Vol. 12, No. 1, pp. 294-306, 2022 DOI: 10.23960/jpp.v12.i1.202223
Jurnal Pendidikan Progresif
e-ISSN: 2550-1313 | p-ISSN: 2087-9849
http://jurnal.fkip.unila.ac.id/index.php/jpp/

Evaluation of Accreditation and National Examination using Multilevel Generalized Structured Component Analysis

Budi Susetyo*, Anwar Fitrianto
Department of Statistics, IPB University, Indonesia
*Corresponding email: budisu@apps.ipb.ac.id

Received: 02 March 2022 *Accepted: 02 April 2022* *Published: 11 April 2022*

Abstract: Evaluation of Accreditation and National Examination using Multilevel Generalized Structured Component Analysis. Hierarchical elements or higher levels often influence school accreditation and the national exam because education units are nested in the characteristics of the province. **Objectives**: This study aims to evaluate the relationship between accreditation and the national exam at the level of Junior high school/Madrasa in Java which are nested in province. **Methods:** The analysis employs multilevel GSCA analysis (MGSCA). **Findings:** UNBK has good convergent validity and it can explain each of the subjects tested in each province up to more than 90%. Concerning the estimates of path coefficients, the study found eight patterns of relationship between SNP and UNBK that have a significant effect in the six provinces. **Conclusion:** The relationship between content and competency standard for UNBK shows that there are significant differences in all provinces in Java island. This shows that provincial characteristics affect school quality. The model can explain the total variability of all variables is 72.44%.

Keywords: multilevel generalized structured component analysis, national education standards, national examination.

Abstrak: Evaluasi Akreditasi dan Ujian Nasional menggunakan Analisis Komponen Terstruktur Umum Bertingkat. *Unsur berhierarki atau tingkat yang lebih tinggi sering mempengaruhi akreditasi dan ujian nasional karena satuan pendidikan bersarang di karakteristik provinsi. Unsur hierarki atau jenjang yang lebih tinggi seringkali mempengaruhi akreditasi sekolah dan ujian nasional karena satuan pendidikan bersarang pada karakteristik provinsi.* ***Tujuan:*** *Penelitian ini bertujuan untuk mengevaluasi hubungan antara akreditasi dengan ujian nasional pada tingkat SMP/Madrasah di Jawa yang bersarang di provinsi.* ***Metode:*** *Analisis menggunakan analisis GSCA bertingkat (MGSCA).* ***Temuan:*** *UNBK memiliki validitas konvergen yang baik dan dapat menjelaskan setiap mata pelajaran yang diujikan di setiap provinsi hingga lebih dari 90%. Mengenai estimasi koefisien jalur, studi menemukan delapan pola hubungan antara SNP dan UNBK yang berpengaruh signifikan di enam provinsi.* ***Kesimpulan:*** *Hubungan antara isi dan standar kompetensi UNBK menunjukkan adanya perbedaan yang signifikan di semua provinsi di pulau Jawa. Hal ini menunjukkan bahwa karakteristik provinsi mempengaruhi kualitas sekolah. Model tersebut dapat menjelaskan total variabilitas semua variabel sebesar 72,44%.*

Kata kunci: *analisis komponen terstruktur umum bertingkat, standar nasional pendidikan, ujian nasional.*

To cite this article:

Susetyo, B. & Fitrianto, A. (2022). Evaluation of Accreditation and National Examination using Multilevel Generalized Structured Component Analysis. *Jurnal Pendidikan Progresif*, *12*(1), 294-306. doi: 10.23960/jpp.v12.i1.202223.

INTRODUCTION

The system of education quality is expected to be organized and managed equally in all regions of Indonesia. The quality of primary and secondary education is a level of compatibility between the implementation of primary and secondary education with national education standards (SNP) in schools. The SNP is a minimum criterion of the education system in all jurisdictions of the Unitary State of the Republic of Indonesia. According to the 2006 Ministry of Education Regulation, the National School/ Madrasah Accreditation Board (BAN-S/M) accredits in assessing the feasibility of educational unit programs concerning the SNP. The SNP consists of eight standards developed by the National Education Standards Agency (BSNP) to control the quality of education. In addition, BSNP is also tasked with holding a national examination (UN). The National Examination is an activity to measure the achievement of graduates' competencies in certain subjects nationally by referring to graduates' competency standards. In recent years, two national exams have been carried out: the national examination using paper and pencil (UNKP) and the computer-based national examination (UNBK). It is implemented to improve the efficiency and credibility of the UN.

Several studies have been conducted related to accreditation and the UN, including applying structural equation modeling (SEM) to see the relationship between standards on SNP. Vita et al. (2015)provided an illustration using SEM-generalized structured component analysis (GSCA) that graduate competency standards (SKL) are directly and indirectly affected by other SNP variables at the senior high school level. Standards that directly influence SKL include the cost standard (SB), management standard (SPL), and education assessment standard (SPN). In contrast, the standard of educators and education personnel (SPT), content standard (SI), process standard (SPR), and facilities and infrastructure standards (SSP) have an indirect effect on the SKL through the SPN. The study was also conducted by Hijrah et al. (2018) used partial least square-path modeling (PLS-PM) at the vocational school level, which illustrated that SKL is directly influenced by SPN and SPR variables. The R^2 value of the structural model except SI and SPL is more than 0.75, which shows that the model used is very adequate. The highest R^2 of 86.3% was constructed on variable SPR, which can be explained by the SSP, SPT, and SI variables. In addition, to observe the relationship between standards, other research was also conducted by Setiawan et al. (2018) used SEM-GSCA to see the relationship between standards in SNP and UNBK at the junior high school level, which concluded that the standards affecting UNBK were SKL, SPR, and SPN. Meanwhile, Susetyo & Wahyuni (2021) carried out a study about developing the GSCA method in evaluating the relationship between accreditation and UN using fuzzy cluster-wise generalized structured component analysis (FCGSCA). In the research, it created two groups in which the first group was characterized by schools that had SNP scores and UNBK scores lower than schools in the second group. However, the first group is better at describing the diversity of data compared to the second group.

Related to multilevel modeling, SEM has been used for a long time. SEM is an analytical method used to test theoretical models quantitatively to see the relationship between latent variables or latent variables and indicator indicators (Crockett, 2012; Henseler, 2017; Purwanto & Sudargini, 2021). In their study, the SNP was a latent variable because it cannot be measured directly, so the indicator items are used to measure it. GSCA is a variant-based SEM developed to overcome problems in covariance-based SEM that must meet parametric assumptions, such as multiple normal distributions of data and independent observations (Hair, 2021; Moore et al., 2021; Suk & Hwang, 2016).

In practice, data often have a hierarchical structure; however, this is still often ignored in the general analysis process that has been done. Ignoring higher levels of information can cause heteroscedasticity in errors (Astivia & Zumbo, 2019; Audigier et al., 2018; Leckie et al., 2014; Wu et al., 2018; Zhu et al., 2016). Therefore, multilevel analysis cannot be ignored because each unit comes from different levels, which can cause problems in context and statistics. Multilevel analysis that has been employed is still limited in modeling analysis, for example, multilevel regression. In the context of data with structural and measurement models, the multilevel element in SEM has not been widely applied. Hwang et al. (2007) developed the SEM-GSCA method in the case of customer satisfaction nested in different companies using a multilevel GSCA analysis (MGSCA) with two levels. It was concluded that there were considerable differences in policy and substantial companies viewed from the value of each loading factor and path coefficient (Hwang & Takane, 2014). This study aims to examine the relationship between the SNP and UNBK at the Junior High School/ Madrasa level in Java island using multilevel analysis. This analysis uses MGSCA in which the accreditation and UNBK are nested in each province. It is expected that standards significantly affect the UNBK are identified.

■ METHODS

Participants

Some quantitative approaches were employed in the study. The population of this study is all junior high schools and madrasa in Indonesia that have been accredited by BAN-S/M. The sample from this study was selected using a purposive sampling technique, namely schools and madrasas in six provinces in Java island, namely West Java, Central Java, East Java, Banten, Yogyakarta and DKI Jakarta. The sample selection was based on the consideration that the characteristics of schools and madrasas on the island of Java had variations that were representative of Indonesia. Accreditation results data for 2017 and 2018 are the results of instrument assessments based on tools and accreditations set by the Minister of Education and Culture in 2017, where for the junior high school level there are 124 statement/indicator items.

Research Design and Procedures

The data used in this study are secondary data of accreditation SNP scores and UNBK scores from 3953 schools at the 2017 and 2018 junior high schools/madrasa levels, which are nested in provinces in Java. Accreditation data consists of eight latent variables with 124 indicators obtained from BAN S/M, while UNBK data consists of four indicators obtained from the Center for Educational Research of the Republic of Indonesia.

The procedures for handling the data were carried out in a few steps. In the first step, a data preprocessing procedure was conducted. In this step, the data was divided into two, namely the accreditation and UNBK data of all provinces in Java island. As a result, there were seven datasets to be analyzed, each dataset from six provinces and a dataset from overall Java island. The next step was data exploration which is to describe the accreditation and UNBK data before conducting the multilevel GSCA.

Instruments

Indicators of accreditation are grouped based on eight national standards, namely content standards (SI), process standards (SPR), graduate competency standards (SKL), educators and education personnel standards (SPT), facilities and infrastructure standards (SSP), management standards (SPL).), financing standards (SB) and valuation standards (SP). The number of indicators per standard is presented in Table 1. Meanwhile, the national exam scores are the average school scores for four subjects, namely Indonesian (BIN), English (ING), Mathematics (MAT) and natural sciences (IPA).

Table 1. Latent variable and its indicators

Latent Variables	Indicator Variables	Measurement Model
SI	item 1 - item 9	Formative
SPR	item 10 - item 30	Formative
SKL	item 31 - item 37	Formative
SPT	item 38 - item 56	Formative
SSP	item 57 - item 80	Formative
SPL	item 81 - item 95	Formative
SB	item 96 - item 111	Formative
SPN	item 112 - item 124	Formative
UNBK	Indonesian (BIN), English (ING), Mathematics (MAT), and Science (IPA)	Reflective

Data Analysis

The data accreditation and the UNBK data were analyzed using the multilevel GSCA method so that the weight estimator, loading factor, and path coefficient will be obtained. Then, comparisons of loading factors and path coefficients between the six provinces were conducted. After the comparisons, it was needed to calculate the mean and standard deviation for loading factors and path coefficients in the six provinces in Java island for standardizing the results. To make sure that the model is appropriate, some methods to evaluate the model were done. The methods used three measures, namely evaluation of measurement models, evaluation of structural models, and evaluation of overall goodness of fit.

The evaluation of measurement models consists of two components, which are evaluation of formative measurement models and evaluation of reflective measurement models. The evaluation of the formative measurement model was carried out by looking at the significance of the indicator weights of the SNP. Evaluation of the weight estimator results can be identified based on the results of the GSCA analysis on data for all provinces. The indicator weights are valid if the CR values are more than 1.96 at the alpha 5%. Meanwhile, the evaluation of reflective measurement models was carried out by looking at the significance of the mean and standard deviation for each loading factor on the UNBK variable against the indicator based on step 4. The UNBK variable has good convergent validity if the estimated value of the loading factors is more than 0.7 and significant. If the standard deviation of loading factors is significant, there are differences between the six provinces in the UNBK variable in describing the subject indicators.

The second method to evaluate the model used in the study was the evaluation of structural models. It was carried out by looking at the significance of the mean and standard deviation of path coefficient between SNP and UNBK based on stage 4. The SNP and UNBK variables significantly affect the CR values of more than 1.96. If the standard deviation of the path coefficients are significant, there are differences between the six provinces in the pattern of relations between SNP and UNBK.

Lastly, the measure to evaluate the model was a measure of the overall goodness of fit. Evaluating the overall goodness of fit of the model was carried out by calculating the FIT and AFIT values (Hwang et al., 2020; Ryoo et al., 2020; Ryoo & Hwang, 2017) based on the following equation:

$$FIT=1-\frac{\sum_{g=1}^{G} SS(\mathbf{Z}_g\mathbf{V}-\mathbf{Z}_g\mathbf{W}\mathbf{A}_g)}{\sum_{g=1}^{G} SS(\mathbf{Z}_g\mathbf{V})}$$

$$AFIT=1-(1-FIT)\frac{d_0}{d_1}, \text{ with } d_0=NJ;\ d_1=NJ-k$$

where d_0 is the degree of freedom of model 0 (W = 0 and Ag = 0), d_{01} is the degree of freedom of the model under test, J is the number of indicators and k is the number of parameters.

■ RESULT AND DISCUSSIONS

Data Exploration

The data used in this study are data on 3953 schools that implement the UNBK system for junior high schools in Java consisting of 978 (24.74%) public junior high schools (SMP), 1796 (45.43%) private junior high schools (SMPS), 117 (2.96%) public madrasa (MTSN) and 1062 (26.87%) private madrasa (MTSS). Figure 1 exhibits the mean of UNBK score based on the accreditation ranking from all provinces in Java island. There were 45.69% were schools that were accredited "A", 36.91% were schools that were accredited "B", 16.11% were schools that were accredited "C", and 1.29% were schools not accredited "TT". Based on the overall accreditation ranking for schools in Java, the highest UNBK score is Indonesian (BIN) and the lowest is Mathematics (MAT). This indicates that the ability of students in learning mathematics in general is still quite low when compared to other subjects. This is in line with research conducted by Han et al. (2015). In the English and Natural Sciences, the UNBK score obtained tends to be almost the same. The graph also shows that schools accredited "A" have a higher UNBK score than the UNBK score of schools accredited "B", and so on, the better the school accreditation, the higher the UNBK score obtained. However, the UNBK score in all schools tended to be low with a range of 33.97 to 70.75, especially in Mathematics, Science, and English. This is because the UNBK is held more frequently in every school than the previous ones that still implemented UNKP. With the implementation of UNBK impacts the correction of values so that forms of cheating are more difficult to occur.

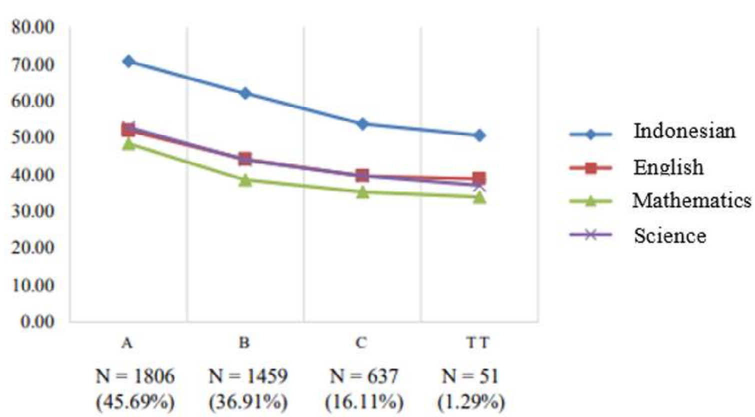

Figure 1: The mean of UNBK in Java island based on accreditation ranking

Figure 2 presents the UNBK scores for all provinces in Java island based on accreditation ranking. The highest UNBK score was in the DIY (57.65) and the lowest was in the East Java (47.12). Research conducted by (Handayani, 2018) also stated that DIY achieved an increase in the average score of the National Examination from UNKP to UNBK.The highest UNBK score in each province is obtained from Bahasa (Indonesian Language) and the lowest is Mathematics. This was possible because the form of Mathematics exam questions began to be inserted into higher-order thinking skills (HOTS) exam questions, which required students to memorize and understand mathematical formulas and have higher reasoning power.

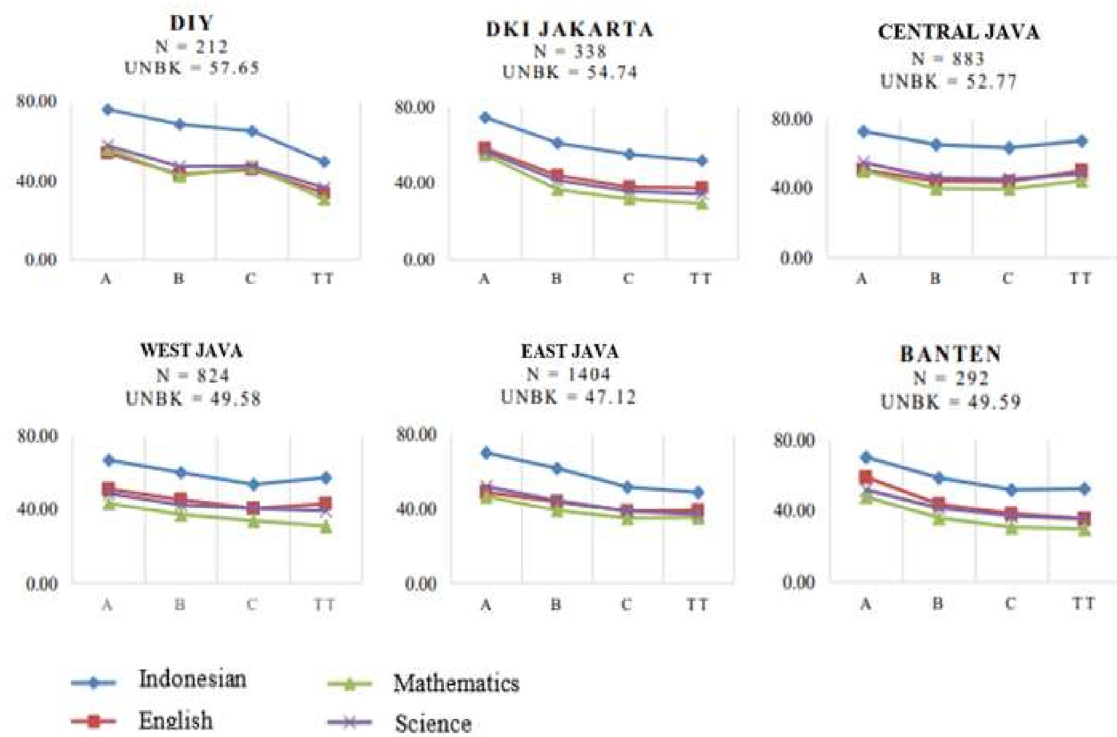

Figure 2: The mean of UNBK all provinces in Hava island based on accreditation ranking

Results of the Inter-province using GSCA

The GSCA analysis on school data in all provinces is intended to see the estimated weight indicators. The weighting model at two levels is the same as at one level GSCA. The measurement and structural two-level model were obtained through the mean, and standard deviation of loading factors and path coefficients from the GSCA analysis of each province. Here are the GSCA analysis results from inter-provinces:

The estimate of indicator weights

Estimating indicator weights are obtained from the GSCA analysis on school data in all Java provinces and serve as indicator weights for multilevel GSCA analysis. At 5% level of significance, based on because the CR value which is less than 1.96, out of the 124 indicators on the SNP variables, four indicators were inappropriate in constructing each of the latent variables of SPT and SSP, namely, item 51, 54, 57, and 76 (Table 2).

Table 2. Inappropriate estimates of SNP indicator weight

Latent variable	Indicator	The estimator that invalid		
		Estimate	SE	CR
SPT	item 51	0.014	0.011	1.27
	item 54	0.006	0.014	0.43
SSP	item 57	0.015	0.014	1.07
	item 76	-0.003	0.013	0.23

The estimated weight of the variables were close to zero, so that these items are inappropriate in describing the latent variable. The indicator items can be used to evaluate the quality of education.

The estimate of loading factors

Table 3 presents the loading factor estimation results for each province. All the estimated loading factors of all provinces in Java island are greater than 0.7. This means that UNBK has good convergent validity and can explain each of the subjects tested in each province up to more than 90%.

The estimate of path coefficients

The path coefficient estimation results for each province is displayed in Table 4. There are 24 structural relationship patterns between SNP and UNBK.

Table 3. The estimates of the loading factors

Latent variable	Indikator	DIY	Banten	DKI Jakarta	Provinces West Jawa	Central Java	East Jawa
UNBK	BIN	0.958	0.952	0.963	0.926	0.952	0.9 03
	ING	0.926	0.973	0.954	0.951	0.937	0.912
	MAT	0.959	0.959	0.97	0.942	0.969	0.915
	IPA	0.966	0.98	0.978	0.957	0.968	0.953

The estimated value of the path coefficient is used to see the effect between latent variables, namely between the eight standards on SNP and UNBK. The sign indicates the direction of the relationship between the two variables. A positive sign means the increase in the score of one variable corresponds to the increase in the score of the other variables according to the estimated value, and vice versa. Eight patterns of relationship between SNP and UNBK significantly affect the six provinces. The include the relationship between SPL to SPT, SPL to SSP, SPL to SB, SPT to SI, SPT to SPR, SI to SPR, SI against SKL, and SPR against SKL. In the six provinces, the pattern of the relationship between SST and SPT is the pattern that has the most considerable influence on the model. For example, in DIY province, the estimated path coefficient value was 0.900. This value is the largest value compared to the estimated value in the pattern of relationships between SNP and other UNBKs in the DIY province. The value means that the larger the score of the management standard, the higher the standard score for educators and teachers by 90% in province.

In the pattern of the relationship between SPN and UNBK as well as the relationship between SPR and UNBK, both showed a pattern of relationship that did not have a significant effect in the six provinces. This means that assessment and process standards do not significantly affect UNBK as school academic achievement. Table 4 also shows significant differences in the pattern of relationships between SNP and UNBK in the six provinces. For example, the relationship between SKL and UNBK has a significant effect in DIY, West Java, and Central Java provinces. Meanwhile, the relationship between SI and UNBK has a significant effect only in DIY. Apart from the difference in the number of significant relationship patterns, the estimated value of the path coefficient of each province in each relationship pattern is quite high and low, thus indicating that the influence of the relationship between SNP and UNBK in the six provinces is quite diverse/different. For example, in the relationship between SKL and UNBK, the influence is quite high (0.765) in DIY province, but the influence in Central Java is only 0.26. In other words, the characteristics of each province affect the quality of its schools. Therefore, the

Table 4. Path coefficient estimator each province

Relation	Provinces					
	DIY	Banten	DKI Jakarta	West Java	Central Java	East Java
SPL→SPT	**0.900**	**0.869**	**0.817**	**0.659**	**0.703**	**0.874**
SPL→SSP	**0.325**	**0.335**	**0.283**	**0.163**	**0.248**	**0.237**
SPL→SB	**0.899**	**0.841**	**0.848**	**0.596**	**0.552**	**0.91**
SPT→SSP	0.29	**0.469**	**0.515**	**0.52**	**0.547**	**0.567**
SPT→SI	**0.61**	**0.205**	**0.287**	**0.148**	**0.215**	**0.151**
SPT→SPN	**0.501**	0.129	0.066	0.116	0.076	0.017
SPT→SPR	**0.346**	**0.344**	**0.3**	**0.237**	**0.519**	**0.163**
SSP→SI	0.056	0.042	0.129	0.093	**0.212**	**0.139**
SSP→SPN	-0.054	0.165	**0.183**	0.08	**0.205**	**0.078**
SSP→SPR	**0.281**	0.004	0.139	**0.136**	0.042	0.161
SB→SSP	**0.331**	0.134	0.094	**0.125**	0.052	0.163
SB→SI	0.244	**0.165**	**0.188**	0.1	**0.15**	**0.227**
SB→SPN	**0.335**	0.125	**0.296**	**0.136**	**0.177**	**0.29**
SB→SPR	0.019	0.113	**0.238**	**0.105**	0.07	0.228
SI→SPR	**0.339**	**0.533**	**0.298**	**0.455**	**0.272**	**0.448**
SI→SKL	**0.58**	**0.272**	**0.176**	**0.126**	**0.192**	**0.345**
SI→UNBK	**-0.507**	-0.382	-0.136	-0.039	0.047	0.014
SPN→SI	0.019	**0.553**	**0.287**	**0.428**	**0.269**	**0.451**
SPN→SKL	0.104	0.175	**0.182**	**0.271**	**0.2**	**0.11**
SPN→UNBK	-0.189	-0.057	0.241	0.067	0.071	0.139
SPR→SPN	0.164	**0.535**	**0.386**	**0.454**	**0.383**	**0.588**
SPR→SKL	**0.272**	**0.48**	**0.556**	**0.429**	**0.449**	**0.508**
SPR→UNBK	0.409	0.427	0.186	0.046	0.112	0.209
SKL→UNBK	**0.765**	0.438	0.228	**0.298**	**0.26**	0.12

Note: numbers in **bold** are significant path coefficients at alpha 5%

GSCA analysis is carried out in a multilevel perspective because schools are nested in the provinces.

Evaluation of Multilevel GSCA Model

Evaluation of the reflective measurement model and the structural model in multilevel is obtained through the mean and standard deviation of the loading factor estimator and path coefficient in six provinces. Estimates of loading factors and path coefficients in the six provinces can be seen in Table 3 and Table 4 previously. Meanwhile, the estimator of the indicator weight is the same as the weighting model on the one-level GSCA. There is no need to look for the mean and standard deviation because the latent variable scores (the SNPs) depend on the groups/provinces.

The mean and standard deviation of the loading factor estimator are used to evaluate the reflective measurement model. It is essential to know whether the UNBK variable has good convergent validity on the subjects tested and whether there are differences between provinces in UNBK in describing each subject. Evaluation of the mean and standard deviation of the estimating loading factor is presented in Table 5. All estimated loading factor values are more than 0.7, indicating that the UNBK has good convergent validity. It can explain each of the subjects tested up to more than 94%. The mean of loading factor is also significant at the 5% real level because it has a CR of more than 1.96. The estimation result of the loading factor standard deviation shows all are significant, which means that there are differences between the six provinces in UNBK in describing each subject. This means that the provincial characteristics affect the schools' academic achievement.

Table 5. Mean and standard deviation of *loading factor* multilevel

Latent variable	Indicator	Mean of *loading factor*			Standard deviation of *loading factor*		
		Estimate	SE	CR	*Estimate*	SE	CR
UNBK	BIN	0.942	0.005	209.41	0.023	0.0010	22.05
	ING	0.942	0.005	182.35	0.022	0.0008	28.90
	MAT	0.952	0.004	219.77	0.021	0.0012	17.24
	IPA	0.967	0.003	305.37	0.011	0.0004	26.56

Furthermore, the mean of path coefficients in evaluating the multilevel structural model can be seen in Figure 3. The mean of path coefficients are said to have a significant effect at alpha of 5% if the CR values are more than 1.96. In Figure 3, the dashed line shows the relationship between variables that is not significant. It can be seen that the path coefficient value from SPL to SPT has the highest value of 0.80. The greater the management standard score, the higher the standard score for educators and teachers by 80% in the six provinces.

In addition, the mean of path coefficient from SST to SB also shows the next highest value, which is 0.77. This also explains that if the management standard score increases, the cost standard score also increases by 77% in the six provinces. Meanwhile, the standard that directly influences UNBK is SKL with an estimated value of 0.35. That is, if the competency standard score of graduates increases, it will significantly affect the academic achievement of the school, namely UNBK by 35% in the six provinces, so in other

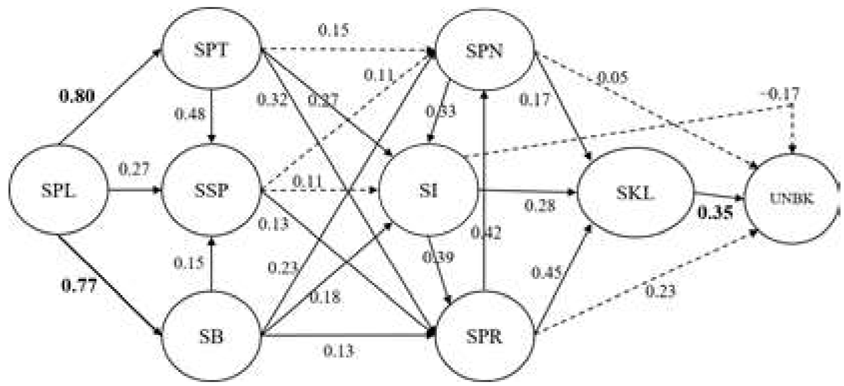

Figure 3: The mean of path coefficients between SNPs and UNBK

words, this standard needs to be considered and improved by each school to improve the quality/school achievement. However, through graduate competency standards, UNBK is indirectly influenced by content, process, and assessment standards.

The multilevel standard deviation of the path coefficients on the relationship between SNP and UNBK is also shown in Figure 4. Based on the standard deviation of the path coefficient, of the 24 patterns of relationship between SNPs and UNBK, 16 patterns have a

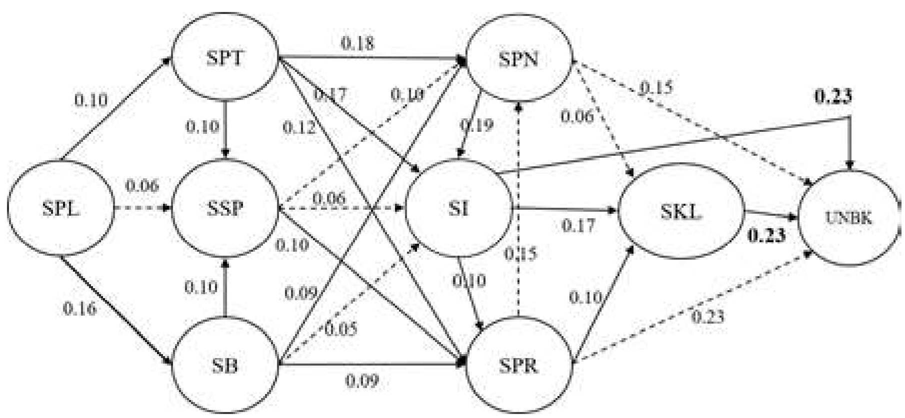

Figure 4: The standard deviation of path coefficients between SNPs and UNBK

significant effect, which means that there is a large difference in the relationship between SNPs and UNBK among the six provinces in the implementation of education quality. The biggest significant difference is in SI and SKL against UNBK of 0.23. There is a significant difference between graduates' content and competency standards that affect academic achievement, namely UNBK between the six provinces. In other words, provincial characteristics can affect the quality of the schools in it.

Based on the results of the estimation of the mean and standard deviation, the multilevel path coefficient is the sum of the mean and standard deviation of the path coefficients so that the multilevel structural model equation can be seen as follows:

$$\begin{bmatrix} \text{SI} \\ \text{SPR} \\ \text{SKL} \\ \text{SPT} \\ \text{SSP} \\ \text{SPL} \\ \text{SB} \\ \text{SPN} \\ \text{UNBK} \end{bmatrix} = \begin{bmatrix} 0 & 0.494 & 0.447 & 0 & 0 & 0 & 0 & 0 & 0.060 \\ 0 & 0 & 0.547 & 0 & 0 & 0 & 0 & 0.567 & 0.387 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.579 \\ 0.444 & 0.439 & 0 & 0 & 0.586 & 0 & 0 & 0.327 & 0 \\ 0.174 & 0.225 & 0 & 0 & 0 & 0 & 0 & 0.206 & 0 \\ 0 & 0 & 0 & 0.903 & 0.329 & 0 & 0.932 & 0 & 0 \\ 0.232 & 0.216 & 0 & 0 & 0.246 & 0 & 0 & 0.318 & 0 \\ 0.522 & 0 & 0.236 & 0 & 0 & 0 & 0 & 0 & 0.196 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \text{SI} \\ \text{SPR} \\ \text{SKL} \\ \text{SPT} \\ \text{SSP} \\ \text{SPL} \\ \text{SB} \\ \text{SPN} \\ \text{UNBK} \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \\ \zeta_5 \\ \text{SPL} \\ \zeta_6 \\ \zeta_7 \\ \zeta_8 \end{bmatrix}$$

The overall goodness of fit of the multilevel GSCA model on the relationship between the SNP and UNBK was carried out using FIT and AFIT values. The FIT and AFIT values are obtained by calculating the matrix of the indicator variables, the weight of the indicators, and the mean and standard deviation of the loading factor and path coefficient. The FIT value obtained from this analysis was 0.7245 and the resulting AFIT value was not much different, namely equal to 0.7244. This means that the total diversity of all variables that the multilevel model can explain is 72.44%.

CONCLUSIONS

Based on the results of the discussion above, it can be concluded that in a multilevel, standard that has a direct effect on UNBK is the graduate competency standard, while the content standard, process standard, and assessment standard have an indirect effect on UNBK through graduate competency standards. In addition, the relationship between content standards and graduate competency standards against UNBK shows that there are significant differences in the six provinces on the island of Java. This suggests that provincial characteristics affect school quality.

The study implies that it can be used as input for the Ministry of Education and Culture to evaluate the curriculum (content standards) and learning processes carried out by teachers in schools (process standards) to improve the quality of graduates.

The limitation of this research is that the indicators of the quality of graduates that can be used at this time are only data on the results of the national exam. With the abolition of the national exam since 2019, and the implementation of a national assessment by the Ministry of Education and Technology, it is possible that in the future there will be many indicators that can characterize the quality of graduates more comprehensively. Therefore, in the next research, it is necessary to re-analyze the results of BAN-S/M accreditation with the results of a national assessment.

REFERENCES

Astivia, O. L. O., & Zumbo, B. D. (2019). Heteroskedasticity in Multiple Regression Analysis: What it is, How to Detect it and How to Solve it with Applications in R and SPSS. *Practical Assessment, Research, and Evaluation*, *24*(1), 1.

Audigier, V., White, I. R., Jolani, S., Debray, T. P. A., Quartagno, M., Carpenter, J., van Buuren, S., & Resche-Rigon, M. (2018). Multiple imputation for multilevel data with continuous and binary variables. *Statistical Science*, *33*(2), 160–183.

Crockett, S. A. (2012). A five-step guide to conducting SEM analysis in counseling research. *Counseling Outcome Research and Evaluation*, *3*(1), 30–47.

Hair, J. F. (2021). Reflections on SEM: An Introspective, Idiosyncratic Journey to Composite-based Structural Equation Modeling. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, *52*(SI), 101–113.

Han, S., Capraro, R., & Capraro, M. M. (2015). How science, technology, engineering, and mathematics (STEM) project-based learning (PBL) affects high, middle, and low achievers differently: The impact of student factors on achievement. *International Journal of Science and Mathematics Education*, *13*(5), 1089–1113.

Handayani, M. (2018). Utilization of national exam results in advancing educational ecosystem to improve the quality of education. *ICEAP 2018*, *1*(1), 21–30.

Henseler, J. (2017). Bridging design and behavioral research with variance-based structural equation modeling. *Journal of Advertising*, *46*(1), 178–192.

Hijrah, M., Susetyo, B., & Sartono, B. (2018). Structural equation modeling of national standard education of vocational high school using partial least square path modeling. *International Journal of Scientific Research in Science Engineering and Technology*, *4*(4), 1418–1422. https://doi.org/10.32628/IJSRSET1844465

Hwang, H., Sarstedt, M., Cheah, J. H., & Ringle, C. M. (2020). A concept analysis of methodological research on composite-based structural equation modeling: bridging PLSPM and GSCA. *Behaviormetrika*, *47*(1), 219–241.

Hwang, H., & Takane, Y. (2014). *Generalized structured component analysis: A component-based approach to structural equation modeling*. CRC Press.

Hwang, H., Takane, Y., & Malhotra, N. (2007). Multilevel generalized structured component analysis. *Behaviormetrika*, *34*(2), 95–109.

Leckie, G., French, R., Charlton, C., & Browne, W. (2014). Modeling heterogeneous variance–covariance components in two-level models. *Journal of Educational and Behavioral Statistics*, *39*(5), 307–332.

Moore, Z., Harrison, D. E., & Hair, J. (2021). Data Quality Assurance Begins Before Data Collection and Never Ends: What Marketing Researchers Absolutely Need to Remember. *International Journal of Market Research*, *63*(6), 693–714.

Purwanto, A., & Sudargini, Y. (2021). Partial Least Squares Structural Squation Modeling (PLS-SEM) Analysis for Social and Management Research: A Literature Review. *Journal of Industrial Engineering & Management Research*, *2*(4), 114–123.

Ryoo, J. H., & Hwang, H. (2017). Model evaluation in generalized structured component analysis using confirmatory tetrad analysis. *Frontiers in Psychology*, *8*, 916.

Ryoo, J. H., Park, S., Kim, S., & Ryoo, H. S. (2020). Efficiency of cluster validity indexes in fuzzy clusterwise generalized structured component analysis. *Symmetry*, *12*(9), 1514.

Setiawan, A. I., Susetyo, B., & Fitrianto, A. (2018). Application of generalized structural component analysis to identify relation between accreditation and national assessment. *International Journal of Scientific Research in Science Engineering and Technology*, *4*(10), 93–97. https://doi.org/10.32628/18410IJSRSET

Suk, H. W., & Hwang, H. (2016). Functional generalized structured component analysis. *Psychometrika*, *81*(4), 940–968.

Susetyo, B., & Wahyuni, R. (2021). Application of the fuzzy clusterwise generalized structured component method to evaluate implementation of national education standard in Indonesia. *Management Science Letters*, *11*(4), 1379–1384. https://doi.org/10.5267/j.msl.2020.11.002

Vita, F., Susetyo, B., & Indriyanto, B. (2015). Generalized structured component analysis for national education standards of secondary school in Indonesia. *Global Journal of Pure and Applied Mathematics*, *11*(4), 2441–2449.

Wu, W., Carroll, I. A., & Chen, P.-Y. (2018). A single-level random-effects cross-lagged panel model for longitudinal mediation analysis. *Behavior Research Methods*, *50*(5), 2111–2124.

Zhu, C., Lopez, R. A., & Liu, X. (2016). Information cost and consumer choices of healthy foods. *American Journal of Agricultural Economics*, *98*(1), 41–53.