

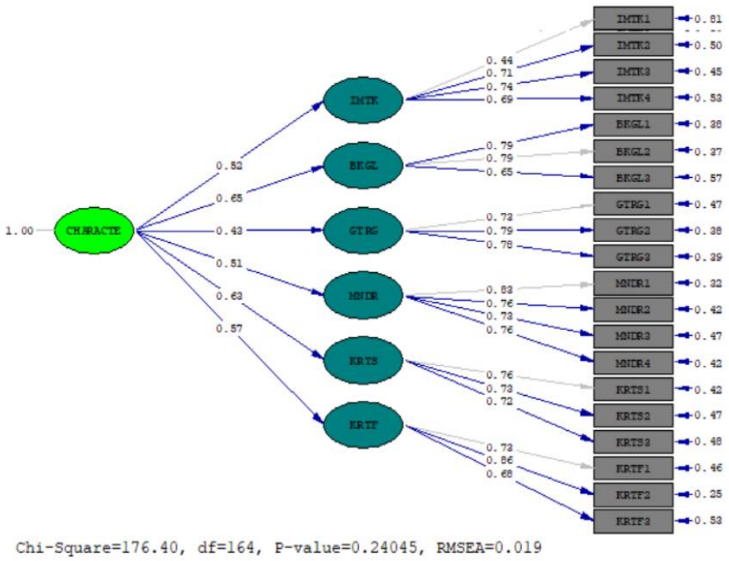

Figure 2. Loading Factor Standardized Solution

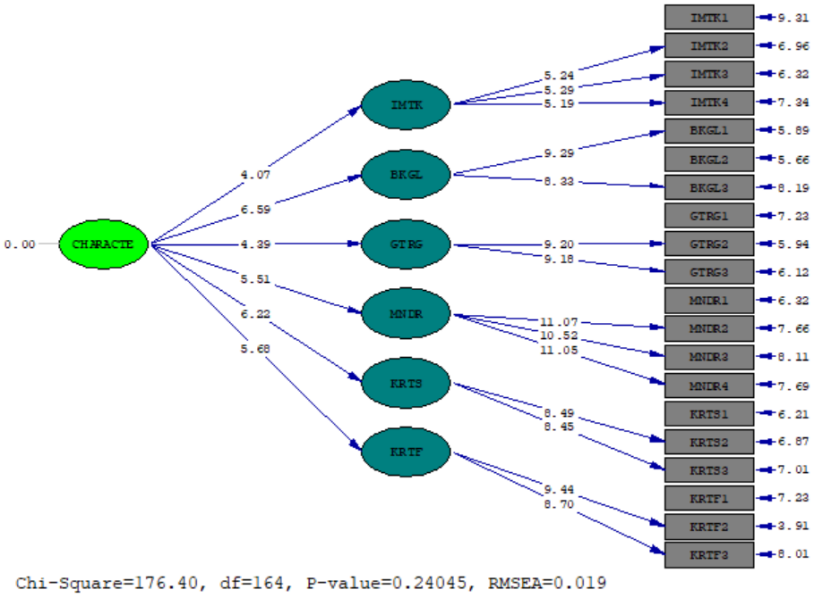

Figure 3. t-Value

Based on Figures 2 and 3, it can be seen that there are 20-character assessment instruments with a significant factor loading with a factor loading value of Standardized Solution .40 and a t-count value > 1.96 both at the first and second levels (Hair et al., 2010). In addition, the result of a p-value that is greater than that is equal to .240 and RMSEA .08 that is equal to = .019. It means that the proposed theoretical model fits the empirical data.

Table 4. Criteria for Model Fit

No.	Fit Index	Score	Standard value	Information
1.	P-Value	.240	.05	Fit
2.	RMSEA	.019	.08	Fit
3.	Chi-Square	176.4	2df (164)	Fit
4.	GFI	.92	.90	Fit
5	AGFI	.90	.90	Fit
6	CFI	.99	.90	Fit
7	IFI	.99	.90	Fit

Based on the table. The p-value is .240 > .05, the RMSEA value is .019 < .08 and the Chi-Square value is less than 2 df (165) and the GFI value is .92 > .90, the AGFI value is .90 >.90, the CFI value is .99 > .90 and the IFI value is .99 >.9. This is in line with the opinion which states that the model is said to be fit if at least the chi-square value is less than or equal to 2 df (Arbuckle, 1997), the p-value is more than or equal to .05 RMSEA less than or equal to .08 and Goodness of Fit Index (GFI) greater than or equal to or close to .9 or close to one (Ferdinand, 2002; Pedhazur, 1997; Sarwono, 2010). Thus, it means that the proposed theoretical model fits the empirical data. These results indicate that the theoretical model of the Character Assessment variable. So, it can be concluded that the character assessment instrument construct consisting of 6 valid and significant aspects to measure the latent variable (character) and 20 item items is said to be valid and significant to its aspects.

Reliability of Character Assessment Instruments for Students in High School

The next unit of analysis is reliability testing. Based on the results of the analysis of construct reliability can be seen in the table as follows.

Table 5. Construct Reliability Analysis

Instrument Code	λ	λ_i^2	1-λ_i^2	ω
IMTK1	.44	.194	.806	.936
IMTK2	.71	.504	.496	
IMTK3	.74	.548	.452	
IMTK4	.69	.476	.524	
BKGL1	.79	.624	.376	
BKGL2	.79	.624	.376	
BKGL3	.65	.423	.578	
GTRG1	.73	.533	.467	
GTRG2	.79	.624	.376	
GTRG3	.78	.608	.392	
MNDR1	.82	.672	.328	
MNDR2	.76	.578	.422	
MNDR3	.73	.533	.467	
MNDR4	.76	.578	.422	
KRTS1	.76	.578	.422	
KRTS2	.73	.533	.467	
KRTS3	.72	.5184	.4816	
KRTF1	.73	.5329	.4671	
KRTF2	.86	.7396	.2604	
KRTF3	.68	.4624	.5376	
Σ	14.66		9.1182	

Based on table 5. The total value of the omega coefficient (ω) or composite reliability is .936, which is greater than .70 which means reliable. Likewise, if you look at each aspect / dimension, the Composite reliability value is as follows.

Table 6. Composite reliability values in each aspect/dimension

No	Aspect/dimension	Composite reliability *(pc)*
1	Faith and fear of God Almighty (IMTK)	.745
2	Global nationality (BKGL)	.789
3	Mutual cooperation (GTRG)	.811
4	Independent (MNDR)	.794
5	Critical reasoning (KRTS)	.781
6	Creative (KRTF)	.803

Table 6. Shows that the composite value of reliability in each aspect is greater than or equal to .70. Although the composite reliability value in each aspect is smaller than the total composite reliability, it is still acceptable. According to Hair et al. (2010) the instrument is said to be reliable if the composite value of reliability is greater than or equal to .70. So that it can guarantee that the 21st century character assessment instrument is reliable, both in terms of total reliability and in every aspect. Instruments that have good reliability will be able to provide the same information even though they are used for several things with the same object.

Meanwhile, based on the results of internal reliability analysis, the value of the Alpha coefficient (Cronbach's Alpha) of .84 is greater than .70 which means reliable. This is in accordance with the opinion of Nunnally (1978) which states that the criterion of a good reliability coefficient is at least .70.

Discussion

Validity and reliability testing on the construct of the 21st century character instrument in this case which focuses on the "profile of Pancasila students" which consists of 6-character aspects, 15 indicators and 20 items are declared valid and reliable. The six aspects of character assessment consist of faith and piety to God Almighty, global nationality, mutual cooperation, independence, critical and creative reasoning. The instrument construct was then tested and analyzed using validity and reliability analysis to obtain a good instrument so that it can be used in assessing the character of students in high school. this is in line with Gronlund and Linn (1990) which states that three main criteria must be met by an assessment instrument to be declared to have good quality, including reliability (validity), constancy (reliability), and practicality. In line with this opinion, Kerlinger (2006) states that if someone does not know the validity and reliability of an instrument used, then the data obtained is not necessarily correct and will affect the conclusions obtained.

The findings from the results of content validity testing whose measurements are based on the results of the 7 expert judgment assessments after being analyzed by V-Aiken's, the 20-character instrument items are declared valid with a

range between .821 to .964. V-Aiken's values with a range of more than .80 are categorized as high. Thus, the developed instrument items meet the valid criteria and can be tested further. This is in line with Aiken (1985) which states that if the V-Aiken's index is less than or equal to .40 it is said to be less valid, .40 -.80 is said to be moderately valid, and if it is greater than .80. Content validity is said to be very high.

The findings on the results of the construct validity analysis in the first stage are unidimensional testing using exploratory factor analysis, the KMO-MSA value is more than .50 and the significance of the Bartlett's test is less than .05. Thus, it can be concluded that all the results of the analysis have been significant so that the instrument is feasible for factor analysis. This is in line with the opinion of Anderson et al. (2001) which states that the main requirement before carrying out the unidimensional test is to look at the standard Kaiser-Meyer Olkin (KMO) - MSA > .50 and the significance of the unidimensional the Bartlett's test < .05.

The findings from the analysis of total variance formed 6 factors that have eigenvalues > 1. Thus, by using this procedure, the number of components with eigenvalues > 1 will be calculated as the number of extracted factors which are then assigned to the model (Fabrigar & Wegener, 2012). Gorsuch proposed a scree plot on the criteria for Eigenvalues > 1 (Pett et al., 2003). The scree plot shows that the distance from component 1 to component 2 is very far or several times the distance between other components. A steep scree plot indicates the presence of a dominant component. That is, the student character instrument only measures one factor or one dimension. This is in line with Hambleton and Swaminathan (1985) assumed that the unidimensional assumption was fulfilled if the test contained one dominant aspect that was used to see a person's ability. The same statement was put forward by Naga (1992) which stated that if in the factor analysis one dominant factor is found, then that factor becomes a single dimension in an instrument. When viewed from the overall value of the total variance of the six factors is 69.15%. Many researchers stop the factor extraction process when the total variance described reaches 50-80%, however, there are no definite guidelines for specific thresholds. Because the variance described is more than 5% or equal to 69.15, it can be concluded that the character assessment instrument is declared valid and able to explain the character of the students (Pett et al., 2003)

Findings in the first aspect, namely the aspect of the character of faith and piety to God Almighty, which is represented by two indicators and four measurement items, it shows that based on the analysis of the seven expert judgment items IMTK3, students actively participate in religious activities at school and stay away. This finding is in line with Education Assessment Center Team, (2019) states that the character of faith and piety to God is marked by the participation of students in various religious activities both at school and outside school.

The findings of the second aspect of global diversity are represented by three indicators and three items. Items BKGL1 and BKGL-2 contribute greatly to the character aspect of global diversity. Item BKGL1 relates to students being proud to wear traditional Indonesian traditional clothes and BKGL2 regarding students not discriminating against friends when hanging out. This is in line with the formulation in the Strategic Plan of the Ministry of Education and Culture 2020-2024 (Wahyuningsih et al., 2021) which states that Indonesian students who have the character of global diversity are students who know and appreciate culture, have intercultural communication skills in interacting with others. including associating and communicating with anyone regardless of social status, religion, customs and class.

The third aspect is the character of the mutual corporation. This aspect relates to the ability to carry out activities together voluntarily so that the activities carried out can run smoothly, easily and lightly. This mutual corporation aspect consists of three indicators and 3 items. Item GTRG2 has the greatest contribution to the mutual corporation aspect. The item relates to the indicator of concern shown by the attitude of students who help friends / other people who experience disasters by holding fundraisers. This is in line with Zuchdi (2011) who explains that caring is an attitude and action that always wants to assist people in need.

The fourth aspect is an independent character. This aspect consists of two indicators and four question items. Item MNDR1 is an item that contributes greatly to the independent character aspect. Item MNDR1 relates to the attitude of students having a target in learning or having a goal in learning. Students who have an independent attitude in this case self-regulation will be able to create behaviour to fulfil a goal or several desired goals (Pintrich et al., 1994).

The fifth character aspect is critical reasoning character. This aspect Critical reasoning is measured based on 3 indicators and 3 items. Item KRTS1 has the greatest contribution to the critical reasoning aspect. Item KRTS1 relates to students asking questions about something they do not know. This is in line with Kurfiss (1988) states that the character of critical reasoning relates to investigations whose purpose is to explore situations, phenomena, questions that integrate all available information and therefore can be justified convincingly.

The last aspect of the character, namely the sixth is the creative character. This aspect is measured based on 2 indicators and 3 items. The KRTF2 item is an item that has a major contribution to the creative character aspect. This item relates to students sparking many ideas, suggestions in solving problems that other people don't think of. This is by the opinion of McGregor (2007) which states that creative thinking is an ability to create or provide a unique idea from an alternative point of view.

Based on the results of the analysis, it was also found that the construct of the student character instrument showed a loading factor of more than .40 and a t-value of more than 1.96, meaning that it met the criteria of construct validity.

According to Prudon, (2015), this indicates that the instrument in the model for measuring the character of 21st-century high school students is significant in measuring latent variables.

Likewise, the measurement of aspects to latent variables of all aspects shows the loading factor is more than .4 and the t-value is more than 1.96. Of the six aspects of character assessment, the global national character (BKGL) aspect has a major contribution to the latent variable (character) of students.

These findings are also supported by the results of the model fit test. The findings of this study indicate that all indicators in measuring the goodness of fit model are met or fit. Such as chi-square value < 2 df (Arbuckle, 1997), p-value > .05, RMSEA < .08 (Sarwono, 2010) and GFI > .90 (Ferdinand, 2002; Pedhazur, 1997) and AGFI .90 (Hair et al., 2010). According to Hoyle, (2004). This indicates that the observed and implied covariance matrices are the same and significant. This means that the theoretical model and the empirical findings are the same or there is no significant difference.

The last unit of analysis is constructed reliability. The results of the analysis of construct reliability estimates found that 20 items measuring the character of students were declared reliable with a composite reliability or omega coefficient of .936 exceeding the predetermined criteria, namely by Hair et al., (2010) which stated that the instrument construct was said to be reliable. If the reliability coefficient is more than or equal to .70. Another opinion is also mentioned by Ebel and Frisbie (1991) which states that if the test/non-test instrument is used as a standard instrument, the reliability coefficient is between .85 and .95, but if it is used for class assessment it is at least .65. According to Widhiarso (2016) High composite reliability (*pc*) indicates that there is internal consistency and homogeneous variance between measurement items with one another. According to Huck (2007), although the items are different, they measure the same construct. Likewise, the results of the reliability analysis of Cronbach's Alpha obtained a value of .84 which is greater than .70 which means reliable. This is in accordance with the opinion of Nunnally (1978) which states that the criterion of a good reliability coefficient is at least .70. So, it can be concluded that the character assessment model for 21st-century students in this study can be used to measure the character constructs of students consistently.

Conclusion

The results of testing the validity of both content validity and construct validity as well as instrument reliability in this study indicate that the six aspects and 15 indicators and 20 items that make up the latent variables of the 21st-century character of students are valid and reliable. In the exploratory factor analysis test, it shows that each item only measures one ability, and the twenty 20 items can explain > 50% of the character of students in high school. In the second-order confirmatory factor analysis (2nd CFA) test, it is seen that there is a relationship between latent variables to aspects and from aspects to items and the relationship is significant. Likewise, in the goodness of fit test, this model has been declared fit. Thus, these findings indicate that the observed and implied covariance matrices are significantly the same. This means that between the theoretical model and the empirical findings there is no difference or the same. Likewise, in reliability testing, it was found that the instrument was reliable. So, it can be concluded that the character assessment model for 21st-century students in this study can be used to measure the character constructs of students consistently. The research implies that with a valid and reliable character assessment instrument the teacher can obtain precise and accurate information so that it can be used as a basis for follow-up in improving the character of students in high school. So, it can be concluded that the character assessment model for 21st-century students in this study can be used to measure the character constructs of students consistently. The research implies that with a valid and reliable character assessment instrument the teacher can obtain precise and accurate information so that it can be used as a basis for follow-up in improving the character of students in high school. It can be concluded that the character assessment model for 21st-century students in this study can be used to measure the character constructs of students consistently. The research implies that with a valid and reliable character assessment instrument the teacher can obtain precise and accurate information so that it can be used as a basis for follow-up in improving the character of students in high school.

Recommendation

This study recommends the use of a 21st-century character measurement scale which consists of aspects of faith and piety to God Almighty, global diversity, independence, mutual cooperation, critical and creative reasoning in character assessment in high school. The integration of character in assessment is very important as a form of monitoring the character development of students in high school. In addition, the application of this assessment is also a means of achieving the vision and mission of a golden Indonesia in 2045, namely as a nation of noble character, advanced and able to compete with other nations. This study also recommends to further researchers to develop instruments and test indicators that have not been discussed in this study and present the results of research on the character of the 21st century using this instrument.

Limitations

The limitation of this study lies in the too many aspects measured in this study, causing few indicators to be measured and only focused on indicators set by the Ministry of Education and Culture 2020-2024. Further research is expected to be able to describe more broadly the indicators and items.

Authorship Contribution Statement

Mistiani: Conceptualization, design, analysis, writing. Istiyono: Editing/reviewing. Syamsudin: Supervision.

References

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement, 45*(1), 131–142. https://doi.org/10.1177/0013164485451012

Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Waveland Press

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing*: *A Revision of bloom's taxonomy of educational objectives*. Longman.

Arbuckle, J. L. (1997). *Amos user's guide version 3.6.* Small Water Corporation.

Azwar, S. (2010). *Penyusunan skala psikologi* [Preparation of the psychological scale]. Pustaka Pelajar.

Bialik, M., & Fadel, C. (2015). *Meta-learning for the 21st century: What should students learn?* Canter for Curriculum Redesign. https://www.curriculumredesign.org

Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. In R. H. Hoyle (Ed.), *Handbook of structural equation modelling* (pp. 361–379). The Guilford Press.

Child, F. (2006). *The essentials of factor analysis* (3rd ed.). Continuum International Publishing Group.

Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement* (5th ed.) Englewoods Cliffs

Education Assessment Center Team. (2019). *Model penilaian karakter* [Character assessment model]. Pusat Penilaian Pendidikan/ Education Assessment Center.

Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory factor analysis*. Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199734177.001.0001

Facione, P., & Gittens, C. A. (2015). *Think critically*. Pearson.

Facione, P. A. (2015). Critical thinking: What it is and why it counts. *Insight Assessment, 28*(1), 1-31. https://bit.ly/3uWv9ul

Ferdinand, A. (2002). *Structural equation modelling in management research*. Diponegoro University Publishing Agency.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.2307/3151312

Ghozali, I., & Fuad. (2014). *Structural equation modeling: Teori, konsep, dan aplikasi dengan program Lisrel 9.1* [Structural equation modelling: Theories, concepts, and applications with the Lisrel 9.10 program]. Badan Penerbit Universitas Diponegoro/ Diponegoro University Publishing Agency.

Gronlund, N. E., & Linn, R. L. (1990). *Measurement and evaluation in teaching.* McMillan Company.

Hadar, L. L., & Tirosh, M. (2019). Creative thinking in mathematics curriculum: An analytic framework*. Thinking Skills and Creativity, 33,* 100585. https://doi.org/10.1016/j.tsc.2019.100585

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.) Pearson Prentice Hall.

Hambleton, R. K., & Swaminathan, H. (1985). *Item respond theory: Principles and applications*. Sage Publications.

Harrington, D. (2009). *Confirmatory factor analysis*: *Pocket guides to social work research methods* (1st ed.). Oxford University Press.

Harun, Kartowagiran, B., & Manaf, A. (2021). Student attitude and mathematics learning success: A meta-analysis. *International Journal of Instruction*, *14*(4), 209–222. https://doi.org/10.29333/iji.2021.14413a

Herwin, H., & Mardapi, D. (2017). An emotion assessment model for elementary school students. *Journal of Educational Research and Evaluation, 21*(1), 80–92. https://doi.org/10.21831/pep.v21i1.14504

Hoyle, R. H. (2004). Confirmatory factor analysis. In M. Lewis-Beck., & T. Liao (Eds.), *Encyclopedia of social science research methods* (pp. 169-174). SAGE Publications, Inc.

Huck, S. W. (2007). *Reading statistics and research.* Allyn & Bacon.

Istiyono, E. (2020). *Pengembangan instrumen penilaian dan analisis hasil belajar fisika, dengan teori tes klasik dan modern* [Development of assessment instruments and analysis of physics learning outcomes, with classical and modern test theory]. UNY Press.

Kaelan. (2005). *Metode penelitian kualitatif bidang filsafat* [Qualitative research methods in philosophy]. Paradigma.

Kerlinger. (2006). *Asas–asas penelitian behaviour* [Principles of research behavior] (3rd ed.). Gadjah Mada University Press.

Klaus, S. (2016). *The global competitiveness report 2016–2017*. World Economic Forum. https://www3.weforum.org

Kline, P. (2014). *An easy guide to factor analysis*. Routledge. https://doi.org/10.4324/9781315788135

Kurfiss, J. G. (1988). *Critical thinking: Theory, research, practice, and possibilities* (ASHE-ERIC Higher Education Report No:2). Association for the Study of Higher Education. https://files.eric.ed.gov/fulltext/ED304041.pdf

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84–99. https://doi.org/10.1037/1082-989X.4.1.84

Mardapi, D. (2017). *Pengukuran, penilaian dan evaluasi Pendidikan* [Measurement, assessment and evaluation of education]. Parama Publising.

McGregor, D. (2007). *Developing thinking; developing learning*. McGraw-Hill Education

Mehrens, W. A., & Lehman, I. J. (1973). *Measurement and evaluation in education and psychology.* Holt, Rinehart, and Winston, Inc.

Mulaik, S. (2010, October 7). *Another look at a cancer research model. Theory and indeterminacy in the BMC model by Hayduk et al.* [Paper presentation]. Annual meeting of the Society for Multivariate Experimental Psychology. Atlanta, GA, United States.

Naga, D. S. (1992). *Pengantar teori skor pada pengukuran pendidikan* [An introduction to scoring theory on educational measurement]. Gunadarma.

Nasir, Sappaile, B. I., & Arafah, K. (2015). Pendeteksian pelaksanaan supervisi akademik pengawas sekolah pada SMA Negeri di Kota Baubau melalui analisis faktor konfirmatori (CFA) [Detection of the implementation of the academic supervision of school supervisors at SMA Negeri in Baubau City through confirmatory factor analysis (CFA)]. In K. Arafah & Patahuddin (Eds.), *Proceedings of the Annual Scientific Conference of Himpunan Evaluasi Pendidikan Indonesia* (pp. 126-138). Universitas Negeri Makassar.

Natalya, L., & Purwanto, C. V. (2018). Exploratory and confirmatory factor analysis of the academic motivation scale (AMS)–Indonesian. *Makara Human Behavior Studies in Asia*, *22*(1), 29-42. https://doi.org/10.7454/hubs.Asia.2130118

Nunnally, J. C. (1978). *Psychometric theory*. McGraw-Hill.

Nurhasanah, N., & Nida, Q. (2016). Character building of students by guidance and counseling teachers through guidance and counseling services. *Jurnal Ilmiah Peuradeun*, *4*(1), 65-76. https://doi.org/10.26811/peuradeun.v4i1.86

O'Rourke, N., & Hatcher, L. (2013). *A step-by-step approach to using SAS® for factor analysis and structural equation modelling*. SAS Institute Inc.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed). Thompson Learning.

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: the use of factor analysis for instrument development in health care research*. Sage. https://doi.org/10.4135/9781412984898

Piirto, J. (2011). *Creativity for 21st century skills how to embed creativity into the curriculum.* Sense Publishers. https://bit.ly/34H8jwh

Pintrich, P. R., Roeser, R. W., & De Groot, E. V. (1994). Classroom and individual differences in early adolescents' motivational and self-regulated learning. *Journal of Early Adolescence*, *14*(2), 139–161. https://doi.org/10.1177/027243169401400204

Prudon, P. (2015). Confirmatory factor analysis as a tool in research using questionnaires: A critique. *Comprehensive Psychology*, *4*(3), 1-19. https://doi.org/10.2466/03.cp.4.10

Retnawati, H. (2017). *Validitas, reliabilitas, dan karakteristik butir*: *Panduan untuk peneliti, mahasiswa, dan psikometrian* [Item validity, reliability and characteristics: A guide for researchers, students, and psychometricians]. Parama Publishing.

Ritter, S. M., & Mostert, N. (2017). Enhancement of creative thinking skills using a cognitive-based creativity training.

Journal of Cognitive Enhancement, *1*(3), 243–253. https://doi.org/10.1007/s41465-016-0002-3

Santoso, S. (2012). *Panduan lengkap SPSS versi 20* [Complete guide to SPSS version 20]. PT Elex Media Komputindo.

Sarwono, Y. (2010). *Pengertian dasar structural equation modeling (SEM)* [Basic understanding of structural equation modelling (SEM)]. Ilmiah Manajemen Bisnis.

Selman, Y. F., & Jaedun, A. (2020). Evaluation of the implementation of 4C Skills in Indonesian subject at senior high schools. *Indonesian Journal of Education/ Jurnal Pendidikan Indonesia, 9*(2), 244–257. https://doi.org/1.23887/jpi-undiksha.v9i2.23459

Solimun. (2002). *Multivariate analysis structural equation modelling (SEM) LISREL dan AMOS: Aplikasi di manajemen, ekonomi pembangunan, psikologi, sosial, kedokteran dan agrokompleks* [Multivariate analysis structural equation modelling (SEM) LISREL and AMOS: Applications in management, development economics, psychology, social, medicine and agrocomplex]. Universitas Negeri Malang

Suyitno, H., Zaenuri, Sugiharti, E., Suyitno, A., & Baba, T. (2019). Integration of character valuesin teaching-learning process of mathematics at elementary school of Japan. *International Journal of Instruction*, *12*(3), 781–794. https://doi.org/10.29333/iji.2019.12347a

Wahyuningsih, S., Zuchron, D., & Purnawinarni, E. (2021). *Tunas pancasila sebagai duta profil pelajar pancasila di sekolah dasar* [Tunas Pancasila as an ambassador for the profile of Pancasila students in elementary schools]. Direktorat Sekolah Dasar, Kementerian Pendidikan Kebudayaan, Riset Dan Teknologi /Directorate of Elementary Schools, Ministry of Education, Culture, Research and Technology. https://bit.ly/3sRmdUk

Widhiarso, W. (2016). Estimasi reliabilitas pengukuran dalam pendekatan model persamaan struktural [Estimation of measurement reliability in the structural equation modeling approach]. *Buletin Psikologi*, *17*(1), 33–38. https://doi.org/10.22146/bpsi.11480

Yadama, G. N., & Pandey, S. (1995). Effect of sample size on goodness-of-fit indices in structural equation models. *Journal of Social Service Research, 20*(3-4), 49–70. https://doi.org/10.1300/J079v20n03_03

Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology, 9*(2), 79-94. https://doi.org/10.20982/tqmp.09.2. p079

Zuchdi, D. (2011). *Pendidikan karakter dalam prespektif teori dan praktek* [Character education in theory and practice perspective]. UNY Press.