Jurnal
Penelitian dan Evaluasi Pendidikan
HIMPUNAN EVALUASI PENDIDIKAN INDONESIA (HEPI)
in cooperation with
GRADUATE SCHOOL OF UNIVERSITAS NEGERI YOGYAKARTA
Kampus Karangmalang, Yogyakarta 55281. Phone. 0274 550836 Fax : 0274 520326

Website: http://journal.uny.ac.id/index.php/jpep
e-mail: jurnalhepi@uny.ac.id

9 772338 606001
9 772685 711007

Jurnal Penelitian dan Evaluasi Pendidikan
Volume 23, No 2, December 2019

Volume 23, No 2, December 2019
ISSN 2685-7111 (print)
ISSN 2338-6061 (online)
Akreditasi No. 040/P/2014

Jurnal
Penelitian dan Evaluasi Pendidikan

Jurnal
Penelitian dan Evaluasi Pendidikan
ISSN 2685-7111 (print) | ISSN 2338-6061 (online)

Publisher
HIMPUNAN EVALUASI PENDIDIKAN INDONESIA
In Cooperation With
PROGRAM PASCASARJANA UNIVERSITAS NEGERI YOGYAKARTA
(MOU Nomor 195 B/J.35.17/LK/04)

Correspondence: Graduate School of Universitas Negeri Yogyakarta
Kampus Karangmalang, Yogyakarta, 55281, Telp. (0274) 550835, Fax. (0274) 520326
Homepage: http://journal.uny.ac.id/index.php/jpep
e-mail: jurnalhepi@uny.ac.id

FOREWORDS

We are very pleased that *Jurnal Penelitian dan Evaluasi Pendidikan* is releasing its issue **Volume 23, No 2, December 2019.** We are also very excited that the journal has been attracting papers from many institutions in Indonesia and many foreign countries. *Jurnal Penelitian dan Evaluasi Pendidikan* was first published in **1998** and since then regularly published online and in print twice a year: June and December.

Jurnal Penelitian dan Evaluasi Pendidikan with ISSN 2338-6061 (*online)* has been **re-accredited** by the Ministry of Research, Technology, and Higher Education of Republic of Indonesa under the Decree Number 30/E/KPT/2018 which is valid for 5 (five) years since enacted on 24 October 2018 (Vol. 20, No 2, 2016 until Vol. 24, No 2, 2021). Jurnal Penelitian dan Evaluasi Pendidikan successfully achieved accreditation in four periods in a row (in 2007, 2010, 2014, & 2018).

Jurnal Penelitian dan Evaluasi Pendidikan is a showcase of original, rigorously conducted educational evaluation, measurement and assessment from primary, secondary, and higher education institutions. Each issue of this journal is not limited to comprehensive syntheses of studies towards developing new understandings of educational evaluation, measurement and assessment only, but also explores scholarly analyses of issues and trends in the field.

Yogyakarta, December 2019

Editor in Chief

Table of Content

Jurnal Penelitian dan Evaluasi Pendidikan
Volume 23, No 2, December 2019 (106-116)
Online: http://journal.uny.ac.id/index.php/jpep

EVALUATION OF ISLAM EDUCATION TEACHERS TRAINING IMPLEMENTATION

Helen Sabera Adib
Universitas Islam Negeri Raden Fatah

Djemari Mardapi
Universitas Negeri Yogyakarta

Zamroni
Universitas Negeri Yogyakarta

Adam Jait
Universiti Islam Sultan Sharif Ali

Abstract

The purpose of this study is to evaluate the implementation of the Islam education teachers' training in the Religion Education and Training Center in Palembang. This study is an evaluation research conducted using the Islam Education Teachers Training model, which consists of Context, Input, Process, and Product aspects. Sources of information in this study include the training organizers, *Widyaiswara* (civil servants assigned as instructors), and the participants of the training. Data collection was conducted using questionnaires, observation, and documentation. The data analysis technique used in this study was quantitative descriptive. The results of the study conclude that the implementation of the Islam Education teachers' training program is very successful, as shown by the success indicators in terms of context, input, process, and product aspects.

Keywords: *evaluation model, teacher training, Islam education*

Permalink/DOI: http://dx.doi.org/10.21831/pep.v23i2.20986

Contact: *Helen Sabera Adib*
oxaliscrgirl@gmail.com
Department of Islam Education, Faculty of Tarbiyah and Educational Science, Universitas Islam Negeri Raden Fatah
Jl. Prof. K.H. Zainal Abidin Fikri Km. 3,5, Pahlawan, Kemuning, Palembang, South Sumatera 30151, Indonesia

Introduction

Assessment in the education curriculum with all its problems is an interesting topic to discuss. The most widely discussed topic in education is the teacher. Teachers are professional educators who must have a variety of abilities or competencies that are needed to carry out their educating tasks. Law No. 14 of 2005 of the Republic of Indonesia on teachers and *Widyaiswara* article 8 paragraph (1) stipulates four types of competencies that must be possessed by teachers, namely pedagogical competence, personal competence, social competence, and professional competence.

Teachers are important in the teaching and learning system so that they have to be able to combine these two things well. The first thing is that teachers must always update their knowledge and abilities in teaching. The second thing is that teachers must equip themselves with the competencies needed and behave positively to do their profession (Bhargava & Pathy, 2011). Teachers are required to be able to convey the lesson well so that students can absorb the lesson well, meaning that teachers are required to have good communication competence. If the teachers play their roles well, then the target of teaching and learning will be achieved because the learning process in the class determines the quality of the school, and teachers are the determining keys for the high or low quality of the school (Ilukena, 1998).

The results of the teacher competency test indicated that the teachers' competency level is still low. There is no significant difference between the teachers' professional competence and the senior level teacher training college students' professsional competence. The teacher pedagogical competence test results also showed less encouraging results. The average ability of teachers in planning lessons, especially in formulating learning indicators, organizing materials, determining learning resources, creating learning scenarios, and ways of assessing students still need to be improved. The weakness lies in the way teachers assess students (Research Center for Educational Policy and Innovation, 2007, pp. 57–58).

Suraji (2010), in the results of his study, states that the pedagogical, personal, and social competence of *Madrasah Ibtidaiyah* teachers in Pekalongan is still not optimal. The results of another study (Zainuri, 2012, p. 3) also explain that the competency of State *Madrasah Ibdtidaiyah* teachers in Palembang still needs to be improved because it is still low. Government efforts to improve the teachers' competence have been carried out by implementing teachers' training programs. The training was not only conducted by the Ministry of National Education but also carried out by the Ministry of Religious Affairs as well as by many other government institutions. The training was carried out in order to increase the teachers' competence.

Teacher training is a form of in-service education for teachers. Training is an activity that is carried out to improve the quality of teachers and acts as a program that helps teachers understand and implement educational policies that are directly related to the teachers' task. The training program aims to equip and improve teachers' abilities and competence. Teachers who have participated in the training program are expected to be able to apply the abilities that they obtain from the training, meaning that after the training, there will be an improvement in the quality of teachers and changes in the form of better learning behavior. Herzberg, Mausner, and Snyderman (2010) state that training is a learning process to improve the teaching ability of teachers to carry out their tasks in the future.

The purpose of the training is to improve the quality of teachers, which in this case are the Islam education teachers, which can be pursued through the Islam education teachers training program that is organized by the education and training organizers, one of which is the religious education and training center. Susatya (2013, pp. 107–108) states that high-quality education and training centers can produce superior, strong, high-tech human resources

who have adequate competence so that they are able to compete with other countries' human resources. Islam education teachers' training is carried out to equip the Islam education teachers to become qualified, strong, and superior teachers.

According to Notoatmodjo (1991, p. 53), the implementation of the training program can be said to be successful if a transformation process occurs within the training participant in the form of increased ability to carry out tasks and changes in behavior which are reflected in their attitudes, discipline and work ethic. The changes and improvements can be seen if an assessment or evaluation of the implementation of the training program is carried out.

Goldstein and Ford (2002, p. 110) state that the readiness to learn, the background of the field of work, and the learning motivation of the participants are decisive in achieving the benefits of the training. Adults are ready to learn when they know that they need to learn about something that is usually related to the problem at hand, work assignments, or the desire to achieve something more in their lives (Knowles, 1980, p. 44).

Program evaluations produce information concerning the design, implementation, and assessment of the results of the efforts that have been made to solve the problems (McDavid & Hawthorn, 2006, p. 3). Program evaluation results can be used as a benchmark to find solutions for the problems, and development can also become the basis in making decisions regarding the program. Evaluation of the training program becomes an important thing to do because it aims to find information about the success rate of the training. According to Arikunto and Jabar (2009, p. 5), program evaluation is an activity to find out how high the success rate of the program is, to find out whether the educational goals have been achieved, and the efforts to provide information to be conveyed to the decision-makers.

The religious education and training center, as the agency in charge of the training program, must have this evaluation tool to assess the success of the training. Besides, the evaluation process can be carried out by applying a suitable evaluation model. The evaluation model can be chosen according to the needs and focus of the evaluation, which is going to be conducted. The Islam education teachers' training program needs a comprehensive evaluation. An evaluation model may be suitable to evaluate a program, but it does not mean that it can meet the needs of other program evaluation implementation (Phillips, 1991).

The best evaluation model is an evaluation model that can evaluate the program according to the needs of the program. Based on the observations conducted at the religious education and training center in Palembang, an Islam education teachers' training program was carried out. The training program aims to improve the competence of Islam education teachers. However, there has been no suitable evaluation model and evaluation instruments used specifically to evaluate the implementation of the program. Therefore, Islam education teachers' training evaluation model becomes an alternative to solve the problem. This study aims to evaluate the implementation of the Islam education teachers training program in the religious education and training center in Palembang by using the Islam education teachers' training evaluation model.

Research Method

This study is evaluation research conducted by applying the Islam education teachers' training (*Pendidikan Guru Pendidikan Agama Islam* or PGPAI) evaluation model. The PGPAI evaluation model consists of four components, namely context, input, process, and product. The source of information in this study was 43 people covering the training organizers, *Widyaiswara*, and the training participants. The purpose of this study was to evaluate the implementation of Islam education teachers' training in the religious education and training center in Palembang.

On the other hand, the PGPAI evaluation model had been developed through the ten steps of model development (Borg & Gall, 1993), which were then simplified into three steps, as suggested by Cennamo and Kalk (2005). The three steps of model development research included (1) the model pre-development stage, (2) the conceptual model for the field trials (model testing) development stage, and (3) the model application stage. As an illustration, the PGPAI evaluation model is described as follows.

The pre-development stage of the PGPAI evaluation model is the initial research stage, and the information collected was related to the evaluation system of the Islam education teachers' training at the religious education and training center in Palembang. The information was about the type of evaluation models that have been used so far. In addition, a theoretical review of the existing evaluation models was also carried out, as well as a review of the results of relevant studies. The result of the pre-development stage was the initial model (prototype). This model refers to the CIPP evaluation model (Context, Input, Process, Product).

The conceptual model development stage was the stage of planning the model and assembling the PGPAI evaluation instruments. The design of the evaluation model involved six experts in the field of educational research and evaluation and the field of Islam education. In addition, the involvement of practitioners and academicians had also been carried out through the Focus Group Discussion (FGD) technique. Furthermore, content validation was conducted through expert judgment using the Aiken formula. The experts assessed the design of the model and the instruments as well as the equipment using the Delphi technique. The experts' assessments were used as the guideline to improve the materials before the models, instruments, and equipment were used. Based on expert judgment, the PGPAI evaluation model instrument was declared valid. It was concluded valid after conducting a review of the Aiken table, which showed that the V Aiken instrument achieved a score above 0.70. In addition, there were also expert judgments on the evaluation procedure, evaluation guideline, and the instruments' clarity. The results of the expert judgment indicated that the evaluation procedure, evaluation guideline, and the instruments' clarity were categorized as very good. The summary of the results of the expert judgment was presented in Table 1 and Table 2. The collected data were analyzed using quantitative descriptive analysis and then compared with the criteria that have been set in Table 3.

Table 1. Content Validity of the PGPAI Evaluation Model Instruments

No	Component	V Aiken	Interpretation
1	Context	0.95	Valid
2	Input	0.95	Valid
3	Process	0.94	Valid
4	Product	0.93	Valid

Table 2. Expert Judgment

Assessment Aspects	Average	Category
Evaluation Procedure	3.66	Very Good
Evaluation Guideline	3.54	Very Good
Intrument Clarity	3.46	Very Good

Table 3. Scoring Category

No	Score	Category
1	$X \geq \bar{X} + 1.SB_x$	Very Successful
2	$X + 1.SB_x > X \geq \bar{X}$	Successful
3	$\bar{X} > X \geq \bar{X} - 1.SB_x$	Less Successful
4	$X < \bar{X} - 1.SB_x$	Unsuccessful

Source: Mardapi (2012, p. 162)

Notes:
X = Score
$\bar{X}$ = Average Score
SB_x = Standard Deviation

Findings and Discussion

The study obtained some information about the implementation of the Islam education teachers' training in the religious education and training center in Palembang,

which was evaluated using the PGPAI model consisting of the aspects of context, input, process, and product. In the aspect of the process, there was a part showing the participants' satisfaction towards the implementation of the training program. The results of the study are described based on their respective aspects. Each of them is elaborated as follows.

Evaluation on the Context of the Islam Education Teachers' Training Implementation

Evaluation of the context consists of several assessed sub-components, including the program planning, program implementation, program needs, the vision and mission distribution, administration of participants and *Widyaiswara*, training facilities, training classroom atmosphere, and training environment atmosphere. The evaluation of the context was assessed by the *Widyaiswara*, the organizers, and the researchers by conducting observation and documentation. The distribution of the context component assessment with the *Widyaiswara* as the respondents is shown in Figure 1.

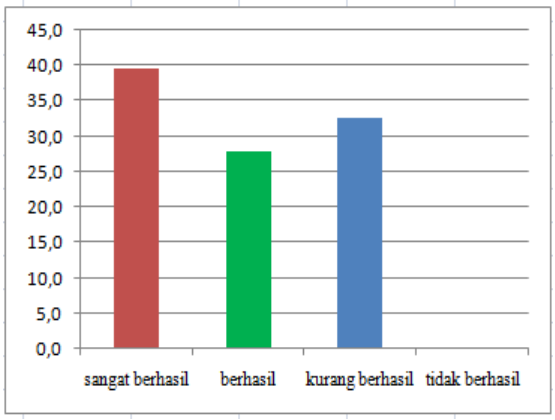

Figure 1. Distribution of Answers on the Context Evaluation (*Widyaiswara* Respondents)

The information of the context components is explained as follows. The term *sangat berhasil* means very successful, *berhasil* means successful, *kurang berhasil* means less successful, and *tidak berhasil* means not successful.

Figure 1 shows the distribution of answers on the context evaluation of the implementation of the Islam education teachers' training program. There are 39.5% of the *Widyaiswara* respondents who assess that the program is very successful. The other 27.9% say it is successful, while there are 32.6% of *Widyaiswara*, who states that the context of the program is not successful. Furthermore, the distribution of answers on the context evaluation based on the questionnaires which were distributed to the program organizers as the respondents is presented in Figure 2.

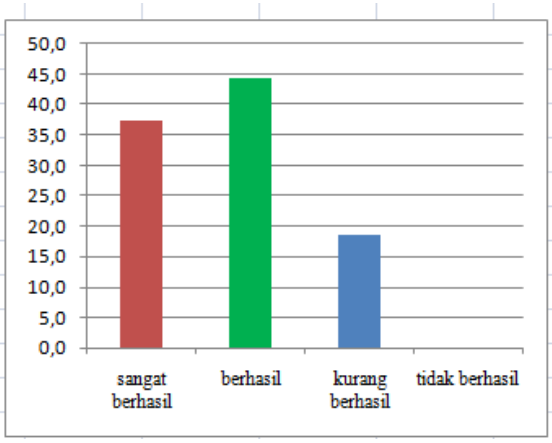

Figure 2. Distribution of Answers on the Context Evaluation (Organizers Respondents)

Figure 2 shows that 44.2% of the organizers state that the implementation of the Islam education teachers' training program in terms of the context aspect is included in the successful category. The success is achieved in the sub-components of program planning, training facilities, and training classroom atmosphere. The very successful category is stated by 37.2% of the program organizers. In comparison, 18.6% of the organizers state that the program was less successful in terms of the context component.

The results of the assessment conducted through the observation and documentation by the researchers show that the implementation of the Islam education teachers' training program is included in the very successful category, obtaining a per-

centage of 41.9% and 32.6%, respectively. In the successful category, the percentage is 41.9% and 46.6%, respectively. The less successful category obtains a percentage of 16.3% and 20.9%, respectively. The distribution of the observation and documentation results can be seen in Figure 3.

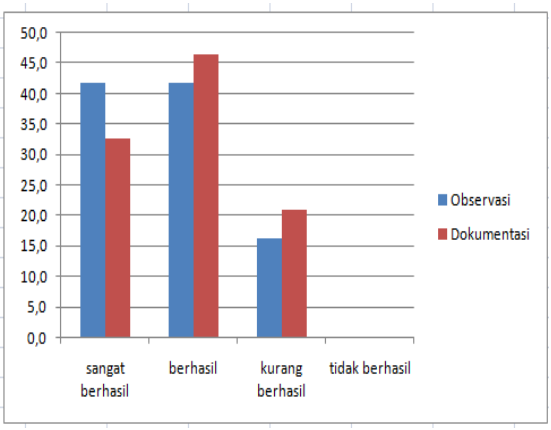

Figure 3. Context Evaluation Results (Observation and Documentation by the Researchers)

In general, the results of the study with the *Widyaiswar*a and program organizers as the respondents and then the observation and documentation activities on the context aspect of the Islam education teachers' training program obtain a mean score of 3.15, which is included in the successful category. Therefore, it can be concluded that the context of the implementation of the Islam education teachers' training program is in accordance with what is expected.

Evaluation on the Input of the Islam Education Teachers' Training Implementation

Evaluation of the input component is obtained from the measurement of several sub-components, including the administration of the participants and *Widyaiswara*, training environment, training facilities, availability of training equipment, and the atmosphere of the training environment. The distribution of answers to the input component from the *Widyaiswara*, organizers, and participants as the respondents is shown in Figure 4.

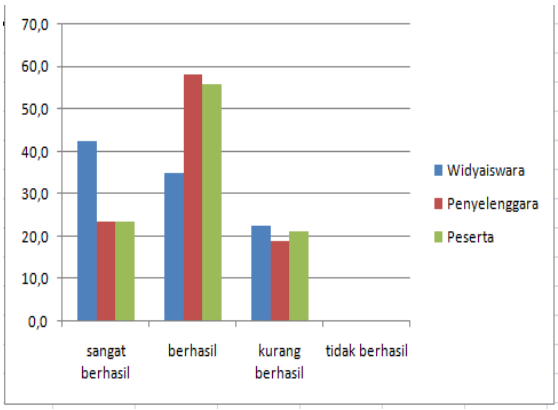

Figure 4. Distribution of Answers on the Input Evaluation

Figure 4 shows the distribution of the respondents' answers regarding the aspect of input evaluation. As many as 42.5% of *Widyaiswara* state that the input aspect of the implementation of the Islam education teachers' training program is very successful. In addition, 23.3% of the organizers (*penyelenggara*) and 23.3% of the training participants (*peserta*) also state that the implementation of the program in terms of the input aspect is very successful. Furthermore, 35.5%, 58.1%, and 55.8% of the three groups of respondents state that the input aspect is classified as successful. The less successful category is chosen by 22.5%, 18.6%, and 20.9% of the respondents from the three different groups. In general, the answers of the respondents obtain an average score of 3.04, which is included in the category of very successful. Thus, it indicates that the administration of participants and *Widyaiswara*, the training environment, training facilities, the availability of training equipment, and also the atmosphere of the training environment in terms of the input aspect are classified as very good.

Evaluation on the Process of the Islam Education Teachers' Training Implementation

The process evaluation is obtained from an analysis of the process of the Islam education teachers' training program implementation. Evaluation in the process component consists of several sub-components covering the *Widyaiswara* motivation, teach-

ing and learning process, development opportunities, *Widyaiswara* performance, use of media, management capabilities, *Widyaiswara* competence, and participant assessment. In addition, the process component can also show the reaction or satisfaction of the participants in the implementation of the Islam education teachers' training program. Evaluation of this component is conducted by using a questionnaire, observation sheet, and documentation sheet. Questionnaires were distributed to the *Widyaiswara*, the organizers, and the participants. Observations and documentation were carried out by the researchers themselves. The results of the assessment of the process component using questionnaires are shown in Figure 5.

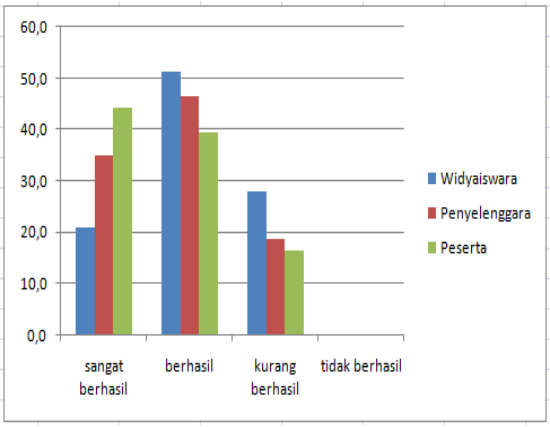

Figure 5. Distribution of Answers on the Process Evaluation Using Questionnaires

Figure 5 shows that the implementation of the Islam education teachers training program in the religious education and training center in Palembang with the *Widyaiswara* as the respondents on the process aspect is included in the successful category with a percentage of 51.25%. The organizers' answers for the four categories in a row obtain a percentage of 34.9%, 46.7%, 18.6%, and 0%. The highest percentage from the organizers is in the successful category. From the participant respondents, the percentage of the very successful category is 44.2%, the successful category is 39.5%, and the less successful category is 16.3%. Process evaluation with the participants as the respondents is included in the very successful category. The success of the process aspect is found in the sub-components of the conducive teaching and learning process, *Widyaiswara* performance, development opportunities, and the use of various media. Meanwhile, the lack of success is in the sub-component of a lack of motivation from the *Widyaiswara*. Overall, the process evaluation is classified as very successful because it obtains an average score of 3.15. In addition, the results of the process evaluation analysis based on the observations and documentation conducted by the researchers can be seen in Figure 6.

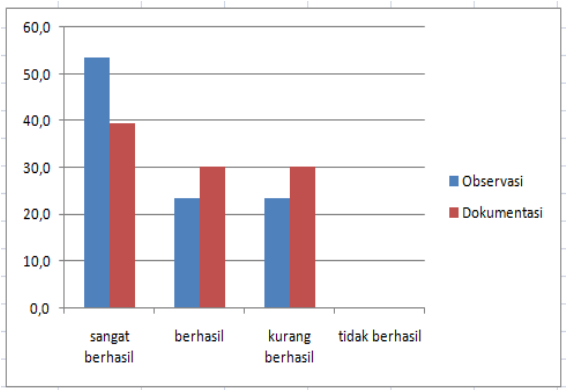

Figure 6. Observation dan Documentation Results of the Training Program

These results indicate that the implementation of Islam education teachers' training program in the religious education and training center in Palembang in the aspect of process using observation and documentation is included in the category of very successful. The observation obtains a greater percentage than the documentation with a percentage of 53.5% for observation and 39.5% for the results of the documentation. Success is found in the subcomponents of good management skills, and the assessment from the participants who fulfill the objective assessment criteria. In the category of successful, the observation obtains a smaller percentage than the documentation with a percentage of 23.3% for observation and 30.2% for documentation. In the less successful category, the percentage obtained is the same as the percentage in the successful category, which is

23.3% and 30.2%. Lack of success is shown in the sub-component of the lack of competence of the *Widyaiswara*. Whereas the unsuccessful category obtains 0%, meaning that no respondent states that the observation and documentation are unsuccessful.

The results of this study indicate that the implementation of the Islam eductaion teachers' training program in the religious education and training center in Palembang in terms of process aspect is included in the category of successful. It means the subcomponents of the process are fulfilled, namely the conducive teaching and learning process, *Widyaiswara* performance, development opportunities, the use of various media, good management skills, and assessment from the participants who fulfill the objective assessment criteria.

In addition, the process component can also show the participants' satisfaction towards the implementation of the Islam education teachers' training program. As many as 39.5% of the training participants express that they are very satisfied with the implementation of the Islam education teachers' training program at the religious education and training center in Palembang. As many as 41.9% of the training participants express that they are satisfied with the implementation of the training. The less satisfied category obtains a percentage of 18.6%. There are no participants who express dissatisfaction with the implementation of Islam education teachers' training. It means that most participants are satisfied with the implementation of the Islam education teachers' training at the religious education and training center in Palembang.

Evaluation on the Product of the Islam Education Teachers' Training Implementation

The product component consists of three subcomponents, namely the production of the teaching materials, the production of the evaluation equipment, and the production of the learning products. The results of the product component data analysis are presented in Table 4.

Table 4. Product Evaluation Distribution

Category	Respondents (%)	
	Participants	Researchers
Very Successful	44.2	48.8
Successful	39.5	41.9
Less Successful	16.3	9.3
Unsuccessful	0.0	0.0

The distribution of product component evaluation by the training participants obtains a percentage of 44.2%, which is classified in the very successful category. The percentage of successful, less successful, and unsuccessful categories is 39.5%, 16.3%, and 0%, respectively. The results of documentation conducted by the researchers obtain 48.8% for the very successful category, 41.9% for the successful category, 9.3% for the less successful category, and 0% for the unsuccessful category. It indicates that the evaluation of the product of the Islam education teachers' training program at the religious education and training center in Palembang is included in the very successful category. This success can be seen from the fulfillment of the three product evaluation sub-components.

Further, an analysis based on the average of each component is done to determine the success of the evaluation using the success criteria of Mardapi (2012, p. 162). The analysis results are shown in Table 5.

Table 5. Evaluation Results of Each Component

Component	Average	Category
Context	3.15	Very Successful
Input	3.04	Very Successful
Process	3.15	Very Successful
Product	3.34	Very Successful

Table 5 shows that the average of the context evaluation was at 3.15, and the average of the input evaluation was at 3.04. The average of the process and product are 3.15 and 3.34, respectively. As the average score of the four components is more than 3, according to the success criteria of the evaluation, then the four components are included in the very successful category.

The model implementation stage is a trial stage in the field where the model and its instruments, as well as the equipment, are applied with the aim of knowing whether or not the evaluation model can be used properly to evaluate the implementation of the Islam education teachers' training program. The trial subjects in this study consist of the organizers of the Islam education teachers' training program, the *Widyaiswara* who taught at the training program, the participants of the training program at the religious education and training center in Semarang (Central Java), the religious education and training center in Jakarta, and the religious education and training center in Palembang (South Sumatera), which consist of 150 people. The data collection instruments consist of questionnaires, observation sheets, and documentation sheets.

The data of the trial results questionnaire, observation sheets, and documentation sheets that had been collected were then analyzed quantitatively using the Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) through the SPSS program and Lisrel 8.54. CFA is used to check the construct validity of the instrument (Miller, 1996).

The CFA results show that the suitability test of the PGPAI evaluation model meets the model's goodness of fit test criteria because the p-value is 0.9996 ($p \geq 0.05$) and the RMSEA is 0.000 (RMSEA $\leq$ 0.08) (Ghozali, 2009, p. 32; Jöreskog & Sörbom, 1996, p. 124). In addition, the goodness of fit index (GFI) value is 0.98, adjusted goodness of fit index (AGFI) value is 0.96, normed fit index (NFI) value is 0.99, and comparative fit index (CFI) 1 has a value of $\geq$ 0.90, which means that the hypothetical PGPAI evaluation model meets the criteria of the model fit test. Further, the relation of component and variables of evaluation that include the PGPAI's context, input, process and product show correlational indicators with variables that have high loading factors of $\geq$ 0.3 (Doll, Xia, & Torkzadeh, 1994; Gorsuch, 1983; Tabachnick & Fidell, 2007). This result can be interpreted as the main indicator of the latent construct of the PGPAI model that has been considered good so that it is feasible to use because it has been tested empirically. The empirical model of the PGPAI evaluation model is presented in Figure 7.

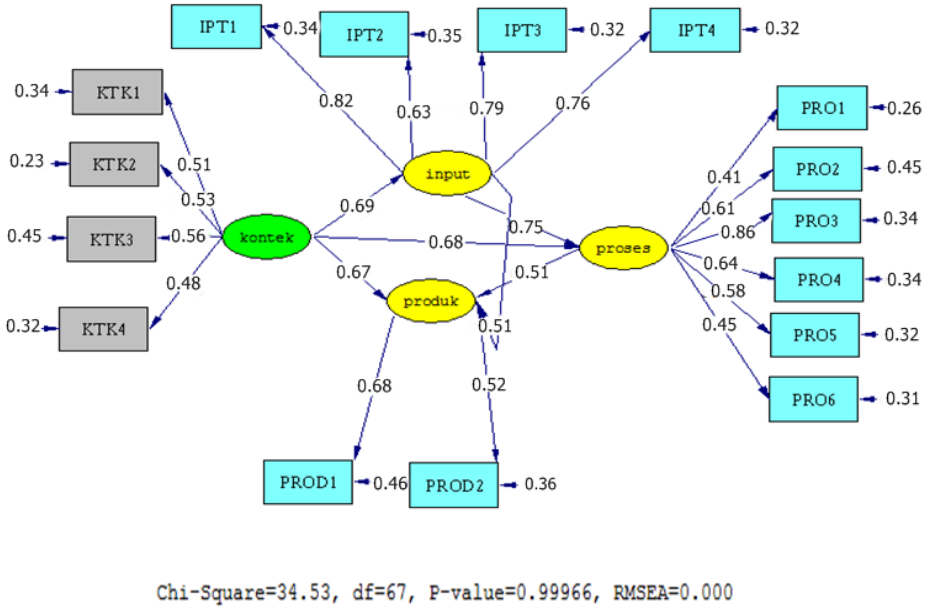

Figure 7. Empirical Model of the PGPAI Evaluation Model

In addition, an analysis was conducted to find out the results of the evaluation of the overall implementation of the Islam education teachers' training program. The analysis results obtain an overall average of 3.17. The average score, which is more than 3, indicates that it is in a very successful category. Based on the good average score from each component and as a whole, the evaluation results are included in the very successful category. It means that the Islam education teachers' training program at the religious education and training center in Palembang is very successful.

Conclusion

Based on the results of the data analysis, several conclusions can be drawn as follows. First, the implementation of the Islam education teachers' training program at the religious education and training center in Palembang in terms of the context component ias categorized as very successful. Second, the implementation of the training program in terms of the input component is categorized as very successful. Third, the implementation of the training program in terms of the process component is categorized as very successful, and the training participants are satisfied with the implementation of the training. Fourth, the implementation of the training program in terms of the product component is categorized as very successful. Fifth, the implementation of the Islam education teachers' training program at the religious education and training center in Palembang as a whole is very successful.

References

Arikunto, S., & Jabar, C. S. A. (2009). *Evaluasi program pendidikan: Pedoman teoritis praktis bagi praktisi pendidikan*. Jakarta: Bumi Aksara.

Bhargava, A., & Pathy, M. (2011). Perception of student teachers about teaching competencies. *American International Journal of Contemporary Research*, *1*(1), 77–81. Retrieved from http://www.aijcrnet.com/journals/Vol._1_No.1_July_2011/10.pdf

Borg, W. R., & Gall, M. D. (1993). *Educational research: An introduction*. New York, NY: Longman.

Cennamo, K., & Kalk, D. (2005). *Real world instructional design*. Belmont, CA: Thomson Wadsworth.

Doll, W. J., Xia, W., & Torkzadeh, G. (1994). A confirmatory factor analysis of the end-user computing satisfaction instrument. *MIS Quarterly*, *18*(4), 453–461. https://doi.org/10.2307/249524

Ghozali, I. (2009). *Aplikasi analisis multivariate dengan program SPSS*. Semarang: Badan Penerbit Universitas Diponegoro.

Goldstein, I. L., & Ford, J. K. (2002). *Training in organizations: Needs assessment, development, and evaluation* (4th ed.). Belmont: Wadsworth/ Thomson Learning.

Gorsuch, R. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Herzberg, F., Mausner, B., & Snyderman, B. B. (2010). *The motivation to work*. New Brunswick, NJ: Transaction Publishers.

Ilukena, A. (1998). Qualities and competencies of the professional teacher. *Journal for Educational Reform in Namibia*, *7*, 1–5. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.559.7304&rep=rep1&type=pdf

Jöreskog, K. G., & Sörbom, D. (1996). *Lisrel 8: User's reference guide*. Englewood Cliffs, NJ: Prentice Hall.

Knowles, M. S. (1980). *The modern practice of adult education: From pedagogy to andragogy*. Englewood Cliffs, NJ: Prentice Hall.

Law No. 14 of 2005 of Republic of Indonesia about Teachers and Lecturers. , (2005).

Mardapi, D. (2012). *Pengukuran, penilaian, dan evaluasi pendidikan*. Yogyakarta: Nuha Medika.

McDavid, J. C., & Hawthorn, L. R. L. (2006). *Program evaluation and performance measurement: An introduction to practice*. Thousand Oaks, CA: Sage Publication.

Miller, V. A. (1996). The history of training. In R. Craig (Ed.), *The ASTD training and development handbook: A guide to human resource development*. New York, NY: McGraw-Hill.

Notoatmodjo, S. (1991). *Pengembangan sumber daya manusia*. Jakarta: Rineka Cipta.

Phillips, J. J. (1991). *Handbook of training evaluation and measurement methods*. Houston, TX: Gulp Publishing Company.

Research Center for Educational Policy and Innovation. (2007). *Kompetensi guru sesuai standar nasional pendidikan (SNP)*. Jakarta: Institute of Research and Development, Department of National Education.

Suraji, I. (2010). *Kompetensi guru Madrasah: analisis kompetensi pedagogik, kepribadian, dan sosial guru Madrasah Ibtidaiyah di Kota Pekalongan*. Doctoral thesis. Universitas Islam Negeri Sunan Kalijaga, Yogyakarta.

Susatya, E. (2013). Pengembangan model pelatihan guru sekolah menengah kejuruan kelompok seni dan budaya. *Jurnal Pendidikan Dan Kebudayaan*, *19*(1), 107–122.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (4th ed.). Hillsdale: Erlbaum.

Zainuri, A. (2012). *Tingkat kompetensi guru Madrasah Ibtidaiyah Negeri (MIN) Kota Palembang*. Doctoral thesis. Universitas Islam Negeri Sunan Kalijaga, Yogyakarta.

Jurnal Penelitian dan Evaluasi Pendidikan
Volume 23, No 2, December 2019 (117-128)
Online: http://journal.uny.ac.id/index.php/jpep

PAPS PREDICTIVE VALIDITY IN PREDICTING THE LEARNING SUCCESS OF MASTER OF PROFESSIONAL PSYCHOLOGY STUDENTS

Flaviana Rinta Ferdian
Universitas Gadjah Mada

Saifuddin Azwar
Universitas Gadjah Mada

Abstract

This study aims to determine the ability of *Tes Potensi Akademik Pascasarjana* (PAPS Test) in predicting the grade point average (GPA) of the graduate students of the Master of Professional Psychology. PAPS test is the graduate entry requirement test in UGM. Graduate students' GPA includes semester grade point average (semester GPA), cumulative grade point average (cumulative GPA), grade point average of master courses, and grade point average of professional courses. The data were obtained in the form of PAPS test scores (verbal subtest, reasoning subtest, quantitative subtest, and total score) and the GPA of the graduate students of Master of Professional Psychology of Universitas Gadjah Mada in the class of 2013, 2014, and 2015. The research subjects were 214 students. The analysis was carried out using correlation analysis. The results of the analysis generally indicate that the PAPS test is able to become a predictor of learning success for the graduate students in the Master of Professional Psychology of Universitas Gadjah Mada. The learning success consists of semester GPA, cumulative GPA, GPA of master courses, and GPA of professional courses. If analyzed per subtest, the quantitative subtest has the highest predictive power. On the other hand, the verbal subtest is only a good predictor for GPA of professional courses.

Keywords: *graduate students, entry requirement test, predictive validity, GPA, PAPS test*

Permalink/DOI: http://dx.doi.org/10.21831/pep.v23i2.25299

Contact: *Flaviana Rinta Ferdian*
flaviana.rinta@gmail.com
Faculty of Psychology, Universitas Gadjah Mada
Jl. Sosio Humaniora Bulaksumur, Karangmalang, Caturtunggal, Depok, Sleman, Yogyakarta 55281, Indonesia

Introduction

Nowadays, continuing to graduate education seems to be a necessity for students. It is indicated by the increasing number of master and doctoral degree students in Indonesia from year to year. Based on the data on education development in 1999/2000-2013/2014 from the Ministry of Education and Culture in 2015, the number of students in the master's degree program has the highest growth of 7.15% per year from 40,309 to 65,366. Likewise, the increase also occurred in the doctoral degree program, which has a growth of 1.54% from 4,349 to 4,839 students (Ministry of Education and Culture, 2015).

The increase in the number of graduate students in Indonesia also happens at Universitas Gadjah Mada (UGM). This increase is indicated by the number of UGM's graduate students in 2002 reaching 12% of the total number of UGM's students, and in 2007, it rose sharply to 28% (Public Relations of UGM, 2007).

Similarly, the increase in the number of graduate students also took place in the Master of Professional Psychology of Universitas Gadjah Mada. Table 1 shows data on enthusiasts and students accepted in the Master of Professional Psychology of Universitas Gadjah Mada in the 2015/2016 until 2017/2018 academic year based on the academic data of the Master of Professional Psychology of Universitas Gadjah Mada.

Table 1. Data of the Enthusiasts and Students Accepted in Master of Professional Psychology of Universitas Gadjah Mada in the Academic Year of 2015/2016 until 2017/2018

Academic Year	Number of Applicants (percentage)	Total Accepted (percentage)
2015/2016	30.31%	30%
2016/2017	31.71%	34.83%
2017/2018	37.98%	35.17%

The vast public interest in continuing to the Master of Professional Psychology of Universitas Gadjah Mada (UGM) indeed cannot be separated from the good quality of education at UGM. UGM is a university providing many departments and study programs in the field of science in Indonesia. The Master of Professional Psychology of Universitas Gadjah Mada consists of 18 faculties and one graduate school with a total of 89 master study programs. Besides being extensive in the field of science, UGM also has good achievements in the international arena. It is indicated by UGM as the top 500 best universities in the world, according to QS World University Ranking in 2017/2018 (Top Universities, 2017). It makes many people interested in continuing their graduate studies at UGM.

Since the even semester of 2008/ 2009, UGM has required the applicants of the Graduate Program to take the *Tes Potensi Akademik* (TPA) or the Academic Potential Test. TPA is developed and implemented by *BAPPENAS* (*Badan Perencanaan Pembangunan Nasional* or National Development Planning Agency)or the *Potensi Akademik Pascasarjana* (PAPS) test, which is developed and implemented by the *Unit Pengembangan Alat Tes Psikodiagnostika* (UPAP) or Psychodiagnostic Test Equipment Development Unit, Faculty of Psychology UGM. TPA is used as a graduate entry requirement because it is specifically designed to explore a person's ability to deal with cognitive problems requiring a strategic and fast solution approach. This ability is a general reasoning ability supporting individual performance in solving problems. General reasoning skills are needed in the learning process in universities, where the acquisition of knowledge and skills is not much more determined by the method of delivery and order following the learning syllabus as it occurs at the middle and primary level of education (Azwar, 2016b).

This entry test requirements must be genuinely effective in determining the students predicted to have a high learning success in graduate studies. Moreover, UGM, with many interested applicants, must predict the learning success so that the decisions taken on the results of the tests do not disadvantage some parties. This test has

serious consequences for at least one stakeholder or often referred to as a high stakes test (Gregory, 2007; Stobart & Eggen, 2012). The intended consequences are for students and the university. For students, the test determines the future of an individual. If an error occurs, unselected prospective applicants (false negative), it affects the future of the students. The student should be able to enroll in a proper department and university according to his ability, but he should not choose another department or university because of the inability of the test in predicting his potential. Likewise, choosing applicants who have no successful prospects (false positives) will be detrimental to the faculty and students. The disadvantage is that choosing students who have no successful prospects will weaken the quality of the faculty, and the students will also have difficulty in attending lectures that have an impact on students' academic inefficiencies (Kuncel et al., 2001). Thus, the test used both for the entry requirements and graduate selection must be absolutely valid to predict students' success during the study.

Therefore, it is important to ensure that the measuring instruments in the student admission process in the graduate programs at UGM are truly able to predict the learning success of students by examining the predictive validity of the measuring instrument. Validity is the degree to which a test performs its measuring function or measures what intends to measure (Azwar, 2015b; Cohen et al., 2013; Gregory, 2007). Meanwhile, predictive validity is the extent to which test scores can be used to predict measure scores administered at a later date (Cohen et al., 2013; Gregory, 2007; Rudner, 1994).

One of the measuring instruments used in the student admission process at Universiats Gadjah Mada is the *Tes Potensi Akademik Pascasarjana* (PAPS test). PAPS test is a cognitive ability test specifically designed to reveal the academic potential underlying the possibility of the applicants to be successful if he gets the opportunity to learn more in university, especially in the master and doctoral programs (Widhiarso, 2017). PAPS test is developed using the intelligence theory foundation of Spearman's g factor (general cognitive ability). This general cognitive ability is needed by students to study in higher education. Therefore, the PAPS test is used as one of the graduate entry requirements and is expected to be highly correlated with the indicators of learning success in university (Azwar, 2008).

Related to the use of TPA in the student admission process, the aspects of validity, especially predictive validity, are important for predictive accuracy (Azwar, 2008). Tests that have been examined for validity are effective tests for the student admission process because they have good predictive functions (Azwar, 2015b; Kuncel et al., 2001). An effective acceptance process can maintain qualified candidates and improve the quality of the department (Kuncel et al., 2001). Thus, it is important to look at the PAPS test predictive validity as an admission requirement for the graduate programs at UGM.

The importance of research on predictive validity, especially for tests determining the individual's potential, is also indicated by many studies related to testing of predictive validity. For the PAPS test, several studies had been conducted on predictive validity. The research by Pradipta et al. (2017) shows that TPA scores in both TPA *Bappenas* and PAPS tests can predict the learning success of the master's degree students at UGM. In addition, the research by Nurhayati and Widhiarso (2017) examines the specific predictive validity of the PAPS test and shows that the test predicts the learning success in the Master of Psychology of UGM. From the research by Nurhayati and Widhiarso (2017), it is suggested to conduct research towards other study programs.

Other research also stated the importance of conducting predictive validity in other study programs. Academic potential test scores, such as Graduate Record Examination (GRE), demonstrate varying pre-

dictive validity, depending on the area of study (Goldberg & Alliger, 1992; Sternberg & Williams, 1997; Thornell & McCoy, 1985). Although there are many similarities in fundamental assignments for the students, there are still many differences in curricula, output, success measures, assessment standards, and evaluation criteria in different academic areas (Mupinga & Mupinga, 2005; Wang, 2013). Therefore, the results of predictive validity research produce valid conclusions for several purposes and several contexts, but cannot be generalized for all purposes and contexts (Powers, 2004). Thus, it is necessary to consider the predictive validity research on TPA in different programs, even on closely related study programs (Wang, 2013). Therefore, this research aims to conduct a PAPS predictive validity test on the learning success of the graduate students of the Master of Professional Psychology of Universitas Gadjah Mada.

More specifically, the PAPS test criteria used in this study have two parameters, namely the semester grade point average (semester GPA) and cumulative grade point average (cumulative GPA). These criteria are used to see the trends in the relation between PAPS test and learning success. In particular, the trend analysis aims to look at PAPS predictability over time during the academic semester (Wao et al., 2016). The criteria are also used with the aim of contributing to the literature focusing on long-term performance, while long-term performance is a more informative criterion than short-term performance (Wao et al., 2016). The GPA referred to in this study is not only the overall GPA from the beginning to the end of the semester but the accumulation of each semester. For example, cumulative GPA 2 is an accumulation of the first semester GPA and the second semester GPA.

In addition, other criteria used are the GPA of master courses and GPA of professional courses. In the Master of Professional Psychology, there are master courses that cover Philosophy, Research Method courses, and Thesis, and then, there are professional courses such as Observation, Interview, and Assessment courses. Professional courses require more empathy and skill, while the PAPS test measures cognitive abilities more. Therefore, the researchers tested the ability of PAPS in predicting GPA of Master and GPA of professional courses.

Based on the aforementioned description, this study aims to examine further the extent to which PAPS test scores can predict learning success as indicated by semester GPA, cumulative GPA, GPA of master courses, and GPA of professional courses for graduate students of Master of Professional Psychology at UGM. The research problem is formulated as follows: (1) how far PAPS test can predict semester GPA courses for graduate students of Master of Professional Psychology at UGM; (2) how far PAPS test can predict cumulative GPA courses for graduate students of Master of Professional Psychology at UGM; (3) how far PAPS test can predict GPA of master courses and GPA of professional courses for graduate students of Master of Professional Psychology at UGM.

Research Method

Research regarding the predictive validity involves at least one test variable as a predictor and one other variable as a validation criterion. In this study, the following variables were determined. The predictive variable in this study consists of PAPS test scores of graduate students of the Master of Professional Psychology at UGM. The PAPS test scores are verbal subtest score, quantitative subtest score, reasoning subtest score, and PAPS test total score. Then, for the criterion variables in this study, there are four parameters, namely semester grade point average (semester GPA), cumulative grade point average (cumulative GPA), GPA of master courses, and GPA of professional courses. Semester GPA consists of first semester GPA (semester GPA 1), second semester GPA (semester GPA 2), third semester GPA (semester GPA 3), fourth

semester GPA (semester GPA 4), and fifth semester GPA (semester GPA 5). Cumulative GPA consists of cumulative GPA 2 (accumulation of semester GPA 1 and semester GPA 2), cumulative GPA 3 (accumulation of semester GPA 1, semester GPA 2, and semester GPA 3), cumulative GPA 4 (accumulation of semester GPA 1, semester GPA 2, semester GPA 3, and semester GPA 4), and cumulative GPA 5 (accumulation of semester GPA 1, semester GPA 2, semester GPA 3, semester GPA 4, and semester GPA 5).

The study data were in the form of a total PAPS test score and GPA of graduate students of Master of Professional Psychology of UGM in 2013, 2014, and 2015. Data were obtained from the *Unit Pengembangan Alat Tes Psikodiagnostika* (UPAP) UGM and Master of Professional Psychology UGM.

To determine the relation of predictor variables, namely PAPS test (verbal subtest score, quantitative subtest score, reasoning subtest score, and total score) with GPA of graduate students of Master of Professional Psychology of UGM (semester GPA, cumulative GPA, GPA of master courses, and GPA of professional courses), correlation analysis was used. The predictor variable is the PAPS test score, and the criteria variable is the Students' GPA. Table 2 presents a guideline for interpreting uncorrected correlation coefficients in the study of predictive validity based on simplified guidelines from the US Department of Labor, Employment Training and Administration cited by Emery (2007) in (Azwar, 2016a).

Table 2. Simplified Guidelines for Interpretation of Uncorrected Correlation Coefficients in Predictive Validity Studies

Coefficient of Validity	Interpretation
> 0.35	Very satisfying
0.3-0.35	Satisfying
<0.3	Less satisfying

Findings and Discussion

A predictive validity test was done by correlating predictors with the criteria to obtain a predictive validity coefficient (Rudner, 1994). The PAPS predictive validity test is indicated by the linear correlation coefficient between verbal subtests, quantitative subtests, reasoning subtests, and PAPS test total scores as predictors with students' GPA as the criteria. The results of simple linear correlation coefficient computation are presented in Table 3.

Table 3. The Results of Correlation Analysis between PAPS and Students' GPA

Criteria	N	Predictors			
		Verbal	Quantitative	Reasoning	Total Score
Semester GPA 1	213	0.047	0.208	0.118	0.177
Cumulative GPA 1	213	0.047	0.208	0.118	0.177
Semester GPA 2	214	0.025	0.226	0.113	0.175
Cumulative GPA 2	214	0.046	0.244	0.142	0.205
Semester GPA 3	173	0.086	0.052	0.110	0.100
Cumulative GPA 3	214	0.045	0.215	0.120	0.180
Semester GPA 4	35	-0.273	0.077	0.294	0.032
Cumulative GPA 4	214	0.041	0.218	0.128	0.183
Semester GPA 5	36	0.063	-0.072	-0.084	-0.041
Cumulative GPA 5	212	0.027	0.206	0.116	0.167
GPA of Master Courses	145	0.075	0.208	0.192	0.216
GPA of Professional Courses	171	0.099	0.139	0.086	0.141

Note:
Semester GPA = Semester Grade Point Average;
Cumulative GPA = Cumulative Grade Point Average

The analysis shows that the highest correlation is the correlation between reasoning subtests and semester GPA 4 of 0.294. On the other hand, the lowest correlation is the correlation between verbal subtests and semester GPA 2 of 0.025. From the results of the analysis, it can also be revealed that the correlation between the PAPS test score and the grade point average generally has a positive correlation up to the third semester. Nevertheless, in the fourth and fifth semesters, there is a negative correlation. In addition, it can also be seen that general quantitative subtests have a higher correlation number than verbal subtests, reasoning subtests, and also the total score. However, verbal subtests generally have lower correlation rates than quantitative subtests, reasoning, subtests, and also total score.

The results of the analysis also reveal that the grade point average of the master courses compared to the grade point average of the professional courses has a higher correlation with quantitative subtests, reasoning subtests, and total scores. On the other hand, the grade point average of the professional courses compared to the grade point average of the master courses, has a higher correlation with verbal subtests.

From the results of the analysis, the correlation coefficient between predictors (the PAPS test and criteria) can be seen, namely, the students' GPA. The correlation coefficient obtained turns out that some of the numbers are so close to zero that there is no linear relation between the GPA and the PAPS test score. In addition, there is a correlation coefficient that has a negative mark.

Then, the researchers correct the restriction of range in order to increase the accuracy of the estimation of the coefficient of actual predictive validity. Positive correlation coefficients in Table 2 were corrected for restriction of range on the criteria, namely students' GPA. The correlation coefficients marked negative, even though the numbers are large, are not included in the analysis because negative numbers have no meaning in interpreting validity (Azwar & Ancok, 2008). Therefore, in this study, corrections for restriction of range on the criteria were carried out because the criteria were more homogeneous than the predictors.

Table 4. Coefficient Validity Predictive (Corrected for Restriction of Range on the Criteria)

Criteria	N	Predictors			
		Verbal	Quantitative	Reasoning	Total Score
Semester GPA 1	213	0.130	0.510	0.315	0.449
Cumulative GPA 1	213	0.130	0.510	0.315	0.449
Semester GPA 2	214	0.080	0.595	0.341	0.493
Cumulative GPA 2	214	0.160	0.664	0.451	0.594
Semester GPA 3	173	0.225	0.138	0.284	0.260
Cumulative GPA 3	214	0.157	0.613	0.392	0.542
Semester GPA 4	35		0.099	0.368	0.041
Cumulative GPA 4	214	0.143	0.619	0.414	0.549
Semester GPA 5	36	0.197			
Cumulative GPA 5	212	0.095	0.596	0.381	0.513
GPA of Master Courses	145	0.190	0.481	0.450	0.495
GPA of Professional Courses	171	0.347	0.463	0.306	0.468

Note:
Semester GPA = Semester Grade Point Average;
Cumulative GPA = Cumulative Grade Point Average;
the bold numbers are coefficient ≥ 0.30

The formula used is as follows:

$$r_U = r_R (s_U/s_R) / \sqrt{[1-r_R^2 + r_R^2(s_U^2/s_R^2)}$$

Note:
s_U = standard deviation predictor score for unrestricted group
s_R = standard deviation predictor score for restricted groups
r_R= validity coefficient obtained from restricted groups (Azwar, 2015b).

r_R was obtained from correlation coefficients in Table 2. s_R was obtained from empirical standard deviations. s_U was obtained from a hypothetical standard deviation calculation. The hypothetical standard deviation was calculated by dividing the GPA range by six units of standard deviation (Azwar, 2015a). The range of GPA is 4 divided by six units of standard deviation resulting in a hypothetical standard deviation of the GPA of 0.67. The computational results of the corrected correlation coefficients are presented in Table 4.

This study is intended to see the predictive power of the PAPS test on the learning success of graduate students of Master of Professional Psychology of UGM. The PAPS test has been used as an entry requirement for the Graduate Programs at UGM. Therefore, it is important to examine the PAPS test's ability to predict the learning success as reflected in the students' GPA. If PAPS test has good predictive power, it will help attract prospective students who have the potential to succeed so that they can maintain the qualified applicants and improve the quality of study programs (Kuncel et al., 2001).

The small correlation coefficient, as presented in Table 3, can be caused by homogeneous groups. Homogeneous groups can be seen from the small differences in the observed scores, showing the narrow spread of numbers and also the lack of diversity in numbers (Azwar, 1996; Furr & Bacharach, 2014; Kurpius & Stafford, 2006). The correlation coefficient indicates how well the point in the scatters plot runs along the line. If the group is homogeneous, then the point in the scatter plot only collects on a certain line and does not run along the line. It has an impact on the low correlation coefficient (Allen & Yen, 1979).

Seeing whether the research data is homogeneous or not can be identified from its variability. In statistics, there are three measures showing the variability or diversity of numbers in a distribution. The three measures are range, standard deviation, and variance. Range, or distribution distance, is the difference between the biggest score and the smallest score. As a variability measure, the range is actually very unstable and easily provides false information about actual variability. It is because if there is one extreme number, the range will change dramatically while the overall variation in numbers is not much affected by only one extreme number. The standard deviation is the average number deviation from the mean. The variance is the sum of squares of deviations from the mean (Allen & Yen, 1979; Furr & Bacharach, 2014; Kurpius & Stafford, 2006).

In the data, it can be seen that the standard deviation for the PAPS test score and the students' GPA value is relatively small. The standard deviation of the PAPS test is 100. In this study, the standard deviation moves between 52.83 until 86.43. The standard deviation in this study is below the standard deviation of the PAPS test. Then, the standard deviation of students' GPA moves between 0.18 until 0.82. A small standard deviation indicating the group of the sample of this study is homogeneous. This homogeneous group influences the magnitude of the coefficient of the validity of measurement results (Azwar, 2016a; Sackett et al., 2002).

In predictive validity, it is often defined as a restriction of range. Restriction of a range is a decrease in the variance of the sample score in both the independent and dependent variables so that the score distribution becomes more homogeneous compared to the previous variant. This restriction results in a low correlation coefficient or predictive validity coefficient (Allen & Yen, 1979; Sackett et al., 2002).

Such cases usually occur when groups that initially have a wide variety of abilities then become "selected" groups (Azwar, 2015b). Likewise, when a large number of prospective students participating in the admission process of Master of Professional Psychology of UGM were selected, only a small number of them were accepted, namely those who met the entry requirements and had the best scores on aspects of the admission test. When undergoing an admission test, the variation in their scores was large because prospective students consist of students with very diverse or varied abilities. Therefore, if these all could be accepted as students, the GPA would also vary greatly. However, only those with the best abilities were selected in the selection, namely those who met the entry requirements and the highest admission test scores. Hence, groups of students who passed the selection are more homogeneous subject groups and resulting in restriction of range. This restriction of range resulted in underestimating the true coefficient of validity. Thus, we must correct the restriction of range to increase the accuracy of the estimation of the coefficient of actual predictive validity. The result of Coefficient Validity Predictive (Corrected for Restriction of Range on the Criteria) can be seen in Table 4.

Considering the categorization of the validity coefficient meaning according to the US Department of Labor, Employment Training and Administration (Emery, 2007) in Azwar (2016b) which had been simplified, all validity coefficients smaller than r = 0.300 is stated not satisfying and the test is declared unable to produce a valid measure. From the calculation of the correction for restriction of range, it can be seen that there is an increase in the correlation coefficient. It further supports that the group in this study is homogeneous. Therefore, the small number of correlation coefficients or predictive power in this study is not because there is a problem with the predictor but is caused by a very homogeneous group. If the group is more heterogeneous, then it will produce higher PAPS test predictive power towards students' GPA. It is in accordance with previous research by Azwar (2008), which found a low relationship between TPA version 07A and the first-year GPA in Master of Professional Psychology of UGM, Class of 2007. It is because the first-year GPA used as a validation criterion is very homogeneous. This condition resulted in no estimation of true predictive validity because computational validity resulted in underestimation. The low coefficient of validity occurred not because of the problem in the predictor but rather from the problem of reliability of the criterion score that was worth questioning and the variability of the predictor and also the criteria score, which is equally homogeneous.

Thus, it appears from the computational results in Table 4 that, in general, the coefficient of validity obtained is greater than r = 0.300, except for verbal subtests. More specifically, the total PAPS test score serves as a valid predictor of the students' GPA, except for semester GPA 3, semester GPA 4, and semester GPA 5. Verbal PAPS test subtitles only function as valid predictors of the GPA of Professional courses. In addition, it can also be seen that the quantitative subtests on the PAPS test are valid predictors for the students' GPA, except for semester GPA 3, semester GPA 4, and semester GPA 5. The reasoning subtests serve as valid predictors for students' GPA, except for semester GPA 3 and semester GPA 5.

In addition, it can be seen that the PAPS test is more consistent if the cumulative GPA is used as a criterion. It can be explained that the learning success of students is reflected by overall achievement, so the use of Cumulative GPA as a criterion is more precise than use semester GPA (Azwar & Koentjoro, 1985). The results of the analysis also show that quantitative subtests are generally the best predictive power compared to other subtests in predicting semester GPA and cumulative GPA. On the other hand, verbal subtests

do not predict semester GPA and cumulative GPA. It can be explained that in previous studies, it was also found that verbal abilities such as those in Indonesian and English test courses did not have predictive power towards learning achievement as well as quantitative abilities (Azwar & Koentjoro, 1985).

The Azwar and Ancok (2008) study of differential predictive validity in the examination in the admission process or *Ujian Masuk* of UGM (UM UGM) also shows that quantitative subtests in the potential academic test at the Faculty of Psychology have a higher predictive power (r = 0.263), compared to verbal subtests (r = 0.152) and reasoning subtests (r = 0.015). The reasoning ability that is revealed through skills in finding more quantitative solutions is needed to undergo success in psychology studies rather than reasoning skills through understanding the meaning of words.

More specifically, in the Master of Professional Psychology of UGM, in semester GPA and cumulative GPA, there are master courses and professional courses. Quantitative abilities are used both for basic master courses and professional courses. Quantitative abilities are used to understand the results of research, especially in the interpretation of the results of quantitative analysis. Quantitative ability is also used to determine statistical analysis used in accordance with research problems in the experimental research design. Then, in the basic professional courses, the students will accumulate behavior from observation, words from interviews, and calculate scores from psychological tests to make a diagnosis. On the other hand, verbal ability is only needed in professional courses that require the students to write reports for all of their subjects.

In understanding references for master and professional subjects, reasoning skills rather than verbal abilities are needed. It is because, in graduate studies, the students need not only understanding reading but also a more critical attitude towards existing references by drawing conclusions from theory and then used to interpret cases. It requires more reasoning skills, not only the ability to interpret words as measured by verbal abilities. Therefore, quantitative ability predicts semester GPA and cumulative GPA in the Master of Professional Psychology of UGM. In contrast, the verbal ability does not predict semester GPA and cumulative GPA in the Master of Professional Psychology of UGM (Azwar, 2016b; Master Program of Professional Psychology, 2013).

However, when distinguished into a master and professional degree course, the total score has the highest predictive validity. Then, the verbal ability can predict professional courses. Verbal ability can predict the learning success in psychology, especially in the Master of Professional Psychology, only in courses that meet directly with clients. When students meet with clients to assess and provide therapy, and verbal abilities predict their success. On the other hand, when they learn about theories in psychology research and carry out theses, verbal abilities do not predict their learning success. In addition, in the professional courses, the students always make psychological reports in each of their subjects. The students describe the conclusions of the assessment results, observations, interviews, and psychological tests in the form of psychological reports. The students also make reports on the results of psychological interventions that have been given. Therefore, verbal ability is used in professional courses to interact with clients, arrange good sentences, and collect the results of interviews. On the other hand, verbal ability is not needed for master courses because, in the master courses, there tends to be more theory and does not meet with clients to collect the results of interviews (Master Program of Professional Psychology, 2013).

This study found that the initial analysis shows a low correlation between the PAPS test and the GPA of Master of Professional Psychology of UGM. However, after correction for the restriction of

range on the criteria shows that the PAPS test has good predictive validity on the grade point average. It shows that the criteria of a homogeneous grade point average have an impact on the low correlation coefficient. The students' GPA of the Master of Professional Psychology of UGM has a relatively high score because, indeed, the students are selected students and have the best scores. Therefore, the low correlation coefficient in the initial analysis is not caused by the inability of the PAPS test to predict students' GPA, but rather due to homogeneous criteria score data. Thus, in general, the PAPS test can be a good predictor for the students of the Master of Professional Psychology of UGM.

The limitation of this study is vulnerability to commission bias. PAPS test is tested both for enrolled applicants and applicants who were not accepted. However, in this study, only the prospective students were accepted. It causes that the predictive validity of the PAPS test cannot be empirically known for all applicants of Master of Professional Psychology of UGM, both those who are accepted and those who are not accepted. In addition, another limitation is that this study only focuses on predictive validity. Therefore, this study can only answer whether the PAPS test is valid or not as a predictor of the learning success of the Master of Professional Psychology of UGM. However, this study did not examine further how far the PAPS test was able to predict the learning success in the Master of Professional Psychology of UGM.

Conclusion

Based on the results of this study, it can be concluded that in general the PAPS test is able to be a predictor of the learning success of the graduate students of Master of Professional Psychology of UGM and can continue its use as an entry requirement for Master of Professional Psychology of UGM. Learning successes referred to in this research include semester GPA, cumulative GPA, GPA of master courses, and GPA of professional courses. If viewed per subtest, quantitative subtest has the highest predictive power. On the other hand, verbal subtests are only good predictors of the GPA of professional courses.

Based on the results of the study, some advices can be given. For the master of Professional Psychology of UGM, as the party using PAPS test as an entry requirement can continue its use because it is proven that the PAPS test is able to be a predictor of the students' learning success of the master of Professional Psychology of UGM. The results of this study also show that quantitative subtests predict the most learning success in the Master of Professional Psychology of UGM. Therefore, in the use of the PAPS test in the admission process in the Master of Professional Psychology of UGM, in addition, to see the total score, it is also seen whether the quantitative subtest score had a high score. Then, for the *Unit Pengembangan Alat Tes Psikodiagnostika* (UPAP) of the Faculty of Psychology UGM, to maintain the quality of the PAPS test both in terms of test content and test administration. For further researchers, it is expected to conduct research related to the PAPS predictive validity in other study programs because the results of this research can only be used for the Master of Professional Psychology of UGM.

References

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

Azwar, S. (1996). *Pengantar psikologi inteligensi*. Pustaka Pelajar.

Azwar, S. (2008). Kualitas Test Potensi Akademik versi 07A. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *12*(2), 231–250. https://doi.org/10.21831/pep.v12i2.1429

Azwar, S. (2015a). *Penyusunan skala psikologi*. Pustaka Pelajar.

Azwar, S. (2015b). *Reliabilitas dan validitas* (4th ed.). Pustaka Pelajar.

Azwar, S. (2016a). *Dasar-dasar psikometrika (II)*. Pustaka Pelajar.

Azwar, S. (2016b). *Konstruksi tes kemampuan kognitif*. Pustaka Pelajar.

Azwar, S., & Ancok, D. (2008). *Validitas prediktif diferensial pada ujian tulis UM UGM*. Universitas Gadjah Mada.

Azwar, S., & Koentjoro, K. (1985). *Validitas prediktif ujian masuk perguruan tinggi negeri tahun 1980 di Fakultas Psikologi Universitas Gadjah Mada*. Faculty of Psychology, Universitas Gadjah Mada. https://repository.ugm.ac.id/digitasi/download.php?file=4342_nurhayati_201307147_saifuddin azwar.pdf

Cohen, R. J., Swerdlik, M. E., & Sturman, E. D. (2013). *Psychological testing and assessment: An introduction to tests and measurement*. McGraw-Hill Education.

Furr, R. M., & Bacharach, V. R. (2014). *Psychometrics: An introduction*. Sage Publications.

Goldberg, E. L., & Alliger, G. M. (1992). Assessing the validity of the Gre for students in psychology: A validity generalization approach. *Educational and Psychological Measurement*, *52*(4), 1019–1027. https://doi.org/10.1177/0013164492052004026

Gregory, R. J. (2007). *Psychological testing: History, principles, and applications* (5th ed.). Allyn & Bacon.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate student selection and performance. *Psychological Bulletin*, *127*(1), 162–181. https://doi.org/10.1037/0033-2909.127.1.162

Kurpius, S. E. R., & Stafford, M. E. (2006). *Testing and measurement: A user-friendly guide*. Sage.

Master Program of Professional Psychology. (2013). *Panduan akademik 2013 program studi Magister Psikologi Profesi*. Faculty of Psychology Universitas Gadjah Mada.

Ministry of Education and Culture. (2015). *Perkembangan pendidikan tinggi tahun 1999/2000—2013/2014*. Center for Data and Statistics of Education and Culture of Republic of Indonesia.

Mupinga, E. E., & Mupinga, D. M. (2005). *Perceptions of international students toward GRE*. College Student Journal. http://link.galegroup.com/apps/doc/A133606109/AONE?sid=googlescholar

Nurhayati, N., & Widhiarso, W. (2017). *Tes Potensi Akademik Pascasarjana (PAPS) sebagai Pprediktor prestasi belajar pada mahasiswa magister psikologi UGM*. Thesis, Universitas Gadjah Mada, Yogyakarta.

Powers, D. E. (2004). Validity of Graduate Record Examinations (GRE) general test scores for admissions to colleges of veterinary medicine. *Journal of Applied Psychology*, *89*(2), 208–219. https://doi.org/10.1037/0021-9010.89.2.208

Pradipta, A. P., Kusumawardani, S. S., & Widyobroto, B. P. (2017). *Hubungan antara nilai Tes Potensi Akademik dengan prestasi akademik mahasiswa program magister di Universitas Gadjah Mada*. Thesis, Universitas Gadjah Mada, Yogyakarta.

Public Relations of UGM. (2007). *Meningkatkan jumlah mahasiswa S2 dan S3 melalui peningkatan jumlah beasiswa*. Universitas Gadjah Mada. https://ugm.ac.id/id/berita/1440-meningkatkan-jumlah-mahasiswa-s2-dan-s3-melalui-peningkatan-jumlah-beasiswa

Rudner, L. M. (1994). *Questions to ask when evaluating tests* (pp. 1–7). ERIC/AE Digest.

Sackett, P. R., Laczo, R. M., & Arvey, R. D. (2002). The effect of range restriction on estimates of criterion interrater reliability: Implications for validation research. *Personnel Psychology*, *55*(4), 807–825. https://doi.org/10.1111/j.1744-6570.2002.tb00130.x

Sternberg, R. J., & Williams, W. M. (1997). Does the Graduate Record Examination predict meaningful success in the graduate training of psychology? A case study. *American Psychologist*, *52*(6), 630–641. https://doi.org/10.1037/0003-066X.52.6.630

Stobart, G., & Eggen, T. (2012). High-stakes testing – Value, fairness and consequences. *Assessment in Education: Principles, Policy & Practice*, *19*(1), 1–6. https://doi.org/10.1080/0969594X.2012.639191

Thornell, J. G., & McCoy, A. (1985). The predictive validity of the Graduate Record Examinations for subgroups of students in different academic disciplines. *Educational and Psychological Measurement*, *45*(2), 415–419. https://doi.org/10.1177/001316448504500229

Top Universities. (2017). *QS World University Rankings 2018*. QS Top Universities. https://www.topuniversities.com/university-rankings/world-university-rankings/2018

Wang, W. (2013). *Testing the validity of GRE scores on predicting graduate performance for engineering students*. Thesis, University of Nebraska, Lincoln, NE.

Wao, J. O., Ries, R., Flood, I., Lavy, S., & Ozbek, M. E. (2016). Relationship between admission GRE scores and graduation GPA scores of construction management graduate students. *International Journal of Construction Education and Research*, *12*(1), 37–53. https://doi.org/10.1080/15578771.2015.1050562

Widhiarso, W. (2017). *Format baru tes PAPS (Seri E) – UPAP Psikologi UGM*. Unit for Psychodiagnostics Tools Development, Faculty of Psychology, Universitas Gadjah Mada. http://upap.psikologi.ugm.ac.id/format-baru-tes-paps-seri-e/

Jurnal Penelitian dan Evaluasi Pendidikan
Volume 23, No 2, December 2019 (129-143)
Online: http://journal.uny.ac.id/index.php/jpep

A DIAGNOSIS OF STUDENTS' ERRORS IN ANSWERING THE MATHEMATICS TEST IN SENIOR HIGH SCHOOL

Evana Gina Shantika
Universitas Nahdlatul Wathan Mataram

Edi Istiyono
Universitas Negeri Yogyakarta

Abstract

This study aims to reveal: (1) the characteristics of diagnostic test items used in mathematics subject of the first grade of senior high school (SHS); (2) the degree of errors from several types of students' errors in answering mathematics test; and (3) the dominant factor causing the students to make mistakes in answering mathematics test. This study used a quantitative approach involving the first graders of eight SHSs in Mataram as the population. The sample was collected by a proportionate random sampling technique, consisting of 350 students for preliminary field testing and 450 students for main field testing. The research instrument was a mathematics diagnostic test, questionnaire, and interview guidelines. The results show that: (1) the diagnostic test instrument meets the qualitative and quantitative content validity, proves empirically fit with Partial Credit Model (PCM), has reliability index of 0.92 (high category), and all items in the diagnostic test instrument are categorized in moderate difficulty; (2) misrepresentation becomes dominant errors if compared with misconceptions, the counting errors, and procedural errors. The errors dominant in the topic domain are Inequalities Linear System Two Variables if compared with Rational and Irrational Inequality One Variable, Equations and Inequalities of Absolute Value, and Linear Equations System Three Variables; and (3) most of the students do not understand how to solve the problem of inequality, determining factor quadratic equations, determining the members of the set completion, problems concerning graphs, and problems that require the capability of language interpretation into the mathematics model.

Keywords: *diagnostic errors, mathematics test, Partial Credit Model*

Permalink/DOI: http://dx.doi.org/10.21831/pep.v23i2.16370

Contact: *Evana Gina Shantika*
evanashantika@gmail.com
Faculty of Teacher Training and Education, Universitas Nahdlatul Wathan Mataram
Jl. Kaktus No. 1-3, Mataram, Nusa Tenggara Barat, Indonesia

Introduction

The measurement and assessment in education are essential, especially towards students. Measurement is an activity which is conducted to describe the characteristics of the object being measured, and it is stated in numbers. The Regulation of the Minister of Education and Culture of Republic of Indonesia No. 23 of 2016 regarding the educational assessment standard, which is used as the foundation in the assessment of students in elementary and middle school, explains that the purpose of learning outcome assessment is to evaluate the process of learning interest, and the remedial of students' learning outcome continuously.

The result of the assessment which has been conducted can become the determinant of the learning outcome, which is based on a proof of measurement result (Mardapi, 2005, p. 75). The assessment of students' learning outcomes is in the form of the test consisting of daily examination, middle semester examination, final semester examination, and national final examination. The government conducts national examination as a medium to equalize the assessment of students' abilities in all areas of Indonesia.

National examination is one kind of formative assessment. The result of students' learning outcome, on a large scale, aims to investigate the achievement of national education based on the national education standard (Mardapi, 2012, p. 223). All assessments which are conducted aim to improve the quality of teachers in teaching and the quality of students in the cognitive, affective, as well as psychomotor aspects. In consequence, the implementation of national examination is expected to be a consideration and indicator in achieving the goals.

Based on the data of the results of national examination in some state senior high schools in Mataram City, the scores of mathematics tests for the last five years are still below the standard. It is presented in Figure 1.

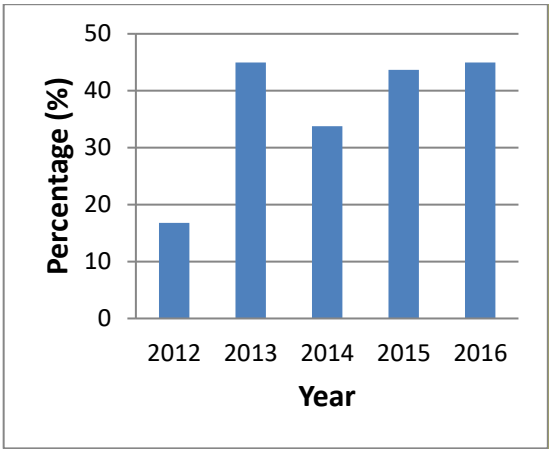

Figure 1. Scores of Mathematics National Examination of State Senior High Schools in Mataram City for the Last Five Years (N < 5.5)

Figure 1 shows that many students obtain the scores far below the national standard, which is 5.5. Based on the analysis of the questions by Isgiyanto (2011), some dominant errors found in studens' answer to the questions, which are categorized into conceptual errors, representation errors, procedural errors, and counting errors.

In realizing the expectation of the national examination, some reformations in the learning process are urgently needed. From the cognitive aspect, students' error in answering the questions were caused by their weakness in some aspects. Therefore, the reformation is expected to improve the result of the mathematics final examination and also can solve the students' weaknesses, especially in certain aspects.

One way to make the learning process more effective is by acknowledging the errors done by students in answering the questions. Acknowledging the errors done by students is one way to analyze the weaknesses in some aspects. Therefore, in the reformation of the learning, teachers should become more direct in emphasizing the students' weaknesses.

The students' errors in answering the questions can be defined as the response of students towards the question item, which is inappropriate with the expected response (Leighton & Gierl, 2007, p. 332). The error which possibly occurs can be identified

using the diagnostic method, especially in analyzing students' strength and weakness in answering questions (Samejima, 1995, p. 402). Diagnosis can be defined as a characteristic or phenomenon in identifying the cause and the decision made through the description or analysis (Embretson, 2007, p. 218). Through a diagnostic method, the information regarding the errors often done by students, the strength and weakness are obtained, and the factors which cause the errors can be acknowledged.

The diagnostic method can be implemented using a polytomous scoring. Polyomous scoring is selected in the diagnosis because it is an item response model that has more than two categories of possibility (Wells, Hambleton, & Purwono, 2008). The difference of polytomous scoring is in the item response where the dichotomous item response has two categories of score, namely the right answer (score 1) and the wrong answer (score 0) (Bond & Fox, 2007, p. 320). It becomes the essential reason why polytomous scoring is used in the diagnostic method.

The more popular model in the initial development of the item response theory is the Partial Credit Model (PCM), the extension of the Rasch model (Retnawati, 2016b, p. 149). In addition, research conducted by Wasis (2011) states that the "partial credit scoring" in physics can produce the estimation of the ability, which is more accurate than the other kinds of scoring, which based on the complexity of every option. It is supported by the research conducted by Isgiyanto (2011) regarding the polytomous scoring with the partial credit model in mathematics learning. The item response theory (IRT) model of PCM polytomous assumes that every question item has the same discrimination index (Embretson & Reise, 2000, p. 125). The use of a diagnostic method based on the polytomous scoring with PCM is expected to give a deeper contribution to diagnosing students' errors.

The consideration of the use of PCM as the extension of the Rasch model, which is a 1-PL model, is because it can utilize the smaller sample than the calibration of the polytomous data using 2-PL or 3-PL model. The characteristic of response in every item follows the PCM, especially in the difficulty level at one stage of the above category, which is not same from each other, so the amount of delta for a stage of the below category and the delta for the stage of the above category is not same to each other (Keeves & Masters, 1999, p. 95). PCM does not require the finishing stage of the test item, which should be ordered or not, and does not need to have the same difficulty level (De Ayala, 1993, p. 175). The test which employs PCM can give information about the difficulty level of students, which can be obtained through the errors they done at every different difficulty level.

Based on the aforementioned description, research to diagnose the students' error in answering mathematics questions is an essential thing, because this research can provide diagnostic information which is needed by teachers in conducting some improvements to obtain the better result for students, especially in mathematics learning. The diagnosis result can also provide information for students to analyze their mathematics skills and motivate them to improve their weak aspects.

Research Method

This study was descriptive-quantitative research based on the polytomous scoring with the Partial Credit Model (PCM) in analyzing the result of the diagnostic test. This research was conducted at eight state senior high schools in Mataram City at the end of the academic year of 2016/2017. The data sampling was conducted in May until June 2017.

The population of this research was 3,598 first graders of those eight senior high schools. The sample was taken through proportionate random sampling. The trial sample comprises of 350 students, and the measurement sample comprises of 450 respondents based on the formula of Krejcie and Morgan (Wagiran, 2013, p. 174).

The results of the diagnostic test were analyzed considering the scoring rubric using polytomous scoring with PCM. Then, it was described to draw a conclusion regarding the characteristic of the mathematic diagnostic test item in the first semester of the first-year students, the percentage of students' error level in various error types determined, and the materials causing errors. The results of questionnaire and interview were described to draw conclusions related to the factors causing students' errors in answering mathematics questions.

The data collection technique used is documentation to obtain data on students' national examination results in mathematics at Mataram City. It is also used to obtain the data of competence and material used in the design of the test. The second technique used is a diagnostic test to diagnose students' errors in answering the questions.

The first step in composing the diagnostic test is determining the aim of the test, that is, to diagnose students' errors in answering math questions in the first year of senior high school. The materials tested are the equation and inequality from the absolute value of the linear form of one variable, rational inequality and irrational one variable, the system of linear equations of three variables, and also the system of inequality of two variables (linear-squared and squared-squared). The materials had been adjusted to the K-13 curriculum used in some state high schools in Mataram.

The next stage is composing the test's outlines, containing the main material, core competence, basic competence, sub material, indicator, types of error, cognitive aspect, and question items. The indicator selected was the one suitable to the domain and ability aspect decided. The diagnostic test item is in the form of multiple choice with every question item having five alternative options with a score of 1-5. Scoring was adjusted to the error level. The distractors were composed in the alternative options through observation in some senior high schools at Mataram. The time given to finish the diagnostic test is 90 minutes.

Based on the outlines, there are 30 indicators in which every indicator is represented by one question. Those questions comprise of 10 question items from the domain of equation and inequality of absolute values in the linear form of one variable, eight items from the domain of rational inequality and irrational variable**,** four items from the domain of the linear equation system of three variables**,** and eight items from the domain of two-variable inequality system (linear-square and square-square).

The composing of the test questions was conducted through expert judgment by three experts consisting of one lecturer of mathematics education, one lecturer of educational research and evaluation, and one mathematics teacher of a senior high school in Yogyakarta. The expert judgment was conducted quantitatively and qualitatively. Quantitatively, every test item was scored and analyzed using the Aiken formula. The score of every test item in the validation sheet is between 1 until 4. The expert judgment was also conducted qualitatively in the form of a conclusion of the item with the criteria based on some aspects, such as material, construction, and the language used. The third data collection technique used in this research is a questionnaire and interview to obtain information regarding the factors that cause errors of students in answering the questions, seen from the aspect of competence needed by students in answering the questions.

Findings and Discussion

The Characteristic of Question Item

Validity and Reliability

The content validity is related to the rational analysis of the domain to be measured to find out instructional representation with the ability to be measured (Retnawati, 2016a, p. 158). The first-year semester one mathematical diagnostic test instrument fulfilled the content validity qualitatively with expert judgment and quantitatively obtained the Aiken index of 0.848 in the high catego-

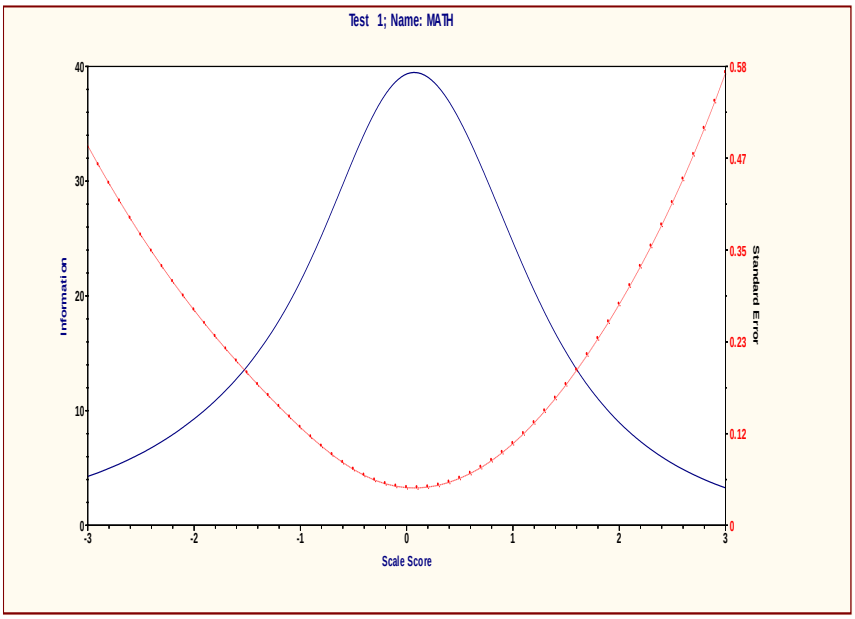

Figure 2. The Information Function and SEM

ry (Retnawati, 2016b, p. 19). The estimated function of the test information and the calculation of the standard error of measurement is shown in Figure 2.

Based on Figure 2, the information and SEM function intersect on a scale of -1.6 and +1.6. It means that the diagnostic test that has been prepared is suitable for test-takers in the ability range of -1.6 to +1.6. The estimated reliability results of 0.92 are classified as a high category.

Item Response Theory

The analysis in this study uses the IRT approach. Before an analysis using the IRT approach, there are several assumptions that must be met, namely unidimensional and local independence (Hambleton, Swaminathan, & Rogers, 1991, pp. 9–12; Mardapi, 2012, p. 201). The unidimensional assumption test is done by exploratory factor analysis. Before further factor analysis (Eigen Value and Scree Plot), the sample adequacy test is performed using Barlett's-Test as presented in Table 1.

Based on Table 1, the result for the KMO MSA diagnostic test instrument >0.5 is 0.895 with a significance value of the Bartlett's Test of Sphericity <0.5, 0.000. It shows that the number of the students' responses have met the requirements for factor analysis. The requirements for the factor analysis are the KMO MSA value >0.5 and the significance value of the Bartlett test <0.05 (Hair, Anderson, Tatham, & Black, 2006, p. 115). In addition, there are also things that need to be considered, namely the MSA value per item in the Anti-Image Correlation matrix. Based on the Anti-Image Correlation matrix, it is known that there is no item that has an MSA value <0.5, so there is no aborted item, and there is no need to re-test (Susetyo, 2015, p. 70).

Table 1. The Result of KMO MSA Test and Bartlett's Test of Sphericity

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.895
Bartlett's Test of Sphericity	Approx. Chi-Square	3255.986
	Df	435
	Sig.	.000

Unidimensional proof can be performed by looking at the Eigen value or scree plot that is formed based on the Eigen value. Furr states that a test is said to be unidimensional if components 1 and 2 in the scree plot have a considerable distance, as shown in Figure 3 (Furr & Bacharach, 2008, p. 74).

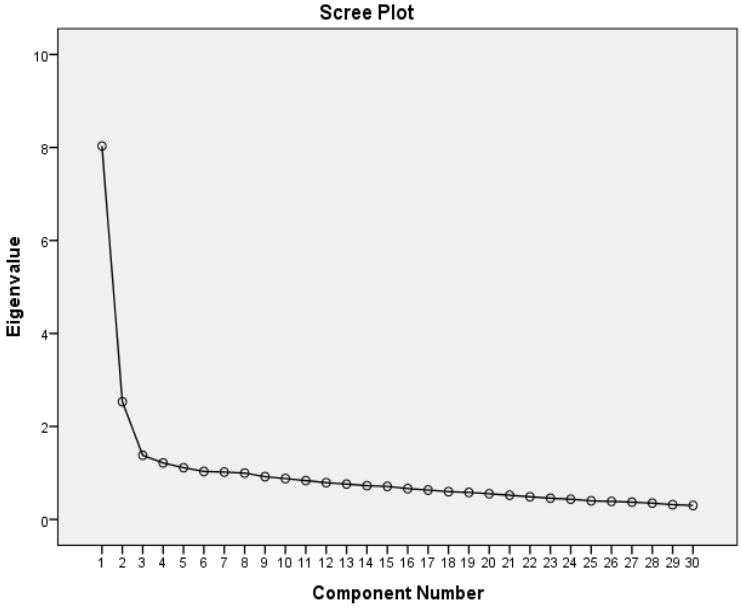

Figure 3. Scree Plot of Eigen Test Value of Mathematics Diagnostic Test

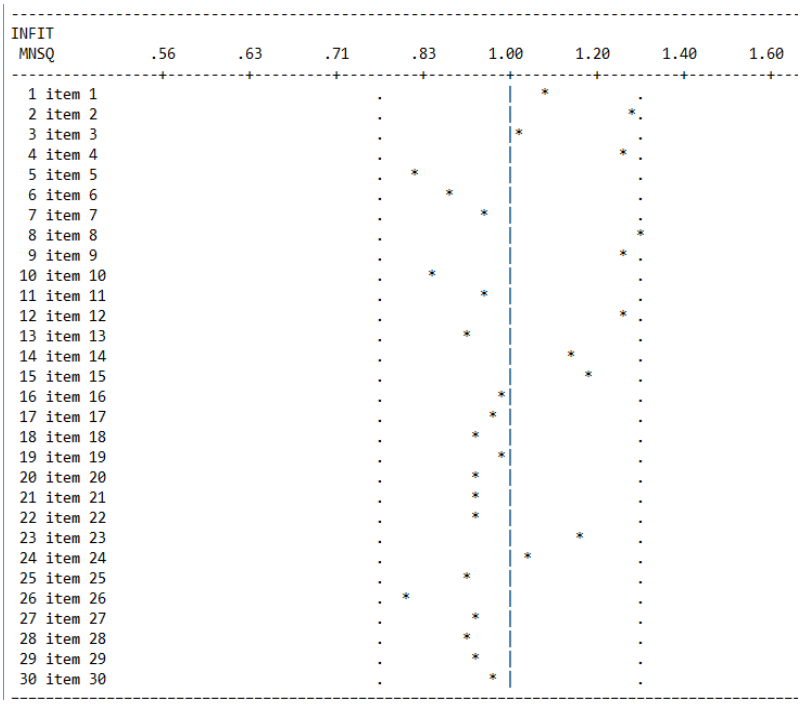

Figure 4. The Output Quest of Infit MNSQ Value in the Determination of Fit Model

Based on Figure 3, the scree plot shows that factor 1 looks steep because it has a great distance with factor 2, while factor 2 to factor 3 and the forth starts to ramps because it has a very close distance. In addition, the scree plot also shows that there is only one steepness, namely in factor 1 to factor 2. It corroborates the results which show that the diagnostic instrument of mathematics for the first year of high school is unidimensional.

The assumption of local independence is automatically proven after being proven by the unidimensionality of the test participant's response data. It means that the test participant's answer to an item has no relationship with other items in a test set (Retnawati, 2014, pp. 3–7).

Model Suitability (Goodness of Fit (GoF))

Determination of overall item fits the model using the Quest program (Adams & Khoo, 1996, p. 90), which is based on the magnitude of the average value of INFIT Mean of Square (INFIT MNSQ) and standard deviation or the average value of INFIT Mean of INFIT t. The Quest program output for fit model analysis is presented in Figure 4.

Based on Figure 4, it can be seen that for each MNSQ INFIT value, it is in the range of values of 0.77 to 1.30. Thus, it indicates that each item on the mathematical diagnostic test instrument for the first graders of senior high school fits the model.

The Level of Test Item Difficulty

Theoretically, the difficulty level index (b_i) lies between -2 and +2. Difficulty value approaches -2 identifies that the item is easier. Moreover, if the difficulty value approaches +2, it shows that the item is a more difficult item (Mardapi, 2012, p. 205). The results of the analysis of the item's level of difficulty are summarized in Table 2.

Table 2. The Summary of Analysis Result of Item Difficulty Level

Question Category	Item Number	Total	Exp.
Concept	1, 4, 7, 11, 14, 19, 23, 26	8	Medium
Procedure	2, 5, 8, 12, 15, 20, 24, 27	8	Medium
Counting	3, 6, 13, 16, 21, 25, 28	7	Medium
Representation	9, 10, 17, 18, 22, 29, 30	7	Medium
Total		**30**	

Table 2 shows that the difficulty level of each item on the diagnostic test instrument is in the moderate category. A good point of difficulty is located at the -2 interval b≤2 interval (Hambleton et al., 1991, p. 13).

Error Level in Every Error Type

Error Concept

Students who make a conceptual mistake means the students are wrong in answering questions about the concept categories that are included in scores 1, 2, 3, and 4. Items included in the concept category are items 1, 4, 7, 11, 14, 19, 23, and 26, which in answering, they only need an understanding of the concept. Conceptual understanding is the ability to use new ideas depending on how to connect to previous ideas and processes. In learning mathematics, students can connect between new material and material that was previously received (prerequisite) (Hasselbring, Lott, & Zydney, 2005). The distribution of the students' scores in answering the question of concept categories is presented in Figure 5.

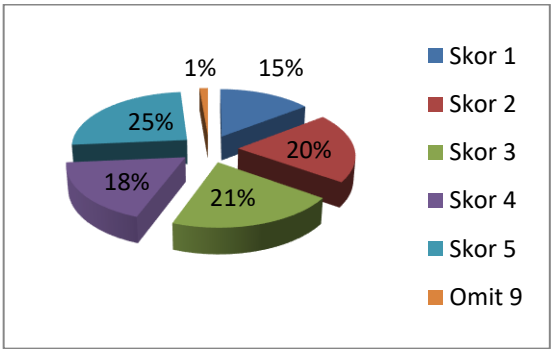

Figure 5. The Distribution of Score in the Conceptual Question

Based on Figure 5, it can be seen that students who made a conceptual error that is getting a score of 1, 2, 3, and 4, are 74%. Students who were correct in answering questions about the concept category or those included in the score of 5 are 25%, while 1% did not answer the questions.

Procedural Error

Students who make mistakes in answering items number 2, 5, 8, 12, 15, 20, 24, and 27 mean that they have made a procedural error, where these items are included in the procedure category, which in answering them, it requires proper understanding of procedures. These students are

wrong in answering the question of procedure categories in scores 1, 2, 3, and 4. This procedural error occurs because students do not know the procedures required to carry out operations accurately, even though students are able to identify the correct operation or sequence of operations (White, 2005, p. 15). The form of questions in the procedure category is to choose the right procedure or not the right one in solving a problem. The distribution of students' scores in answering the question of procedure categories is presented in Figure 6.

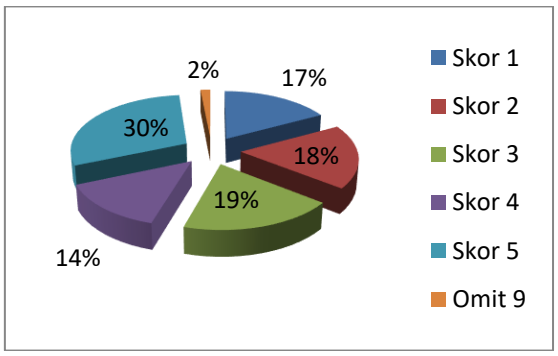

Figure 6. The Distribution of Score in Procedural Question

Figure 6 shows that students who made a procedural error that is included in scores 1, 2, 3, and 4 are 68%. Students who are correct in answering questions about the procedure category or who belong to a score of 5 are 30%, while 2% do not answer the questions.

Counting Error

Items included in the counting category are items number 3, 6, 13, 16, 21, 25, and 28 where the answer choice distractor in the calculation problem category is arranged by placing an error on the calculation error in the arithmetic operations of addition, subtraction, multiplication, and division. Indicators of difficulty in numeracy skills are difficulty in computing, difficulty in manipulating operations, and not rechecking the results of calculations (Retnawati, Pardi, & Prastowo, 1994, pp. 12–13). Students who make a numeric mistake means the students are wrong in answering the numeracy category questions in scores 1, 2, 3, and 4. The distribution of students' scores in answering numeracy category questions is presented in Figure 7.

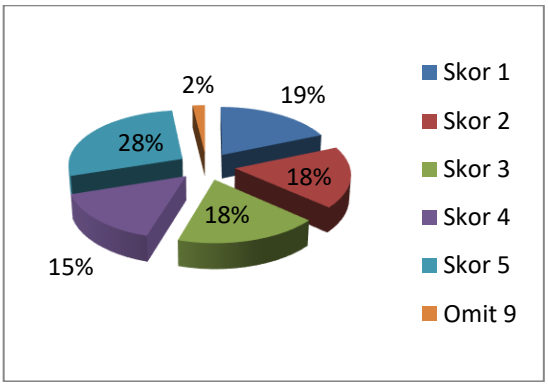

Figure 7. The Distribution of Score in Counting Questions

Based on Figure 7, students who make mistakes in counting are those included in scores 1, 2, 3, and 4 by 70%. Students who were correct in answering questions on the numeracy category or who included in the score of 5 were 28%, while 2% did not answer the questions.

Representation Error

The form of category representation questions is about a story or application problem. Indicators of weakness in language interpretation are difficulty in expressing everyday language in mathematical language, difficulty in interpreting graphs or tables into mathematical language, and difficulty in expressing mathematical language in everyday language (Retnawati et al., 1994, pp. 12–13). The items included in the category of representation are items 9, 10, 17, 18, 22, 29, and 30, which in answering, it requires the ability to represent the problem language into mathematical language. Students who make a misrepresentation means that students are wrong in answering questions about the categories of representtation, including in scores 1, 2, 3, and 4. The distribution of students' scores in answering questions about representation is presented in Figure 8.

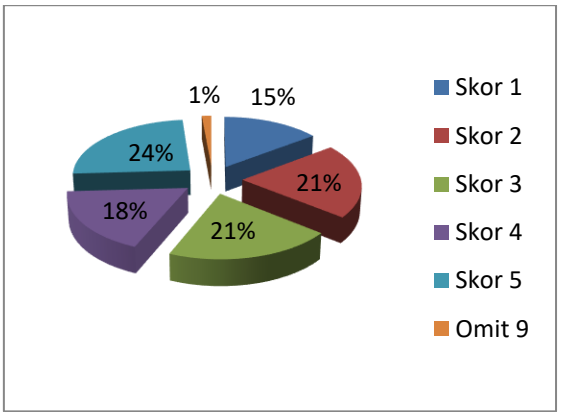

Figure 8. The Distribution of Score in Representation Questions

Figure 8 shows that students who made a misrepresentation were in scores 1, 2, 3, and 4 by 75%. Students who were correct in answering questions about the category of representation or included in the score of 5 were 24%, while 1% did not answer the questions. The comparison of error rates for each type of error can be seen in Figure 9.

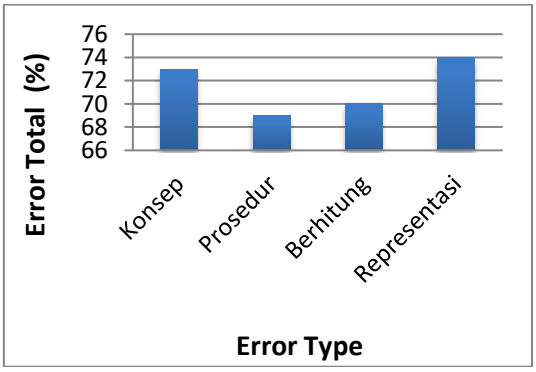

Figure 9. Error Level in Every Error Type

Based on Figure 9 and exposure to the level of errors in each type of error, it can be seen that the errors of students in answering math problems for the first year of high school in the first semester dominantly lie in the error of representation with the highest percentage of errors of 75% among the other percentage of errors. The order of error rates of the types of errors starting from the highest are the error of representation, then followed by concept errors, counting errors, and the smallest percentage is in procedural errors.

The Error Level in Every Material Tested

The Equation and Inequality of Absolute Value of One Linear Variable Form

The item questions which are included in the Equality and Inequality Domains Absolute Value Form of Linear One Variable are items number 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10. The distribution of students' scores in answering the questions about the Equation domain and the Inequality of Absolute Value of Linear Forms of One Variable is presented in Figure 10.

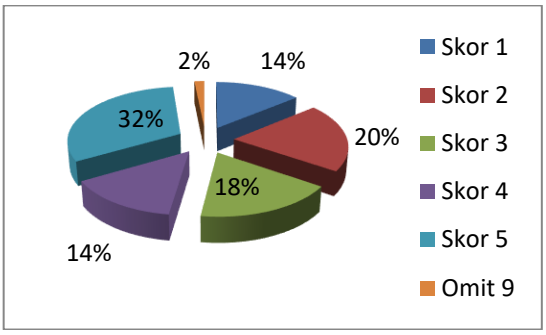

Figure 10. The Distribution of Score in the Questions of Equation and Absolute Inequality Value Domain

Based on Figure 10, it can be seen that students who make mistakes in answering questions about the Equation and Inequality Absolute Value Form of Linear One Variable, namely in scores 1, 2 3, and 4, are 66%. Students who are correct in answering questions about the domain of Equality and Inequality Absolute Value in the Form of Linear One Variable or included in the score of 5 are 32%. Meanwhile, 2% of them do not answer the questions.

The Rational and Irrational Inequality of One Variable

The question items which are included in the Rational and Irrational Inequality domain of One Variable are items number 11, 12, 13, 14, 15, 16, 17, and 18. The distribution of students' scores in answering questions about the Rational and Irrational Inequality Variable One variable is presented in Figure 11.

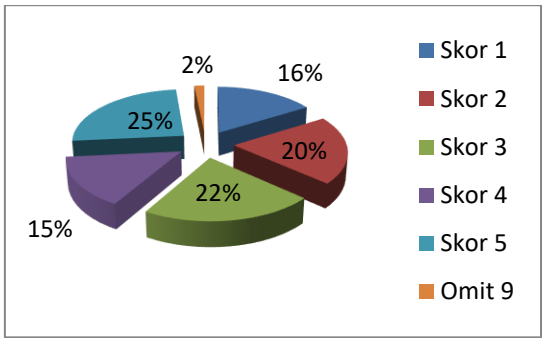

Figure 11. The Distribution of Score in the Questions of Rational and Irrational Inequality of One Variable

Figure 11 shows that students who made mistakes in answering questions about the domain of Rational and Irrational Inequality of One Variable, namely in scores 1, 2, 3, and 4 are 73%. Students who are correct in answering questions about the Rational and Irrational Inequality Variable Variables or those that score 5 are 25%, while 2% do not answer the questions.

Linear Equation System of Three Variable

The items included in the Three Variable Linear Equation System domain are items number 19, 20, 21, and 22. The distribution of students' scores in answering the problem domain of the Three Variable Linear Equation System is presented in Figure 12.

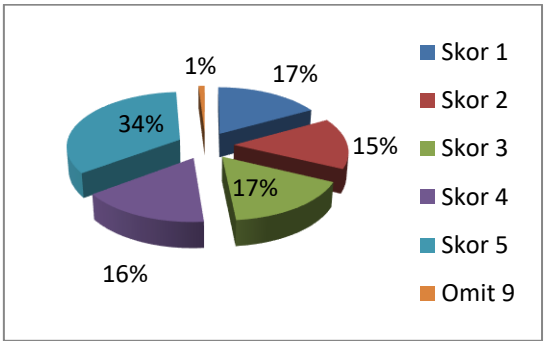

Figure 12. The Distribution of Score in the Questions of Linear Equation System of Three Variables

Based on Figure 12, students who make mistakes in answering the problem domain of the Three Variable Linear Equation System are in scores 1, 2, 3, and 4 by 65%. Students who are correct in answering questions about the domain of the Three Variable Linear Equation System or those included in the score of 5 are 34%, while 1% does not answer the questions. Research conducted by Blanco and Garrote (2007, p. 228) shows that students who are not able to master the basic concepts in arithmetic will result in these students not being able to solve linear equations.

The Inequality System of Two Variables (Linear-Square andSquare-Square)

The items included in the Two-Variable Inequality System domain (Linear-Squares and Squares-Squares) are items number 23, 24, 25, 26, 27, 28, 29, and 30. The distribution of students' scores in answering the Inequality System domain questions Two Variables (Linear-Squared and Squares-Squares) is presented in Figure 13.

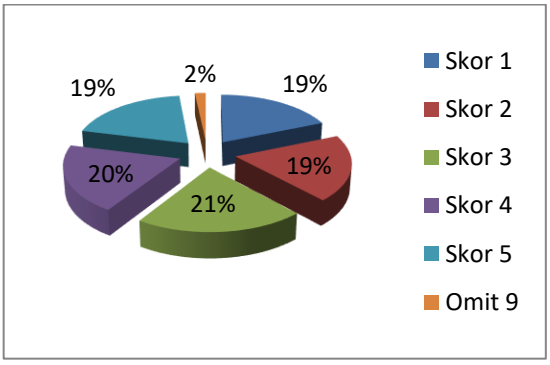

Figure 13. The Distribution of Score in the Questions of Inequality System of Two Variables

Figure 13 presents that students who made mistakes in answering the problem of the Inequality System of Two Variables (Linear-Square and Square-Squared), namely, in a score of 1, 2, 3, and 4 are 79%. Students who are correct in answering domain Inequality Systems Two Variables (Linear-Squares and Squares-Squares) or those included in the score of 5 are 19%, while 2% do not answer the questions. The comparison of error rates for each material is shown in Figure 14.

Based on Figure 14 and exposure to the error rate on each material, the errors of

students in answering math problems for the first graders of senior high school in the first semester dominantly lie in the material of Linear Inequality System Two Variables with the highest percentage of errors of 79% among the highest percentage of errors. The order of the error rate of the material being tested starting from the highest is the error in the Linear Inequality Two Variable System domain, then followed by an error in the Rational and Irrational Inequality domain of One Variable, the Equation domain and the Inequality of the Absolute Form of a Linear Variable, and the smallest percentage of errors is in the domain of the Three Variable Linear Equation System.

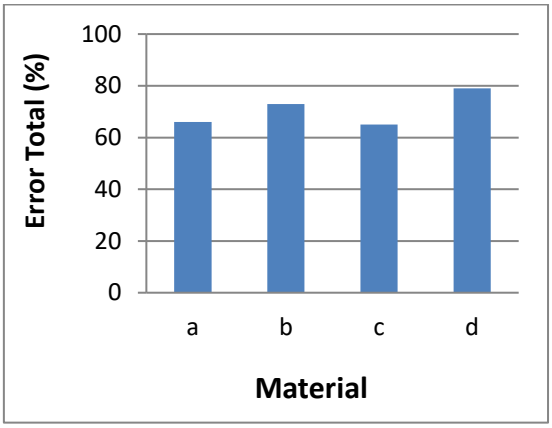

Figure 14. Error Level in Every Material

The Factors Causing Errors when Students Answer the Mathematics Questions

Interview

Through a structured interview conducted on eight mathematics-subject teachers in each of the state high schools in the city of Mataram, some information about the factors causing students to make mistakes in answering math problems is obtained. In terms of the material aspect, students still have difficulty in understanding the linear inequality two-variable system material, which contains material about linear-squared inequality systems and quadratic-inequality inequality systems. This material includes a material with a high degree of difficulty because understanding this material requires a deep understanding of quadratic equations and inequalities, both in its operation, determining factors, and graphs about quadratic equations and inequalities. It causes the linear inequality two-variable system is more difficult to understand. Moreover, in solving problems related to the system of linear inequality two variables, it requires a fairly long solution.

The results of research conducted by Kusaeri (2012) shows that the basic operations or basic concepts in arithmetic contribute to the mastery of students in inequality. Understanding the basic concepts of algebra and equality will influence students' mastery of concepts on the concept of inequality.

Previewed from the aspects of the prerequisite material, students are still weak in terms of working on inequality questions. There are still many students who do not understand the problems related to graphics, do not understand the material concepts related to the problems and concepts of prerequisite materials needed, and the students' weak ability in interpreting the intentions of the problem, especially in the case of a story or application.

From the aspect of implementing mathematics learning in the classroom, the teacher explained that students were not too tense and more active in learning mathematics. Many learning methods have been conducted, including learning using lecture and group discussion methods, sometimes giving students worksheets, as much as possible creating a relaxed learning atmosphere by joking occasionally so that students are not too tense in learning mathematics, increasing the number of questions and answers to students during learning so that learning activities are more active, and carrying out a remedial for students who did not pass the test.

Viewed from the aspect of the availability of materials/tools and media for learning mathematics, the teacher explained the availability was quite good, such as books for each student, and LCD installed in each class because they often use power

point media and videos in learning mathematics in class. Power point media are often used in learning mathematics, whereas media for teaching aids in learning mathematics depends on the material being taught whether or not it needs media.

The factors explained by the teachers who are considered to be able to inhibit the mathematics learning in the classroom include (1) students' lack of focus in learning, especially when there are math courses at the seventh and eighth sessions, (2) a large number of students in one class which makes it difficult for teachers to control students, (3) school activities that lead to reduced effective learning time of mathematics so that often, when approaching to the semester exam, many materials have not been completed yet, and at the end of the learning, the classroom is accelerated to be able to finish all the material before the semester ends, (4) there are still many students who are afraid to ask questions, (5) students' interest in learning math tends to be very lacking, and (6) the students' mindset that mathematics is difficult.

The teacher further explained the efforts made to overcome obstacles in learning mathematics in the classroom, including by (1) providing motivation to students to be more enthusiastic in learning mathematics, (2) forming learning groups by appointing one student as a group leader in each group responsible to the group members in order to understand the learning material, (3) giving punishment to students who make learning conditions uncomfortable, (4) providing additional lessons outside of school hours or sometimes asking for additional hours from other subject teachers whose subject material is almost complete, (5) occasionally joking with students to create the fun atmosphere of learning if students begin to look tired and lose the focus, (6) approaching the students outside of the class hours so that when learning in class they are not afraid to express opinions or ask questions directly to the teacher, and (7) changing the mindset about mathematics as a difficult subject.

In addition, if viewed from the aspect of the relationship between teacher and students, the teacher explained that the relationship is considered to be quite good. It is because the teacher realizes that fostering a good relationship between the teacher and students can facilitate learning activities in the classroom. One of the advantages is that when learning, students are not afraid to ask questions directly to the teacher.

Open-Ended Questionnaire

The conclusion obtained from the open-ended questionnaire is about the factors that can cause students to make mistakes in answering mathematics questions, in terms of the ability needed by students to answer the questions correctly. In the material of Equality and Inequality Absolute Value Form Linear One Variable, most students still do not understand the steps of solving absolute inequality values, determining the members of the settlement set, using inequality marks, the steps in drawing graphs of absolute value, and interpreting the language of the problem into mathematical models. There are still many students in the domain of Rational and Irrational Inequality One Variable who do not understand the use of inequality signs, how to determine the members of the settlement set, and how to interpret the language of questions into rational/irrational inequalities. Viewed from the material of the Three-Variable Linear Equation System, most students do not understand how to use the elimination method and interpret the problem language into a mathematical model.

The last material domain is the Inequality System of Two Variables (Linear-Square and Square-Squared), where, in this material, there are still many students who do not understand how to determine the quadratic equation factor, determine linear and quadratic inequality, and determine the intersection of quadratic charts. Mastery of the material in the first semester among the first graders of high school is very closely related to the mastery of the material at the

previous level of education. It is in line with Warren (2003) that students experience the process of thinking transition in junior high school that is the transfer of knowledge needed to solve arithmetic equations (operations involving numbers or numbers) to the knowledge needed to solve algebraic equations (operations involving variables) which bring changes in the shape of the object of study from numbers to variables, equations, and so on.

Conclusion

Based on the results of the study, several conclusions are drawn. First, concerning the characteristics of diagnostic test items, the mathematical diagnostic test instrument for the senior high school first graders in the first semester fulfills the content validity qualitatively and quantitatively, and it has been proven empirically fit with the Partial Credit Model (PCM) based on five categories of polytomous data. The instrument reliability estimate is 0.92 in the high category. Based on information and SEM functions, this diagnostic test instrument is very appropriate to be used for test participants in the ability range of -1.6 to +1.6. All items contained in the diagnostic test instrument used in the study have a moderate level of difficulty.

Second, it is concluded that the mistake of representation is the most dominant error made by students in answering math problems compared to concept errors, calculation errors, and also procedural errors. Viewed from the material aspect, it is concluded that the most dominant mistakes made by students are in the domain problems of the Linear Inequalities of Two Variables (Linear-Square and Square-Squared) when compared to the domain aspects of the Rational and Irrational Inequality Problems of One Variable, Inequality Value Absolute, and Three Variable Linear Equation System.

Third, the dominant factor causing students' errors in answering the math problems is in the aspect of the ability needed by students in answering questions. This aspect is elaborated as follows: most students still do not understand how to solve problems about inequality, determine the quadratic equation factor, determine the set members solving, solve problems with graphics, and solve problems that require the ability to interpret the language of problems into mathematical models.

References

Adams, R. J., & Khoo, S.-T. (1996). *Acer quest version 2.1*. Camberwell, Victoria: Australian Council for Educational Research.

Blanco, L. J., & Garrote, M. (2007). Difficulties in learning inequalities in students of the first year of pre-university education in Spain. *EURASIA Journal of Mathematics, Science and Technology Education*, *3*(3), 221–229. https://doi.org/10.12973/ejmste/75401

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

De Ayala, R. J. (1993). Methods, plainly speaking: An introduction to polytomous item response theory models. *Measurement and Evaluation in Counseling and Development*, *25*(4), 172–189.

Embretson, S. E. (2007). Construct validity: A universal validity system or just another test evaluation procedure? *Educational Researcher*, *36*(8), 449–455. https://doi.org/10.3102/0013189X07311600

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologist*. London: Lawrence Erlbaum Associates.

Furr, R. M., & Bacharach, V. R. (2008). *Psychometrics: An introduction*. Thousand Oaks, CA: Sage Publications.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2006). *Multivariate*

data analysis (6th ed.). Upper Saddle River, NJ: Prentice-Hall.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications.

Hasselbring, T. S., Lott, A. C., & Zydney, J. M. (2005). Technology-supported math instruction for students with disabilities: Two decades of research and development. Retrieved June 10, 2017, from LD Online website: http://www.ldonline.org/article/6291/

Isgiyanto, A. (2011). Diagnosis kesalahan siswa berbasis penskoran politomus model partial credit pada matematika. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *15*(2), 308–325. Retrieved from https://journal.uny.ac.id/index.php/jpep/article/view/1099/1151

Keeves, J. P., & Masters, G. N. (1999). Partial credit model. In G. N. Masters & J. P. Kevees (Eds.), *Advances in measurement in educational research and assesment*. Amsterdam: Pergamon-Elsevier Science.

Kusaeri, K. (2012). *Pengembangan tes diagnostik dengan menggunakan model DINA untuk mendapatkan informasi salah konsepsi dalam aljabar*. Doctoral dissertation, Universitas negeri Yogyakarta, Yogyakarta.

Leighton, J. P., & Gierl, M. J. (2007). Defining and evaluating models of cognition used in educational measurement to make inferences about examinees' thinking processes. *Educational Measurement: Issues and Practice*, *26*(2), 3–16. https://doi.org/10.1111/j.1745-3992.2007.00090.x

Mardapi, D. (2005). Pengembangan sistem penilaian berbasis kompetensi. In Himpunan Evaluasi Pendidikan Indonesia (Ed.), *Rekayasa sistem penilaian dalam rangka meningkatkan kualitas pendidikan* (pp. 71–85). Yogyakarta: Himpunan Evaluasi Pendidikan Indonesia.

Mardapi, D. (2012). *Pengukuran, penilaian, dan evaluasi pendidikan*. Yogyakarta: Nuha Medika.

Regulation of the Minister of Education and Culture of Republic of Indonesia No. 23 of 2016 on Educational Assessment Standard, (2016).

Retnawati, H. (2014). *Teori respons butir dan penerapannya: Untuk peneliti, praktisi pengukuran dan pengujian, mahasiswa pascasarjana*. Yogyakarta: Nuha Medika.

Retnawati, H. (2016a). Proving content validity of self-regulated learning scale (The comparison of Aiken index and expanded Gregory index). *REiD (Research and Evaluation in Education)*, *2*(2), 155–164. https://doi.org/10.21831/reid.v2i2.11029

Retnawati, H. (2016b). *Validitas reliabilitas & karakteristik butir (Panduan untuk peneliti, mahasiswa, dan psikometrian)*. Yogyakarta: Nuha Medika.

Retnawati, H., Pardi, P., & Prastowo, R. (1994). *Identifikasi kesulitan calon guru matematika untuk menyelesaikan persoalan matematika DIY*. Yogyakarta: Universitas Negeri Yogyakarta.

Samejima, F. (1995). A cognitive method using latent trait models: Competency space approach and its relationship with Dibello and Stout's unified cognitivepsychometric diagnosis model. In P. D. Nichols, S. F. Chipman, & R. L. Brennan (Eds.), *Cognitively diagnostic assessment* (pp. 391–410). Mahwah, NJ: Lawrence Erlbaum Associates.

Susetyo, B. (2015). *Prosedur penyusunan dan teknik analisis tes*. Bandung: Refika Aditama.

Wagiran, W. (2013). *Metodologi penelitian pendidikan: Teori dan imlementasi*. Yogyakarta: Budi Utama.

Warren, E. (2003). The role of arithmetic structure in the transition from arithmetic to algebra. *Mathematics Education Research Journal*, *15*(2), 122–137. https://doi.org/10.1007/BF03217374

Wasis, W. (2011). *Penskoran model partial credit pada item multiple true false bidang fisika*. Doctoral dissertation, Universitas Negeri Yogyakarta, Yogyakarta.

Wells, C. S., Hambleton, R. K., & Purwono, U. (2008). Polytomous response IRT models and applications. *Pelatihan Asesmen Pendidikan Dan Psikologi (Psikometri)*, pp. 2–10. Yogyakarta: Program Pascasarjana Universitas Negeri Yogyakarta.

White, A. L. (2005). Active mathematics in classrooms: Finding out why children make mistakes - and then doing something to help them. *Square One*, *4*, 15–19.

Jurnal Penelitian dan Evaluasi Pendidikan
Volume 23, No 2, December 2019 (144-155)
Online: http://journal.uny.ac.id/index.php/jpep

A DIAGNOSIS OF DIFFICULTIES IN ANSWERING QUESTIONS OF CIRCLE MATERIAL ON JUNIOR HIGH SCHOOL STUDENTS

Minhajjul Abidin
Universitas Negeri Yogyakarta

Heri Retnawati
Universitas Negeri Yogyakarta

Abstract

This research is aimed at describing (1) the characteristics of items about a diagnostic test of circle material in mathematics, (2) how significant the percentage of students' types of errors in answering the questions, and (3) the diagnosis of students' difficulties in answering the questions based on DINA model. This research is quantitative descriptive research involving eighth graders of junior high school in East Lombok regency as the population. The sample was chosen by a proportionate random sampling technique, consisting of 105 students for preliminary field testing and 416 students for main field testing. The instrument of this study was a diagnostic test using a four-option multiple-choice format. Data on students' responses were analyzed using the R program with CDM (Cognitive Diagnostic Model) DINA model, which requires underlying attributes for each item. The results show that: (1) the diagnostic test instrument had met the qualitative and quantitative content validity; (2) the percentage of students' answers retrieved from the conceptual error is 18.47%, 9.99% is interpretation error, the procedural error is 7.80%, and counting error is 14.57%; (3) based on the results of the analysis with DINA model, students' error in answering the questions of the circle material in mathematics are dominantly caused by students' lack of mastery on solving problems associated to the circumference of a circle, that is, (A28) the ability to calculate the length of a path, and (A29) the ability to calculate many rounds on wheels.

Keywords: *diagnosis of difficulties, DINA model, attributes*

Permalink/DOI: http://dx.doi.org/10.21831/pep.v23i2.16454

Contact: *Minhajjul Abidin*
hajjulabidin@gmail.com
Department of Educational Research and Evaluation, Graduate School,
Universitas Negeri Yogyakarta
Jl. Colombo No. 1, Karangmalang, Depok, Sleman, Yogyakarta 55281, Indonesia

Introduction

Mathematics is an obligatory subject to be taught at school. Mathematics learning is expected to be able to equip students with the skill to face every mathematics problem in daily life. Mathematics learning in the level of junior high school is conducted based on the curriculum. According to the Regulation of Minister of National Education, No. 22 of 2006, the purpose of mathematics learning in school is to facilitate students' understanding on mathematical concept, explaining the correlation between concept and the application of concept or algorithm flexibly, accurately, and efficiently in solving problems, using reasoning in pattern and feature, doing some mathematical manipulation in making a generalization, composing proof, or explaining ideas and mathematical statements.

Isgiyanto (2011) states that in the level of junior high school, mathematics has a function to develop some skills, such as counting, measuring, formulating, and using mathematics formula needed in daily life through the giving of some learning materials, such as number, algebra, geometry and measurement, and also statistics and chance. Students' mastery of mathematics can be seen from their accuracy in answering questions. Students are expected to solve mathematics problems contained in the questions.

In measuring the students' learning outcome and level of understanding regarding the subject studied, especially mathematics, teachers frequently use a test. The test is in the form of questions which requires the right answer from several options, and its implementation is conducted systematically to have the quantification of students' skill (Kartowagiran, 2013, p. 16). The test is composed based on some indicators which aim to observe the material mastery of students, especially in mathematics. Abadyo and Bastari (2015) state that the purpose of mathematics test is to access participants' ability to transfer qualitative reasoning and problem-solving skill from one context to another context.

Information towards students' learning outcomes can be investigated by the giving of assignments, exercises, and tests, which consist of daily examination, middle semester examination, final semester examination, and national examination. The government conducts national examination as a way to equalize the assessment of the ability of all students in Indonesia. National examination is a kind of formative assessment, the assessment of students' learning outcomes on a broad scale, which aims to analyze the achievement of national education based on the national standard (Mardapi, 2012, p. 223).

In practice, students' understanding of the learning material is still low. The problem in mathematics learning is that the competence mastered by students in mathematics learning is still low, and they tend to consider mathematics as an abstract and difficult subject to understand. Westwood (2000, p. 21) states that the difficulties found in mathematics learning are caused by their weakness, especially in understanding the symbolic abstract.

One of the topics studied in the material about the circle is the circumference and area of a circle. In fact, the topic is not difficult when it is presented in a direct question. Yet, it turns to be complicated when it is provided in the form of a story because it requires a deeper understanding. Besides, errors are also found in the calculation and when entering the formula. Besides those mistakes, it is possible that other errors affect students' learning outcomes.

Based on the observation and interview conducted with students, some students still have difficulty in understanding the basic concept of mathematics, such as problem-solving procedures, they still carry out incorrect calculations, and many of them still obtain scores below the minimum completeness standards in certain tests such as the results of the daily test, midterm test, and national examination. Based on the results of the national examination, the score of mathematics is still the lowest

score obtained by students. The average score of mathematics at the junior high school level based on the 2014/2015 national examination results in East Lombok Regency is 49.09. The average mathematics score is lower than the average value of other subjects such as Indonesian Language 63.47, English 55.33, and Science 54.18. These results indicate the condition that students have difficulty in answering mathematics questions.

Mulyadi (2010, p. 9) states that to mark individuals who have difficulty in learning, then a benchmark is needed to determine the symptoms of learning difficulties themselves. Isgiyanto (2011) states that the information on completeness attributes achieved by students can be obtained by examining items and student response items. Westwood (2008, p. 1) explains that learning difficulties are referring to obstacles that limit access to participation and results in a learning plan. In line with this opinion, Suwarto (2013, p. 87) insists that the difficulties experienced by students mean a failure in achieving learning goals, characterized by low learning achievement (the score obtained is less than 75). Based on the aforementioned definition, it is clear that difficulties are all forms of obstacles or deficiencies experienced by students in mastering subject matter, which makes them unable to reach the level of mastery needed as a prerequisite for learning further materials.

The characteristics of students who have difficulty in learning mathematics, according to Runtukahu and Kandou (2014, p. 49), include having difficulties in understanding the concept of spatial relationships, the concept of direction and time, visual concepts - spatial, the concept of arithmetic, the concept of symbols, the concept of geometry, the concepts of language and writing, and the concept of prerequisites. Difficulties in learning mathematics are students' difficulties that can be expressed from the pattern of mistakes or errors made by students in solving problems.

The students' difficulties in answering questions can be determined by a diagnostic method. Suwarto (2013, p. 116) states that diagnostic in education is a broad concept that involves the identification of students' strengths and weaknesses. A diagnostic is conducted to assist teachers in determining where teaching and learning processes are or have not been mastered by students. Thus, a diagnosis is an attempt to study the condition of an individual so that it can be classified into certain groups, namely, the group of those who have mastered or of those who have not mastered the given subject matter. In order to find out the difficulties, a diagnostic test can be used to obtain information on students' weaknesses in solving mathematics questions.

Hughes (1986) states that a diagnostic test can be used to determine the strength and weaknesses of students in learning, with the aim that teachers can determine the model and learning strategies that need to be implemented in the future. Mardapi (2012, p. 89) states that a diagnostic test can also be used to find out the learning difficulties faced by students. It is in line with Kusaeri (2012) who states that diagnostic test has two main functions, namely, identifying problems or errors experienced by students, and planning the follow-up in the form of efforts to solve the problem or error that is identified. Thus, a diagnostic test is a test used to find out the weaknesses of students in answering the questions and finding solutions to overcome them.

The diagnostic test used in this study is multiple choice. Osterlind (1998, p. 163) states that multiple-choice forms have the potential to open high-level thinking processes and can provide diagnostic information if the items are carefully constructed. Therefore, it is necessary to make a distinction between distractors on multiple-choice items. Deceivers must be able to provide information and conclusions about what is mastered. Thus, the multiple-choice test form is appropriate, and many previous researchers have used it to diagnose the difficulties in mathematics learning.

This study uses the DINA model. Further, de la Torre (2009) says that the DINA model assigns each student in a latent class that shows the attributes of the students mastering and, thus, provides the cognitive profile of students according to the attributes assessed by the test. In order to use the DINA model, it is necessary to determine the attributes of the attributes in the item. According to Tatsuoka (2009), attributes are knowledge and cognitive abilities. The binary attribute pattern expresses the mastery or non-mastery of the knowledge states, and latent knowledge states attributes.

Furthermore, these attributes are used to compile matrix-Q. Chen and de la Torre (2014) state that the matrix-Q required for each assessment will provide diagnostic information. The matrix-Q plays an important role because it provides attribute specifications for items. Based on the description and characteristics possessed by the DINA model, this research is used to diagnose the location of mathematic learning difficulties in the material of circle in the students of Junior High School at East Lombok Regency. Thus, the purpose of this study is to obtain information on (1) the validity and reliability of the mathematical diagnostic test questions regarding the material of circle, (2) the characteristics of mathematical diagnostic test items regarding the material of circle, (3) the difficulties experienced by students in answering mathematical problems tested, and (4) the description of the diagnosis of the students' difficulty in answering questions based on the DINA model.

Research Method

This research was a quantitative-descriptive study because it describes the learning difficulties in mathematics, in order to obtain strong and complementary conclusions about the difficulties in learning mathematics. This research was conducted at some junior high schools in the Regency of East Lombok. Data collection was conducted in May and July 2017. The sampling technique used in this study was proportionate stratified random sampling. Various considerations were used as the basis for choosing the schools, including school distribution, variations in the level of students' ability, and the schools' readiness. Based on these considerations, six schools were then chosen. The number of samples used in this study was 416 students of class VIII.

The research instrument was in the form of a multiple-choice test with four answer choices, consisting of 30 items in circle material. The instrument trial was conducted at a junior high school with a sample consisting of 105 eighth graders. A trial was conducted to determine the characteristics of the items on the test.

The analysis used in this study was the DINA model. Analysis of the DINA model requires a decrease in the attributes of each item. Attributes for each item consist of content attributes and process skill attributes. Thus, an example of a 3x3 ordo-Q matrix is presented as follows.

$$Q_{3x3} = \begin{array}{c} \\ l_1 \\ l_2 \\ l_3 \end{array} \begin{array}{c} \begin{array}{ccc} A1 & A2 & A3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \end{array}$$

In the matrix-Q modeled (Tatsuoka, 2009), attribute 1 must be mastered by students in order to complete item 1, attribute 3 is needed to be able to complete item 2, and attribute 2 and attribute 3 are needed to be able to complete item 3. Thus, item 1 is used to measure attribute 1, item 2 is used to measure attribute 3, and item 3 is used to measure attributes 2 and 3, so all attributes must be mastered by students. These attributes are process attributes.

The results of the students' answers were further analyzed using three computer software, namely SPSS, MG Bilog, and the R program. The SPSS program was used for the analysis of construct factors, MG Bilog was used to observe the characteristics of the test items used, and DINA model using the R program with CDM packages was used to analyze the diagnosis of learning difficulties in mathematics.

Findings and Discussion

The Analysis of Question Items Characteristics

Validity and Reliability

The content validity is related to the rational analysis of the domain to be measured in order to find out the instructional representation with the ability to be measured (Retnawati, 2016, p. 158). The diagnostic test instrument fulfills the content validity qualitatively with the expert judgment, and quantitatively obtains the Aiken index of 0.78 in the category of medium (Retnawati, 2016, p. 19). Reliability was analyzed using SPSS. The results of the analysis show that, based on the test results, the reliability coefficient obtained with the Cronbach Alpha formula of 0.809 means that the test instrument used is reliable to measure students' mathematical abilities on the material being tested.

Item Response Theory

The empirical analysis of items in this study used the item response theory (IRT) approach with the help of Bilog-MG software. Before analyzing the items, first, the model compatibility test was conducted. The aim was to see a model that matched the characteristics of students' answers. In order to see a suitable model, the decision criteria probability value must be > 0.05. Based on the Grain Response Theory approach, a suitable model for 30 items is the 2-PL model. Of the 30 items, 27 are good because they meet the criteria of the 2-PL model. From the results of the analysis, it shows that the difference in power (a) of the whole test items ranges from 0 to +1, included in the good category. The level of difficulty (b) shows that there are three items, namely 3, 14, and 30, that do not meet the requirements. All three items have a difficulty level which is not good, namely item 3 is classified as an easy item (b <-2) with a value of -2.003, and items 14 and 30 are classified as difficult items (b_i + 2) with values of 2.084 and 2.107, respectively. Based on the information function curve, the circle material diagnostic test is most suitable for students with a range of abilities between -1.4 to +2.8.

Percentage of Students' Error

The percentage of students' error was observed based on their responses/answers. Errors found in students' answers occur because of the difficulties experienced by students when answering items. Triana (2016), in her research, states that errors in answering questions are an indication of difficulties in solving the problem being tested. Students' difficulty in answering items occurs because they do not master all abilities needed. In this study, there are four types of errors made by students, namely concept error, language interpretation error, procedural error, and numeracy error.

Based on students' answers, it is found that the average students' error based on their answers, is 18.47% of concept errors, 9.99% of interpretation errors, 7.80% of procedure errors, and 14.57% of calculation errors. Based on the aforementioned percentage of errors, the most dominant error is concept errors, which still occur in students when answering questions about the circle material. Those results are supported by research conducted by Isgiyanto (2011), which states that many junior high school students experience misconceptions in answering items about geometry material.

The student errors are related to the slip and guessing parameters produced by the DINA model. Then, the percentage of wrong answers from students can occur because of slips and guessing. Kusaeri, Suryanto, and Kumaidi (2012) in their study state that an item is said to be easy if it has a high guessing parameter and a low slip, an item is difficult if it has a low guessing parameter and a high slip, and an item is medium if it has a low guessing and slip parameter. The highest slip obtained on item number 20 is 0.996, which means students slip when choosing answers on item number 20. Items with a slip value equal to zero are item number 4. Items on which

slip does not occur can be interpreted that there are no students who are slipped in answering the item.

In the DINA model, a correct answer from a student cannot be interpreted as a sign that the student is not experiencing difficulties. To see whether the correct answer given by students is because they master the intended attributes or not, the guessing parameters can become the key. The highest guessing value on item 8 is 0.366. The higher the value of the slip parameter will reduce the power of different items, the smaller difference in power of items will result in the item not being able to differentiate optimally between the students who are capable and those who are incapable to master the intended attributes. It also applies to parameter guessing. The higher the guessing of an item about the power index, the more different the item will decrease. Thus, the smaller the slip and guessing parameters of the items, the greater the different grain power index and the greater the different grain power index, the better the item distinguishes the ability between students who are capable and students who are incapable of answering the questions.

The Diagnosis of Students' Difficulty

The analysis is used to see the location of students' difficulties in answering questions about the circle material by using the DINA model analysis based on the errors on students' answer. It is in accordance with the opinion of Ozaki (2015) in his research that the DINA model is used to obtain information about students' abilities. The first step is to reduce the ability attribute on each item to identify the attributes that underlie the item by designing or constructing questions from the start for diagnostic purposes in accordance with one of the approaches in the research of Ravand and Robitzsch (2015). The resulting attributes are 43 attributes which consist of nine content attributes and 36 attributes of process skills. The content attributes and process skill attributes are then used to compile the matrix-Q column. Matrix-Q formed by items and some attributes on the test is eight matrixes. Items 1 and 2 form matrix-Q1 with two attributes. Items 4 to 8 form matrix-Q2 with four attributes. Items 10 to 13 form matrix-Q3 with four attributes. Items 15 to 17 form matrix-Q4 with three attributes. Items 18 to 21 form the matrix-Q5 with three attributes. Items 22 to 24 form matrix-Q6 with three attributes. Items 25 to 26 form matrix-Q7 with two attributes and Items 27 to 29 form matrix-Q8 with three attributes.

The analysis in this study is consistent with a research conducted by Retnawati (2017), which looks at the location of difficulties based on mastery of attributes in the matrix-Q. Mistakes made by students can be seen from the mastery of attributes. It can be seen based on the results of the analysis of students' answers using the matrix-Q. Besides, the results of this analysis support the research conducted by Wulansari (2014) in her research looking at students' mistakes in solving UN questions based on the matrix-Q, and the research analysis conducted by Triana (2016) who analyzed the location of students' difficulties with the dominant percentage of latent classes not controlled by students. Therefore, this study views the location of student difficulties based on the difficulty of each attribute and the dominant percentage of latent classes results from the matrix-Q analysis with the DINA model.

Based on the results of the analysis using the DINA model with the help of the R program, the results obtained by the difficulties of students in answering the circle material questions are as follows. Item number 1 through number 2 has the compiler attributes of matrix-Q1. The ordo of matrix- Q1 is 2 x 2, which means that there are three attributes in the matrix-Q1. Because there are two attributes, then there are likely four latent classes. The most dominant percentage is based on the likelihood of latent class. Hence, the items and attributes that make up the matrix-Q1 are clearly presented in Table 1.

Table 1. Attribute and Item Which Compose Matrix-Q_1

Item	Attributes	
	A9	A11
1	1	0
2	0	1

Based on Table 1, the attributes in the matrix-Q1 are (A9) the ability to define the elements and parts of a circle and (A11) the ability to determine the parts of a circle. The most dominant attribute not mastered by students is attribute A9, with a percentage of 54.2%, which means that attribute A9 becomes the hardest attribute controlled by students. Therefore, it indicates that students still have difficulty mastering the ability to determine the circle.

The dominant percentage based on latent classes in items 1 to 2 presents that students in the latent class 0 0 illustrate that they have not mastered all abilities, while the other latent classes have varying attribute completeness, incomplete incompleteness that can be caused by slip factors. It can be seen from the large percentage of the slip in item 1 of 15.1%. It indicates that students still have difficulty mastering the ability to determine the elements and parts of a circle.

Items 4 through 8 have the matrix-Q2 compiler attributes. The order of the matrix-Q2 is 5 x 4, meaning that there are four attributes in the matrix-Q2. Because there are four attributes, there are probably 16 latent classes. The items and attributes that make up the matrix-Q2 are presented in Table 2.

Table 2. Attribute and Item Which Compose Matrix-Q_2

Item	Attribute			
	A13	A14	A15	A16
4	1	0	0	0
5	0	1	0	0
6	0	0	1	1
7	0	0	0	1
8	0	0	1	1

Based on Table 2, the four attributes that make up the matrix-Q2 are (A13) the ability to calculate the circumference of a circle, (A14) the ability to calculate the area of a circle, (A15) the ability to determine the length of a circle's radius, and (A16) the ability to determine the length of a circle's diameter. In each latent class profile, the most dominant attribute not mastered by students is the A14 attribute, with a percentage of 34.8%, which means that the A14 attribute becomes the hardest attribute controlled by students.

Based on the latent class, the dominant percentage in items 4 to 8, it is found that students in the latent class 1 0 1 1 and 0 0 0 1, illustrate that students have not mastered the ability attributes, while the other latent classes have varying attribute completeness, incomplete incompleteness which varies bias caused by slip factors. It can be seen from the magnitude of the percentage of the slip in point 7 of 66%. Based on the location of the difficulties of each latent class in items 4 and 8, it indicates that students still have difficulty mastering the ability to calculate the area of a circle.

Items 10 through 13 have the matrix-Q3 compilation attributes. The order of matrix-Q3 is 4 x 4, meaning that there are four attributes in matrix-Q3. Because there are four attributes, there are probably six latent classes. The most dominant percentage is based on the possibility of latent class. The items and attributes that make up the matrix-Q3 are presented in Table 3.

Table 3. Attribute and Item Which Compose Matrix-Q_3

Item	Attribute			
	A18	A24	A25	A26
10	1	0	0	0
11	0	0	1	0
12	0	1	1	0
13	0	0	0	1

Based on Table 3, the four attributes of the matrix-Q3 compiler are (A18) ability to calculate half-circle outer, (A24) ability to calculate the area of a triangle, (A25) ability to calculate the shaded area, and (A26) ability to calculate the perimeter of the shaded

area. The most dominant attribute not mastered by students is attribute A26, with a percentage of 31.5%, which means that attribute A26 becomes the hardest attribute not mastered by students.

The dominant percentage based on latent classes in items 10 to 13, obtained that students in the latent class 0 0 0 0, 0 1 0 0 and 1 1 1 0, illustrate that students have not mastered the ability attributes, while the other latent classes have varying completeness attributes, slip factors can cause incomplete incompatibilities that vary. It can be seen from the large percentage of slips in point 13 of 43.8%. Based on the location of the difficulties of each latent class in items 10 and 13, it indicates that students still have difficulty mastering the ability to calculate the length of the cross and the ability to calculate the circumference of a half-circle.

Items number 15 to number 17 have the compositions of matrix-Q4. The order of matrix-Q4 is 3 x 3, meaning that there are three attributes in matrix-Q4. Because there are three attributes, there are probably eight latent classes. The items and attributes that make matrix-Q4 are presented in Table 4.

Table 4. Attribute and Item Which Compose Matrix-Q_4

Item	Attribute		
	A28	**A29**	**A30**
15	1	0	0
16	0	1	0
17	0	0	1

Based on Table 4, the three attributes forming matrix-Q4 are (A28) ability to calculate the length of a track, (A29) ability to count multiple turns on a wheel, and (A30) ability to count many trees planted. The most dominant attributes not controlled by students are attribute A28 with a percentage of 25.5% and A29 with a percentage of 27.7%, which means that the attributes A28 and A29 become the hardest attributes controlled by students.

According to the dominant percentage based on latent classes in items 15 to 17, it indicates that students are in the latent class 0 0 0 and 1 0 1, which illustrate that students have not mastered the ability attributes, while the other latent classes have varying attribute completeness, incomplete incompleteness that varies can be possibly caused by slip factor. It can be seen from the large percentage of the slip in item 15 of 61.2%. Based on the location of the difficulties of each latent class in items 15 and 17, it indicates that students still have difficulty mastering the ability to calculate the length of the cross and the ability to count how many trees are planted.

Items number 18 through number 21 have the compositions of the matrix-Q5. The order of matrix-Q5 is 3 x 4, meaning that there are three attributes in matrix-Q5. Because there are three attributes, there are probably eight latent classes. The items and attributes that make matrix-Q5 are shown in Table 5.

Table 5. Attribute and Item Which Compose Matrix-Q_5

Item	Attribute		
	A32	**A37**	**A38**
18	1	1	0
19	0	1	0
20	0	0	1
21	0	0	1

Based on Table 5, the three attributes forming a matrix-Q5 are (A32) ability to determine the angle's size on a circle, (A37) ability to determine the length of an arc, and (A38) ability to calculate the area of a juring. The most dominant attributes not controlled by students are attribute A37 with a percentage of 36.4% and A38 with a percentage of 37.3%, which means that the attributes of A38 and A37 are the hardest attributes mastered by students.

The dominant percentage based on latent classes in items 10 to 13, obtained that students in the latent class 0 0 0 and 0 1 0, illustrate that students have not mastered the ability attributes, while the other latent classes have varying attribute completeness, incomplete incompleteness that varies bias caused by slip factor. It can be seen from

the magnitude of the percentage of slip-on item 20 of 99.6%. Based on the location of the difficulties of each latent class on items 18 and 21, it indicates that students still have difficulty mastering the ability to determine the angle at the circumference and the ability to calculate the juring area.

Items number 22 to number 24 have the compositions of matrix-Q6. The order of matrix-Q6 is 3 x 3, meaning that there are three attributes in matrix-Q6. Because there are three attributes, then there are probably eight latent classes. The items and attributes that make matrix-Q6 are presented in Table 6.

Table 6. Attribute and Item Which Compose Matrix-Q_6

Item	Attribute		
	A34	A35	A36
22	1	0	0
23	0	1	0
24	0	0	1

Based on Table 6, the three attributes forming matrix-Q6 are (A34) the ability to determine the angle size in a siki-angular triangle, (A35) the ability to determine the angle's circumference, and (A36) the ability to determine the magnitude of the center of a circle. The most dominant attribute not controlled by students is attribute A35, with a percentage of 27.9%, which means that attribute A35 becomes the hardest attribute mastered by students.

Based on the latent class the dominant percentage in items 22 to 24, it is found that students in the latent class 0 0 0 and 1 0 0, illustrate that students have not mastered the ability attributes, while the other latent classes have varying attribute completeness, incomplete incompleteness that varies bias caused by slip factor. It can be seen from the large percentage of the slip in point 23 of 34.4%. Based on the location of the difficulties of each latent class in items 22 and 24, it indicates that students still have difficulty mastering the ability to determine the angle around the circle and the ability to determine the magnitude of the center angle circle.

Items 25 and 26 have matrix-Q7 compiler attributes. The order of matrix-Q7 is 2 x 2, meaning that there are two attributes in matrix-Q7. Because there are two attributes, then there are probably four latent classes, namely 0 0, 0 1, 1 0, and 1 1. The percentage is based on the likelihood of the latent class. The items and attributes that make matrix-Q7 are presented in Table 7.

Table 7. Attribute and Item Which Compose Matrix-Q_7

Item	Attribute	
	A39	A41
25	1	0
26	0	1

Based on Table 7, the attributes on the Q7 matrix are (A40) the ability to determine the length of the center of the circle and also (A41) the ability to determine the length of the tangent of the circle. The most dominant attribute not mastered by students is the A41 attribute, with a percentage of 30.8%, meaning that the A41 attribute becomes the hardest attribute controlled by students. The dominant percentage based on latent classes in items 25 and 26, it is found that students in the latent class 0 0 and 0 1 illustrate that students have not mastered the ability attributes, while the other latent classes have varying attribute completeness, incomplete incompleteness that is biased due to factors slip. It can be seen from the large percentage of the slip in point 26 of 52.8%. Based on the location of the difficulties of each latent class in items 25 and 26, it indicates that students still have difficulty mastering the ability to determine the length of the center point of the circle and the length of the tangent circle.

Items number 27 through number 29 have the compositions of matrix-Q8. The order of matrix-Q8 is 3 x 3, meaning that there are three attributes in matrix-Q8. Because there are three attributes, then there are probably eight latent classes. The items and attributes that make matrix-Q8 are presented in Table 8.

Table 8. Attribute and Item Which Compose Matrix-Q_8

Item	Attribute		
	A43	A44	A45
27	1	0	0
28	0	1	0
29	0	0	1

Based on Table 8, the three attributes which form matrix-Q8 are (A43) the ability to determine the length of tangents along two circles, (A44) the ability to determine the length of tangents in two circles, and (A46) the ability to count the fingers of a small circle on the tangents of a circle. In each latent class profile, the most dominant attribute not controlled by students is the A43 attribute with a percentage of 32.6%, which means that the A43 attribute becomes the hardest attribute controlled by students.

Based on the latent class the dominant percentage in items 27 to 28, it is found that students in the latent class 0 0 0, 0 1 1 and 0 0 1 illustrate that students have not mastered the ability attributes, while the other latent classes have varying attribute completeness, incomplete incompleteness which varies bias caused by slip factors. It can be seen from the magnitude of the percentage of the slip in point 27 of 8.8%. Based on the location of the difficulties of each latent class in items 27 to 28, it indicates that students still have difficulty mastering the ability to determine the tangent length of the outer circle of two circles.

From the analysis of the results of students' answers to the circle material, the dominant errors are caused by students' concept error. It is in accordance with the results of research conducted by Isgiyanto (2011) on junior high school students that the highest types of errors in numbers, algebra, and geometry and measurements are misconceptions, and research conducted by Wulansari (2014) that the main cause of concept errors in numbers is the incorrect application of the counting operations correctly, and the main cause of concept errors in geometry and measurements is not applying the formulas. Thus, to find out students' difficulties in answering DINA model mathematical problems, as stated by Ozaki (2015), the deterministic-input, noisy, and DINA model can assess whether an individual has mastered every skill needed to answer a matter. This information is useful for students to know their weaknesses and so that teachers can teach them effectively.

Conclusion

Based on the results of the analysis using the DINA model in this study, the following conclusions are obtained. Based on the item response theory approach, a suitable model for 30 items is the 2-PL model. Of the 30 items, 27 items are good and meet the criteria of the 2-PL model. The most suitable diagnostic test is used on students with a range of abilities between -1.4 to +2.8.

Based on students' answers obtained, the causes of errors from each type are as follows. The percentage of concept error is 18.47%, the interpretation error is 9.99%, procedure error is 7.80%, and calculation error is 14.57%. Based on the percentage of errors, the most dominant cause is concept error of 18.47%, which means that there are 18.47% of students from a total sample of 416 students who still have misconceptions in answering questions about the circle material.

References

Abadyo, A., & Bastari, B. (2015). Estimation of ability and item parameters in mathematics testing by using the combination of 3PLM/ GRM and MCM/ GPCM scoring model. *REiD (Research and Evaluation in Education)*, *1*(1), 55–72. https://doi.org/10.21831/reid.v1i1.4898

Chen, J., & de la Torre, J. (2014). A procedure for diagnostically modeling extant large-scale assessment data: The case of the programme for international student assessment in

reading. *Psychology*, *5*(18), 1967–1978. https://doi.org/10.4236/psych.2014.518200

de la Torre, J. (2009). DINA model and parameter estimation: A didactic. *Journal of Educational and Behavioral Statistics*, *34*(1), 115–130. https://doi.org/10.3102/1076998607309474

Hughes, M. (1986). *Children and number: Difficulties in learning mathematics*. Oxford: Basil Blackwell.

Isgiyanto, A. (2011). Diagnosis kesalahan siswa berbasis penskoran politomus model partial credit pada matematika. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *15*(2), 308–325. https://doi.org/10.21831/pep.v15i2.1099

Kartowagiran, B. (2013). *Optimalisasi evaluasi pembelajaran teknik mesin melalui logic model untuk meningkatkan soft skills lulusan*. Retrieved from https://studylibid.com/doc/395347/optimalisasi-evaluasi-pembelajaran-teknik-mesin

Kusaeri, K. (2012). Menggunakan model DINA dalam pengembangan tes diagnostik untuk mendeteksi salah konsepsi. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *16*(1), 281–306. https://doi.org/10.21831/pep.v16i1.1118

Kusaeri, K., Suryanto, S., & Kumaidi, K. (2012). Estimasi parameter item dan latent class dengan model DINA untuk diagnosis kesulitan belajar. *Jurnal Ilmu Pendidikan*, *18*(2), 187–193. https://doi.org/10.17977/jip.v18i2.3620

Mardapi, D. (2012). *Pengukuran, penilaian, dan evaluasi pendidikan*. Yogyakarta: Nuha Medika.

Mulyadi, M. (2010). *Diagnosis kesulitan belajar dan bimbingan terhadap kesulitan belajar khusus*. Yogyakarta: Nuha Litera.

Osterlind, S. J. (1998). *Constructing test items: Multiple-choice, constructed-response, performance and other formats*. Boston, MA: Kluwer Academic Publishers.

Ozaki, K. (2015). DINA models for multiple-choice items with few parameters. *Applied Psychological Measurement*, *39*(6), 431–447. https://doi.org/10.1177/0146621615574693

Ravand, H., & Robitzsch, A. (2015). Cognitive diagnostic modeling using R. *Practical Assessment, Research, and Evaluation*, *20*(1), 1–12. https://doi.org/10.7275/5g6f-ak15

Regulation of Minister of National Education, No. 22 of 2006 on Content Standard. , (2006).

Retnawati, H. (2016). *Validitas reliabilitas dan karakteristik butir (Panduan untuk peneliti, mahasiswa, dan psikometrian)*. Yogyakarta: Nuha Medika.

Retnawati, Heri. (2017). Diagnosing the junior high school students' difficulties in learning mathematics. *International Journal on New Trends in Education and Their Implications*, *8*(1), 33–50. Retrieved from http://www.ijonte.org/FileUpload/ks63207/File/04.heri_retnawati.pdf

Runtukahu, J. T., & Kandou, S. (2014). *Pembelajaran matematika dasar bagi anak berkesulitan belajar*. Yogyakarta: Ar-Ruzz Media.

Suwarto, S. (2013). *Pengembangan tes diagnostik dalam pembelajaran: Panduan praktis bagi pendidik dan calon pendidik*. Yogyakarta: Pustaka Pelajar.

Tatsuoka, K. K. (2009). *Cognitive assessment: An introduction to the rule space method*. New York, NY: Routledge/Taylor & Francis.

Triana, A. (2016). *Pengembangan tes diagnostik kesulitan belajar matematika siswa SMA kelas XII IPA*. Master thesis, Universitas Negeri Yogyakarta, Yogyakarta.

Westwood, P. S. (2000). *Numeracy and learning difficulties: Approaches to teaching and assessment*. Camberwell, Victoria:

Australian Council for Educational Research.

Westwood, P. S. (2008). *What teachers need to know about teaching methods*. Camberwell, Victoria: Australian Council for Educational Research.

Wulansari, W. (2014). *Analisis untuk menemukan informasi kesalahan konsep siswa dalam menyelesaikan soal berdasarkan hasil Ujian Nasional (UN) matematika SD tahun 2012/2013 di Kota Kediri*. Master thesis, Universitas Negeri Yogyakarta, Yogyakarta.

Jurnal Penelitian dan Evaluasi Pendidikan
Volume 23, No 2, December 2019 (156-169)
Online: http://journal.uny.ac.id/index.php/jpep

AN EVALUATION OF MATHEMATICS LEARNING PROGRAM AT PRIMARY EDUCATION USING COUNTENANCE STAKE EVALUATION MODEL

Bayuk Nusantara Karaeng Jannang Tompong
Universitas Negeri Yogyakarta

Jailani
Universitas Negeri Yogyakarta

Abstract

The quality of mathematics learning in Bantaeng Regency, South Sulawesi were in a low category based on the research findings from the Institute of Educational Quality Assurance of South Sulawesi in 2011. It affects students to be unwilling to be involved in the process of mathematics learning. This study aims to evaluate the process of learning mathematics in the elementary educational level at Bantaeng Regency. The model of evaluation used in this study is the Stake Countenance Model consisting of three steps of evaluation, namely antecedent, transaction, and outcomes. The subjects of this study were 12 teachers at state elementary schools in Bantaeng Regency assessed by three raters. There were 363 midterm score tests from 12 schools. The instruments used in this study were observation sheet, lesson plan sheet, assessment documentation sheet, and interview guideline. This study used a quantitative research approach supported by qualitative data. The result of the study shows that (1) the lesson plan of mathematics subject is in a good category (93.45%), (2) the learning process is in the good enough category (67.07%), (3) the document of students assessment is in the good enough category (71.34%), and (4) most students in each school do not pass the school standard.

Keywords: *evaluation, mathematics learning, Countenance Stake model*

Permalink/DOI: http://dx.doi.org/10.21831/pep.v23i2.16473

Contact: *Bayuk Nusantara Karaeng Jannang Tompong*
bayuknusantara@gmail.com
Department of Educational Research and Evaluation, Graduate School, Universitas Negeri Yogyakarta
Jl. Colombo No. 1, Karangmalang, Depok, Sleman, Yogyakarta 55281, Indonesia

Introduction

Education is the success key to any aspect because you can improve yourself through education. In Indonesia, there are two kinds of education namely formal and informal education. Formal education is divided into three steps which is primary, secondary, and higher education. Education at the elementary level is a program that involves some components in order to achieve the aim of the program. As a program, education is conscious and deliberate activity directed to achieve the goal (Salmayzuri, Ruslan, & Pristiwaluyo, 2015). Thus, it can be concluded that the education program in primary education is a program that has some components to achieve an educational goal.

In elementary education, some subjects have been taught, one of which is mathematics. Mathematics is taught both in primary and secondary education. Students know how important mathematics is, but some of them find it difficult so that their awareness in learning math is still not enough (Rosnani, Sugiyono, & Tampubolon, 2015). They think that math is difficult to learn then they are burdened to learn math (Mardapi, 2009). The difficulty is not only about the subject but also the teacher.

Teachers play an important role in learning mathematics. How students view mathematics is influenced by the teacher. Bahri and Alimuddin (2016) state that teachers should know students' needs so that they can plan the learning as fun as possible then students will enjoy the learning process. Teachers can find a solution if there is a problem in learning.

Mathematics learning is a process in building a structured mathematics concept for students so that they get mathematic knowledge through experience in the teaching and learning process. There are three steps of mathematics learning, namely preparation, implementation, and assessment. These steps should run well so that students get maximum learning experience. Mathematics learning is regulated by the government.

The learning process is designed by applicable standards. These standards ease teachers in planning, implementing, and assessing. Burton and Kappenberg (2013, p. 9) state that standard is divided into three, elaborated as follows. (1) It should provide information about learning content that should be mastered by students specifically. (2) It helps teachers in determining knowledge based on students' needs. (3) It should help the country and province in assessing program effectiveness and learning method.

Furthermore, standards for the learning process consist of content, process, assessment, and graduate competence standards. The content standard is a fundamental thing in developing a curriculum or lesson plan. Burton and Kappenberg (2013, p. 14) insist that "the content standards set priorities for broad topics by grade level". In other word, content standard is the basis for curriculum development based on educational level. The content standard is used as the basis of the arrangement, implementation, and also assessment in the learning process. Content standard in K*urikulum Tingkat Satuan Pendidikan* (KTSP) or school-based curriculum is arranged in the Regulation of Minister of National Education No. 22 of 2006 which contains the minimum scope of the material and minimum level of competence to achieve minimum competencies for graduates at certain levels and types of education. The implementation of learning process is regulated in process standard.

Process standard is a national standard of education related to the learning process in an education unit to achieve graduate competence standard (Regulation of the Minister of National Education No. 19 of 2005, chapter 1 article 1 verse 6). Process standards are a reference for education implementers in conducting or implementing learning in the classroom. Thus, the teacher has guidelines for implementing learning. The standard process includes three activities related to each other, namely, planning the learning process, implementing the learning process, and evalu-

ating the learning outcomes. The process standard refers to the Regulation the Minister of National Education of Republic of Indonesia No. 41 of 2007 which contains the minimum criteria for the learning process in primary and secondary education units in the entire jurisdiction of the Republic of Indonesia. This standard is also related to the implementation of learning in education units to achieve graduate competence. This standard applies to primary and secondary education at the formal track, both in the package system and in the semester credit system. Process standards can be used as a reference for teachers in developing the learning process, but process standards are not an obligation. Process standards are a reference, but the teacher must develop the process in accordance with the conditions of each student so that students' learning outcomes are better.

Teachers' activities in conducting assessments are regulated in the assessment standards. Assessment standards contain knowledge about the philosophy and purpose of the assessment (van de Walle, 2008, p. 6). Thus, Walle reports that in the assessment standards, the form of assessment along with its purpose is mentioned. In contrast to these standards, assessment standards, according to the National Education Standard Agency or *Badan Standar Nasional Pendidikan* (BSNP), include mechanisms, procedures, and instruments for assessing student learning outcomes. The assessment standards referred to by teachers are based on the Regulation of the Minister of National Education No. 20 of 2007 which contains the standards for conducting assessments. With the existence of assessment standards, teachers are expected to be able to develop forms of assessment following students' needs.

Therefore, teachers in Indonesia have a reference in implementing the learning process, namely: content standards, process standards, and assessment standards. The three standards are set by the National Education Standards Agency (BSNP). Content standards are used as a basis for the preparation of learning designs, learning processes, and assessment processes. The implementation of learning is regulated in a standard process. Process standards are implemented to achieve graduate competency standards. Teachers' guidelines for conducting assessments are contained in the assessment standards. Assessment standards include mechanisms, procedures, and instruments for assessing student learning outcomes.

Learning mathematics in elementary school consists of three stages (Singh, 2008, p. 28). The first stage is the preparation phase. The preparation phase is everything that is done by the teacher before learning begins. Teaching preparation includes the making and preparation of a learning plan, the material to be used, the learning method chosen, the learning objectives themselves, and determining what activities are carried out to meet the learning objectives. A teacher is required to make a learning plan that is intended to facilitate the teacher in implementing the learning process. The making of lesson plans by teachers should be based on content standards that have been determined by the government. This statement is supported by Maryani and Fatmawati (2015, p. 75) who explain that the lesson plan is designed based on content standards. The importance of the learning plan was expressed by Niwaz, Shah, and Rajper (2016) who reveal that learning planning which is carried out regularly can improve the effectiveness of mathematics learning, because the learning plan is a document that has the idealism of the teacher in implementing the learning process so that when the lesson plan is implemented properly by the teacher, then, it can improve the effectiveness of mathematics learning.

The second stage is the process of learning implementation. The learning process will be more meaningful when the teacher implements a lesson plan that has been made. The learning process has three stages, namely the preliminary, main, and closing stages. Andriani (2015) says that

there are three activities in the implementation of the mathematics learning process, namely preliminary, core, and closing activities. The results of research conducted by Andriani show that the teacher had carried out the learning process but some activities were considered to be not optimal. Every stage in the learning process should refer to the standard process. The appropriateness between the learning plans that are made and what happens on the field deserves attention.

The last step in implementing mathematics learning is assessment. Assessment is defined as an activity that can be done either before learning, when learning, or when learning ends to get information both quantitatively and qualitatively with specific objectives (Kasih & Purnomo, 2016). Thus, assessment is done to know students' understanding of the material so that it can be done anytime. In the learning process, teachers are required to conduct an assessment to find out how far students can absorb what has been taught in the learning process. The complete assessment documents are contained in the assessment standards making it easier for teachers to complete the assessment documents. Assessments made by teachers can be in the form of daily assessments, midterm tests, and end of semester tests.

A program is said to have achieved its goals if it obtains information about the usefulness of a program. In line with this idea, Fitzpatrick, Sanders, and Worthen (2011) say that evaluation is an activity to determine the usefulness of an evaluation object. In other words, evaluation is an activity to identify, clarify, and apply several criteria to determine whether a program is useful or not. Evaluation leads to a decision. Thus, the result of the evaluation is a decision whether the program is useful or not so that policymakers can decide the sustainability of the program.

Evaluation is a systematic activity. Consequently, evaluation has a procedure of activities in its implementation. Rossi and Freeman (1985, p. 19) explain that evaluation is an activity that has certain procedures in assessing the application of a program. Therefore, the evaluation uses certain procedures to provide an assessment to increase the planning, monitoring, effectiveness, and efficiency of a program.

Evaluation Model of Countenance Stake

The Stake's Countenance Model is an evaluation model which was developed by Robert Stake in 1967. Stake (1996) argues that evaluation can provide an overall picture of the implementation of a program so that it can be given consideration afterward. It means that an evaluation model must be able to provide comprehensive information about what is being evaluated, both the measurement results and judgment.

Stakes' Countenance Model consists of two activities: description and assessment. Each activity consists of three aspects that are of concern to an evaluator in evaluating a program, namely, Antecedents (Context), Transactions (Process), and also Outcomes (Output) (Kaufman & Thomas, 1980).

The Countenance Stake evaluation model is widely used in research evaluating the learning process. Lukum (2015) evaluates the junior high school natural science learning program in Bone Bolango Regency using this evaluation model. Observation, interview, and documentation techniques were used in gathering information. The results of this study mention that learning planning is still not in accordance with the standards. The second finding is about the learning process that is still not in accordance with the standard process. It results in students' grades that still do not meet the specified minimum completion criteria or *Kriteria Ketuntasan Minimal* (KKM) grades. This study is in line with research conducted by Dole and Wibowo (2013) which says that teachers' understanding of School-Based Curriculum (SBC) is very low so that it results in the implementation of SBC in the classroom. Dole and Wibowo (2013) have found that the ability of teach-

ers in the implementation of SBC (planning, implementation, and assessment) was still in the low category. This research was conducted at several elementary schools in Ende.

In contrast to research conducted by Lukum, Waluyati (2012) has found that learning planning in junior high school or Islamic-based junior high school in the City of Bima is categorized as good or in accordance with the standards. Furthermore, student learning outcomes already meet the KKM or good grades. This is different from the previous study which said that the learning process was not appropriate, so that resulted in student learning outcomes at the time of the study conducted by Waluyati, the learning process was included in the quite good category with good learning outcomes.

Wibowo and Wutsqa (2014) obtain the same results, namely, a compilation of good planning, then, will produce a good learning process and student learning outcomes. This research is evaluation research on the implementation of *Kurikulum Tingkat Satuan Pendidikan* (KTSP) or School-Based Curriculum for mathematics in junior high school in Yogyakarta City. The results show that the planning made by the teacher is in a good category, the learning process is in a good category, and the learning outcomes are in a good category. This third thing happens because teachers' understanding of SBC is also in the good category. Thus, the teacher has sufficient knowledge in conducting learning.

One of the causes of students' difficulty in learning mathematics is the challenge. Ramirez, Chang, Maloney, Levine, and Beilock (2016) state that teachers must become compatible with mathematics. In other words, the teacher must be able to solve every mathematical problem that is done by students so the teacher must have some references to it.

A research conducted by Setiawan (2016) is classroom action research. The research aims to improve student learning outcomes with the paper props method. The results show that the ability of teachers and students improve after using paper props. The teacher improves the way of teaching given to the results of reflection so that students' abilities increase. Thus, it can be concluded that student learning outcomes can be improved by improvements made by the teacher during the learning process. In other words, the learning process must run according to the applicable minimum standards.

It was stated in the report of the results of school self-evaluation conducted by the Institute of Educational Quality Assurance or *Lembaga Penjaminan Mutu Pendidikan* (LPMP) of South Sulawesi Province in 2011, that the achievement of national education standards in Bantaeng Regency is still in the low category. Some of them are still at minimum service standards. In the data, it is stated that the highest achievement is only in the implementation of the assessment. It can be interpreted that the quality of education delivery in Bantaeng Regency is still lacking. This assessment is carried out by each school as a need to improve school performance and quality. However, this evaluation is an internal evaluation that might be biased because it was done by the school principal. Therefore, it is necessary to hold an evaluation of the learning process in schools conducted by evaluators who are not from the school environment.

Based on the background and some theoretical studies regarding evaluation and learning, the research objectives are (1) identifying the suitability of learning planning; (2) identifying the suitability of the learning process; (3) identifying appropriateness of learning assessment; and (4) identifying student learning outcomes.

Research Method

This research was descriptive evaluation research with a quantitative approach supported by qualitative data, where the object to be evaluated was the implementation of elementary school mathematics

learning. There were several stages in the implementation of mathematics learning, starting from the planning, implementation, to evaluation stages in Bantaeng Regency, South Sulawesi.

The evaluation model used is the Countenance Evaluation Model. The model is an evaluation model developed by Stake which emphasizes the implementation of two main components, namely: (1) description (description) and (2) considerations (judgments), and dividing the evaluation object into three things, namely: (a) antecedents (context evaluation); (b) transactions (evaluation of the process); and (c) outcomes (evaluation of outputs and outcomes) (Fernandes, 1984). The selection of this evaluation model is based on the formulation of the research problem, which is to find out the implementation of mathematics learning for elementary school students as a whole (planning, implementing, and evaluating learning) where the three evaluation objects are in the Countenance Evaluation Model.

This research was conducted in 12 public elementary schools in Bantaeng District, Bantaeng Regency, South Sulawesi. The selected school is a public school that uses a school-based curriculum or KTSP. At the research site, only classes V and VI still use KTSP. The selection of class V is based on the ongoing learning process. The determination of the sample from this study used a purposive sampling technique in which the research sample was taken on the basis that the class V's teachers were not fixated with the National-Based School Final Examination so that observations of the learning process could be made.

In this study, the data to be collected was a description of planning, implementing, and also evaluating elementary school mathematics learning. Data collection in this study was conducted using non-test techniques using observation, documentation, and interview techniques. First, an observation aims to see the implementation of learning undertaken by the teacher. Second, documentation aims to obtain information about the Learning Program Plan or *Rencana Pelaksanaan Pembelajaran* (RPP) prepared by the teacher, assessment documents, and student mathematics learning outcomes. Finally, interviews with teachers are used to obtain secondary information regarding the implementation of elementary school mathematics learning.

Validity is an index that shows how far the accuracy of the instrument is in carrying out its functions (Azwar, 2016, p. 8). A research instrument is said to be valid if it is able to measure what should be measured. Therefore, research instruments need to be estimated for validity before going into the field. The validity that was used in this research is content validity. Content validity is validity where evidence of validity is obtained from estimates made by competent experts in the measured field (Mardapi, 2012, p. 39). Validity estimation of the observation sheet was done with the help of experts (expert judgment) who later provide an assessment of each instrument item.

The results of the estimation of content validity by using the Aiken index formula for 21 items in the lesson plan review sheet instrument indicate that four items (items 12, 17, 19, and 20) have a high coefficient of validity, while the other 17 items have a medium coefficient of validity. The results of the estimation of the content validity of 34 items in the observation sheet instrument using the Aiken index formula indicate that there are five items (items 1, 2, 15, 32, and 33) that have a high coefficient of validity. Meanwhile, 29 of the 34 observation sheet instrument items are in the medium category. The results of the estimation of the content validity of 15 items in the assessment sheet instrument of the assessment document using the Aiken index formula show that there are six items (items 8, 10, 11, 12, 13, and 14) having a high coefficient of validity. Nine of the 15 items of the instrument review sheet of assessment documents fall into the medium category.

Reliability is an index that shows the extent to which a measuring instrument can be trusted or reliable. It means that the instrument can be said to be reliable if the instrument is used to measure something with the same symptoms twice or more and the results are relatively consistent. The reliability of an instrument is determined by the reliability coefficient. Observation sheet is used to observe the process of implementing mathematics learning in class. The lesson plan review sheet was used to determine the suitability of the lesson plan with the syllabus, while the assessment document review sheet is used to determine the suitability of the assessment document with the assessment standard. To estimate the reliability of the observation sheet, the lesson plan review sheet, and the evaluation document review sheet, interrater reliability was used. The reliability of the three instruments used the Fleiss' Kappa scale which states the coefficient > 0.40 is reliable (Gwet, 2012, p. 125). Reliability for each instrument is 0.931, 0.966, and 0.977.

The main data in this study was quantitative data supported by qualitative data. Data analysis techniques in this study used quantitative descriptive. Information obtained from observations and documentation was analyzed using quantitative analysis. The quantitative data were analyzed with three stages of data scoring, data tabulation, and data applications. In this study, data analysis was determined based on the ideal mean and ideal standard deviation. Achievement of respondents' scores on each instrument was processed and compared with the category formula by using the ideal mean and ideal standard deviation and then was used as interpretation material. The categorization for processing quantitative data referred to the categorization proposed by (Azwar, 2017, p. 148), as presented in Table 1.

The antecedent aspects of this study were measured through the components of the learning plan (RPP). To get the results of an assessment of the learning planning, the analysis technique of the learning planning documents was used. In this component, there were 21 grading statements that must be filled by three raters with a range of scores from 1 to 4. Then it could be seen, the highest ideal score is 84 and the lowest ideal score is 21. Thus, the ideal average (M) = ½ (84 +21) = 52.5 and for ideal standard deviation (SD) = 1/6 (84-21) = 10.5. Evaluation criteria according to the ideal formula are presented in Table 2.

In the aspect of the transaction (program implementation), assessment is carried out on the implementation of learning and also the study of assessment documents that have been made by the teacher. The learning process is measured through observation of the learning activities in class. In these components, there are 34 items with a range of scores from 1 to 5. Then it can be seen, the highest ideal score is 170 and the

Table 1. Categorization of Quantitative Data Processing

X score	Category
$X > M + 1.5\ SD$	Very good/ very effective
$M+ 0.5\ SD < X \leq M + 1.5\ SD$	Good / effective
$M - 0.5\ SD < X \leq M + 0.5\ SD$	Pretty good / effective enough
$M - 1.5\ SD < X \leq M - 0.5\ SD$	Poor / less effective
$X \leq M - 1.5\ SD$	Not good / ineffective

Table 2. Evaluation Criteria of Lesson Plan

Interval	Criteria
$X > 68.25$	Very good
$57.75 < X \leq 68.25$	Good
$47.25 < X \leq 57.75$	Pretty good
$36.75 < X \leq 47.25$	Poor
$X \leq 36.75$	Not good

lowest ideal score is 34. Thus, the ideal average (M) = ½ (170+34) = 102, and for ideal standard deviation (SD) = 1/6 (170-34) = 22.67. Evaluation criteria according to the ideal formula are clearly presented in Table 3.

Table 3. Evaluation Criteria of Learning Process

Interval	Criteria
X > 136	Very good
113.33 < X ≤ 136	Good
90.67 < X ≤ 133.33	Pretty good
68 < X ≤ 90.67	Poor
X ≤ 68	Not good

Implementation of the assessment is measured using an assessment document, using a data aggregation technique, namely document analysis. In these components, there were 15 items with a range of scores from 1 to 4. From the calculations, it could be seen that the highest ideal score is 60 and the lowest ideal score is 15. Thus, ideal mean (M) = ½ (60+15) = 37.5 and for ideal standard deviation (SD) = 1/6 (60+15) = 7.5. Evaluation criteria according to the ideal formula are presented in Table 4.

Table 4. Evaluation Criteria of Process Assessment

Interval	Criteria
X > 48.75	Very good
41.25 < X ≤ 48.75	Good
33.75 < X ≤ 41.25	Pretty good
26.25 < X ≤ 33.75	Poor
X ≤ 26.25	Not good

The outcome aspect of this study was measured through the results of students' midterm examinations. The score which was obtained from the teacher was then compared with the graduate competency standard or *Standar Kompetensi Lulusan* (SKL) of each school. After being compared with the SKL, the results were then categorized as graduated and not graduated for each school.

The evaluation criteria were used to analyze and interpret the results of data processing obtained. In other words, the judgment or categorization of the average score in the field in each aspect referred to the ideal average criteria except for the outcome aspect. The average achievement score in the field was then converted to the percentage of achievement with the following formula.

$$Achievement\ Percentage = \frac{\text{average achievement score}}{\text{maximum ideal average for each aspect}} \times 100\%$$

Table 5. Criteria of Learning Program Success

Stages	Aspects	Success criteria
Antecedent	Lesson Plan	100%
Transaction	Learning process Assessing process	100%
Outcomes	Students' score	80%

The result of the percentage was then compared to the standard based on Table 5. Success criteria for learning programs are obtained from the content standards, process standards, and also assessment standards. For content standards and process standards refer to the suitability of standards set by the government, then the success criteria must be 100%. In contrast to the content standards and process standards, the aspect of results only stands at 80% due to differences in graduation standards per school.

Findings and Discussion

Implementation of learning in the classroom should become a concern for teachers, especially mathematics teachers because the good implementation of learning gives impact to student learning outcomes. In this study, there are three stages, namely the antecedent, transaction, and outcome stages. Of the three stages, the antecedent and transaction stages use five categories, which are very good, good, good enough, not good, and not good, while for the outcome phase, only the graduated and non-graduated categories are used. The results of this study indicate that the assessment of lesson plan documents is included in the good category, while the assessment documents are included in the category of

quite good. It is because of the percentage obtained from the two variables, which is, more than 50%. The implementation of learning got a pretty good category with a percentage of 67.07%. In addition, for the outcomes phase, the midterm grades for 10 schools do not meet the minimum completion criteria.

Assessment for the learning plan document made by teachers was conducted using instruments that refer to the Regulation of Minister of National Education No. 22 of 2006 concerning the content standards for primary and secondary education units. Learning planning is said to be good if the teacher makes a learning/lesson plan in accordance with the established standards. The results of the analysis are then compared with the success criteria set at 100%.

Antecedent

A lesson plan is a very important document for both the teacher and other teaching staffs, so whenever a teacher implements learning, he/she already understands what will be taught in class. In addition, the learning planning document will also help the teacher to make them go back to the track whenever the learning process does not run as planned. In this study, the learning plan or lesson plan documents that have been prepared by teachers are then assessed using document review. Overall, the teacher makes the design of the learning plan very good. It is evidenced by 93.45% of learning plans made by teachers which are included in the excellent category. In other words, in planning learning, the teacher has carried out good planning activities. Based on interviews with teachers, they receive training from the government regarding the learning process, making it easier to create lesson plans for learning implementation even though some teachers are still difficult to apply the knowledge gained through the training. It can be concluded that the evaluation of the learning plan is included in the good category, but consideration remains. The consideration referred here is that the teacher must pay attention to the suitability of the formulation of indicators with the competence standard and also basic competence, as well as the selection of learning models that are in accordance with the characteristics of students. The results show that the two items are included in the quite good category. Therefore, teachers have not considered the competence standard and basic competence in the formulation of indicators and the selection of learning models. Whereas, according to Salmayzuri et al. (2015), learning planning is conducted in accordance with the established standards. In addition, research conducted by Morris and Hiebert (2017) reveals that planning is very important in learning so that it will have a positive influence on the students' achievement.

Transaction Aspect of Learning Implementation

Evaluation at the transaction stage is an evaluation of the learning process in the classroom and assessment documents. The researcher and rater conduct an assessment of the learning process conducted by the teacher in the classroom. Observation of the learning process in the classroom is done by using an observation sheet instrument that refers to the Regulation the Minister of National Education of Republic of Indonesia No. 41 of 2007 concerning the standard processes for primary and secondary education units. In the learning process, there are three stages, namely the preliminary, core, and closing stages. Observations of the three activities refer to the established process standards. The learning process is said to be good if the teacher carries out the learning process in accordance with the stages in the standard process. The results of the analysis are then compared with the success criteria set at 100%.

In contrast to the learning planning described earlier, evaluation of learning observations obtains a quite good category. Based on observations, some teachers do not do what is written in the document. There are three sub-indicators in this stage,

namely, preliminary, core, and closing activities. Thus, it can be concluded that in the implementation of learning, there are things that are not in accordance with the planning of learning.

When confirmed to the teacher by using an interview, it was found that there are some obstacles felt by the teacher when teaching, namely, lack of learning resources, and feeling burdened when teaching mathematics because the teacher's educational background is not from mathematics education. Both of these obstacles result in the teacher feeling difficulties in implementing learning in accordance with the standard process. When compared with the standard, there is a gap of 32.93%. Improvements must be made to all aspects, namely preliminary, core, and closing activities because, in all three activities, the teacher does not carry out the stages according to the standard, for example, the teacher does not carry out preliminary activities such as providing motivation to students before the lesson begins.

The results of this study are in line with research conducted by Lukum (2015) where the learning process is not in accordance with the established process standards. In addition, Waluyati (2012) also states that the learning process carried out in junior high school/Islamic-based junior high school in the city of Bima was included in the quite good category. Thus, it can be concluded that the learning process has not been carried out in accordance with existing process standards. It is because the expected standard of learning is 100%. Research conducted by Bolkan, Goodboy, and Myers (2017) is in line with the results of this study which states that explanations from teachers affect students' mathematics learning outcomes. Therefore, the decision taken is the consideration to improve the process of implementing learning. The consideration that can be given is that the teacher is more able to apply what has been written in the learning plan. In addition, the teacher must master the material first before delivering information to students.

Transaction Aspect of Assessment

The assessment of assessment documents made by the teacher was done using the document review sheet instrument made based on the Regulation of Minister of National Education No. 22 of 2006 concerning the assessment standards for primary and secondary education. In evaluating students' learning outcomes, teachers are required to make several instruments or assessment documents which have been determined by the government. Assessment documents are said to be good if the teacher has made an assessment document in line with applicable assessment standards. The results of the analysis are then compared with the success criteria set at 100%.

Assessments assessed in this study are assessment documents that have been written by the teacher. The assessment document was then assessed using a review sheet of the assessment document to see its conformity with the assessment standard. From the results of the study, it is found that the assessment documents are included in the category of quite good with a percentage of 71.34%. In the assessment document review sheet, there are five sub-indicators, divided into sub-indicators of the completeness of the set of assessment documents, set of techniques, preparation of assessment strategies, processing assessment, and conducting the assessment. Thus, it indicates that the assessment documents are in a quite good category with consideration. Activities that need to be improved in carrying out the assessment are the instrument grids, the use of non-test assessment instruments, the multiple-choice assessment guidelines, the use of multiple-choice tests, the assessment using observation sheets, and the assessment is carried out in the form of daily tests. Teachers are expected to complete the assessment documents in accordance with the established standards. Interviews were conducted with the teacher during the observation process. The interview results found that the teacher applies knowledge when training to conduct assessments in the classroom. When observing, the teacher as-

sesses students as questions and answers as stated in the learning implementation plan, however, the completeness of the assessment documents is not given enough attention. It is similar to what was expressed by Noviana and Kartowagiran (2015) whose research results reveal that the teacher has carried out the assessment very well because they not only did the assessment during the test but also during the learning process. However, the lack of the assessment document causes the assessment is included in a quite good category. The use of grading rubrics will help the teacher to grade consistently. It is in accordance with Jönsson and Panadero (2017) who report that there are two advantages if the teacher makes an assessment rubric, namely, helping the teacher in terms of the consistency of assessment, and it can improve the instruction done by the teacher after giving feedback in learning. To improve the implementation of assessments conducted by teachers, the government should disseminate information to teachers to complete the assessment documents such as grading instrument assessment, rubric assessment essay questions, and conduct training to analyze the results of the assessment (Setiadi, 2016).

Outcomes

Evaluation of outcome is an assessment of the results of students' midterm test grades. The resulting score is then matched with the graduate competence standard or *Standar Kompetensi Lulusan* (SKL) value, so that the pass and not pass categories will appear. Based on the results of the study, it is recognized that 43% of students are declared to have passed or exceeded the SKL score, while 57% are declared not to pass. We can ignore the majority of students not fulfilling the SKL grades because the implementation of learning is still not very good for many students who have not exceeded the SKL score. The results of the study are in line with the research conducted by Lukum (2015) which agrees to the learning conducted by the teacher that is not in accordance with the existing standards so that students are unable to meet the SKL. It is because the specified standard is 100%, whereas, at the specified time, students who meet the SKL score are less than 50%. Therefore, teachers need to consider what is given by researchers in order to improve the learning process. Ngware, Ciera, Musyoka, and Oketch (2015) also state the quality of the teacher positively contributes to improving student learning achievement. Moreover, Ottmar, Decker, Cameron, Curby, and Rimm-Kaufman (2014) reveal different research results that there is no relationship between the quality of learning with an increase in student achievement. The same thing was agreed by Fung et al. (2017) that the pedagogical ability of teachers supports student achievement. Thus, it can determine the number of students who cannot meet the SKL to determine the quality of learning that is not in accordance with predetermined process standards. Student learning outcomes can be improved by the methods used by teachers. Therefore, the teacher must recognize several learning strategies so that students have the motivation to learn so as to improve student achievement. The learning strategy used in a research conducted by Wahyuni and Jailani (2017) is realistic mathematics learning. The result of the study is that realistic learning can increase motivation and learning achievement of elementary school students.

After a description of the stages in the implementation of the learning program, then a comparison is made between the ideal conditions in accordance with the established standards and actual situation when on the field, then the gap between stages is seen vertically. The results show that the teacher could make plans for implementing learning very well in accordance with the school-based curriculum. Based on the results of observations made during the learning process, there is a mismatch between planning and learning implementation. The assessment documents made by the teacher are included in the quite good

category. The teacher has not made an assessment document in accordance with the specified standards. It makes student learning outcomes less favorable because the tendency in every school is not yet to meet the minimum completion criteria.

Overall, the teacher has done a very good planning because they have received training from the government which makes them to create lesson plans for the learning implementation, even though some of them are still difficult to apply the knowledge gained through training. Ideally, good planning will facilitate the teacher in implementing the learning process. However, it did not happen. The implementation of learning in this study is only included in the category of quite good. This discrepancy occurs because teachers have difficulty in implementing planning that has been made such as providing motivation in preliminary activities not done by the teacher. The assessment documents made by the teacher still need a lot of improvement because the results of the analysis of the assessment documents are included in the quite good category. The discrepancy has an impact on student learning outcomes. The results of the analysis on the aspects of student learning outcomes indicate that the tendency in each school is not yet fulfilling the minimum completion criteria score. It might be the result of a mismatch between the learning plan, the assessment documents, and the learning implementation. Besides, the high minimum completion criteria score in each school is also one of the consequences of many students who do not meet the minimum completion criteria score.

Conclusion

Based on the results of the study, it is concluded that the evaluation of the antecedent component, namely the planning of learning programs, is in the good category of 93.45%. Evaluation of the transaction component, namely the implementation of the learning process aspects of the process is included in the category of quite good by 67.07%. Meanwhile, evaluation of the transaction component, namely aspects of the assessment of learning, is included in the quite well category by 71.34%. Finally, the results of the evaluation on the outcome component are more than 50% of students included in the category of not passing.

References

Andriani, S. (2015). Evaluasi CSE-UCLA pada studi proses pembelajaran matematika. *Al-Jabar: Jurnal Pendidikan Matematika*, *6*(2), 167–175. https://doi.org/10.24042/ajpm.v6i2.46

Azwar, S. (2016). *Reliabilitas & validitas (IV)*. Yogyakarta: Pustaka Pelajar.

Azwar, S. (2017). *Penyusunan skala psikologi*. Yogyakarta: Pustaka Pelajar.

Bahri, H. A., & Alimuddin, A. (2016). The profile of teacher's understanding on student's mathematics ability based on teacher's teaching experience at SMPN 1 Gantarangkeke. *Jurnal Daya Matematis*, *4*(2), 119–128. Retrieved from https://ojs.unm.ac.id/JDM/article/view/2892/pdf_31

Bolkan, S., Goodboy, A. K., & Myers, S. A. (2017). Conditional processes of effective instructor communication and increases in students' cognitive learning. *Communication Education*, *66*(2), 129–147. https://doi.org/10.1080/03634523.2016.1241889

Burton, D. T., & Kappenberg, J. W. (2013). *Mathematics, the common core, and RTI: An integrated approach to teaching in today's classrooms*. Newbury Park, CA: Corwin Press.

Dole, F. E., & Wibowo, U. B. (2013). Kemampuan guru mengimplementasikan KTSP pada SD di kecamatan Detukeli Kabupaten Ende. *Jurnal Prima Edukasia*, *1*(2), 147–161. https://doi.org/10.21831/jpe.v1i2.2632

Fernandes, H. J. X. (1984). *Evaluation of educational programs*. Jakarta: National Education Planning, Evaluation and Curriculum Development.

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2011). *Program evaluation: Alternative approaches and practical guidelines* (4th ed.). Boston, MA: Pearson Education.

Fung, D., Kutnick, P., Mok, I., Leung, F., Lee, B. P.-Y., Mai, Y. Y., & Tyler, M. T. (2017). Relationships between teachers' background, their subject knowledge and pedagogic efficacy, and pupil achievement in primary school mathematics in Hong Kong: An indicative study. *International Journal of Educational Research*, *81*, 119–130. https://doi.org/10.1016/j.ijer.2016.11.003

Gwet, K. L. (2012). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among multiple raters* (3rd ed.). Gaithersburg, MD: Advanced Analytics.

Jönsson, A., & Panadero, E. (2017). The use and design of rubrics to support assessment for learning. In D. Carless, S. M. Bridges, C. K. Y. Chan, & R. Glofcheski (Eds.), *Scaling up assessment for learning in higher education* (pp. 99–111). https://doi.org/10.1007/978-981-10-3045-1_7

Kasih, P. A., & Purnomo, Y. W. (2016). Peningkatan hasil belajar matematika siswa sekolah dasar melalui pembelajaran berbasis penilaian. *Journal of Research and Advances in Mathematics Education*, *1*(1), 69–78. https://doi.org/10.23917/jramathedu.v1i1.1776

Kaufman, R. A., & Thomas, S. (1980). *Evaluation without fear*. Newberry, FL: New Viewpoints.

Lukum, A. (2015). Evaluasi program pembelajaran IPA SMP menggunakan model Countenance Stake. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *19*(1), 25–37. https://doi.org/10.21831/pep.v19i1.4552

Mardapi, D. (2009). Evaluasi penerapan ujian akhir sekolah dasar berbasis standar nasional. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *13*(2), 227–245. https://doi.org/10.21831/pep.v13i2.1411

Mardapi, D. (2012). *Pengukuran, penilaian, dan evaluasi pendidikan*. Yogyakarta: Nuha Medika.

Maryani, I., & Fatmawati, L. (2015). *Pendekatan scientific dalam pembelajaran di sekolah dasar: Teori dan praktik*. Yogyakarta: Deepublish.

Morris, A. K., & Hiebert, J. (2017). Effects of teacher preparation courses: Do graduates use what they learned to plan mathematics lessons? *American Educational Research Journal*, *54*(3), 524–567. https://doi.org/10.3102/0002831217695217

Ngware, M. W., Ciera, J., Musyoka, P. K., & Oketch, M. (2015). Quality of teaching mathematics and learning achievement gains: Evidence from primary schools in Kenya. *Educational Studies in Mathematics*, *89*, 111–131. https://doi.org/10.1007/s10649-015-9594-2

Niwaz, A., Shah, S. Z. H., & Rajper, A. K. (2016). Lesson planning in teaching of mathematics: Teachers' perceptions and current practices. *The Sindh University Journal of Education*, *45*(1), 45–66.

Noviana, N., & Kartowagiran, B. (2015). Evaluasi implementasi penilaian otentik pembelajaran Bahasa Indonesia pada Kurikulum 2013 di SMA. *Jurnal Evaluasi Pendidikan*, *3*(2), 157–166. Retrieved from http://journal.student.uny.ac.id/ojs/index.php/jep/article/download/1271/1146

Ottmar, E. R., Decker, L. E., Cameron, C. E., Curby, T. W., & Rimm-Kaufman, S. E. (2014). Classroom instructional quality, exposure to mathematics instruction and mathematics achievement in fifth grade. *Learning Environments Research*, *17*, 243–262.

https://doi.org/10.1007/s10984-013-9146-6

Ramirez, G., Chang, H., Maloney, E. A., Levine, S. C., & Beilock, S. L. (2016). On the relationship between math anxiety and math achievement in early elementary school: The role of problem solving strategies. *Journal of Experimental Child Psychology*, *141*, 83–100. https://doi.org/10.1016/j.jecp.2015.07.014

Regulation of Minister of National Education No. 22 of 2006 on Content Standard. , (2006).

Regulation of the Minister of National Education No. 19 of 2005, on National Standard of Education. , (2005).

Regulation of the Minister of National Education No. 20 of 2007 on the educational assessment standard. , (2007).

Regulation the Minister of National Education of Republic of Indonesia No. 41 of 2007 on Process Standard for Primary and Secondary Educational Units. , (2007).

Rosnani, R., Sugiyono, S., & Tampubolon, B. (2015). Peningkatan hasil belajar pada pembelajaran matematika menggunakan alat peraga realita di kelas 1 sekolah dasar. *Jurnal Pendidikan Dan Pembelajaran Khatulistiwa*, *4*(1), 1–9. Retrieved from http://jurnal.untan.ac.id/index.php/jpdpb/article/view/8529/pdf

Rossi, P. H., & Freeman, H. E. (1985). *Evaluation: A systematic approach*. Beverly Hills, CA: Sage.

Salmayzuri, S., Ruslan, R., & Pristiwaluyo, T. (2015). Evaluasi program pembelajaran matematika di SMA Negeri Watansoppeng. *Riset Assesmen: Jurnal Penelitian Dan Evaluasi Pendidikan*, *1*(1), 1–8. Retrieved from https://ojs.unm.ac.id/assesment/article/view/1778

Setiadi, H. (2016). Pelaksanaan penilaian pada kurikulum 2013. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *20*(2), 166–178. https://doi.org/10.21831/pep.v20i2.7173

Setiawan, I. (2016). Peningkatan hasil belajar matematika pada materi pecahan siswa kelas III SD Negeri Bendungan III dengan alat peraga. *Basic Education: Jurnal Elektronik Mahasiswa Prodi PGSD*, *5*(3), 214–226. Retrieved from http://journal.student.uny.ac.id/ojs/index.php/pgsd/article/view/873

Singh, Y. K. (2008). *Teaching practice: Lesson planning*. New Delhi: APH Publishing.

Stake, R. E. (1996). For all program evaluations there's a criterion problem. *Evaluation - The International Journal of Theory, Research and Practice*, *2*(1), 99–103.

van de Walle, J. A. (2008). *Matematika sekolah dasar dan menengah: Pengembangan pengajaran jilid 1*. Jakarta: Erlangga.

Wahyuni, N. D., & Jailani, J. (2017). Pengaruh pendekatan matematika realistik terhadap motivasi dan prestasi belajar siswa SD. *Jurnal Prima Edukasia*, *5*(2), 151–159. https://doi.org/10.21831/jpe.v5i2.7785

Waluyati, I. (2012). Evaluasi program pembelajaran ilmu pengetahuan sosial SMP/MTs di Kota Bima. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *16*(1), 260–280. https://doi.org/10.21831/pep.v16i1.1117

Wibowo, R., & Wutsqa, D. U. (2014). Evaluasi pelaksanaan Kurikulum Tingkat Satuan Pendidikan (KTSP) mata pelajaran matematika SMP di Kota Yogyakarta. *Jurnal Riset Pendidikan Matematika*, *1*(1), 58–68. https://doi.org/10.21831/jrpm.v1i1.2664

Jurnal Penelitian dan Evaluasi Pendidikan
Volume 23, No 2, December 2019 (170-183)
Online: http://journal.uny.ac.id/index.php/jpep

THE EFFECTIVENESS AND CONSTRAINTS OF LEARNING IN POLYTECHNIC EDUCATION

I Nyoman Gunung
Politeknik Negeri Bali

I Ketut Darma
Politeknik Negeri Bali

Abstract

This study is aimed at finding out: (1) the level of effectiveness of the teaching implementation at a polytechnic education institution viewed from the interrelatedness of the components of context, input, process, and product, and (2) the constraints in the teaching implementation at the institution. This study used the Contex Input Process Product (CIPP) model from Stufflebeam. The data were collected using a questionnaire, interview guide, and related documents. The subjects of the evaluation were the students, lecturers, and the top management staff of the departments at the polytechnic. The data were analyzed using a quantitative descriptive method. The effectiveness of the learning program is determined based on the results of the prototype verification in the Glickman quadrant. The results of the analysis show that the teaching program implementation at the polytechnic viewed from the interrelatedness of context, input, process, and product is considered effective enough. The constraints in the implementation of teaching at the polytechnic are in teaching planning, implementation of semester teaching plans, curriculum, and infrastructure as well as facilities. Those constraints create other barriers to the achievement of learning outcomes, both academic and non-academic.

Keywords: *program evaluation, effectiveness, constraints, CIPP*

Permalink/DOI: http://dx.doi.org/10.21831/pep.v23i2.23526

Contact: *I Nyoman Gunung*
nyomangunung@pnb.ac.id
Department of Mechanical Engineering, Politeknik Negeri Bali
Jl. Kampus Bukit Jimbaran, Kuta Selatan, Badung 80364, Bali, Indonesia

Introduction

Education is known as a deliberate activity process done to the input according to the predetermined goal. As a process, the results must be evaluated to find out the effectiveness of the teaching and learning process that has been implemented.

In the educational process, evaluation is one of the essential components that play an essential role in identifying the success of an educational program. The evaluation is aimed at obtaining data on the distance between the existing situation and the expected situation using certain criteria. Gronlund and Linn (1990) state that the evaluation of instruction is a process of collecting, analyzing, and interpreting information systematically to determine to what extent the objective of education has been attained. The components that need to be evaluated are students, teachers/lecturers, course materials, curriculum, infrastructure, and facilities.

Evaluation is an effort in deciding the attainment quality of the goal of the program and the students' performance and ability (Ebel & Frisbie, 1991). To find out an appropriate evaluation system or model, it is important to provide accurate information for the stakeholders, especially the top management of the institution and to optimally enhance the teaching program.

Many evaluation models have been developed by experts and can be used to evaluate the teaching program, one of which is the CIPP evaluation model that was developed by Stufflebeam, *et al.* 1971 (Arikunto, 2009). Each of the models has a different orientation and approach. To know if the selection of a particular model is appropriate or not, it depends on the type of activity evaluated, whether the program is in the form of activity process, services, or general services.

Stufflebeam, et al. developed the context, input, process, and product (CIPP) model in 1971 (Stufflebeam, 2003). This model is an evaluation model that is oriented toward decision making (a decision-oriented evaluation approach) to help the administrators or leaders in the decision making in the process. Its point of departure is the view that success in an educational program is determined by various factors such as the learner and environmental characteristics, the equipment used, the procedure, and the mechanism of the program implementation. The result will give an alternative solution to the problems faced by the decision-makers (Stufflebeam & Coryn, 2014).

The CIPP model considers four dimensions: context, input, process, and product. The uniqueness of the model is that it is a type of evaluation related to the decision-makers which concerns with the planning and operation of a program. This model also views the program under evaluation as a system and forms an evaluation model that has some advantages compared to other evaluation models (Forouzandeh, Riazi, & Sadighi, 2008; Sukardi, 2014; Zhang et al., 2011). The advantages are: (1) it gives a very detailed description of a project, (2) it has the potentiality for use in the area of formative and summative evaluation, (3) it is more comprehensive in filtering information, and (4) it is capable of giving a good basis in making decisions and policies in writing a further program. Another advantage is that the CIPP model provides a comprehensive evaluation format in every stage of evaluation.

The CIPP evaluation model is designed to evaluate just certain aspects, but it can be used comprehensively to see various things related to a program to improve programs, including a development program. The CIPP evaluation model approach is representative enough in probing things related to the implementation of a program. The CIPP evaluation model can be used structurally and significantly. It can evaluate the effectiveness of the assessment formatively and summatively, and also has the ability in solving problems that occur (Hasan, Yasin, & Yunus, 2015).

The result of the CIPP evaluation model can be used as the basis for making decisions in four types of decision making:

(1) planning (that influences the selection of the objective of the activity), (2) structuring (that influences the optimal strategy and design of the procedures in attaining the objective), (3) implementation (that provides tools for implementing the program and improving the existing program), and (4) recycling (whether a program of activity needs to be continued, changed, or stopped).

The evaluation of the setting (context evaluation) will produce information on the need (to what extent deviations have occurred between what is expected and what has been realized through the activity program). The evaluation of the supporting capacity (input evaluation) stresses on the provision of information on the strengths and weaknesses of the strategy and procedure that have been selected in the effort to achieve the predetermined objective. Process evaluation stresses on the "what" activity that has been done in the program, "who" the persons appointed as the persons in charge, and "when" the activity will end. On the other hand, product evaluation stresses on to what extent the result that has been attained is in accordance with the desired objective, and whether an activity needs to be stopped, continued, improved, etc. Product evaluation is aimed at relating information on the result to the objective, setting, and process that has been determined, supporting capacity, and process that has been determined beforehand in the program implementation (Arikunto, 2009; Stufflebeam & Shinkfield, 2007). Many researchers have shown the effectiveness of the CIPP evaluation model for evaluating educational program and training program (Ariawan, Sanjaya, & Divayana, 2016; Arikunto, 2009; Waluyati, 2012). The result of the CIPP evaluation can give some guidance on what has been done and what has not been done, whether the teaching process that has been designed and the materials given by the teacher can be understood by the students, and are in accordance with the content standard of the teaching implementation.

A school or college as a system is composed of the components of context, input, process, output, and outcome. Context influences input, input process, process output, and output outcome. In a system, there are formed some subsystems that are in a synergy supporting each other in the attainment of a program (Nasution, 2001).

Politeknik Negeri Bali, or Bali State Polytechnic (BSP) is a vocational educational institution in Bali. Its vision is to become a leading college that produces professional graduates with an international competitive advantage in 2025, while one of its missions is to prepare reliable human resources oriented toward the market demand in engineering and commerce with tourism as its superior field, in an effort to improve, modify, and terminate the program. In its progress, there is a need to evaluate in a comprehensive, systematical, and diagnostically way the planning and implementation of the program, involving the teaching staffs at BSP. This evaluation finds out whether the implementation of the teaching has been oriented toward the national standard of education as stipulated by the Regulation of the Minister of Education and Culture No. 49 of 2014 about the national standard of higher education. The appropriate model of evaluation uses the one developed by Stufflebeam, that is, on the components of context, input, process, and product (CIPP).

Research Method

This study is an evaluation research using the CIPP model conducted at BSP in 2017-2018. The subjects of the study consisted of all the polytechnic members: heads of departments, lecturers, students, administration staffs, and technicians, determined by Krejcie and Nomogram Harry King's table based on a 5% error (Sugiyono, 2012) The sample for the context, input, and process is 172 people. As for the product, there were 337 students.

The data were collected using context, input, process, and also product instruments that have been proven to have

high levels of validity and reliability. The level of the content validity of each instrument at the minimum was 0.60, and that of the highest was 0.90. The data that have been collected were analyzed quantitative-descriptively using Z-Score and T-Score as follows (Arikunto, 2012) T*score* = 50 + 10Z and $Z_{scor} = \frac{x - M}{SD}$ (Sudjana, 2012).

The data from the result of analysis from each variable were compared to the real mean. Then, they were classified based on their tendency into five categories with the ideal normal curve theoretic norm, as follows (Mardapi, 2012).

Mi + 1.5 SDi < x ≤ Mi + 3SDi (very high)
Mi + 0.5 SDi < x ≤ Mi +1.5 SDi (high)
Mi – 0.5 SDi< x ≤ Mi + 0.5 SDi (medium)
Mi – 1.5 SDi < x ≤ Mi – 0.5 SDi (low)
Mi – 3 SDi < x ≤ Mi – 1.5 SDi (very low)

Notes:
Mi = ½ (maximum score + minimum score)
SDi = 1/6 (maximum score – minimum score)

The effectiveness of the program implementation in each variable was classified with the percentage of the attainment. The classification criteria are based on the following hierarchy (Ebel & Frisbie, 1972).

A = (81 - 100) % : very high
B = (61 - 80) % : high
C = (41 - 60) % : fair
D = (21 - 40) % : low
E = (0 - 20) % : very low

The T-Score determines the qualification of each component. If the T-Score > 50, it is positive (+), and if the T-Score < 50, it is negative (-). To find out the result of each component, the number of positive (+) scores and that of the negative (-) ones are computed. The number of the positive scores is more or the same as that of the negative scores, the result is positive or ∑ the (+) scores ≥ the (-) scores = (+), and vice versa, if the number of the positive scores is less than that of the negative scores, the result is negative or ∑ the (+) scores < the (-) scores = (-).

To find out the level of effectiveness, an analysis was made of the components of context, input, process, and product using Glickman's quadrant model (Glickman, Gordon, & Ross-Gordon, 2001) which was divided into four quadrants (Gregory, 2004; Sahertian, 2000). If the analysis indicates that the results of all the components are positive (+), then it is located in quadrant I, meaning that the implementation is effective. On the other hand, if it shows that all of the results are negative (-) and located in quadrant IV, it means that it is not effective. If the analysis shows that one or two components have a negative (-) result, then it is located in quadrant II, and it means the implementation is effective enough. Meanwhile, if the analysis shows that three of the components are negative (-), then it is located in quadrant III, which means the implementation is less effective (Gregory, 2004).

Thus, there will be four models of the classification of the level of effectiveness of teaching at BSP, namely (1) if the CIPP position shows + + + +, then the implementation of the teaching program will fall into the effective category; (2) if the CIPP position shows - - - - , then the implementation of the teaching will belong to the ineffective category; (3) if the CIPP position shows - + + + or + - + + or + + - + or + + + - or + + - - or - + - + or - - + + or + - + - or - + + - or + - - +, then the implementation of the teaching program will fall into the effective enough category; and (4) if the CIPP position shows + - - - or - + - - or - - + - or - - - +, then the implementation of the teaching program will fall into the less effective category.

Findings and Discussion

The result of the measurement of the components of context, input, process, and product is presented in Table 1. Based on Table 1, the data of the result of measurement of the context variable tend to center around the score of 104.72 (high). The score with the highest frequency is 104, and the median is 105, the highest score is 116,

the lowest is 89, the standard deviation is 4.4, with the score variation of 16.44. The data of the input variable tend to center around the score of 142.41 (very high). The score with the highest frequency is 137, and the median is 140, the highest score is 167, the lowest is 121, the standard deviation is 8.28, with the score variation of 68.58. The data of the process variable show that they tend to center around 118.15 (very high). The score with the highest frequency is 112, and the median is 119, the highest score is 134, the lowest is 99, the standard deviation is 7.20, with the score variation of 51.79. On the other hand, the data of the product variable tend to center around the score of 14.93 (enough). The score with the highest frequency is 15, and the median is 14.6, the highest is 20.92, the lowest is 11.13, the standard deviation is 2.04, with the score variation of 4.17.

The results of the T-score are presented in Table 2. Based on Table 2, all the components of the context variable produce positive (+) scores. It shows that the indicators that cover the teaching plan, vision, and missions, and learning environment have met the expectation. The difference between the (+) score and the (-) score is positive (+), the score of vision and missions is positive (+), and that of the learning environment is positive (+). There is a match between reality and theory. Hence, in general, the context variable has supported the implementation of the teaching program. The components of the teaching plan, vision and missions, and learning environment contribute significantly to the implementation of the teaching program. The teaching plan obtains a score of 79.74% (good). The construction has followed the principles, goals, and functions of the teaching plan. An appropriate and effective teaching plan affects the success of the teaching program implementation (Majid, 2009).

Table 1. Description of the Results of the Measurement of the Components of Context, Input, Process and Product of the Teaching Implementation at BSP

Statistic	Context	Input	Process	Product
N	172	172	172	337
Mean	104.72	142.41	118.15	14.93
Median	105.00	140.00	119.00	14.6
Mode	104	137	112	15
Std. Deviation	4.40	8.281	7.197	2.04
Variance	19.38	68.582	51.790	4.173
Range	27	46	35	9.79
Minimum	89	121	99	1113
Maximum	116	167	134	20.92

Table 2. Recapitulation of the Result of the T Score Analysis of the Context Variable Component

No	Component	Category of Frequency			Result
		Σf (+)	Σf (-)	Σf (+) %	
1.	Teaching Plan	106	66	61.63	positive (+)
2.	Vision and Missions	89	83	51.74	positive(+)
3.	Learning Environment	93	79	54.06	positive (+)
Total		288	228		positive (+)
Result					**Effective**

The institution's vision and missions obtain a score of 88.89% (very good). The vision and missions have been able to be identified, understood, and comprehended by all of the staff members of BSP. This result supports the opinion of Hamdan (2001) that good vision and missions bring about a conducive organizational atmosphere, enthusiasm, and desire to participate in an organization. Such kind of condition encourages the staff to become more productive without any feeling of strain. Achievement learning is 83.99% (very good).

The learning environment at BSP is very good and conducive, both in the laboratory and classroom. This result supports and plays an essential role in a successful education (Purwanto, 2004). This result also confirms the finding of some researches (Latief, 2014; Menrisal, 2014), that the learning environment contributes positively and significantly to success in learning.

An appropriate and effective teaching plan directs the learning process toward the attainment of the learning objectives. Good vision and missions which are supported by a good and conducive learning environment support the staff to work more productively, so that, viewed from the components of context, the teaching program at BSP is effective.

The evaluation of the input variable covers four indicators: curriculum, semester teaching plan, infrastructure and facilities, and human resources. The recap of the result of the T-scores can be seen in Table 3.

Based on Table 3, it is clear that the input components, including: the curriculum, semester teaching plan, and infrastructure and facilities obtain negative scores. On the other hand, human resources components obtain positive scores. Not all of the indicators of the input variable meet the expectation. The components of the curriculum, semester teaching plan, infrastructure, and facilities have not met the expectation by the fact that although all the study programs have got a lesson plan for all courses. Besides, there is no guideline on how to develop the curriculum, the statements of the attainment of the learning achievement tend not to refer to the statement of the graduate's standard competencies stipulated in the national standard of higher education. The semester teaching plans which are designed by the lecturers do not satisfactorily meet the national standard of higher education.

Semester teaching plans play an important role in interactions between the students and lecturer in the learning process (Amador & Lamberg, 2013). Furthermore, semester teaching plans are efforts to estimate what actions will be made by the lecturers in the teaching activities. When they are designed well, one can guarantee that half of the activities have been completed (Mulyasa, 2008). The writing of the teaching plans should meet the principles of the semester teaching plans and be implemented in accordance with the characteristics of the courses as well as the students. The procedure of how to write teaching plans have been stipulated in the Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015 on the national standard of higher education.

Table 3. Recapitulation of the Result of the T-Score Analysis of the Input Variable

No	Component	Category Frequency			Result
		Σf (+)	**Σf** (-)	**Σf** (+) %	
1	Curriculum	83	89	48.26	negative (-)
2	Semester Teaching Plan	81	91	47.09	negative (-)
3	Infrastructure and Facilities	80	92	46.51	negative (-)
4	Human resourcees	90	82	52,32	positive(+)
	Total	334	354		negative (-)
	Result				**Less effective**

The infrastructure and facilities obtain a negative score. They have not met the expectation, because, in every department, students have not got the printed learning materials in the form of textbooks and job sheets as the guides for them to learn. Part of the learning materials is available online through BSP e-learning, but it is difficult to access this facility, both by the students and the lecturers. The media that support the curriculum, such as LCD projector, in every classroom, do not work well.

The educational infrastructure and facilities are required in teaching and learning process, both movable and fixed ones. This component is one of the essential supports in implementing the teaching process. If the infrastructure and facilities are available, adequate, and made use of effectively, they can contribute to the students' academic achievement. Thus, adequate infrastructure and facilities are the keys to lecturers' performance success. Some research findings (Akomolafe & Adesua (2016); Ayeni & Adelabu (2012); Jannah & Sontani (2018)) prove that infrastructure and facilities affect and contribute significantly to the lecturers' performances, the students' learning motivation, and academic achievement. Hence, when the infrastructure and facilities are not adequate or have not been used effectively, they can become a hindrance to the lecturers' performances and the improvement of the students' achievement.

The human resources component obtains a positive score. It means that human resources, particularly the lecturers at BSP, have met the expectation. They have contributed positively in implementing the teaching program. It supports the finding in the study conducted by Heyneman and Loxley in 1983 (Supriadi, 2005; Widoyoko & Rinawati, 2012), that teachers give a contribution to learning achievement by 34%. Sudjana (2002) states that the teacher's performance influences 6.6% of the learning achievement. Thus, human resources correlate positively with the quality of the teaching process. The input variable does not support the implementation of the teaching program implementation at BSP since the input components like semester teaching plans, and infrastructure and facilities do not contribute satisfactorily to the teaching program implementation.

The evaluation of the process variable covers the teaching contents, the teaching and learning activities, and assessment. The recap of the result of the T-score analysis is presented in Table 4. Based on Table 4, the analysis of the T-score of the process variable components (contents, teaching activities) obtains a positive score, while the assessment component obtains a negative score. The contents and teaching activities have met the expectation and contributed to the process variable. However, the assessment component has not met the expectation and contributed less satisfactorily to the process variable. In addition, 51.4% of the assessment the lecturers implemented has not met the evaluation standard of teaching outlined in the Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015, especially concerning the principles, techniques,

Table 4. Recapitulation of the Result of the Analysis of the Components of the Process Variable

Component	Category of Frequency			Result
	Σf (+)	Σf (-)	Σf (+) %	
Content	92	80	53.48	positive (+)
Learning activities	101	71	58.72	positive (+)
Assessment	83	89	48.6	negative (-)
Total	276	240		positive (+)
Result				**Effective enough**

and instruments as well as the procedures of assessment. Assessment is done separately with the teaching activities, namely in the middle of the semester and the end of the semester. This condition is in contradiction to the statement of Badmus (2007) and Sani (2016) that an implementation of a proper assessment should follow a good teaching quality. Assessment is an integral part of the teaching process, and should be implemented in integration with the teaching process and conducted continually.

The contents and activities of good teaching, when it is unsupported by a good and well-planned evaluation, will make the teaching ineffective. In support of the findings of Darma (2018) and Darma, Candiasa, Sadia, and Dantes (2018), learning achievement is significantly influenced by an evaluation system implemented. Hence, viewed from the input component, the teaching program at BSP is sufficient enough.

The evaluation of the product variable covers the academic area of the students' learning achievements as shown by cumulative grade point average (GPA) and non-academic achievement. The recap of the result of the T-score of the product components is presented in Table 5.

Based on Table 5, all of the components of the product obtain a negative score. In general, the product variable does not support the teaching implementation at BSP. The academic and the non-academic components have not met the expectation and contributed less significantly to the product variable. The objective of the teaching program at BSP has not attained optimally. Yusuf (2015) explains that product evaluation is done at the end of an activity and is intended to measure the achievement of the objectives that have been determined before based on a particular standard or criteria. This finding is supported by theories (Arikunto, 2009; Stufflebeam & Shinkfield, 2007) and findings (Warju, 2016; Zhang et al., 2011) that the final evaluation (product) relates information on the final achievement to the objectives, context, input, and process that have been implemented before it.

The failure to achieve the objectives of teaching at BSP is since the teaching is less supported by a good and appropriate curriculum, semester teaching plan, and infrastructure as well as facilities suitable for students' needs. The recap of the T-scores related to the context, input, process, and product variables is presented in Table 6.

Table 5. Recapitulation of the Result of the Analysis of the Components of the Product Variable

Component	Category of Frequency			Result
	Σf f (+)	Σf (-)	Σf f (+) (%)	
Academic	146	191	43.32	negative (-)
Non-academic	166	171	49.26	negative (-)
Total	312	362		negative (-)
Result				**Ineffective**

Table 6. Relation Among the Variables of Context, Input, Input and Product

Variable	Category of Frequency					Result	Remarks
	f (+)		f (-)				
	Σf (+)	%	Σf (-)	%	Diff. %		
Context	89	51.74	83	48.26	3.48	+	**Positive**
Input	78	45.33	94	54.67	- 9.34	-	**Negative**
Process	104	60.47	68	39.53	20.94	+	**Positive**
Product	166	49.26	171	50.74	-1.48	-	**Negative**
Result						**+ - + -**	**Effective enough**

Notes: T ≥ 50 means that the component = (+); T < 50 means that the score component = (-).

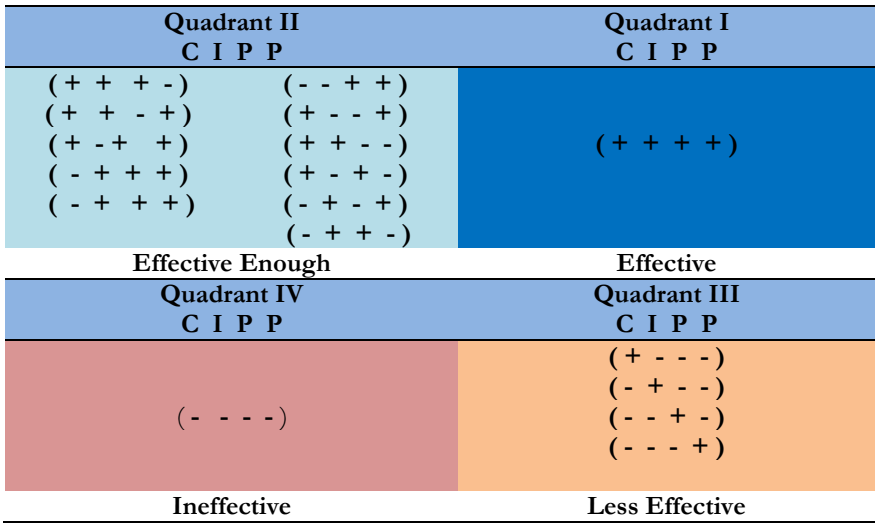

Figure 1. Glickman's Quadrants of the Effectiveness of the Implementation of Teaching at BSP

Table 6 shows that the T-score of the context variable is positive (+), that of input is negative (-), that of the process is positive (+), and that of the product is negative (-). The whole the scores are (+ - + -). These results are verified into the Glickman quadrant, as shown in Figure 1.

The pattern of the CIPP scores of **+ - + -** in Figure 1 falls into Quadrant II. This position shows that the condition of teaching at BSP is sufficient enough. It is because, simultaneously, the variables of context, input, process, and product have given different supports to the effectiveness of the teaching. The context and process have strong supports, while the input and product have weak supports. It supports the result of an evaluation by Gunung and Darma (2018) that the implementation of the teaching program at BSP, viewed from the variables of context, input, process, and the product, is sufficient enough. It confirms the studies by Kavgaoglu and Alci (2016) and Riptiani, Manuaba, and Putra (2015) that the effectiveness of each component of context, input, process, and product influences the level of effectiveness of the program evaluated.

The result of the evaluation agrees with Kaufman and Thomas (in Mukhadis, 2013) that the aspect of the context in the program evaluation can affect the process of the implementation and attainment of the objectives of a program that has been designed. Komariah and Triatna (2005) state that effectiveness is a dimension of the objective of the management that focuses on the result, objectives, and target expected. A competent institution of higher education is an institution that determines its success in terms of input, process, context, and product characterized by the high quality of the components of the system. Arikunto (2009) states that the CIPP model of evaluation sees the program evaluated as a system. The integration of context, input, and process supports and at the same time, influences the product or output. The level of effectiveness of each component influences the effectiveness of the program. The presence of constraints in each component will, at the same time, influence the effectiveness of the program. Hence, the level of effectiveness of the teaching program at BSP is effective enough because of the presence of constraints in the components of context, input, and product.

The effectiveness of the teaching at BSP viewed from context is effective. The factors that are related to the context variable are teaching plan, vision and missions, learning environment which has met the expectation and contributed to the implementation of the teaching program. There is no constraint on the context variable. In relation to this absence of constraint, the

institution needs to maintain and even to improve the effectiveness of the context both in relation to the teaching plan, vision, and missions and learning environment and to enhance the effectiveness of the implementation of the teaching program.

The planning of the teaching program is something that needs much improvement. The improvement of the quality of the teaching can be started from the teaching plan (Dwiyogo, 2018) since the teaching plan serves as the reference for the lecturers in implementing the teaching activities that are more oriented, which will proceed efficiently and effectively. Işman (2011) states that the planning of a teaching program is an essential stage in improving the quality of teaching. Its design has to be based on active learning. The success of a teaching process is determined by a mature plan. When a plan is made well, it means that half of the success has been attained, the rest lies in its implementation.

The implementation of the teaching program at BSP viewed from the input is low active. This condition is caused by the fact that some of its components are not effective. The curriculum, semester teaching plan, infrastructure, and facilities do not contribute significantly to the implementation of the teaching program. Thus, these components have some constraints in the input variable.

The constraints that are originated from the curriculum include the contents presented, namely the relation between the graduates' learning achievements and the course learning achievements. The learning activities that are designed do not support the graduates' learning achievements and the course learning achievements, the statements of the indicators of achievement in the course learning achievements are not clear, and the types of evaluation implemented are still oriented toward the product evaluation.

The constraint which is originated from the semester teaching plan is related to its implementation. The level of the impracticality of the semester teaching plan is still high or 69.5%. In implementing the teaching process, the lecturers tend not to bring the semester teaching plans that they have prepared. On the other hand, the semester teaching plan is a projection of the activities which will be done by the lecturer with the students in the teaching and learning process in the classroom and is an integral part that cannot be separated from the teaching. Every lecturer who will teach should write a semester teaching plan first. A well written semester teaching plan can help the lecturer in implementing the teaching in the classroom effectively and efficiently.

The curriculum, as a design, consists of four elements: learning achievements (learning objectives), materials that have to be acquired, teaching strategies to achieve the objectives, and the evaluation system to evaluate the achievement of the objectives. The statement of the compulsory graduate's learning achievement refers to the description of the learning achievements of the Indonesian National Qualifications Framework graduates, and the achievements are equivalent to the level of qualification at the Indonesian National Qualifications Framework. The statement contains three elements, namely attitude, knowledge, and skills, and the latter is divided into general and specific skills by the graduates of the institute of higher education (Regulation of the Minister of Research, Technology, and Higher Education No. 44, 2015).

On the other hand, the constraints also come from the infrastructure and facilities as well as the media. The e-learning facility does not function maximally. The textbooks as the handbooks for the lecturers and students do not satisfactorily meet the students' needs. In this context, the institution needs to improve the effectiveness of the input concerning the curriculum, semester teaching plan, infrastructure and facilities, and human resources so that it can enhance the effectiveness of teaching.

The process variable, in general, has supported the implementation of the teach-

ing program. The effectiveness of the teaching program at BSP viewed from the process component is effective. The factors that are related to the process variable are the contents of the teaching program, teaching activities, and assessment which have contributed to the implementation of the teaching program. However, there is a component that is still less effective, namely evaluation. The assessment component does not contribute much to the teaching implementation. The assessment system implemented has not met the expectation. It has not met the standard of teaching evaluation and become a constraint that is originated from the process variable.

The standard of teaching assessment is the minimal criterion of process and product evaluation in the effort of meeting the graduate's learning achievement. The principles of assessment include educative, authentic, objective, accountable, and transparent principles that are used integrated. The authentic principle orientations are toward a sustainable learning process and learning achievement, which reflect the students' ability (Regulation of the Minister of Research, Technology, and Higher Education No. 44, 2015).

Assessment and teaching are two activities that support each other. Improvement in teaching quality can be made by improving its evaluation system (Mardapi, 2014). A proper assessment system will encourage the lecturers in determining a good teaching strategy in motivating students to learn better.

In this context, the institution needs to maintain and even to improve the effectiveness of the process, especially in relation to the implementation of the assessment so that in the end, the teaching program can be improved. To improve the quality of education, one needs to improve the assessment system implemented. Good-quality teaching should be followed by the implementation of a good assessment (Sani, 2016).

The teaching at BSP viewed from the product component is not effective since, in general, the product variable less satisfactorily supports the teaching program implementation. The components or academic and non-academic learning achievements contribute less satisfactorily to the implementation of the teaching program. Hence, these components become the constraints originated from the product variable in the teaching program implementation.

The less satisfactory contribution of the learning achievement in the academic and non-academic activities to the effectiveness of the teaching program is caused by the ineffectiveness of the input variable or curriculum, semester teaching plan, and also infrastructure and facility components. Another cause of less effectiveness comes from the evaluation component. Therefore, these components became the constraints in the product variable in supporting the teaching program implementation. In relation to these constraints, the institution needs to improve the effectiveness of the product, thereby the effectiveness of the teaching program.

Conclusion

The teaching program at BSP viewed from the interrelatedness of context, input, process, and product components is effective enough. The constraints in the teaching implementation at BSP consist of those in the context, input, process, and product components. The context variable consists of the teaching plan. The input variable consists of the curriculum, the semester teaching plan, and the infrastructure and facilities. The process variable consists of the assessment system. The product variable consists of the quality of learning achievement, both academic and non-academic.

It is recommended to the decision-makers at BSP to (1) review the implemented curriculum, especially concerning the statement of learning achievement; (2) adapt the activities of teaching in the classroom and laboratory to the semester teaching plan that has been written; (3) encourage the lecturers to prepare printed learning materials according to the students'

learning needs; and (4) change the evaluation system implemented into the evaluation system oriented to the process as stipulated in the Regulation of the Minister of Education and Culture No. 49 of 2014 on the National Standard of Higher Education.

Acknowledgment

The researchers would like to express their gratitude to the Directorate General of the Strengthening of Research and Development, the Ministry of Research, Technology, and Higher Education who financed this research.

References

Akomolafe, C. O., & Adesua, V. O. (2016). The impact of physical facilities on students' level of motivation and academic performance in senior secondary schools in South West Nigeria. *Journal of Education and Practice*, *7*(4), 38–42. Retrieved from https://iiste.org/Journals/index.php/JEP/article/view/28745

Amador, J., & Lamberg, T. (2013). Learning trajectories, lesson planning, affordances, and constraints in the design and enactment of mathematics teaching. *Mathematical Thinking and Learning*, *15*(2), 146–170. https://doi.org/10.1080/10986065.2013.770719

Ariawan, I. P. W., Sanjaya, D. B., & Divayana, D. G. H. (2016). An evaluation of the implementation of practice teaching program for prospective teachers at Ganesha University of Education based on CIPP-forward chaining. *International Journal of Advanced Research in Artificial Intelligence (IJARAI)*, *5*(2), 1–5. https://doi.org/10.14569/IJARAI.2016.050201

Arikunto, S. (2009). *Evaluasi program pendidikan*. Jakarta: Bumi Aksara.

Arikunto, S. (2012). *Dasar-dasar evaluasi pendidikan*. Jakarta: Bumi Aksara.

Ayeni, A. J., & Adelabu, M. A. (2012). Improving learning infrastructure and environment for sustainable quality assurance practice in secondary schools in Ondo State, South-West, Nigeria. *International Journal of Research Studies in Education*, *1*(1), 61–68. https://doi.org/10.5861/ijrse.2012.v1i1.20

Badmus, G. A. (2007). *Changing nature of technical and vocational education and students' assessment methods*. Retrieved from https://www.iaea.info/abstract_files/

Darma, I. K. (2018). Improving mathematical problem solving ability through problem-based learning and authentic assessment for the students of Bali State Polytechnic. *Journal of Physics: Conference Series*, *953*, 012099. https://doi.org/10.1088/1742-6596/953/1/012099

Darma, I. K., Candiasa, I. M., Sadia, I. W., & Dantes, N. (2018). The effect of problem based learning model and authentic assessment on mathematical problem solving ability by using numeric ability as the covariable. *Journal of Physics: Conference Series*, *1040*, 012035. https://doi.org/10.1088/1742-6596/1040/1/012035

Dwiyogo, W. D. (2018). *Pembelajaran berbasis blended learning*. Depok: PT Rajagrafindo Persada.

Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Ebel, R. L., & Frisbie, D. A. (1972). *Essentials of educational measurement*. Englewood Cliffs, NJ: Prentice-Hall.

Forouzandeh, E., Riazi, A. M., & Sadighi, F. (2008). TEFL program evaluation at master's level in Iran. *Teaching English Language (Teaching English Language and Literature Society of Iran)*, *2*(6), 71–100. Retrieved from https://www.sid.ir/en/journal/ViewPaper.aspx?ID=135337

Glickman, C. D., Gordon, S. P., & Ross-Gordon, J. M. (2001). *Supervision and instructional leadership: A developmental approach* (5th ed.). Needham Heights, MA: Allyn & Bacon.

Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Boston, MA: Allyn & Bacon.

Gronlund, N. E., & Linn, R. L. (1990). *Measurement and evaluation in teaching* (6th ed.). New York, NY: Macmillan.

Gunung, I. N., & Darma, I. K. (2018). Evaluation of learning programs at Bali State Polytechnic (BSP). *Proceedings of the International Conference on Science and Technology (ICST 2018)*, 1129–1134. https://doi.org/10.2991/icst-18.2018.228

Hamdan, Y. (2001). Pernyataan visi dan misi perguruan tinggi. *Mimbar: Jurnal Sosial Dan Pembangunan*, *17*(1), 90–103. https://doi.org/10.29313/mimbar.v17i1.34

Hasan, A., Yasin, S. N. T. M., & Yunus, M. F. M. (2015). A conceptual framework for mechatronics curriculum using Stufflebeam CIPP evaluation model. *Procedia - Social and Behavioral Sciences*, *195*, 844–849. https://doi.org/10.1016/j.sbspro.2015.06.324

Işman, A. (2011). Instructional design in education: New model. *Turkish Online Journal of Educational Technology - TOJET*, *10*(1), 136–142. Retrieved from http://www.tojet.net/articles/v10i1/10114.pdf

Jannah, S. N., & Sontani, U. T. (2018). Sarana dan prasarana pembelajaran sebagai faktor determinan terhadap motivasi belajar siswa. *Jurnal Pendidikan Manajemen Perkantoran (JPManper)*, *3*(1), 63–70. Retrieved from https://ejournal.upi.edu/index.php/jpmanper/issue/view/952

Kavgaoglu, D., & Alci, B. (2016). Application of context input process and product model in curriculum evaluation: Case study of a call centre. *Educational Research and Reviews*, *11*(17), 1659–1669. https://doi.org/10.5897/ERR2016.2911

Komariah, A., & Triatna, C. (2005). *Visionary leadership menuju sekolah efektif*. Jakarta: PT Bumi Aksara.

Latief, A. (2014). Pengaruh lingkungan sekolah terhadap hasil belajar Pendidikan Kewarganegaraan pada peserta didik di SMK Negeri Paku Kecamatan Binuang Kabupaten Polewali Mandar. *Pepatudzu: Media Pendidikan Dan Sosial Kemasyarakatan*, *7*(1), 13–26. https://doi.org/10.35329/fkip.v7i1.11

Majid, A. (2009). *Perencanaan pembelajaran*. Bandung: PT Remaja Rosdakarya.

Mardapi, D. (2012). *Pengukuran, penilaian, dan evaluasi pendidikan*. Yogyakarta: Nuha Medika.

Mardapi, D. (2014). Authentic assessment. *The Paper Presented at HEPI Conference*. Denpasar, Bali.

Menrisal, M. (2014). Kontribusi lingkungan belajar terhadap hasil belajar sistem operasi siswa kelas X Teknik Komputer dan Jaringan (TKJ) di Sekolah Menengah Kejuruan (SMK) Negeri 3 Pariaman semester ganjil tahun ajaran 2014/2015. *Jurnal KomTekInfo Fakultas Ilmu Komputer*, *1*(2), 77—83. Retrieved from http://lppm.upiyptk.ac.id/komtekinfo/index.php/KOMTEKINFO/article/download/10/10

Mukhadis, A. (2013). *Evaluasi program pembelajaran bidang teknologi*. Malang: Bayumedia.

Mulyasa, M. (2008). *Menjadi guru profesional*. Bandung: PT Remaja Rosdakarya.

Nasution, M. N. (2001). *Manajemen mutu terpadu (Total quality management)*. Jakarta: Ghalia Indonesia.

Purwanto, M. N. (2004). *Psikologi pendidikan*. Bandung: PT Remaja Rosdakarya.

Regulation of the Minister of Education and Culture No. 49 of 2014 on the National Standard of Higher Education. , (2014).

Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015 on the National Standard of Higher Education. , (2015).

Riptiani, K. M., Manuaba, I. B. S., & Putra, M. (2015). Studi evaluasi implementasi Kurikulum 2013 ditinjau dari CIPP pada sekolah dasar negeri di wilayah pedesaan Kabupaten Badung. *Mimbar PGSD Undiksha*, *3*(1). https://doi.org/10.23887/jjpgsd.v3i1.5015

Sahertian, P. A. (2000). *Konsep dasar dan teknik supervisi pendidikan dalam rangka pengembangan sumber daya manusia*. Jakarta: Rineka Cipta.

Sani, R. A. (2016). *Penilaian autentik*. Jakarta: PT Bumi Aksara.

Stufflebeam, D. L. (2003). The CIPP model for evaluation. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International Handbook of Educational Evaluation* (vol. 9, pp. 31–62). Dordrecht: Springer.

Stufflebeam, D. L., & Coryn, C. L. S. (2014). *Evaluation, theory, models, and applications* (2nd ed.). San Francisco, CA: John Wiley & Sons.

Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, & applications*. San Francisco, CA: Jossey-Bass.

Sudjana, N. (2002). *Dasar-dasar proses belajar dan mengajar*. Bandung: Sinar Baru Algesindo.

Sudjana, N. (2012). *Penilain hasil proses belajar mengajar*. Bandung: Remaja Rosda Karya.

Sugiyono. (2012). *Metode penelitian pendidikan: Penelitian kuantitatif, kualitatif, dan R&D*. Bandung: Alfabeta.

Sukardi, S. (2014). *Evaluasi program pendidikan dan pelatihan*. Jakarta: Bumi Aksara.

Supriadi, D. (2005). *Mengangkat citra dan martabat guru*. Yogyakarta: Adicita Karya Nusa.

Waluyati, I. (2012). Evaluasi program pembelajaran ilmu pengetahuan sosial SMP/MTs di Kota Bima. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *16*(1), 260–280. https://doi.org/10.21831/pep.v16i1.1117

Warju, W. (2016). Educational program evaluation using CIPP model. *Innovation of Vocational Technology Education*, *12*(1), 36–42. https://doi.org/10.17509/invotec.v12i1.4502

Widoyoko, S. E. P., & Rinawati, A. (2012). Pengaruh kinerja guru terhadap motivasi belajar siswa. *Cakrawala Pendidikan*, *31*(2), 278–289. https://doi.org/10.21831/cp.v5i2.1563

Yusuf, A. M. (2015). *Asesmen dan evaluasi pendidikan: Pilar penyedia informasi dan kegiatan pengendalian mutu pendidikan*. Jakarta: Prenada Media Group.

Zhang, G., Zeller, N., Griffith, R., Metcalf, D., Williams, J., Shea, C., & Misulis, K. (2011). Using the context, input, process, and product evaluation model (CIPP) as a comprehensive framework to guide the planning, implementation, and assessment of service-learning programs. *Journal of Higher Education Outreach and Engagement*, *15*(4), 57–83. Retrieved from https://openjournals.libs.uga.edu/jheoe/article/view/901

Jurnal Penelitian dan Evaluasi Pendidikan
Volume 23, No 2, December 2019 (184-196)
Online: http://journal.uny.ac.id/index.php/jpep

EVALUATION OF THE IMPLEMENTATION OF INTEGRATIVE THEMATIC LEARNING: A QUALITATIVE RESEARCH APPROACH PHENOMENOLOGY

Fitri Indriani
Universitas Ahmad Dahlan

Atiaturrahmaniah
Universitas Hamzanwadi

Abstract

This research aims at evaluating the implementation of integrative thematic learning. This research is an evaluation of the CIPP model with a qualitative approach phenomenology. Data were collected through observations, in-depth interviews, and documentation. The informants in the study were principals, teachers, and students. Data were analyzed using a qualitative method that is descriptive and critical. The results show that the implementation of integrative thematic learning in Muhammadiyah Suronatan Elementary School Yogyakarta has been carried out quite well, both in planning, implementation, and assessment of processes and learning outcomes. The implementation is success because it has been supported by excellent teachers' competence and adequate facilities and infrastructure supporting the learning process. Besides, teachers also have the motivation and spirit of learning. The implementation of integrative thematic learning of curriculum 2013 has a positive impact on student learning outcomes from both aspects of knowledge, attitude, and skills. From the knowledge aspect, the integrative thematic learning of the curriculum 2013 can deliver the best value to the national exam in the top 10 of Yogyakarta. The attitude aspect of the student spirit in learning shows an increase in students' confidence. As for the skill aspect, students become creative and innovative in both thinking and working. However, constraints are still found: handbooks for students have not been distributed equally by the government, and the lack of teachers' understanding of the way of the subject complaint that corresponds to the sub-theme.

Keywords: *evaluation, integrative thematic learning, phenomenology*

Permalink/DOI: http://dx.doi.org/10.21831/pep.v23i2.27431

Contact: *Fitri Indriani*
fitrimusi2013@gmail.com
Department of Elementary School Teacher Education, Faculty of Education and Teacher Tranining, Universitas Ahmad Dahlan
Jl. Ki Ageng Pemanahan No. 19, Sorosutan, Umbulharjo, Yogyakarta 55162, Indonesia

Introduction

National education is one of the development sectors that has a vision of the realization of education system as a strong social institution, to empower all Indonesians to develop into high-quality human that is intelligent and merciful to God the Almighty, noble, healthy, knowledgeable, capable, creative, independent, democratic, and responsible (Murtikasari, 2013, p. 107).

In realizing quality human beings, the Indonesian government is continually striving to improve the quality of education by implementing policies through curriculum 2013. The curriculum 2013 has been applied in response to today's demands and challenges, both internal and external challenges. The internal challenges are related to the gold generation estimated in 2020-2035 as a demographic bonus. External challenges are related to the current existence of globalization and Indonesian's participation in the international study of Trens in International Mathematics and Science Study (TIMSS) and Program for International Student Assessment (PISA), which are successful (Suharto, 2013, pp. 66–67).

With the changing of the curriculum, it then impacted the learning system, in elementary schools, such as the teaching system implemented in integrative thematic. Thematic learning is essentially not a new thing in the Indonesian education world. In Curriculum 2006 (*Kurikulum Tingkat Satuan Pendidikan* or KTSP), thematic learning has already been applied. Only the system and rules of implementation are different. In Curriculum 2006 (KTSP), thematic learning is only implemented in low-grade, with systems and rules, and teachers are required to design their thematic learning device, both in the syllabus, lesson plan, and theme mapping. In practice, some teachers are confused about implementing thematic learning, so the thematic learning in KTSP has not been done correctly, and tends to be implemented in fragmentary. It has been strengthened by some research results, one of which is by Aryasi (Kristiantari, 2014, p. 463) which reveals that in carrying out thematic learning in KTSP, teachers have not been able to translate the lesson plan and syllabus in the field. Thus, learning is both partial and conventional. Similarly, in judgment, the gems still do lessons instead of each theme.

The problem arises from the implementation of the curriculum 2013 to answer the difficulties of teachers in the field by implementing integrative thematic learning. Integrative thematic learning is carried out at all levels, both low and high class with the system and the rules that the teachers no longer develop their own learning devices as found in the curriculum 2006 (KTSP), but they need a complete handbook with learning tools, from planning to evaluation. It is expected that integrative thematic learning in the curriculum 2013 can be implemented with well.

However, in reality, the teacher is still having trouble. It is because, in the integrative thematic learning of the curriculum 2013, teachers are required to apply active learning with scientific approaches and authentic assessments, while teachers are accustomed to implementing conventional learning and give the students score in the form of numbers.

The curriculum 2013 has been tested on several elementary schools throughout Indonesia, especially in the Special Region of Yogyakarta. There are about 52 elementary schools in both public and private schools, one of them is Muhammadiyah Suronatan Elementary School in Yogyakarta City. Muhammadiyah Suronatan Elementary School implemented the curriculum 2013 since the curriculum's enactment until now. In order to see the success in its implementation, an evaluation in the framework of improvement is critical to be carried out.

Therefore, this research is important to provide information on the implementtation of the integrative thematic learning with curriculum 2013, which includes; (1) context, (2) input, (3) the learning process, (4) output of learning. In addition to these four aspects, this research also discusses the

factors supporting and inhibiting the implementation of integrative thematic learning with curriculum 2013 in Muhammadiyah Suronatan Elementary School, Yogyakarta City.

Curriculum Evaluation

Stufflebeam views evaluation as "a process of delineating, obtaining, and providing useful information for judging decision alternatives" (Oliva, 1992, p. 489). In addition, many experts have essentially developed a curriculum evaluation. According to Suharsimi (2007, p. 24), there are approximately seven models of evaluation expressed by the six experts, namely: (a) Tyler with a purpose-oriented model, (b) Scriven with a freelance model of purpose, (c) Stake with countenance, (d) Michael Scriven with formative–summative model, (e) the CSE-UCLA (CSE stands for the Center for the Study of Evaluation, and UCLA stands for the University of California in Los Angeles), (f) Stufflebeam and friends with CIPP, and (g) Malcolm Provus with gaps.

From those evaluation models, this study uses a CIPP evaluation model proposed by Stufflebeam. The CIPP model has been chosen on the grounds. This model is a more comprehensive, easy, and practical model. The CIPP model is the most widely-known and applied model by evaluators. The CIPP model used in the study was developed by Stufflebeam and friends in 1967 at Ohio State University. The evaluation of the CIPP model consists of four aspects, namely, context, inputs, processes, and products (Darodjat & Wahyudhiana, 2015, p. 5).

Context evaluation is emphasized on the grounds as a basis for determining the objectives of the curriculum, whether it is general, institutional, curricular, or general instructional purposes. Input evaluation has been emphasized on resources (teachers, media, materials, and modules), and the use of strategies (learning strategies, learning experiences, and learning environments) to achieve goals. The evaluation process is aimed at determining the weaknesses of the plan and implementation of the curriculum, as well as at acquiring the information of various program activities as an ingredient in making decisions, such as improvement, enhancement, and development of the curriculum. Product evaluation aims to determine the success of the program in the form of learning outcomes, both attitudes, knowledge, and skills. These elements can be illustrated in the chart, as presented in Figure 1.

Furthermore, each of these elements is the Context, Inputs, Processes, and Products that have the criteria as the relevant standards, as follows. On the Context element, in the vision, mission, and program developed, there is conformity to implement the integrative thematic learning with curriculum 2013. On the Input element, the availability of resources, such as competent teachers in the field, is adequate to succeed in the learning. Facilities and infrastructure,

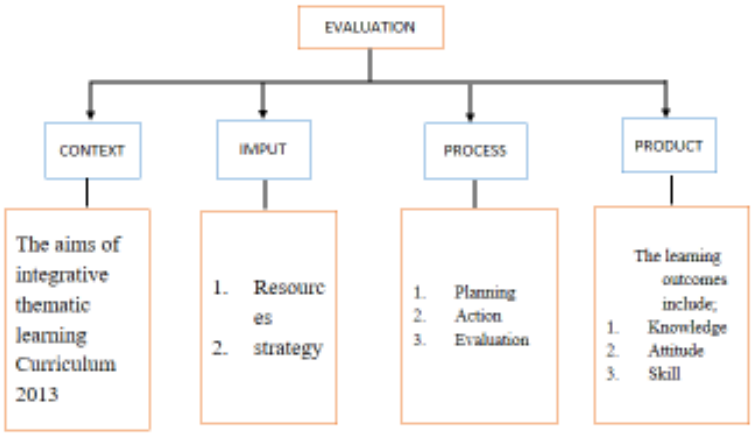

Figure 1. CIPP Evaluation Model

such as learning media, learning environment, and learning resources, are adequate. Besides, the learning process is actively and pleasantly conducted by using scientific approaches, promoting character education, and using authentic judgment. On the Product element, students are not only proficient in knowledge but also good both in morals and skills, so that the graduates produced are not only intelligent intellectually, but also morally to avoid immoral action.

Integrative Thematic Learning

The concept of integrative thematic learning is the development of two educational figures, namely Jacob in 1989, with the concept of interdisciplinary learning, and Fogarty in 1991 with the concept of integrated learning (Majid, 2014, p. 85). Related to integrated learning, Fogarty offers ten models, namely (1) fragmented (fragments), (2) connected (connectedness), (3) nested (nest), (4) sequenced (sequence), (5) shared (part), (6) webbed (spider webs), (7) threaded (strain), (8) integrated (alignment), (9) immersed (cress), and (10) networked (network) (Fogarty, 1991, p. xv).

One of the ten models, which deals with an integrative thematic learning model, is the webbed model. The webbed model is an integrated learning model that uses a thematic approach. Thematic approach, according to Hidayah (2015, p. 37), is a learning approach that integrates several competencies into certain subjects. Ministry of Education and Culture also conveys that integrative thematic learning is a learning approach that combines various competencies from learning content based on the theme in which the system allows students to learn in active, holistic, meaningful, and authentic ways (Julianti & Mawardi, 2018, p. 206). From those two opinions, it can be inferred that integrative thematic learning is a Webb model learning approach that integrates the various competencies of learning content in an active teaching system, which is holistic, meaningful, and authentic.

Integrative thematic learning has several foundations, one of which is a philosophical foundation. It is the first foundation as a design built up in integrative thematic learning supported by three traditions of philosophy, namely: progressivism, constructivism, and humanism.

Progressivism Stream

Progressivism stream has the principle that learning is centered on learners, that emphasizes creativity and activity, and are problem-solving (Assegaf, 2011, p. 20). It goes from thinking that education should be "life" itself, not preparation for life, learning should be directly associated with the child's interest, learning through problem-solving should be precedence from the strict repetition of the subject, the role of the teacher is not to demonstrate, but to guide, the school must enhance the cooperation efforts, not competing, only the treatment that has democratized can actually improve the role of the idea, and child's personality has been freely expressed, and it is necessary to pay attention to the correct growth conditions (Assegaf, 2011, pp. 204–205). This flow rejects the authoritarian learning systems, book-oriented texts, memorization, learning that is limited in class so that the learners are isolated to real life, and promoting punishment in building student discipline. Based on those principles of learning, an educator from the progressivism stream, in this case, a teacher, should have some views on education, including (1) the curriculum is developed tailored to the needs of learners, (2) learners basically have a learning interest, if not frustrated by adults, (3) the task of the teacher is not as a class ruler, but rather as a mentor and advisor for students' learning, (4) learning is conducted not only in classrooms but also outside classes, (5) learning activities should focus on problem solving, (6) the school climate is created democratically and cooperatively (Indriani, 2019, p. 41).

Constructivism Stream

The constructivism stream has the view that knowledge is not something that is already finished but continues to process.

In the learning system, learners give a direct experience of conjugant knowledge gained through contextual study (Majid, 2014, pp. 87–88).

Humanism Stream

This stream looks at learners as a unique creature. The purpose of learning, according to humanistic theory, is human humanizing. The task of educators is to help students develop their unique self-potential. One of the humanistic figures is Abraham Maslow, who is famous for his motivational theory about the hierarchy of basic human needs: physical necessity, sense of security, love and compassion, appreciation, and self-actualization. Regarding this view, then in conducting learning, teachers should pay attention to the basic needs of learners so that motivation and attention are well developed (Maslow, 2017, pp. 70–80).

From those foundations, integrative thematic learning has different characteristics with other subjects. The different characteristics are, among others, (1) being student-centered, (2) providing a direct experience, (3) the separation of subjects is not very clear, (4) presenting the concepts of various subjects, (5) being flexible; (6) using learning principles while playing and having fun (Ahmadi & Amri, 2014, pp. 192–193).

Curriculum 2013

The term curriculum was firstly used in sports at the ancient Greek, derived from the word *curio* and *curere*. At that time, the word curriculum was interpreted as the distance to be traveled by a runner. People termed it with a race place or a running place from start to finish (Raharjo, 2012, p. 15). In its development, then the curriculum is used in education. In this field of education, the curriculum has expanded. The first curriculum sees as a series of subjects, just the material to be given to learners (Asifudin, 2009, p. 93).

The curriculum is now more broadly defined. In Law of Republic of Indonesia No. 20 of 2003 on national education system, it is said that the curriculum is a set of plans and arrangements regarding content and materials and the way it is used as a guideline for organizing activities of learning and teaching (Fadlillah, 2014, p. 15). Understanding the curriculum has two dimensions, namely, as a plan and arrangement about the content and material, and as a way to teach learning activities.

Regulation of the Government No. 32 of 2013 explains that the enforcement of the curriculum 2013 requires a legal basis for changing to the Government Regulation No. 19 of 2005 about national standards of education. The national standards of education in the Regulation of the Minister of Education and Culture consist of eight standards, one of which is the process standard.

The process standard is the criteria for the implementation of learning in the unit of education to achieve the competency standards of graduates. The process standard is developed referring to the standard competency of the graduates and the predefined content standards by the provisions in the Government Regulation Number 19 of 2005 concerning education standards. The learning process in the education unit is organized to be interactive, inspiring, fun, challenging, and motivating to make learners actively participate, to provide ample space, improve creativity, and become independent according to their talent, interests, and physical and psychological development (Suharto, 2013, pp. 208–209).

Thus, the principles of learning used include encouraging learners to find out the problem-solving, encouraging learners to learn with various sources, applying learning with a scientific approach, applying learning with competency-based, learning in integratedly-organized environment, balancing between softskills and hard skills, develop the creativity of learners, utilization of information technology and communication in learning, and paying attention to individual characteristics and the cultural background of learners (Suharto, 2013, pp. 208–209).

Research Method

This research is an evaluation of CIPP model with a qualitative approach of phenomenology that uses the principles of qualitative research of Bogdan and Biklen, namely, done in a natural condition, searching directly to the data source, and the researchers become the key instrument, being descriptive, data collected in the form of words or images, putting more emphasis on the process than the outcome, the analysis is conducted inductively, and putting more emphasis on meaning (Bogdan & Biklen, 1982, pp. 27–29). In collecting the data, the researchers were in the field or at the research site for a relatively long time so that the problems studied related to the issues can be revealed more clearly, accurately, objectively, and in-depth. Meanwhile, for the phenomenological approach used to understand the meaning of human behavior based on understanding, thought subjects were researched instead of the thought researchers.

The research took place at the elementary school of Muhammadiyah Suronatan in Yogyakarta City. Data collection techniques used in the study were observations, in-depth interviews, and also documentation. The informant in this study was principals, teachers, and students. Triangulation was done to obtain data validity. Data were analyzed interactively until they were saturated, referring to Miles and Huberman's technique. The activities carried out were data collection, data reduction, data presentation, and data conclusions drawing or verification (Miles & Huberman, 1994, pp. 10–12) as presented in Figure 2.

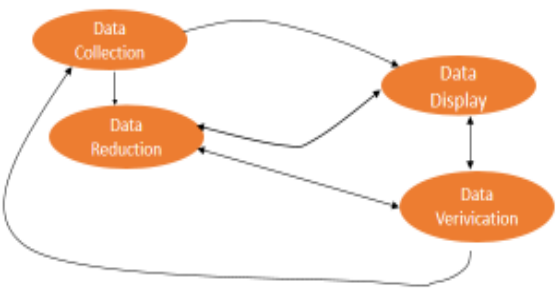

Figure 2. Miles and Huberman Analysis Model

Findings and Discussion

This evaluation study examines four aspects of context, input, process, and product. In addition to these four contexts, supporting factors and the implementation of the integrative thematic learning curriculum 2013 are also discussed.

From the context aspect, there is conformity between the vision, mission, and program developed by the school with the objectives of the implemented integrative thematic learning with curriculum 2013. From the input aspect, the availability of resources is adequate to succeed in the learning, such as competent teachers in the field, facilities, and infrastructure (learning media, learning environment, and learning resources). From the aspect of the learning process, it is conducted actively and pleasantly, using scientific approaches, promoting character education, and using authentic judgment. In terms of the product, not only does it result in intellectually intelligent students, but it is also able to produce morally brilliant students.

Context

In the aspect of context, the implementation of the integrative thematic learning in Muhammadiyah Suronatan Elementary School in Yogyakarta is in line with the curriculum 2013 preparation, which seeks to develop the gold generation to have competence. The main points of the 21st-century learning are collaboration, high-level thinking, creativity, and communication. Therefore, the teacher is encouraged to organize learning with a new paradigm oriented towards the development of creativity, activity, and noble character with scientific approaches and authentic judgment.

It is important to note that the vision of Muhammadiyah Suronatan Elementary School is "forming superior students based on faith, sciences, and technology." The missions conducted include (1) cultivating Islamic values in everyday life, (2) implementing active learning to develop students' maximum potential, (3) developing the entire potential of school citizens to achieve a

level of excellence, (4) improving the faith and mastery of sciences and technology by working with the school citizen and related parties, and (5) improving discipline in various aspects so that students become superior and dignified humans (Documentation, 2018). As an effort to support the vision and missions, the school develops various programs in the form of extracurricular activities, including robotics, computer skills, dance, painting, futsal, music skills, vocal, swimming, archery, holy site, drum band, and cooking class (Interview result, 2018).

Special character-based learning, according to the headmaster, has been implemented first by the school, but the implementation is not maximized yet. The new policy, through the curriculum 2013 that is in the process of teaching the character education, precisely has given a reformation for the school because it can strengthen the program developed by Muhammadiyah Suronatan Elementary School, especially in the assessment of the attitude aspect. Previously, the policy implementation is only limited to integration in the learning process and has not created any format of attitude assessment, such as observation sheets, self-assessment, peer assessment, and journals (Interview result, 2018).

Based on the results of the interview, it is noted that the implementation of the integrative thematic learning of the curriculum 2013 is tailored to the needs and developments of the times, in this case, is the 21st-century development that demands various competencies, namely collaboration, communication, high-level thinking, and creativity. It was done to prepare the golden generation, which is in 2045, to make Indonesia has a demographic bonus. Therefore, through the policy of curriculum 2013, the nation is expected to compete in the era of globalization as it is known today as the era of the Industrial Revolution 4.0.

Input

The evaluation of inputs focuses on teachers, learning resources, learning media, and learning strategies because some elements play an essential role in the success of the integrative thematic implementation of the curriculum 2013. Teachers at Muhammadiyah Suronatan Elementary School have an understanding of the implementation of integrative thematic learning of curriculum 2013. The results of interviews with some teachers, both in the low class and in high class, reveal that they understand the change in the learning with curriculum 2013. According to them, in the implementation of integrative thematic learning, teachers must apply active learning by promoting character education to develop attitudes, knowledge, and skills. Also, teachers prosecuted in the process of learning to implement scientific measures and conduct an authentic judgment (Interview results, 2018). The teacher of training activities has owned the understanding. Meanwhile, the headmaster said that almost all teachers in Muhammadiyah Suronatan Elementary School Yogyakarta had attended training, both the training organized by the government and held internally by the inviting experts. Thus, the teachers are better prepared to implement the integrative thematic learning in curriculum 2013.

The learning resources in Muhammadiyah Suronatan Elementary School, such as the library, internet, laboratory, and good school environment in the form of a mini garden, are quite adequate. The mini garden was used by teachers and students as a learning resource, especially in the theme of plants and the environment. The media available is also quite adequate, such as LCD, which has already been installed in every class. Besides, the learning media also include a human skeleton, globe, and map, which are in good condition. The learning strategy applied is varied by implementing four learning models: problem-based learning, project-based learning, discovery learning, and cooperative learning (Observation result, 2018).

The availability of various facilities and infrastructure, especially internet and LCD, is because the school develops not only students' faith, but also their knowl-

edge and skills of science and technology. It means that the carrying capacity of the school is to prepare millennials to master technology, as we enter the digital era.

Process

The Learning Plan

Based on the results of a lesson plan documentation designed by the teacher, the following data are obtained. A lesson plan designed by the teacher is a design of integrative thematic learning of curriculum 2013 using scientific approaches. The components developed are, among others, identity, core competencies, basic competencies, indicators, learning objectives, learning materials, learning approaches and methods, media, tools and learning resources, learning steps, and assessments.

Indicators develop the core competencies and basic competencies using operational verbs, and various levels of the domain (cognitive, affective, and psychomotor), meaning that in the indicators developed by the teacher, the balance between soft skills and hard skills is shown. The purpose of learning is designed to describe student activity following the method to be used. The learning methods designed have varied, such as questions and answers, discussions, lectures, and experiments. With the varied methods of learning, the teacher has no direct effort to apply student-centered learning, as required by the curriculum 2013. Learning with curriculum 2013, according to Prasetyo (2017, p. 102), has several characteristics: (1) the studies use scientific approaches, (2) learning is performed not only in classrooms but also in the environment and society, (3) teachers are not the only learning resources; instead, teachers act as facilitators and motivators.

Steps-designed learning already describes active learning, and each step demonstrates scientific activities. In this case, teachers' role in designing learning is not only limited to exploration activities, elaborations, and confirmation but also observing activities, solicitation, collecting information, and communicating. Furthermore, the evaluation aspect has been designed by the teacher well with various domains, namely cognitive, affective, and psychomotor, by advancing the process assessment (written, observation of attitudes and performance).

Based on the analysis of the results of the study, a related learning plan can say that the teacher has both designed the learning with scientific approaches and by the principles of learning the curriculum 2013. Nevertheless, the teacher still needs to learn in terms of combining subjects that correspond to the theme and sub-theme of learning. Meanwhile, the researchers examine the theme and sub-theme in one of the action plan samples designed by the teacher. In this case, the theme is learning about the ecosystem, with the sub-theme of the relationship of living creatures in the ecosystem. Teacher-combined subjects include sports education, civilization, *bahasa* (language), and mathematics. The themes are studies related to the ecosystem and the subthemes of the relationship of life in the ecosystem. The subjects that correspond to the theme are science, social science, language and arts, and skill.

Besides, the researchers also see the inconsistency of material designed with basic competencies in each subject, in which the material included in the lesson plan, among others, are food chains, building a space in the form of a beam, and a plane figure that is a circle. In addition, the basic competencies in each subject are (1) physical education, sport, and health competencies essentially understand and practice variations and combinations of locomotor and non-locomotor basic motion patterns, (2) the basic competence to understand the diversity of social, cultural, and economic body in the frame of *Bhineka Tunggal Ika* (unity in diversity) concept, (3) Indonesian language's competence to explore text information on book reports on food and food chains, human health, ecosystem balance, as well as nature and influence of human activities, (4) mathematical competence

is essentially the concept of the generation and withdrawal of root of the rank of two and a modest number of three ranks. Exposure has apparent inconsistency between the materials that are listed on the action plan with each basic competency. These materials should derive from any basic competencies.

Due to the inconsistency between the theme and subtheme with the combined subjects, then the inconsistency of learning materials with basic competencies will impact the learning access and learning system. It is possible for teachers to implement it, but they will certainly have difficulty in moving from one subject to another and will seem to feel tired. The later implementation of the learning does not conform to integrative thematic learning characteristics. Frasandy (2017, p. 310) conveys that one of the characteristics of integrative thematic learning is the removal of one subject with the other subject not so obvious and visible. Moving is done subtly, so students do not feel that learning has changed, in which his studies focus on themes that are close to the student's life.

Based on the aforementioned explanation, it can be inferred that the teachers have understood well how to develop an action plan for integrative thematic learning in curriculum 2013, but they have not understood well yet about how to integrate the integrative thematic learning curriculum 2013 that corresponds to themes and subthemes including material suitability with basic competencies.

Furthermore, the learning media designed by teachers are quite varied, such as LCD, projector, props, both the existing one and the ones made by the teachers. The worksheet that teachers will use is the available worksheet or the ones that are created by the teachers. Teacher-designed learning resources are also varied; teachers do not only use books but also sources from the surrounding environment such as students, teachers, school gardens, labs, libraries, and other places that are relevant to learning.

The Implementation of Learning

Based on the observation results, the following information is obtained. Integrative thematic learning curriculum 2013 is implemented by exploration, elaboration, and confirmation, including the activities of observing, solicitation, collecting information, and communicating. For each teacher, learning activity is documented in a video, where the video is made as a teacher's report document to students' parents on a CD and are shared in the class by the teachers at the end of the semester.

In conducting learning, there are three activities conducted by the teachers, namely preliminary activities, core activities, and closing activities. In the preliminary activities, the teacher conducts the activities of preconception, motivation, and pre-test. In the activity, the teacher asks students to observe, sing, and tell the story according to the sub-theme to be studied. Various media are used by teachers, such as projectors, LCD, pictures, and songs, so that the students are excited.

In core activities, teachers employ various learning methods with student-centered learning. In this case, students in groups conduct experiments by observing, examining, collecting information, and also communicating. The teacher-developed material is not only from the book but also associated with the students' social context both at home, in school, and in the community. The atmosphere of the class is very dynamic and conducive, and the students enjoy learning, look joyful, and are happy to learn. Teachers assess the process of learning by observing each student, both on his attitude and work. In this case, the teacher uses the pre-designed assessment section.

Next, in the closing activity, the teacher invites students to conclude the material that has been learned together, then the teacher quizzes with some questions using the problems. At the end of the activities, the teacher gives the students a duty to ask the parents for help with the material learned as a follow-up, or material that will be studied at the next meeting.

Based on the observation of the learning activities conducted by the teacher, it can be inferred that the teacher has implemented integrative thematic learning in accordance with the curriculum 2013 principles, where the teacher, in organizing the learning, has encouraged students to find out problem-solving, to learn from various sources, has applied learning with a scientific approach, learning with competency-based, learning organized in an integrated system, has balanced the proportion of soft skills and hard skills mastery, has developed the learners' creativity, has maximized the use information, communication, and technology in learning, has concerned about individual characteristics and the cultural background of the learners.

Nevertheless, the researchers see that the alignment in one subject to another has not been so visible, and the new thing to come is the students' complaint. The complaint is due to the mistake/lack of carefulness in the selection of themes and sub-themes with the combined subjects, and the and conformity of materials with the basic competencies. It can, therefore, be understood that the teachers can apply learning by the curriculum 2013 principles well, but they have not been able to implement learning yet by the characteristics of integrative thematic learning because of unsynchronized compactions with sub-theme.

Process Assessment and Learning Outcomes

Based on the results of interviews, observations, and documentation, it is revealed that the process assessment and implementation of integrative thematic learning outcomes of the curriculum 2013 involves an authentic approach that includes three domains, i.e., knowledge, attitudes, and skills. The knowledge assessment is carried out in a written test in varied problems such as multiple-choice, a marriage, a short description, and an essay. The attitude assessment is done by observation, either in class or outside class, using an observation sheet, journal, and self-assessment. Skills assessment is carried out through the performance and projects. The results of the eval--uation have been made in the form of narrative reports. The narrative reports made reveal students' success with positive words. Also, learning results are taught by teachers in the form of CDs shared with students. Students who have not succeeded in mastering certain competencies can carry out remedial efforts by editing material that has not been mastered, accompanied by their class teacher (Interview result, August 2018).

Therefore, it can be said that the evaluation system conducted by the teachers in the Muhammadiyah Suronatan in Elementary School in Yogyakarta City is good because the judgment is not only dominated by cognitive values, but also by the domain of attitudes and skills. With the assessment spanning all three domains, students are not only intellectually intelligent, but also morally intelligent.

Product

Based on the results of observation, interviews, and also documentation, information on the impact of integrative thematic learning implementation of the curriculum 2013 towards students' learning outcomes in Muhammadiyah Suronatan Elementary School Yogyakarta City is obtained as follows. (1) It encourages students to be eager to learn, because the lesson activates students, make them learn from various sources, and fun. (2) It is able to increase students' confidence because, in learning, they are encouraged to convey the results of the study in front of the class. (3) Achievement in various fields, both common and religious such as winning races and achieving the first rank of national exam results, is possible. It is in line with the information given by the B-class homeroom teacher that, since the curriculum 2013 has been established with an active learning model, students grow and become confident in performing publicly. Even some students offer themselves to perform when there are activities (Interview results, July 2018).

The Supporting and Inhibiting Factors of the Implementation of Integrative Thematic Learning of Curriculum 2013

Curriculum 2013 is a new program developed by the government. As a new program, in its implementation, it has obstacles in addition to supporters. Based on the results of interviews, there are several supporting factors and barriers in implementing integrative thematic learning with curriculum 2013.

The Supporting Factors

The integrative thematic learning with curriculum 2013 implementation in elementary school of Muhammadiyah Suronatan Yogyakarta can run well. It is because it is supported by many factors, including: (1) the teacher's competency, in which most teachers have been certified with the professional teacher's predicate, (2) most teachers' educational background, in which the teachers are graduates of elementary school teacher training education, (3) adequate facilities and infrastructure in supporting the learning process, such as labs, libraries, IT, school environment, etc., (4) the teachers' average age and their easy and energetic personality, which make them have the motivation and passion for learning and wanting to be invited to make a change through a learning pattern with a new paradigm.

It is important to note that before the curriculum 2013 was developed, teachers in the elementary school of Muhammadiyah Suronatan Yogyakarta have already been accustomed to applying active learning. Thus, it is very easy for the teachers there to adapt to the learning model corresponding to the demands of the curriculum 2013.

The Inhibiting Factors

Based on the results of the interview, the implementation of integrative thematic learning with curriculum 2013 has faced several obstacles, including (1) most students do not have a student handbook yet, and (2) teachers still strive to use other references that are still relevant to the integrative thematic learning of the curriculum 2013 (Interview result, July 2018). In addition, related to the barriers, the books have not been distributed evenly. It is also in line with the results of research by Dewi et al. (2015, p. 8) performed on a state elementary school in the suburbs of Bandung. It reveals that the teachers' book based on the theme and students' books based on the subject in the state elementary school in the suburb of Bandung had not been distributed well. The next obstacle is that teachers have not understood how to integrate competencies and subjects according to the theme and sub-theme.

As previously mentioned in the aspect of the process, in the planning step, the teacher has not understood how to integrate integrative thematic learning in curriculum 2013 that corresponds to the theme and sub-theme, including material conformity with essential competencies. Thus, it impacts the implementation of knowledge, in which the alignment between one subject with the other topics has not been so visible, and the new ones seem to be the reason for which students often complain about. The complaint is due to a mistake or the lack of carefulness in the selection of themes and sub-themes. It will combine subjects and conformity of materials with the essential competencies. This process of learning is also demonstrated by Rasidi and Setiawati (2015, p. 163) that one factor of teacher difficulty in integrative thematic learning is the material that suits the theme.

Based on further searches through interviews and document study, the condition is because, during this time, the teacher only refers to the teacher handbook provided by the government, wherein it has themes and sub-themes. In each sub-theme, there are six learnings, while in the researchers' handbook, teachers see not all the competencies and subjects combined with the themes and sub-themes, especially in each of the studies of learning 1-6. Sometimes, the competencies and subjects that correspond to the theme and sub-themes are found in learning two, while teachers teach using learning 4.

Conclusion

Based on the analysis of research results, it can be concluded that the implementation of integrative thematic learning using curriculum 2013 in Muhammadiyah Suronatan Elementary School in Yogyakarta has been able to run quite well, both in planning, implementation, assessment process, and learning outcomes. However, there are some obstacles faced by teachers of the school in carrying out the integrative thematic learning of the curriculum 2013, including: (1) students' handbook which has not been distributed evenly by the government, and (2) the lack of understanding related to the subject being complained that corresponds to the sub-theme. The integrative thematic learning of the curriculum 2013 has a positive impact on student learning outcomes from both the aspects of knowledge, attitudes, and skills. In the aspect of knowledge, the integrative thematic learning of the curriculum 2013 can deliver the best grades in the national exam by getting a top 10 in the city of Yogyakarta. From the attitude aspect of students' spirit in learning, it can increase student confidence and cooperation. As for the skill aspect, students become creative and innovative in both thinking and working.

Acknowledgments

The authors send gratitude to the Institute of Research and Community Service of Universitas Ahmad Dahlan for the establishment of the research grant. They also thank the related parties who have participated in completing the writing of this article.

References

Ahmadi, I. K., & Amri, S. (2014). *Pengembangan dan model pembelajaran tematik integratif*. Prestasi Pustaka.

Asifudin, A. J. (2009). *Mengungkit pilar-pilar pendidikan Islam: Tinjauan filosofis*. Sunan Kalijaga Press.

Assegaf, A. R. (2011). *Filsafat pendidikan Islan: Paradigma baru pendidikan Hadhari berbasis integratif interkonektif*. Rajawali Press.

Bogdan, R., & Biklen, S. K. (1982). *Qulaitative research for education: An introduction to theory and methods*. Allyn and Bacon.

Darodjat, D., & Wahyudhiana, W. (2015). Model evaluasi program pendidikan. *Islamadina: Jurnal Pemikiran Islam*, *XIV*(1), 1–28. https://doi.org/10.30595/islamadina.v0i0.1665

Dewi, N. L. K., Manuaba, I. B. S., & Putra, M. (2015). Studi evaluasi implementasi kurikulum 2013 ditinjau dari context, input, process, dan product (CIPP) pada sekolah dasar negeri di wilayah pinggiran Kabupaten Badung. *E-Journal PGSD Universitas Pendidikan Ganesha*, *3*(1). https://ejournal.undiksha.ac.id/index.php/JJPGSD/article/download/5034/3798

Fadlillah, F. (2014). *Implementasi kurikulum 2013 dalam pembelajaran SD/MI, SMP/MTs, SMA/MA*. Ar-Ruzz Media.

Fogarty, R. J. (1991). *How to integrate the curricula*. IRI/Skylight.

Frasandy, R. N. (2017). Pembelajaran tematik integratif: Model integrasi mata pelajaran umum SD/MI dengan nilai Agama. *Jurnal Elementary*, *5*(2).

Hidayah, N. (2015). Pembelajaran tematik integratif di sekolah dasar. *Terampil: Jurnal Pendidikan Dan Keterampilan Dasar*, *2*(1), 34–49. https://doi.org/10.24042/terampil.v2i1.1280

Indriani, F. (2019). *Pembelajaran tematik integratif berbasis humanis religius dalam pendidikan karakter di sekolah dasar Yogyakarta*. Doctoral Thesis. Universitas Islam Negeri Sunan Kalijaga Yogyakarta.

Julianti, I. A. R., & Mawardi, M. (2018). Penerapan desain pembelajaran tematik integratif alternatif berbasis sub-subtema untuk meningkatkan kebermaknaan dan hasil belajar. *Publikasi Pendidikan: Jurnal Pemikiran, Penelitian Dan Pengabdian Masyarakat Bidang Pendidikan*, *8*(3), 206–215. https://doi.org/10.26858/publikan.v8i3.6598

Kristiantari, M. R. (2014). Analisis kesiapan guru sekolah dasar dalam mengimplementasikan pembelajaran tematik integratif menyongsong kurikulum 2013. *JPI (Jurnal Pendidikan Indonesia)*, *3*(2), 460–470. https://doi.org/10.23887/jpi-undiksha.v3i2.4462

Law of Republic of Indonesia No. 20 of 2003 on National Education System, (2003).

Majid, A. (2014). *Pembelajaran tematik terpadu*. Remaja Rosda Karya.

Maslow, A. H. (2017). *Motivation and personality* (A. Fawaid & M. Maufur (Trans.)). Cantrik Pustaka.

Miles, M. B., & Huberman, M. (1994). *Qualitive data analiysis*. Sage.

Murtikasari, K. (2013). Menyongsong kurikulum 2013: Sebuah kajian dan rekomendasi implementasi kebijakan. In *National Seminar Held by IKA (Ikatan Alumni) of Universitas Negeri Yogyakarta*.

Oliva, P. F. (1992). *Depeloving the curriculum*. Harper Collins.

Prasetyo, T. (2017). Pengembangan perangkat penilaian hasil belajar dalam pembelajaran tematik integratif kelas V SD. *Jurnal Prima Edukasia*, *5*(1), 102–111. https://doi.org/10.21831/jpe.v5i1.7528

Raharjo, R. (2012). *Pengembangan & inovasi kurikulum*. Baituna.

Rasidi, M. A., & Setiawati, F. A. (2015). Faktor-faktor kesulitan guru pada pembelajaran tematik integratif di SD Kota Mataram. *Jurnal Prima Edukasia*, *3*(2), 155–165. https://doi.org/10.21831/jpe.v3i2.6504

Regulation of the Government, No. 32 of 2013, on National Education Standard, (2013).

Suharsimi, S. (2007). *Evaluasi program pendidikan: Pedoman teoritis praktis bagi praktisi pendidikan*. PT Bumi Aksara.

Suharto, F. S. (Ed.). (2013). *8 (Delapan) standar nasional pendidikan tingkat SD/MI versi kurikulum 2013*. PT Teratai Emas Indah.

Jurnal Penelitian dan Evaluasi Pendidikan
Volume 23, No 2, December 2019 (197-205)
Online: http://journal.uny.ac.id/index.php/jpep

ANALYSIS OF SCIENCE PROCESS SKILLS IN PHYSICS EDUCATION STUDENTS

Maison
Universitas Jambi

Darmaji
Universitas Jambi

Dwi Agus Kurniawan
Universitas Jambi

Astalini
Universitas Jambi

Utari Prisma Dewi
Universitas Jambi

Lia Kartina
Universitas Jambi

Abstract

This study aims to analyze students' science process skills on specific heat material by reviewing two aspects of basic science process skills indicators (observation and classification), and two indicators of integrated science process skills (identifying variables and making hypotheses). This research uses a descriptive quantitative method. In this study, the sample used was 35 students of physics education of batch 2018 who were randomly selected. The assessment instrument used was the science process skills observation sheet with the skill score used in the form of a Likert scale. The results of the study show that the students' mastery of basic science process skills on the observation indicator is 65% in the good category, 30% in the high category, and 10% in the category of not good to low, whereas, the classification indicators obtained are 54.3% and 37.1% of students have mastered classification skills in both good and high categories. The remaining 8.6% are classified in the not good category. For the mastery of integrated science process skills in the variable identification indicator, 60% of them are in the good category and 14.3% in the high category. The rest are in the category of not good and low. For the indicators of skills in making hypotheses, 65.7% and 14.3% are in the good and high categories. It proves that physics education students have mastery of science process skills that are in the good category.

Keywords: *education, science process skills, prospective teachers*

Permalink/DOI: http://dx.doi.org/10.21831/pep.v23i2.28123

Contact: *Lia Kartina*
liakartina98@gmail.com
Department of Physics Education, Faculty of Teacher Training and Education, Universitas Jambi
Kampus Pinang Masak, Jl. Raya Jambi – Muara Bulian Km. 15, Mendalo Darat, Jambi 36361, Indonesia

Introduction

Education, as a conscious effort to realize fully Indonesian people, namely, people who believe and have faith in God the Almighty, are noble, healthy, knowledgeable, capable, creative, independent, responsible, and become citizens who are democratic, is a process of improving the lives of Indonesian people (Kurniawan, Astalini, & Anggraini, 2018, p. 125). Education is a conscious and planned effort to create an atmosphere of learning and learning process so that students actively develop their potential.

Education in Indonesia has been well-integrated and needs further development. Continuing education should be closely related to the curriculum used. The curriculum has two aspects, namely as a plan that must be used as a guideline for the implementation of the teaching and learning process, and as a tool to achieve educational goals (Astalini, Darmaji, Kurniawan, & Destianti, 2019, p. 3). As a process, education must be done in stages starting from basic education to higher education. Each level of education has different characteristics in the knowledge learned, in accordance with the development of students themselves, for instance, high school students will not be the same as elementary and junior high school students (Astalini, Kurniawan, & Sumaryanti, 2018, p. 59). The frequent change of curriculum in schools is due to the fact that it is not appropriate or the learning application is not according to the curriculum. The current curriculum used is the 2013 revised curriculum, so that students are expected to have a scientific attitude in learning (Astalini, Kurniawan, Melsayanti, & Destianti, 2018, p. 216).

In accordance with the 2013 curriculum, high school students have been set on future career choices. It is shown by the selection of majors, namely, Mathematics and Natural Sciences or Social Sciences. The selection of these majors is in accordance with the abilities seen from the interests and talents of the students. In the Mathematics and Natural Sciences major, they will receive lessons focusing on Physics, Biology, and Chemistry (Reza, Syukur, & Soeleman, 2017, p. 57). Usually, the selection of these majors also determines the selection of study programs in tertiary institutions.

Physics is usually always a difficult subject for students so that not many people are interested in taking courses related to physics. Many factors make physics difficult for students. The most frequently-occur factor is that the students must memorize the whole formula given by the teacher, which is a conventional way of learning (Astalini, Kurniawan, & Sumaryanti, 2018, p. 59). One third of students carry it over until they study in higher education (Astalini, Kurniawan, Perdana, & Kurniasari, 2018, p. 475). The method of memorizing learning is the lowest level of the learning sphere raised by Bloom, known as Bloom's Taxonomy. Memorizing is in the lowest realm of knowledge (Netriwati, 2018, p. 349). In fact, the curriculum of all countries is based on Bloom's taxonomy.

The domains discussed in this knowledge are six levels, namely C1 to C6, from the lowest, that is, knowing to create works, while the memorization part is only in the first cognitive domain or C1 (Prasetya, 2012, p. 108). Higher education certainly leads to how theory is obtained. Thus, the physics education study program is required in the field of natural sciences to carry out practical activities which refer to the Indonesian National Qualification Framework (Darmaji, Kurniawan, Parasdila, & Irdianti, 2018b, p. 346). One of the references to be achieved is skills that are specialized in process skills (Darmaji, Kurniawan, & Suryani, 2019, p. 2). Skills also include behavior, which is an important process of one's self. Process skills are the skills of students to carry out activities related to practice, where students are required to experience themselves and discover and connect the results of practicum with theory to write the formula mathematically (Lestari & Diana, 2018, p. 49). Direct student involvement will make it easier to understand the concepts of the theory being studied.

When choosing majors related to natural science, especially physics, students must take part in practical activities. Practicum activities will help students understand the courses taken better. Basic physics is one of the compulsory subjects for physics education students (Septiyanto & Darman, 2018, p. 14). In practicum activities to provide process skills for observational indicators, it must be ensured that students already have information and knowledge about the material to be practiced (Supahar & Prasetyo, 2015, p. 98) through the practicum guides. The material in two basic physics courses includes metal heat. Metal heat cannot be taught only by giving the material in class. It also has to go through experimental activities in the laboratory (Nugroho & Suliyanah, 2018, p. 355). By conducting experiments in the laboratory, students' scientific processes can be improved.

Practicum activities that prioritize process skills are also indicated in 21st-century education, where the practicum will provide a place for students to analyze and evaluate with a critical mind of the results of the experiment (Pop, 2014, p. 1). The 2013 curriculum certainly has a striking difference compared to the curriculum that has been previously set (Setiadi, 2016, p. 167), where the core competencies are classified into four, namely social attitudes, spiritual attitudes, knowledge, and skills. 21st Century education is an education that guarantees students to be able to improve their soft skills (Andrian & Rusman, 2019, p. 15). The 2013 curriculum has been integrated with 21st-century education wherein the 2013 curriculum, there is a scientific approach (Bancong & Putra, 2015, p. 28). The scientific approach is the *5M* approach, namely: observing (*mengamati*), asking (*menanyakan*), gathering information or data (*mengumpulkan informasi atau data*), associating (*mengasosiasi*), and also communicating (*mengomunikasikan*) (Wahyuni, Indrawati, Sudarti, & Suana, 2017, p. 166). This scientific approach is a very natural approach in the learning process.

For these demands, students must have science process skills. Science process skills are a skill that provides a means in science learning, research, and active learning, building a sense of responsibility when learning and increasing knowledge (Darmaji, Kurniawan, & Irdianti, 2019). Science process skills include: asking questions, observing, predicting, and simulating (Lestari & Suliyanah, 2014, p. 61). Those elements will increase the professionalism of students as future teacher candidates.

As future teacher candidates, it is necessary to analyze how the science process skills possessed by physics education students (Siswono, 2017, p. 84). By direct observation in practicum activities undertaken, especially as a prospective physics teacher or natural science teacher, he/she must have scientific process skills.

In this process skills, two scopes can be known, namely basic process skills and integrated process skills (Khaerunnisa, 2017, p. 342). Based on research conducted by Darmaji, Kurniawan, Parasdila, and Irdianti (2018a, p. 497), the science process skills of students in University of Jambi are good enough for the 2017 class.

Based on the aforementioned explanation, the researchers tried to analyze the process skills in the 2018 physics education students of University of Jambi, especially in the practicum method. What will be seen are indicators of students' process skills for the basic skills of the indicators observed and classification and skills, while the integrated indicator is the identification of variables and creating hypotheses. Good science process skills can illustrate the students' knowledge (Srirahayu & Arty, 2018, p. 169). Through good science process skills, students certainly have a strong motivation to learn and have critical thinking skills seen from the classification indicators and making hypotheses.

Research Method

This research uses a descriptive quantitative research method. Quantitative methods are a numerical method of research. This type of research usually connects the attachment between variables (Creswell, 2014, p. 288). The purpose of this study is to know

the mastery of science process skills, especially on basic process skills indicators (observation and classification) and integrity skills indicators (variable identification and hypothesis). This study involved 35 students of the Physics Education Study program of batch 2018 in Jambi, which was currently taking the Primary Physics Course 2.

Data retrieval was conducted with an electronic-based observation score sheet. This observation sheet refers to the value of the Linkert scale which employs the highest Strongswan 4 (Astalini, Darmaji, Kurniawan, Anwar, & Kurniawan, 2019, p. 25). The selection of scale four is intended to adjust the criteria desired by the researchers, namely: (1) low; (2) not good; (3) good; and (4) high. To find out the mastery of science process skills of students who have known the minimum score (X min) and the maximum score (X max), then the ideal average score can be found by using the formula Mi = 1/2 (X max + X min) and looking for ideal standard deviation using the formula SDi = 1/6 (X max – X min). Based on these references, the ideal mean is 2.5 and the standard deviation is 0.5. Thus, students' mastery of science process skills can be grouped into four categories as presented in Table 1.

The data of students' science process skills were analyzed using the statistic descriptive method so that the score of the mean, median, mode, minimum score, maximum score, deviation standard, and percentage, which is viewed from the assessment indicator, results in the sheet observation of the science process skills (Darmaji, Kurniawan, Astalini, Lumbantoruan, & Samosir, 2019, p. 11). Then the data were presented in the table of the distribution.

Findings and Discussion

After analyzing the assessment sheet, the results of mastery of basic science process skills and students' integrated science process skills were found. The results of the mastery analysis of the students' basic science processes on observation indicators and classification can be seen in Table 2.

Table 1. Level of the Score of Process Skills

Score Interval Formula	Score Interval	Categories
$X > M + 1.5\ SD$	$X > 3.25$	High
$M + 0.5\ SD < X \leq M + 1.5\ SD$	$2.75 < X \leq 3.25$	Not Good
$M - 1.5\ SD < X \leq M + 0.5\ SD$	$3.25 < X \leq 2.75$	Good
$X \leq M + 0.5\ SD$	$X \leq 2.75$	Low

(Widoyoko, 2011, p. 238)

Table 2. Analysis of the Students' Science Processe Skills on Observation and Classification Indicators

Classification				Mean	Min	Max	Std. Deviation	%
Indicator	Score Interval	Category	Total					
Observation	$X > 3.25$	High	12	3.34	3.00	4.00	0.48	34.3
	$2.75 < X \leq 3.25$	Good	23					65.7
	$3.25 < X \leq 2.75$	Not Good	0					0
	$X \leq 2.75$	Low	0					0
Total			35					100
Classification	$X > 3.25$	High	13	3.29	2.00	4.00	0.62	37.1
	$2.75 < X \leq 3.25$	Good	19					54.3
	$3.25 < X \leq 2.75$	Not Good	3					8.6
	$X \leq 2.75$	Low	0					0
Total			35					100

Based on Table 2, it can be seen that from the mastery of science process skills of 35 students in the observation indicator, 23 students are in the good category (65.7%) and 12 students are in the high category (34.4%). Besides, a mean of 3.34, a minimum score of 3.00, a maximum score of 4.00, and a deviation standard of 0.48 are obtained. Meanwhile, in the classification skills indicator, 19 students are in the good category (54.3%), 13 students are in the high category (37.1%), and three students are in the not good category (8.6%). A mean of 3.29, a minimum score of 2.00, a maximum score of 4.00, and a deviation standard of 0.62 are obtained.

Based on the data, 23 students have a basic science process on the observation indicator that is already at a good level, while the other 12 people have skills of observation science process at an excellent level. Basically, observation is the first step of an activity, so it must be strengthened at the beginning (Darmaji, Kurniawan, Suryani, & Lestari, 2018, p. 242). The aforementioned results show that the level of the process skills observation skill for observation indicator is well-owned by the physics education students at the University of Jambi.

Furthermore, in the classification indicator, three physics education students are still in the category of not good (8.6%). Then, 19 students have science process skills for the good category (54.3%), while the rest 13 students have excellent science process skills (37.1%). Thus, physics education students at the University of Jambi have mastered the science process skills in the indicator of classification well. Classification is part of the basic science process skills, a second indicator that must be strengthened by the students of physics education of Jambi University after observation skill (Puspita, 2016, p. 31). Classification is also an important thing to do before experimenting. When students are able to classify according to the order or practical measures, it proves that the students are able to distinguish which ones to measure and which ones to weigh.

The information in Table 2 indicates that the physics education students of University of Jambi have a good mastery of basic science process skills. However, students are more proficient in the basic science process skills on the observation indicator with the largest percentage of 65.7%.

Having basic skills is also necessary to know the mastery of the integrated science process skills, where integrated process skills are the enhancement skills of the basic skills in which the indicators are interconnected. There are two indicators that are examined in the science process skills of physical education students. The first integrity indicator is the variable identification and the second indicator is hypothesized experiment results. To know the mastery of the students' integrated science process skills on the indicators of variable identification and creating hypotheses, we can see Table 3.

Table 3. The Mastery Analysis of Student Science Process Skills on Variable Identification and Creating Hypotheses Indicators

Classification				Mean	Min	Max	Std. Dviation	%
Indicator	**Score Interval**	**Category**	**Total**					
Variable Identification	$X > 3.25$	High	3	2.80	1.00	4.00	0.80	14.3
	$2.75 < X \leq 3.25$	Good	6					60.0
	$3.25 < X \leq 2.75$	Not Good	21					17.1
	$X \leq 2.75$	Low	5					8.6
Total			35					100
Creating Hypotheses	$X > 3.25$	High	5	2.91	1.00	4.00	0.66	14.3
	$2.75 < X \leq 3.25$	Good	23					65.7
	$3.25 < X \leq 2.75$	Not Good	6					17.1
	$X \leq 2.75$	Low	1					2.9
Total			35					100

Table 3 provides an overview of the science process skills of 35 physical education students for a metal-type test of a variable identification indicator. It is apparent that there is a variable identification skill level. At the lowest level, 8.6% or three students have a high category of variable identification skills. In the next level, 17.1% or six students have not a good category in variable identification skills, and 60% or 21 students have a well-categorized variable identification. Meanwhile, the remaining 14.3% or five students have skills of variable identification in the science process in the category of very well. Further, a mean score of 2.80, a minimum score of 1.00, a maximum score of 4.00, and a standard deviation of 0.80 are obtained.

Table 3 also provides information related to the hypothesized indicator of a metal-type Calor experiment. From Table 3, it is known that there are one or 2.9% of students who have the skills to make the category hypothesis very unwell. Then, 17.1% or six students have the skills of creating hypotheses in a not good category. A total of 23 students or 65.7% have a science process skill to create hypotheses, or it is categorized into a good category. The remaining 14.3% or five students have the skills to create hypotheses that belong to the category of very well. Further, a mean score of 2.91, a minimum score of 1.00, a maximum score of 4.00, and a standard deviation of 0.66 are obtained.

Based on the information in Table 3, it can be concluded that the physics education students of University of Jambi have a good mastery of integrated science process skills. It is indicated by the percentage in each indicator which is in a good category. However, the students master the integrated science process skills on the indicator of creating hypotheses more, shown by the largest percentage of 65.7%.

After the variable identification, of course, in the practicum, a temporary hypothesis would be made for the results of experiments conducted from the theory studied. After the theory was studied, then, it is attempted to prove the theory. This indicator is very important in a process skill. Creating hypotheses will make students think about the influence of each existing variable in the experiment practice of metal type heat. This process is part of the results of collecting information. When students create hypotheses, it means that the learning has been in progress with a scientific approach, where learners are able to make a temporary suspicion after observing the variables in the metal-type Calor experiments. Based on those data, the physics education students of University of Jambi have mastered the science process skills well.

Conclusion

The remoteness of the science process greatly helps students conduct scientific activities, one of which is a physics education students who are obliged to perform a practicum on a basic physics course. One of the basic physical materials is temperature and heat. Temperature and heat are material that is mandatory for the experiment. In the experiments, the skills of the process especially the science process are certainly needed. Based on the research conducted, there are two kinds of science process skills: basic science process skills and integrated science process skills. This research has focused on the observation and classification indicators for basic science process skills, while variable identification and creating hypotheses indicators are for the integrated science process skills.

The result of this study concludes that in the basic science process skills of physics education students, few students still have ungood observation skills and it proves that the observation skills of physics education students are already in the good category. For the classification indicator, there are still very few students who have high classification skills and very few of them do not have good classification skills. Likewise, in the integrated science process skills, for the variable identification and creating hypotheses, the average physics education students have the skills of identifying variables and

creating hypotheses. Therefore, the physics education students of University of Jambi in batch 2018 have mastered science process skills in the good category.

Acknowledgment

The authors would like to thank the chairman of physics education of the University of Jambi who has given us the opportunity to work together. The gratitude is also sent to the chairman of the physics education laboratory at the University of Jambi who has provided the opportunity to be able to conduct research at the university physics education laboratory and to all of the physics education students who have participated in this research.

References

Andrian, Y., & Rusman, R. (2019). Implementasi pembelajaran abad 21 dalam kurikulum 2013. *Jurnal Penelitian Ilmu Pendidikan*, *12*(1), 14–23. https://doi.org/10.21831/jpipfip.v12i1.20116

Astalini, A., Darmaji, D., Kurniawan, D. A., & Destianti, A. (2019). Description of the dimensions attitudes towards science in junior high school at Muaro Jambi. *International Journal of Sciences: Basic and Applied Research (IJSBAR)*, *47*(1), 1–11. Retrieved from https://www.gssrr.org/index.php/JournalOfBasicAndApplied/article/view/10022

Astalini, A., Darmaji, D., Kurniawan, W., Anwar, K., & Kurniawan, D. A. (2019). Effectiveness of using e-module and e-assessment. *International Journal of Interactive Mobile Technologies (IJIM)*, *13*(09), 21–39. https://doi.org/10.3991/ijim.v13i09.11016

Astalini, A., Kurniawan, D. A., Melsayanti, R., & Destianti, A. (2018). Sikap terhadap mata pelajarn IPA di SMP se-Kabupaten Muaro Jambi. *Lentera Pendidikan: Jurnal Ilmu Tarbiyah Dan Keguruan*, *21*(2), 214–227. https://doi.org/10.24252/lp.2018v21n2i7

Astalini, A., Kurniawan, D. A., Perdana, R., & Kurniasari, D. (2018). Identification of student attitudes toward physics learning at Batanghari District high school. *The Educational Review, USA*, *2*(9), 475–484. https://doi.org/10.26855/er.2018.09.003

Astalini, A., Kurniawan, D. A., & Sumaryanti, S. (2018). Sikap siswa terhadap pelajaran fisika di SMAN Kabupaten Batanghari. *JIPF (Jurnal Ilmu Pendidikan Fisika)*, *3*(2), 59–64. https://doi.org/10.26737/jipf.v3i2.694

Bancong, H., & Putra, D. P. (2015). Analisis proses keterampilan proses sains mahasiswa berdasarkan gaya berpikir dan kecerdasan jamak pada praktikum fisika modern di Universitas Muhammadiyah Makassar. *Jurnal Pendidikan Fisika*, *3*(1), 27–33. https://doi.org/10.26618/jpf.v3i1.248

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: SAGE Publications.

Darmaji, D., Kurniawan, D. A., Astalini, A., Lumbantoruan, A., & Samosir, S. C. (2019). Mobile learning in higher education for the industrial revolution 4.0: Perception and response of physics practicum. *International Journal of Interactive Mobile Technologies (IJIM)*, *13*(09), 4–20. https://doi.org/10.3991/ijim.v13i09.10948

Darmaji, D., Kurniawan, D. A., & Irdianti, I. (2019). Physics education students' science process skills. *International Journal of Evaluation and Research in Education (IJERE)*, *8*(2), 293–298. https://doi.org/10.11591/ijere.v8i2.28646

Darmaji, D., Kurniawan, D. A., Parasdila, H., & Irdianti, I. (2018a). Description of science process skills' physics education students at Jambi University in temperature and heat materials. *The Educational Review, USA*, *2*(9), 485–

498. https://doi.org/10.26855/er.2018.09.005

Darmaji, D., Kurniawan, D. A., Parasdila, H., & Irdianti, I. (2018b). Deskripsi keterampilan proses sains mahasiswa pada materi termodinamika. *Berkala Ilmiah Pendidikan Fisika*, *6*(3), 345–353. https://doi.org/10.20527/bipf.v6i3.5290

Darmaji, D., Kurniawan, D. A., & Suryani, A. (2019). Effectiveness of basic physics II practicum guidelines based on science process skills. *JIPF (Jurnal Ilmu Pendidikan Fisika)*, *4*(1), 1–7. https://doi.org/10.26737/jipf.v4i1.693

Darmaji, D., Kurniawan, D. A., Suryani, A., & Lestari, A. (2018). An identification of physics pre-service teachers' science process skills through science process skills-based practicum guidebook. *Jurnal Ilmiah Pendidikan Fisika Al-Biruni*, *7*(2), 239–245. https://doi.org/10.24042/jipfalbiruni.v7i2.2690

Khaerunnisa, K. (2017). Analisis keterampilan proses sains (fisika) SMA di Kabupaten Jeneponto. *Jurnal Pendidikan Fisika*, *5*(3), 340–350. https://doi.org/10.26618/jpf.v5i3.855

Kurniawan, D. A., Astalini, A., & Anggraini, L. (2018). Evaluasi sikap siswa SMP terhadap IPA di Kabupaten Muaro Jambi. *Jurnal Ilmiah Didaktika: Media Ilmiah Pendidikan Dan Pengajaran*, *19*(1), 124–139. https://doi.org/10.22373/jid.v19i1.4198

Lestari, M. Y., & Diana, N. (2018). Keterampilan proses sains (KPS) pada pelaksanaan praktikum Fisika Dasar I. *Indonesian Journal of Science and Mathematics Education*, *1*(1), 49–54. https://doi.org/10.24042/ijsme.v1i1.2474

Lestari, R. P., & Suliyanah, S. (2014). Penerapan pendekatan keterampilan proses sains dalam model pembelajaran guided discovery pada materi suhu dan kalor terhadap hasil belajar siswa di SMAN 1 Sukomoro. *IPF: Inovasi Pendidikan Fisika*, *3*(2), 60–64. Retrieved from https://jurnalmahasiswa.unesa.ac.id/index.php/inovasi-pendidikan-fisika/article/view/7446

Netriwati, N. (2018). Penerapan Taksonomi Bloom revisi untuk meningkatkan kemampuan pemahaman konsep matematis. *Desimal: Jurnal Matematika*, *1*(3), 347–352. https://doi.org/10.24042/djm.v1i3.3238

Nugroho, A., & Suliyanah, S. (2018). Pengembangan kit praktikum pegas berbasis pembelajaran guided inquiry pada materi elastisitas sebagai media pembelajaran siswa SMA. *IPF: Inovasi Pendidikan Fisika*, *7*(2), 353–360. Retrieved from https://jurnalmahasiswa.unesa.ac.id/index.php/inovasi-pendidikan-fisika/article/view/24871

Pop, F. (2014). High performance numerical computing for high energy physics: A new challenge for big data science. *Advances in High Energy Physics*, (Special Issue), 1–13. https://doi.org/10.1155/2014/507690

Prasetya, T. I. (2012). Meningkatkan keterampilan menyusun instrumen hasil belajar berbasis modul interaktif bagi guru-guru IPA SMP N Kota Magelang. *Journal of Educational Research and Evaluation*, *1*(2), 106–112. Retrieved from https://journal.unnes.ac.id/sju/index.php/jere/article/view/873

Puspita, S. A. (2016). Analisis keterampilan proses sains yang dikembangkan dalam LKS Biologi kelas X yang digunakan oleh siswa MAN di Kota Yogyakarta. *Jurnal Pendidikan Biologi*, *5*(1), 30–39.

Reza, A., Syukur, A., & Soeleman, M. A. (2017). Penentuan jurusan siswa sekolah menengah atas disesuaikan dengan minat siswa menggunakan Algoritma Fuzzy C-Means. *Jurnal Teknologi Informasi CyberKU*, *13*(1), 57–

68. Retrieved from http://research.pps.dinus.ac.id/index.php/Cyberku/article/view/9

Septiyanto, R. F., & Darman, D. R. (2018). Pengaruh model pembelajaran berbasis proyek pada mata kuliah fisika dasar terhadap high order thinking skill mahasiswa. *Gravity: Jurnal Ilmiah Penelitian Dan Pembelajaran Fisika*, *4*(1), 13–22. https://doi.org/10.30870/gravity.v4i1.3114

Setiadi, H. (2016). Pelaksanaan penilaian pada kurikulum 2013. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *20*(2), 166–178. https://doi.org/10.21831/pep.v20i2.7173

Siswono, H. (2017). Analisis pengaruh keterampilan proses sains terhadap penguasaan konsep fisika siswa. *Momentum: Physics Education Journal*, *1*(2), 83–90. https://doi.org/10.21067/mpej.v1i2.1967

Srirahayu, R. R. Y., & Arty, I. S. (2018). Validitas dan reliabilitas instrumen asesmen kinerja literasi sains pelajaran Fisika berbasis STEM. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *22*(2), 168–181. https://doi.org/10.21831/pep.v22i2.20270

Supahar, S., & Prasetyo, Z. K. (2015). Pengembangan instrumen penilaian kinerja kemampuan inkuiri peserta didik pada mata pelajaran fisika SMA. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *19*(1), 96–108. https://doi.org/10.21831/pep.v19i1.4560

Wahyuni, S., Indrawati, I., Sudarti, S., & Suana, W. (2017). Developing science process skills and problem solving abilities based on outdoor learning in junior high school. *Jurnal Pendidikan IPA Indonesia*, *6*(1), 165–169. https://doi.org/10.15294/jpii.v6i1.6849

Widoyoko, E. P. (2011). *Evaluasi program pembelajaran*. Yogyakarta: Pustaka Pelajar.

Jurnal Penelitian dan Evaluasi Pendidikan
Volume 23, No 2, December 2019 (206-218)
Online: http://journal.uny.ac.id/index.php/jpep

THE EFFECT OF PARENTING STYLE AND GENETIC PERSONALITY ON CHILDREN CHARACTER DEVELOPMENT

Masduki Asbari
Universitas Pelita Harapan

Wakhida Nurhayati
Aya Sophia Islamic School

Agus Purwanto
Universitas Pelita Harapan

Abstract

Character building is very important to become a key for educational mainstream in Indonesia, either in a formal, non-formal, and informal setting. Several factors influence character building in children, namely parenting style and genetic personality. The purpose of this research is to identify the influence of parenting style and genetic personality to children's character development. This research is a correlational study using a survey and quantitative method. The population in this study were parents of students in Aya Sophia Foundation Islamic School. Data collection was carried out by simple random sampling to a population of 1,243 students' parents using an electronic questionnaire. The returned and valid questionnaire results were 398 samples. The data collection technique used was a questionnaire with a Likert scale. The instrument of this study is a modification from the Parenting Style Questionnaire (PSQ) of Robinson et al. and Roman et al. To measure the genetic personality, the instrument was adapted from STIFIn Personality Concept. Another instrument used to measure children's character development is an adaptation from Poniman et. al. The analysis in this study used SEM (Structural Equation Model) with SmartPLS version 3.0 as a statistic tool. The result of this study shows that parenting style and genetic personality have a positive influence and significantly contribute to children's character building.

Keywords: *children character building, parenting style, personality genetic, STIFIn personality*

Permalink/DOI: http://dx.doi.org/10.21831/pep.v23i2.28151

Contact: *Agus Purwanto*
agozpor@gmail.com
Department of Management, Universitas Pelita Harapan
Jl. M. H. Thamrin Boulevard 1100 Lippo Village, Tangerang 15811, Indonesia

Introduction

Indonesia currently faces major challenges, which is decentralization and the era of total globalization. The key to success in facing these challenges is to prepare the quality of human resources that is complete, reliable, and cultured (Puspitawati & Sarma, 2012), meaning that those who can save the nation's condition are human resources with character. *Bung* Karno, one of the founding fathers of the nation, stressed: "This nation must be built by prioritizing character building because this character will make Indonesia a great, advanced, and glorious, dignified nation" (Samani & Hariyanto, 2011).

Character education has been coloring the curriculum in Indonesia since the old order, using the term character education with an emphasis on relationships between people, between students and teachers, between students and parents, and between students. Until now, the implementation of character education is still the main mainstream. At the peak of the National Education Day commemoration on May 2, 2010, Susilo Bambang Yudhoyono, as the President of the Republic of Indonesia, launched the National Movement for Nation Character Building (Samani & Hariyanto, 2011).

Curriculum Center for the Research and Development Agency of the Ministry of National Education (2011) states that character education aims to form a strong, competitive, noble, moral, tolerant, mutual cooperation, patriotic-spirited, dynamic-developing, science-and-technology oriented human resources, all of which are imbued with faith and piety to God Almighty. This character education must take place both in formal education (in foundations, primary school/*Madrasah Ibtidaiyah*, junior secondary school/*Madrasah Tsanawiyah*, senior secondary school/vocational secondary school/ *Madrasah Aliyah,* and college), non-formal education, and informal education in the family.

Although character education has become a common concern, it turns out that the picture of the community's situation and even the world of education in Indonesia is still concerning. The cases of brawls between students and other forms of juvenile delinquency, bullying, promiscuity, and also drug use are increasing. Corruption cases are even more widespread. In the country, discipline culture, a clean and healthy life, and also respect for the environment are still far below the standard.

Setiawan (2018) reported that the Indonesian Child Protection Commission noted that many Indonesian children were used as drug couriers during 2017. Children became vulnerable because Indonesia was an easy target for drug trafficking. Commissioner of the Center of Indonesian Child Protection Commission (*Komisi Perlindungan Anak Indonesia* or KPAI), Putu Elvina, said many cases of children dealing with the law include drugs. She mentioned, the number was quite good, namely, during 2017, it is around 22 cases of children who became drug couriers. Then in the same year, there were around 46 children who became victims of drug abuse. Meanwhile, Afifah (2019) said that KPAI Chair, Susanto, said that in 2018, cases of children dealing with law ranked first, which is 1,434 cases, then followed by cases related to family and childcare as many as 857 cases.

Puspitawati and Sarma (2012) state that to solve the problem of quality human resources in this country, especially related to the quality of children, a holistic approach that combines family and education systems is needed. The condition of the family is very dependent on the surrounding environment, and vice versa, that the family also affects the surrounding environment. Soedarsono in Samani and Hariyanto (2011) explained that the synergy between home (family), school, and community in terms of character education has not been realized properly so that it has not yet had a multidimensional impact. Table 1 illustrates a portrait building character that is still neglected.

Family is the smallest unit in society that is a pillar supporting the existence of a nation. Family institutions become important centers of activity from various aspects of life. In it, there is a family leader who is

Table 1. Portrait Builds Neglected Characters

	Home	School	Society
Wisdom old age	Increased spiritual approach	?	Many are apathetic
Consolidation of adult age	?	!	• *Low trust society* • No mutual respect • No role model
Development of adolescence	?	!	• Not conducive • Orientation on money, material and worldly
Early age formation	Many are left to helpers	!	• Not conducive

Information: ? = be questioned
! = need attention
Source: (Soedarsono in Samani & Hariyanto, 2011)

usually attached to a man (head of the family), a household manager who is usually attached to a woman (housewife), and children who have the right to education and teaching of both character, religious, and socio-cultural values. The importance of the role of the family is because the family becomes the first and foremost school for children.

The implementation of family education has a strong legal foundation in Indonesia. In the Law of Republic of Indonesia No. 20 of 2003 concerning the National Education System, it is stated that the education unit is an education service group that organizes education on formal, non-formal, and informal channels at every level and type of education. Informal education is the path of family and environmental education. Government Regulation No. 21 of 1994 states that there are eight functions that must be carried out by the family, including functions to fulfill physical and non-physical needs consisting of religious, social, culture, love, protection, reproduction, socialization and education, economy, and environmental development. The mandate of concerning population and family development aims to improve the quality of the family so that the family members find a sense of safe, peaceful, and hope for a better future in realizing physical welfare and inner happiness. In fact, the Minister of Education and Culture of the Republic of Indonesia formed a new echelon II work unit, namely the Directorate of Family Education Development. The duties of the Directorate are regulated in the Regulation of the Minister of Education and Culture No. 11 of 2015, which is to carry out the preparation of the formulation and implementation policies in the field of fostering family education.

The success of education in the family cannot be separated from the role of parents. Interaction in the early years with parents exerts a persistent and long-term influence on the developmental maturity and success of children's education, especially related to character. Thus, parenting patterns become things that need to be studied and developed continuously. Oktafiany, Solihatin, and Japar (2013) state that there is a correlation between parenting patterns and children's emotional intelligence. Misbach (2010) states that in addition to being influenced by parenting patterns, the characters that appear in children are also influenced by genetic factors. It is reinforced by the opinion of Poniman and Mangussara (2012) that the phenotype is influenced by genetics and the environment (genetic heredity (inheritance) and non-heredity (given)). Genetic personality is a non-heredity nature and genetic structure which is a blueprint of a person's strengths and weaknesses and becomes a "code" of each individual.

The character development of children in Aya Sophia Foundation Islamic School is carried out by considering factors that influence the character of both genetic

and environmental factors. Children's genetic personality types are identified at the beginning of the school year. The parenting program is carried out in the hope that the parenting applied by parents at home can be in harmony with the conditions of the school environment. However, it is not yet known whether there is a relationship between genetic personality identification and parenting with the character development of children in Aya Sophia Foundation Islamic School.

Based on the aforementioned description, it is important to conduct research in order to examine more deeply the relationship between character development of children with parenting style and genetic personality. The results of the study become an evaluation material for children's character development programs at Aya Sophia Foundation Islamic School. The study of these factors is also expected to provide enrichment regarding character education in the family. In addition, there are still gaps in previous studies that discuss the relationship of parenting (parenting style) and genetic personality as a whole to the child's character development. Several research results related to this study are as follows.

Huver, Otten, de Vries, and Engels (2010) conduct research on the relationship of personality and parenting style in parents of adolescents in the Netherlands, showing that parenting style influences the development of adolescent personality (Huver et al., 2010). In addition, Nyarko (2011) conducts research on the relationship between the authoritative parenting style and adolescents' academic achievement. The components used to measure authoritative parenting style variables include parent's acceptance/support of children, parental supervision/control of children, and parents' psychological aspects. The results show that there is a positive and significant relationship between the authoritative parenting style and children's learning motivation (Nyarko, 2011).

A research by Talib, Mohamad, and Mamat (2011) on the effects of parenting style on children's development shows that the authoritative style has a positive effect on children's attitudes and learning motivation (Talib et al., 2011). Besides, Jonynienė and Kern (2012) studied the psychological lifestyle and parenting in Lithuanian. The research shows that authoritative parenting is carried out by parents who have an optimistic attitude, focus on solutions, have low stress levels, and are responsible and cooperative (Jonynienė & Kern, 2012). Moreover, Othman and Khairollah (2013) see that there is a positive and significant relationship between the authoritative parenting style and Islamic personality, including worship, trust, and knowledge (Othman & Khairollah, 2013). Pratiwi and Ekowarni (2015) studied the moral intelligence of Chinese ethnic preschoolers in parenting style. The results show that the moral intelligence of children who receive an authoritative parenting style is higher than the authoritarian parenting style, permissive, and uninvolved (neglectful) (Pratiwi & Ekowarni, 2015). Hasnain, Faraz, and Adlakha (2013) agree that parenting style influences children's self-esteem and happiness. Through a one-way ANOVA and Turkey Test analysis, it is known that children with authoritative parenting patterns have higher self-esteem and happiness than children with authoritarian and permissive parenting patterns (Hasnain et al., 2013).

On the other hand, Oktafiany et al. (2013) believe that there is a positive correlation between parenting and the emotional intelligence of children, and the best parenting pattern is democratic (Oktafiany et al., 2013). Studying the intention of child prosocial behavior in terms of parenting style, Utomo (2014) reveals that there is a positive correlation between parenting style and children's prosocial behavior (Utomo, 2014).

Most of those previous studies only partially analyze the influence of parenting style and personality on character development. In addition, none of the studies has analyzed genetic personality.

This study consisted of two independent variables, namely parenting style (X1) and genetic personality (X2), and the dependent variable was the development of children's character (Y). Based on the previous

studies, a research model was developed that illustrates the relationship between variables. The relationship between variables in this research model and the theoretical basis is explained as follows.

The Influence of Parenting Style on the Development of Children's Character

There is an influence between education in the family to the character development of their children. Family is the foundation of the social cooperation unit by involving parents (fathers and mothers) to work together in educating their children (Coleman in Puspitawati & Sarma, 2012). Huver et al. (2010) state that parenting style influences adolescent personality. In their research, Puspitawati and Sarma (2012) explained that there is a correlation between parenting and the child's self-control abilities. In other words, it is stated that the child's behavior is influenced by the parents' treatment of him. The research results of Abidin (2011) also show that parenting style based on love and high positive acceptance, not ignored, not attacked and not rejected, has a positive effect on children's social behavior. It means that parents who adopt a good parenting style will make the child's social behavior better. It is in line with the research of Oktafiany et al. (2013) that there is a positive correlation between parenting and the emotional intelligence of children. It was further explained that the best parenting is democratic. Based on those explanations, it is suspected that parenting style is related to the character development of children in Aya Sophia Foundation Islamic School.

The Influence of Personality Genetic on Children's Character Development

In the context of nature, a person's character is influenced by a genetic structure which is a blueprint of a person's strengths and weaknesses and becomes a "code" for each individual that is permanent (Misbach, 2010). Genetic personality is an innate character associated with the dominance of the brain's work system. Genetic personality can be determined by the biometry method and the method used in this study is STIFIn fingerprint analysis. In addition, Poniman and Mangussara (2012) state that the phenotype is influenced by genotype and environmental factors. One of the determining genotype factors is genetic personality. Thus, it is suspected that genetic personality influences the character development of children in Aya Sophia Foundation Islamic School.

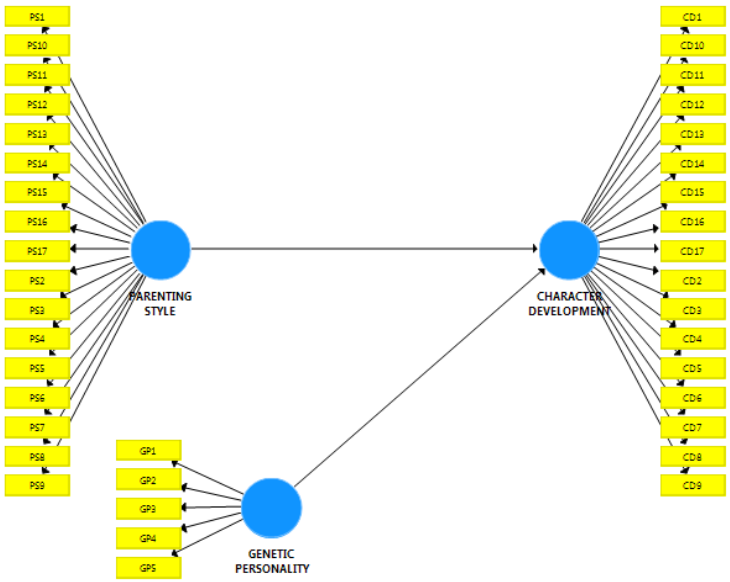

Figure 1. Research Model
Source: Internal data processed by SmartPLS 3.0

Based on those aforementioned relationships, a research model was designed as illustrated in Figure 1. The research model structures the formulation of the hypothesis proposed in this study, elaborated as follows.
H1: Parenting style has a positive and significant effect on the development of children's character in Aya Sophia Foundation Islamic School.
H2: Personality genetic has a positive and significant effect on the character development of children in Aya Sophia Foundation Islamic School.

Research Method

This study aims to determine the effect of parenting style and genetic personality on children's character development at Aya Sophia Foundation Islamic School. Based on the research hypotheses, the research objectives are formulated as follows: (1) knowing the effect of parenting style (X1) on the development of children's character at Aya Sophia Foundation Islamic School (Y), and (2) knowing the influence of personality genetic (X2) on the development of children's character at Aya Sophia Foundation Islamic School (Y).

The research was conducted from September to October 2019 using a survey method with a correlational research approach. Data were collected by distributing questionnaires to students' parents. A modification of the Parenting Style Questionnaire (PSQ) of Robinson, Mandleco, Olsen, and Hart (1995) and Roman et al. (2015) was used to measure the parenting style. The instrument for measuring genetic personality was adapted from Poniman and Mangussara (2013). Adaptations of Poniman, Nugroho, and Azzaini (2014) were used for measuring children's character development. The questionnaire was designed as closed questions, except for questions/statements about the respondent's identity in the form of a semi-open questionnaire. Each closed question/ statement item gives five options: strongly agree (*Sangat Setuju* or SS) with score 5, agree (*Setuju* or S) with score 4, less agree (*Kurang Setuju* or KS) with score 3, disagree (*Tidak Setuju* or TS) with score 2, and strongly disagree (*Sangat Tidak Setuju* or STS) with score 1.

The population in this study was 1,234 parents of students in Aya Sophia Educational Foundation who already know the type of genetic personality of their children through STIFIn fingerprint analysis and are involved in the care of their children. Data were collected by simple random sampling via an electronic questionnaire. The returned and valid questionnaire results were 398 samples. SmartPLS software version 3.0 was used for data analysis since PLS is an analytical method that is soft modeling because it does not assume the data must be of a certain scale measurement (Ghozali, 2014).

Findings and Discussion

Sample Description

Several criteria were established in elaborating the characteristics of the sample. Information on the characteristics of the sample is presented in Table 2.

Table 2. Sample Descriptive Information

Criteria		Amount	Percentage
Age of Parent Couple	< 40 years old	294	73.87%
	≥ 40 years old	104	26.13%
Parent Couple Education	Bachelor	183	45.98%
	Bachelor not yet	215	54.02%
Typology of Student Genetic Personality	Sensing	107	26.13%
	Thinking	99	24.87%
	Intuiting	79	19.85%
	Feeling	78	19.60%
	Instinct	38	9.55%

Source: Internal data processed

Outer Model Testing

The testing phase of the measurement model includes testing for Convergent Validity, Discriminant Validity, and Composite Reliability. The results of the PLS analysis can be used to test the research hypothesis if all indicators in the PLS model have met requirements of those three testing phases.

Convergent Validity Testing

Convergent validity test was done by looking at the loading factor value of each indicator to the construct. For confirmatory research, the loading factor limit used is 0.7, while 0.6 is for exploratory research, and 0.5 is for development research (Ghozali, 2014). Because this research is a confirmatory study, the loading factor limit used is 0.7. The estimated PLS model are shown in Figure 2.

Several indicators have a loading factor below 0.7 so it is declared invalid and must be dropped from the model. Estimation results of the model after an invalid indicator is dropped are shown in Figure 3.

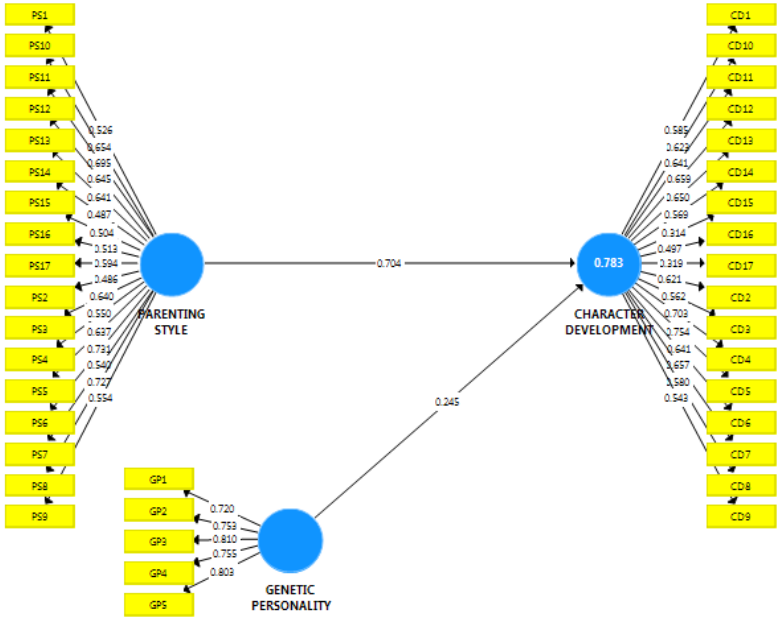

Figure 2. PLS Model Estimation Measurement Model
Source: Internal data processed by SmartPLS 3.0

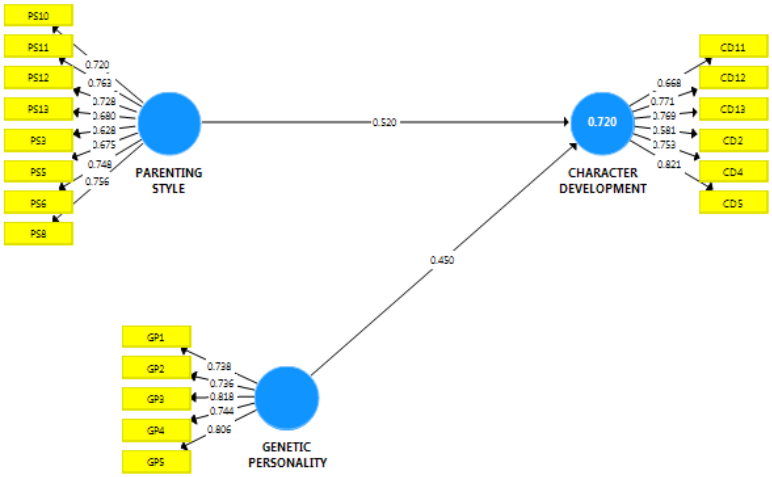

Figure 3. PLS Model Estimation Valid Model
Source: Internal data processed by SmartPLS 3.0

Based on the estimation results of the PLS model presented in Figure 3, all indicators already have a loading factor value above 0.7 so that the model meets convergent validity requirements. Besides looking at the loading factor value of each indicator, convergent validity is also assessed from the AVE value of each construct. The PLS model is declared to have met convergent validity if the AVE value of each construct is > 0.5 (Ghozali, 2014). The AVE value for each construct can be seen in Table 3.

Table 3. Average Variance Extracted Value (AVE)

	Average Variance Extracted (AVE)
CD	0.535
GP	0.592
PS	0.509

Source: Internal data processed by SmartPLS 3.0

Based on the results of PLS analysis in Table 3, the value of AVE for all constructs both in the form of dimensions and variables has exceeded 0.5. It indicates that all indicators in each construct have met the required convergent validity criteria.

Descriminant Validity Testing

Discriminant validity was carried out to ensure that each concept of each latent variable is different from the other variables. The model has good discriminant validity if the AVE squared value of each exogenous construct (the value on the diagonal) exceeds the correlation between construct and other constructs (values below the diagonal) or the values are above of 0.7 (>0.7) (Ghozali, 2014). The discriminant validity test results are obtained, as presented in Table 4.

Table 4. Discriminant Validity Value

	CD	GP	PS
CD	0.732		
GP	0.725	0.769	
PS	0.757	0.528	0.714

Source: Internal data processed by SmartPLS 3.0

The results of the discriminant validity test in Table 4 show that all constructs have the AVE square root value above the correlation value with other latent constructs. Thus, it can be concluded that the model meets discriminant validity.

Composit Reliabitily Testing

Construct reliability can be assessed from the Alpha Cronbach value and the Composite Reliability value of each construct. The recommended composite reliability and Cronbach's alpha values are more than 0.7. However, in development research, because the loading factor limit is low (0.5), low composite reliability and Cronbach's Alpha values can still be accepted as long as convergent and discriminant validity requirements have been met (Ghozali, 2014).

Table 5. Composite Reliability Value

	Cronbach's Alpha	Composite Reliability
CD	0.822	0.872
GP	0.827	0.878
PS	0.862	0.892

Source: Internal data processed by SmartPLS 3.0

Table 5 show that all constructs have composite reliability and Cronbach's Alpha values> 0.7. In conclusion, all constructs have met the required reliability.

Inner Model Testing

Testing the inner model includes testing the significance of direct effects and measuring the magnitude of the influence exogenous on endogenous variables. With the bootstrapping technique, R Square values and significance test values are obtained, presented in Table 6.

Table 6. R Square Value

	R Square	R Square Adjusted
CD	0.720	0.718

Source: Internal data processed by SmartPLS 3.0

Based on Table 6, the R Square value of 0.720 means that the child character development variable can be explained by parenting style and genetic personality variables of 72.0%. Meanwhile, the remaining 28% is explained by other variables not discussed in this study.

Table 7. Value of Significance Test Results

	Original Sample (O)	T Statistics (|O/STDEV|)	P Values
GP -> CD	0.450	9.298	0.000
PS -> CD	0.520	13.585	0.000

Source: Internal data processed by SmartPLS 3.0

Table 7 illustrates the findings of the study. Each point is elaborated as follows.

The Influence of Parenting Style on the Development of Children's Character

The research findings show that parenting style has a positive and significant influence on the character development of children in Aya Sophia Foundation Islamic School. It is evidenced by the p-value of 0.000 which is smaller than 0.05 and the statistic T value of 13.585 which is greater than 1.96. The original sample value is 0.520 which is positive, so, it is an indication that the H1 hypothesis is accepted.

The findings of this study are in line with the Bronfenbrenner Ecological Theory which states that child development is influenced by five environmental systems that range from interpersonal interactions to broader cultural influences. Bronfenbrenner refers to these systems as microsystems, mesosystems, ecosystems, macrosystems, and chronosystems. In early childhood, the most dominant influence is the microsystem in which children spend a lot of time. Some contexts in this system include family, peers, school, and neighbors (Santrock, 2008). However, it does not mean that other environmental systems do not contribute to children's development. Mesosystems are inter-microsystem links such as experiences in families and schools, ecosystems are stakeholder policies related to child development, and macrosystems such as community culture also influence children's development. The condition of sociohistoric (chronosystem) is undeniably influential, where the 21st-century children are the generation of Z which is inseparable from the influence of media and technology development. Aziz (2012) states that the family plays a vital role in the formation and development of character for each member, especially children.

The findings in this study are also in line with the results of Huver et al. which show that parenting style influences the personality development of children (Huver et al., 2010). In addition, it is also in line with the research results of Talib et al. which state that parenting style influences children's attitudes (Talib et al., 2011) and Efobi and Nwokolo's research results which show that parenting style has an impact on children's development (Efobi & Nwokolo, 2014).

Based on the results of this study, character development in schools, especially in Early Childhood Education institutions, involves the role of parents. Parenting education programs in the form of training and workshops in an effort to align parenting in family and the process of character building in schools need to be a major concern.

The Influence of Genetic Personality on Children's Character Development

Research findings indicate that genetic personality has a positive and significant influence on the development of children's character at Aya Sophia Foundation Islamic School. It is evidenced by the p-value of 0.000 which is smaller than 0.05 and the statistic T value of 9.298 which is greater than 1.96. The original sample value is 0.450 which is positive, so the H2 hypothesis is accepted.

The findings of this study corroborate the Dryden and Vos research in Musrofi (2011) which states that each child has potentially a unique character. It is in accordance with the opinion of Murakami (2013) which states that each person is unique. No two sets of genes are exactly the same, no two people are exactly the same. The difference in each person is not only manifested in the face or appearance, but also in nature and abilities. Misbach (2010) reinforces opinion above that in the context

of nature, a person's character is influenced by a genetic structure which is a blueprint of a person's strengths and weaknesses and becomes a "code" for each individual that is permanent. Poniman and Mangussara (2012) state that phenotype is influenced by genotype and environmental factors. One of the determining genotypes is genetic personality, which is an innate character associated with the dominance of the work system of the brain. The results of data analysis show that the genetic personality of children in Aya Sophia Foundation Islamic School is dominated by 27% sensing, then 25% feeling, 20% thinking and intuition respectively, and 8% instinct. Based on the results of this study, it should be the development of children's character in schools, especially in institutions of Early Childhood Education also pay attention to the genetic personality factors of children. Genetic personality identification can be done at the beginning of the school year and schools can collaborate with psychological institutions or human resource development institutions that are affordable.

The findings in this study are also in line with the results of neuroscience research which states that there is a very close relationship between a person's psychological condition and the work system of his brain structure. Meanwhile, the development of dermatoglyphics and dactyloscopy-research related to the structure of fingerprints - illustrates a relationship between biological structures, in this case, fingerprinting with the brain's work system so it can be revealed the relationship between fingerprint patterns with interdisciplinary various fields of science including psychology and education. Dr. Mary Lai, Ph.D., MME from Taiwan is one of the educators who use the benefits of fingerprinting and dermatoglyphics research in parental counseling (Misbach, 2010).

The results of this study support the opinion of Poniman and Mangussara (2012) that phenotype is influenced by genotype and environmental factors. In this case, the character is a phenotype, parenting style is part of environmental factors, whereas genetic personality is part of genotype factors. Murakami (2012) states that each gene contains a lot of information. Genetics greatly influences the behavior of a species, though on the other hand, the environment is also believed to play an important role. Research on genetic links to character is still being conducted until now.

The results of this study corroborate several previous studies, one of which is by Ferguson (2010) that genetics influence antisocial character and behavior by 56%. Miles and Carey (1997) stated that genetic and environmental factors influence children's aggressive behavior by 50% each. Environmental factors that have a strong influence are a family environment. Meanwhile, Carey and DiLalla (1994) stated that genetic factors influence character with a percentage between 30% to 60%.

Based on the results of this study, it should develop the character of children in schools, especially in Early Childhood Education institutions to pay attention together between the factors of parenting style of children's genetic personality. Educational programs about developing children's character based on parenting and genetic personality should continue to be done both to teachers as educators in the school environment and parents as educators in the family environment.

The characters developed at the Aya Sophia Foundation Islamic School refer to the institution's vision, focusing on *sholeh*, intelligent, and independent characters. The three characters are also in accordance with character education guidelines set by the Ministry of Education and Culture about the 18 character values that need to be developed. Religious, honest, tolerance, care for the environment, and peace-loving are included in the character of *sholeh*. Meanwhile, curiosity, interest in reading, and appreciating achievements are included in intelligent characters. Independence also includes discipline, hard work, care for the environment, and responsibility. Equalization of perception between school and parents (family) related to the character developed was car-

ried out intensively since the beginning of the new school year in the hope that there is harmony between the school and family environment.

Conclusion

This research is a survey conducted at Aya Sophia Foundation Islamic School in Tangerang Regency to get a picture of the influence of parenting style and genetic personality on the development of children's character. Based on the results of the study, the following conclusions can be drawn. (1) Parenting style provides a positive and significant influence on the development of children's character at Aya Sophia Foundation Islamic School. It means that the more positive the parenting pattern gets, the better the character development process of the child. (2) Genetic personality has a positive and significant influence on the development of children's character at Aya Sophia Foundation Islamic School. It means that the better the identification of genetic personality gets, the better the character development process of children. The results of data analysis show that the genetic personality of children in Aya Sophia Foundation Islamic School is dominated by sensing by 26.13%, then thinking 24.87%, intuiting 19.85%, feeling 19.60, and instincts 9.55%.

References

Abidin, A. R. (2011). Analisis gender pada gaya pengasuhan, proses pembelajaran di kelas, perilaku sosial dan prestasi belajar siswa SMA di Kota Bogor. Master Thesis, Institut Pertanian Bogor, Bogor.

Afifah, A. (2019). KPAI terima pengaduan 4.885 kasus anak. Retrieved October 15, 2019, from Dakta.com: Bijak dan cerdas website: http://www.dakta.com/news/17920/selama-2018-kpai-terima-pengaduan-4885-kasus-anak

Aziz, H. A. (2012). Pendidikan karakter berpusat pada hati. Jakarta: Al Mawardi Prima.

Carey, G., & DiLalla, D. L. (1994). Personality and psychopathology: Genetic perspectives. Journal of Abnormal Psychology, 103(1), 32–43. https://doi.org/10.1037/0021-843X.103.1.32

Curriculum Center for the Research and Development Agency of the Ministry of National Education. (2011). Pedoman pelaksanaan pendidikan karakter. Jakarta: Curriculum Center for the Research and Development Agency of the Ministry of National Education.

Efobi, A., & Nwokolo, C. (2014). Relationship between parenting styles and tendency to bullying behaviour among adolescents. Journal of Education & Human Development, 3(1), 507–521. Retrieved from http://jehdnet.com/journals/jehd/Vol_3_No_1_March_2014/27.pdf

Ferguson, C. J. (2010). Genetic contributions to antisocial personality and behavior: A meta-analytic review from an evolutionary perspective. The Journal of Social Psychology, 150(2), 160–180. Retrieved from http://www.christopherjferguson.com/evmeta.pdf

Ghozali, I. (2014). Structural Equation Modeling: Metode alternatif dengan Partial Least Square. Semarang: Badan Penerbit Universitas Diponegoro.

Government Regulation No. 21 of 1994 on the the Implementation of Family Welfare Development. , (1994).

Hasnain, N., Faraz, B., & Adlakha, P. (2013). Self-esteem and happiness of children and mothers of different parental authority. The International Journal Of Humanities & Social Studies, 1(3), 1–6. Retrieved from http://internationaljournalcorner.com/index.php/theijhss/article/view/128036/0

Huver, R. M. E., Otten, R., de Vries, H., & Engels, R. C. M. E. (2010). Personality and parenting style in parents of

adolescents. Journal of Adolescence, 33(3), 395–402. https://doi.org/10.1016/j.adolescence.2009.07.012

Jonynienė, J., & Kern, R. M. (2012). Individual psychology lifestyles and parenting style in Lithuanian parents of 6- to 12-year-olds. International Journal of Psychology: A Biopsychosocial Approach / Tarptautinis Psichilogijos Žurnalas: Biopsichosocialinis Požiūris, 11(11), 89–117. https://doi.org/10.7220/1941-7233.11.5

Law of Republic of Indonesia No. 20 of 2003 on National Education System. , (2003).

Miles, D. R., & Carey, G. (1997). Genetic and environmental architecture on human aggression. Journal of Personality and Social Psychology, 72(1), 207–217. https://doi.org/10.1037/0022-3514.72.1.207

Misbach, I. H. (2010). Dahsyatnya sidik jari: Menguak bakat dan potensi untuk merancang masa depan melalui fingerprint analysis. Jakarta: Visimedia.

Murakami, K. (2012). The miracle of DNA. Bandung: Mizan Media Utama.

Murakami, K. (2013). Misteri DNA. Jakarta: Gramedia.

Musrofi, M. (2011). Potensi preneur. Surakarta: Talents Center.

Nyarko, K. (2011). The influence of authoritative parenting style on adolescents' academic achievement. American Journal of Social and Management Sciences, 2(3), 278–282. https://doi.org/10.5251/ajsms.2011.2.3.278.282

Oktafiany, N. D., Solihatin, E., & Japar, M. (2013). Hubungan pola asuh orang tua dengan kecerdasan emosional siswa di SMP Diponegoro 1 Jakarta. Jurnal PPKN UNJ Online, 1(2), 1–15. Retrieved from https://dokumen.tips/documents/hubungan-pola-asuh-orangtua-dengan-kecerdasan-emosional-siswa-di-smp-diponegoro.html

Othman, N., & Khairollah, S. (2013). Exploring the relationship between Islamic personality and parenting style. International Journal of Islamic Thought (IJIT), 4, 48–57. Retrieved from http://journalarticle.ukm.my/6784/

Poniman, F., & Mangussara, R. A. (2012). Konsep PALUGADA. Jakarta: STIFIn Institute.

Poniman, F., & Mangussara, R. A. (2013). STIFIn personality. Jakarta: STIFIn Institute.

Poniman, F., Nugroho, I., & Azzaini, J. (2014). Kubik leadership: Solusi esensial meraih sukses dan hidup mulia. Jakarta: PT Gramedia Pustaka Utama.

Pratiwi, S., & Ekowarni, E. (2015). Kecerdasan moral anak usia prasekolah etnis Cina ditinjau dari gaya pengasuhan orangtua. Proceeding Seminar Nasional Positive Psychology 2015. Surabaya: Unika Widya Mandala.

Puspitawati, H., & Sarma, M. (2012). Sinergisme keluarga dan sekolah. Bogor: IPB Press.

Regulation of the Minister of Education and Culture No. 11 of 2015 on the Organization and Implementation of the Ministry of Education and Culture. , (2015).

Robinson, C. C., Mandleco, B., Olsen, S. F., & Hart, C. H. (1995). Authoritative, authoritarian, and permissive parenting practices: Development of a new measure. Psychological Reports, 77(3), 819–830. https://doi.org/10.2466/pr0.1995.77.3.819

Roman, N. V., Davids, E. L., Moyo, A., Schilder, L., Lacante, M., & Lens, W. (2015). Parenting styles and

psychological needs influences on adolescent life goals and aspirations in a South African setting. Journal of Psychology in Africa, 25(4), 305–312. https://doi.org/10.1080/14330237.2015.1078087

Samani, M., & Hariyanto, H. (2011). Konsep dan model pendidikan karakter. Bandung: Remaja Rosdakarya.

Santrock, J. W. (2008). Educational psychology (3rd ed.). New York, NY: McGraw-Hill.

Setiawan, D. (2018). KPAI catat anak dimanfaatkan jadi kurir narkoba. Retrieved October 15, 2019, from Komisi Perlindungan Anak Indonesia website: https://www.kpai.go.id/berita/kpai-catat-anak-dimanfaatkan-jadi-kurir-narkoba

Talib, J., Mohamad, Z., & Mamat, M. (2011). Effects of parenting style on children development. World Journal of Social Sciences, 1(2), 14–35.

Utomo, D. (2014). Intensi perilaku prososial anak ditinjau dari gaya pengasuhan. Cognicia, 2(1). Retrieved from http://ejournal.umm.ac.id/index.php/cognicia/article/view/1817

Jurnal Penelitian dan Evaluasi Pendidikan
Volume 23, No 2, December 2019 (219-229)
Online: http://journal.uny.ac.id/index.php/jpep

EVALUATION OF LECTURER IN HIGHER EDUCATION CURRICULUM BASED ON THE NATIONAL STANDARDS OF HIGHER EDUCATION NO. 44 OF 2015

I Made Sundayana
Sekolah Tinggi Ilmu Kesehatan Buleleng

Putu Dian Prima Kusuma Dewi
Sekolah Tinggi Ilmu Kesehatan Buleleng

Putu Sukma Megaputri
Sekolah Tinggi Ilmu Kesehatan Buleleng

Abstract

The quality of education is still a benchmark of human resources (HR) in each country. The current era of revolution 4.0 requires a curriculum that is innovative, creative, and based on the needs of market share and the world of work with the addition of several new literacies. This study aims to identify and evaluate the tertiary curriculum (higher education curriculum) that has been applied in a higher education institution of health sciences in Buleleng, which refers to the national standard of higher education no. 44 of 2015. The evaluation was based on input, process, and output. The method used was the mixed method by using questionnaires and focus group discussions on 32 lecturers from the three knowledge fields of midwifery, nursing, and pharmacy. The results show that of the four standards evaluated, the learning process standard and assessment standard need to get priority in improving the academic system at the institution. Both of these components must be immediately improved in terms of concepts and understanding so that the goals, achievements, and quality of graduates can still be improved and maintained. Internal discussions and workshops should be done regularly each semester to refresh pedagogical ability and understanding of the regulatory faculty curriculum.

Keywords: *evaluation, curriculum, national standard of higher education, STIKes Buleleng*

Permalink/DOI: http://dx.doi.org/10.21831/pep.v23i2.28141

Contact: *I Made Sundayana*
madesundayana@gmail.com
Department of Nursing, Sekolah Tinggi Ilmu Kesehatan Buleleng
Jl. Raya Air Sanih Km. 11, Bungkulan, Singaraja, Bali, Indonesia

Introduction

Curriculum is a program that can be planned and implemented to achieve a number of specific educational goals. Some very important curriculum components include goals, materials, methods, students, educators, media, environment, learning resources, and others. This curriculum component must be further developed so that the objectives or profile of the department in each institution can be achieved as appropriate (Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015). The purpose of the existence of Law No. 12 of 2012 of Republic of Indonesia on higher education is that higher education curriculum can be developed by each tertiary institution by referring to the National Higher Education Standards (*Standar Nasional-Dikti*) for each study program covering knowledge, attitudes, general skills and special skills (Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015). The issuance of the Presidential Regulation No. 8 of 2012 on the Indonesian National Qualification Framework (*Kerangka Kualifikasi Nasional Indonesia* or KKNI) which was later also issued by the Regulation of the Minister of Education and Culture No. 49 of 2014 which was updated to the Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015 urged higher education so that curriculum reconstruction must refer to the national standard of the Ministry of Higher Education.

A curriculum that has been prepared by the study program can be periodically reviewed for the achievement of social and scientific relevance in accordance with what is needed by the community and the development of science and technology (National Accreditation Agency for Higher Education, 2008). Thus, the curriculum can be run by education practitioners and education targets well. Lecturers are the spearhead of implementing actors giving methods to achieve *learning outcomes* desired. As a lecturer must run the *Tri Dharma* properly and correctly, the curriculum review process carried out in stages is also related to the process of monitoring and evaluation of the *Tri Dharma* of Higher Education to improve the quality of community life (Wijanto, 2009).

The quality of education is still a benchmark of human resources (HR) in each country. The current era of revolution 4.0 requires a curriculum that is innovative, creative, and based on the needs of market share and the world of work with the addition of several new literacies. Some developed countries evaluate education as an initial evaluation of the quality of their human resources (Kazimirov, 2018; Tiwari, 2018). Various regulations are strived to support the achievement and improvement of the quality of Indonesian human resources. Measurement of the quality of a study program greatly determines the quality of graduates to be produced, one of the most important components is the curriculum. The curriculum is a set of plans and arrangements regarding graduate learning outcomes, study materials, processes, and assessments that are used as guidelines for organizing study programs (Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015).

The National Higher Education Standard in the Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015 is the minimum criterion in determining higher education standards. Of course, the *tri dharma* is a benchmark in its achievement, namely education, research, and community service. National education standards state that eight standards must be met. There are four important standards in an educational process, namely graduate competency standards, content standards, learning process standards, and assessment standards. New literacy according to the demands of the revolution-based higher education curriculum (*Kurikulum Perguruan Tinggi* or KPT) 4.0 expects data literacy, technological literacy, and human literacy to be important components in the preparation of educational standards. The industrial, market, and tertiary education sectors should establish a good partnership, so that the

graduates produced can suit the needs in the field (Kaklauskas et al., 2018). Therefore, lecturers in this 4.0 era are required to have innovations and new learning methods not just lectures but simulations that are in line with competencies expected in the world of work (Cassano, Costa, & Fornasari, 2019; Kazimirov, 2018). Answering the challenge, of course, there must be an appropriate curriculum and framework in the education process so that graduates are not only able to work in the world of work but also are of good quality.

When the curriculum has been prepared properly according to the needs of stakeholders, its process does not merely stop on the shape of the curriculum. Instead, it requires evaluation. Evaluation is needed in assessing the success of a standard that has been applied. Evaluation is divided into internal and external evaluations. Evaluation of modern education places an assessment of learning outcomes at the center in the evaluation process, and evaluation of the "teacher" as a role model of change (Tiwari, 2018). The big challenge faced by tertiary education in Indonesia today is developing learning outcomes that fit the needs of the community and industry. Higher education in the health field often have difficulties in this implementation because most of the products are in the form of services of health workers. Before compiling a graduate learning achievement (*Capaian Pembelajaran Lulusan* or CPL), the most important input is the understanding and innovation of resources, one of which is the lecturer. The ability of lecturers to master CPL and related regulations will largely determine the direction and outputs of graduates (Cassano et al., 2019; Pribadi, 2019).

This study aims to identify and evaluate the tertiary curriculum (KPT) that has been applied in Sekolah Tinggi Ilmu Kesehatan (STIKes) Buleleng, which refers to the National Standard of Higher Education in the Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015, which is important as a benchmark and internal quality assurance. STIKes Buleleng, as one of the health higher education institutions in the North Bali region, must always be able to maintain the quality of health workers produced.

Research Method

This research took place in August 2019 involving all lecturers in STIKes of Buleleng from six Study Programs and three fields of science, namely, midwifery, nursing, and pharmacy. The number of lecturers involved in this study was 32 people who were spread in those three fields of science. This study used a mixed-method (quantitative-qualitative) to assess the evaluation and implementation of a curriculum that has been implemented, including the level of understanding of STIKes Buleleng lecturers of the higher education curriculum referring to the national standard of Higher Education in the Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015. Data were collected through the questionnaire and focus group discussion (FGD). The questionnaire was distributed online via the WhatsApp group. The online instrument can be accessed through this link http://bit.ly/Eval_Pend.

The questionnaire compiled consisted of several questions divided into four components to evaluate input, process, and output, namely competency standard (four questions), content standard (10 questions), learning process standard (14 questions), and assessment standard (five questions). The questionnaire was compiled based on references from the national standard of higher education and then developed according to the needs. The content analysis, validity, and reliability processes were previously carried out to the questionnaire to assess the validity of the questions given. The questions were compiled using Bloom's taxonomic reference to evaluate the components of national standards and the higher education curriculum. The questions for the FGD were divided into 16 questions to assess qualitatively. FGD activities were divided into two groups within two days.

The data obtained were subsequently through the tabulating, editing, and clearing stages before further analysis. The data obtained were then analyzed using software assistance, namely NVIVO 12 Plus and SPSS version 17. The NVIVO 12 Plus software was used to help see the essence and mindset of each respondent for the questions given, while SPSS version 17 software was used to analyze the data quantitatively.

Findings and Discussion

The results of the evaluation and implementation of the higher education curriculum are viewed from the four main standards by looking at the inputs, namely the lecturer. These four standards refer to the national standard of Higher Education No. 44 of 2015 concerning national education standards. Demographic analysis can be seen in Table 1.

Table 1. Respondent Demographic Conditions

Variable	f (%)
Age (Mean ± SD)	30.8 ± 5.48
Length of working (Mean ± SD)	5.34 ± 3.55
Position	
Yes	16 (50.0)
No	16 (50.0)
Study Program	
Midwife	14 (43.8)
Nurses	16 (50.0)
Pharmacy	2 (6.3)
Last Education	
S1	4 (12.5)
S2	27 (84.3)
S3	1 (3.2)

Table 1 shows that the average age of lecturers at STIKes Buleleng is 30.8 or 31 years old, with the working period of more than five years, and largely comes from the field of obstetrics and nursing. In addition, 50% of the respondents are also in structural and functional positions. These results indicate that the majority of lecturers are in the productive age range, in which this age range has high potential and enthusiasm for work, especially has a high level of innovation as well as creativity owned (Kazimirov, 2018; Setyawati, 2015).

An educator plays an important role as facilitator, motivator, and inspiration for the learners in the learning process. Continuous innovation is needed to produce smart, creative, independent, and responsible students (Setyawati, 2015). The results of this study also found that the last education of the most human resources involved in this study was the master's degree, as many as 84.3%.

Input

Graduate Competency Standard

In this standard, there are many indicators arranged. Those indicators are listed in Table 2.

Table 2. Indicators of Graduates Competency Standard

Variable	f (%)
SN Dikti Knowledge	
Yes	31 (96.9)
No	1 (3.1)
Compilation of Competencies of Graduates	
Correct	30 (93.8)
Incorrect	2 (9.4)

The aforementioned results show that almost all lecturers know about the national standard of higher education stated in the Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015 and know that graduate competencies are prepared based on the stakeholder input or the demand of the industry. The quantitative results above are different from the qualitative study results. As illustrated in Figure 1, in the FGD, some respondents are still ambiguous between graduate competency or graduate profile with the CPL courses, evidenced by the results obtained that almost all lectures answer that the CPL subjects that are taught are not CPL from the study program. The mindset found that the study program CPL and course CPL are regarded as the same thing despite having different meanings both in content and formulation. Study program CPLs are compiled from agreements and input from stakeholders, while course CPLs are derived from

curriculum maps based on study material (Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015). Only two respondents named study program CPL that "produces care provider midwives and professional nurses in the field of HIV/AIDS".

Figure 1. FGD Process Step One

Content Standard

In this phase, FGD step two was carried out, illustrated in Figure 2. In addition, Table 3 shows that 90.6% of lecturers never reviewed the syllabus that had been prepared with the process of preparing incorrect syllabus to 53.1%. Nevertheless, most of the lecturers (84.4%) knew that the syllabus is the responsibility of lecturers, with minimal components that are suitable with the national standard of Higher Education as stated in the Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015 (56.3%). This content standard includes criteria for a minimum level of depth and breadth of learning material (Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015).

Figure 2. FGD Process Step Two

Table 3. Criteria for Content Standard

Variable	f (%)
Have reviewed RPS	
Yes	3 (9.4)
No	29 (90.6)
Completing RPS	
Correct	27 (84.4)
Incorrect	5 (15.6)
Minimum Components	
Right	18 (56.3)
Wrong	14 (43.8)
RPS Sequence	
Correct	15 (46.9)
Incorrect	17 (53.1)

The stages of drafting the Higher Education Curriculum (KPT) include (1) determining graduate profiles and learning outcomes (CPL), (2) evaluating each course on the old curriculum structure, (3) selecting and arranging learning materials, (4) composing courses, curriculum structure, and determining semester credits (*Sistem Kredit Semester* or SKS), and (5) developing learning plans (Pribadi, 2019; Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015). The levels of depth and breadth of learning material are decisive in the success of a student's learning and achievements (Luttenberger et al., 2018).

The preparation of the syllabus is an important thing to do. The syllabus is one of the guidelines in the whole semester's learning process so the process of preparing the syllabus requires the team's assistance in doing everything. Oftenly, the syllabus cannot be used as a guide by students and lecturers to design learning and assignment due to its too narrow, too general, or too rigid arrangement so it does not give flexibility to lecturers and students. Such learning experiences will complicate and weaken students' motivation to learn (Hussey & Smith, 2010).

Learning Process Standard

FGD step three was carried out as illustrated in Figure 3. Besides, Table 4 shows that most lecturers have a good understanding of the learning process standard as seen from the percentage of correct answers from the qualitative learning and determination of effective weeks in each semester.

Figure 3. FGD Process Step Three

Table 4. Criteria for Learning Process Standard

Variable	f (%)
Learning Method	
Correct	26 (81.3)
Incorrect	6 (18.8)
Form ofLearning	
Correct	21 (65.6)
Incorrect	11 (34.4)
Amount of 1 SKS	
Correct	24 (75.0)
Incorrect	8 (25.0)
Amount Effective Week	
Correct	25 (78.1)
Incorrect	7 (21.9)
Amount Seminar Time	
Correct	24 (75.0)
Incorrect	8 (25.0)
Amount Practical Time	
Correct	27 (84.4)
Incorrect	5 (15.6)
Amount of Structured Assignment Time	
True	25 (78.1)
False	7 (21.9)

According to the national standard of higher education no. 44 of 2015, the principles of the assessment refer-red to in Article 19 paragraph (2) letter 'a' include the principles of educative, authentic, objective, accountable, and transparent which are carried out in an integrated manner (Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015).

Each form of learning should be determined by regulations in accordance with the targets and levels of education or achievements that have been prepared (Kazimirov, 2018). In the national standard of higher education number 44 of 2015, the learning process standard is the third important standard applied in the KPT. The learning process standard is a criterion for determining each minimum allocation of forms on the implementation of learning in the study program to obtain the learning outcomes of graduates (Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015).

The initial lecture conducted by lecturers is mostly done by giving syllabus to students at the initial meeting. Thus, it is found that very few lecturers did not make or give syllabus to their students. It indicates that the learning process becomes more relevant, motivates students to learn, focuses on student needs, presents meaningful learning, and helps students to be independent and supports lifelong learning (Sitepu & Lestari, 2018).

Assessment Standard

"The principle of learning assessment done is objective and transparent" is the answer most oftenly raised by lecturers. This answer is not an incorrect thing, but it does not cover all the principles of minimum assessment on the assessment standards at the national standard of higher education. The principles in the assessment include the principles of educative, authentic, objective, accountable, and transparent assessment carried out in an integrated manner (Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015). The educational principle referred to is an assessment that motivates students to be able to improve planning and learning methods; and achieve learning outcomes of the graduates.

Authentic educative principle is implicit in the rubric assessment arranged either form of essays and portofolio. The most oftenly mentioned assessment rubric is the cognitive and psychomotor assessment rubric in addition to the essay rubric and portfolio. Assessment as a tool to evaluate students' progress on the main dimensions of learning so that a process needs to be sustainable and comprehensive, which has been

highly emphasized in all the latest educational documents, especially in the National Curriculum (Tiwari, 2018).

Human Resource Standard

Input process in the implementation of learning requires human resources aspect. Human resources include lecturers and educational staff. In accordance with the national standard of higher education number 44 of 2015, it is explained that the criteria for a lecturer is to have a minimum master degree education (Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015). Education resources become an important factor which will provide input, new knowledge, and new experiences for students they educate. The results of this study report that (Table 1) the majority of lecturers fit the criteria of the national standard of higher education, which is, having the master degree education. The results of previous studies also found that the qualifications of an educator must meet the desired competencies so that the learning process runs smoothly and the CPL can be achieved (Rahmawati & Anggraini, 2017).

Process

The process that has been carried out by lecturers based on the national standard of higher education in the Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015 are including competency standard, content standard, learning process standard, and assessment standard. The results are described qualitatively and quantitatively as follows.

Graduate Competency Standard

The results of the FGD show that the CPL component had been prepared and established clearly, in detail, and was contained in the syllabus that had been prepared by the needs and profile of graduates in each study program. Most lecturers (84.3%), before conducting lectures, had given syllabus and lecture contracts. There are 15.6% of lecturers who consider syllabus is less important than the academic calendar.

Competency standards are the main point in starting a process in the Study Program. This standard is like a map so that all components do not lose direction of the objectives to be achieved (Kazimirov, 2018; Pribadi, 2019; Rodríguez-Conde, Olmos-Migueláñez, Gamazo, & O'Hara, 2018; Tiwari, 2018).

Content Standard

Content standards include minimal components that must be designed and stated in the curriculum. The content standard becomes the next reference after the competency standard which has been determined from the beginning.

Based on the results of the questionnaire that almost all lecturers (97.8%) mentioned that in the process of preparing the syllabus conducted by lecturers supporting the course, only 3.2% stated that the syllabus was prepared according to the authorization of the head of the study program. The process of preparing the syllabus conducted at the beginning of the semester also includes the quality control process of the curriculum development unit (*Unit Pengembangan Kurikulum* or UPK) in STIKes Buleleng.

The results of the FGDs conclude that the existence of the the UPK was very beneficial for the lecturers, especially to be able to become an external assessment of the syllabus and the tools that had been prepared. Besides, evaluation and monitoring are also carried out both in the middle and end of the semester, by looking back at the lecturers' teaching journals. This process is carried out by the course coordinator lecturer, which is then reported to the chair of the study program.

Related to the content in the syllabus, it is not 100% done through a face-to-face meeting, but also done online between 10-20% of the learning material provided. The blended learning process is still in the trial process that is implemented at STIKes Buleleng. Various components and readiness in terms of lecturers' abilities, content developed, and students' readiness to accept this process also need to be reviewed. The sem-

ester learning plan (syllabus), task design, Learning Achievement of Courses (*Capaian Pembelajaran Mata Kuliah* or CPMK), learning materials, and learning methods are components that are often out of sync in the learning tools that have been prepared so that the monitoring process needs to be done.

Learning Process Standard

After competency standards and content standards are met, the next is the standard of the learning process. In this case, it is not only centered on lecturers but also students. The process and form of learning are often two things that are difficult to distinguish. The form of learning is a method used to fulfill the specified learning process. The learning process is bound to the unit of time specified in the regulation (Pribadi, 2019; Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015).

Some lecturers mentioned that the learning process is the same as the form of learning. The process and form of learning are considered as a single entity that cannot be distinguished. This statement also affects the time allocation process specified in each learning process.

The results obtained at the process stage are that the method applied is in line with the study material that has been prepared, including the learning outcomes to be achieved. All components of the CPMK are translated in the form of Bloom's taxonomy from cognitive, affective, and psychomotor aspects.

The results of observations on this aspect of the process were also seen from the existing lecturer teaching journal documents which were then checked with each student who was responsible for the course. Based on the results of the interview, there is no learning process which is not by the syllabus or study material from the three fields of science. Face to face meeting must have been conducted at least 14 times outside of the guidance and discussion process which is always opened through media facilities, such as WhatsApp and group e-mails.

The synchronization between the syllabus, task design, description of methods, and study material in input will determine the process carried out next (Hartini & Suryanti, 2019). In addition to synchronization, a pedagogical ability of lecturers has also been proven to affect the learning motivation of students (Rahman, Mutiani, & Putra, 2019). Lecturers become role models indirectly for students, then their attitude is demanded to be good so they can also give good examples.

Assessment Standard

The results of lecturer evaluations on the standard components of this assessment indicate that each course has a clear assessment rubric and is listed in the syllabus. All the assessment principles specified in the Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015 are contained in the rubric of assessment and syllabus prepared by lecturers. The most widely applied components are fair, objective, accountable and transparent. The principle of assessment as referred to in article 19 paragraph (2) letter 'a' includes the principles of education, authentic, objective, accountable, and transparent which are carried out in an integrated manner.

The results of observations in this study show that the compiled syllabus shows the assessment components included in the final syllabus sheet and refers to the academic guidelines that apply to STIKes Buleleng. Remedial is an educational evaluation principle carried out transparently and accountably by announcing the score obtained and then evaluating the components that have not yet been achieved. In the component of the education standard at the national standard of education number 44 of 2015, the assessment standard is the most important component at the end of an educational process. Learning appraisal standard is a minimum criterion on the assessment of student learning processes and outcomes in the context of fulfilling graduate learning outcomes (Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015).

The FGD process related to the assessment standards showed that the final results of the assessment that had been carried out included the remedial process, reported to the lecturer coordinator of the course and the head of the study program at the end of the semester. The assessment process is the end of the evaluation process of a learning process by considering all aspects assessed (Bashri, Puspitawati, & Ibrahim, 2018; Hartini & Suryanti, 2019; Rahman et al., 2019; Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015).

Output

Graduate Competency Standard

The syllabus is by the requirements of the national standard of higher education wherein its preparation should follow the existing minimum standards according to the the standards listed there. In the syllabus, the name of the study program, courses, semesters, credits, lecturers' names applicants, graduate learning outcomes, expected end skills, study materials, learning methods, the time provided, student learning experiences, criteria, indicators and assessment weight, and the list of available references are arranged (Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015).

The lecturers' syllabus study compiled with the team is also evaluated in terms of output. Most of the lecturers, in the preparation of the syllabus, have collaborated with the team following the various decree of the chairperson. When the syllabus drafting team and its teaching team have been determined, the lecturer starts a discussion related to the study material included to obtain the CPL that the study program wants.

Adequate and standardized curriculum documents now also exist in universities. This curriculum document is used by each study program as a guide in learning. The existence of a Curriculum Development Team that has carried out an internal evaluation of the quality, quantity, and quality of the syllabus prepared is also very useful. The curriculum development team has an external monitoring and evaluation function for this learning activity. After the learning process is complete, the curriculum developer unit will evaluate the learning process that has already taken place. This result is good enough to be done so that the desired competency standard output will be easily met.

Content Standard

The components in a syllabus are following the demands of the higher education national standard and the KPT that lead to the revolution 4.0. Revolution 4.0 emphasizes more on students' creativity and reasoning power. Almost all of the study programs owned by STIKes Buleleng include reasoning and creativity as a form of application of mental revolution 4.0. Students are required to have good critical thinking and critical reasoning by including entrepreneurship courses as an initial capital. The output of this course is to produce a model, prototype, or another thing that focuses on decreasing the degree of pain and death.

Learning Process Standard

The method used is following the content presented. Judging from the lecturers' journals, they were following the implementation time, the methods used, and the material presented. This result is seen when the quality assurance institute and study programs, as quality control groups, monitor the learning process. The absent journal that was made at STIKes Buleleng gave rise to the sections of learning material, student attendance, learning methods, and lecturer attendance. To evaluate the suitability of the syllabus with the correct occurrence, it is listed in the journal of lecturer and student activities. Most of the lecturers at STIKes Buleleng have implemented learning methods that are following those listed in the syllabus.

Assessment Standard

The student assessment process is the result of a learning process. The resulting output is a form of assessment which can

validly and reliably collect student grades according to the actual assessment. The assessment technique uses a rubric. The rubric used is following the material or content in the syllabus. The type of assessment rubric should be following the evaluation technique used, so that objectivity in the assessment process can be achieved. The rubric used is also standardized according to national standard of higher eeducation so that the assessment, assessment process, and the results obtained are the results of an objective assessment.

Conclusion

Based on the research findings, it is concluded that the evaluation and implementation of the higher education curriculum based on the national standard of higher education in the Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015 needs to be improved from the standard components of learning content and assessment. Both of these components must be immediately improved in terms of concepts and understanding so that the goals, achievements, and quality of graduates can still be improved and maintained. Internal discussions and workshops should be done regularly each semester to refresh pedagogical ability and understanding of the regulatory faculty curriculum.

References

Bashri, A., Puspitawati, R. P., & Ibrahim, M. (2018). Efektivitas pelaksanaan kurikulum di program studi Pendidikan Biologi Universitas Negeri Surabaya. *Jurnal Penelitian Pendidikan Matematika Dan Sains*, *2*(1), 1–5. Retrieved from https://journal.unesa.ac.id/index.php/jppms/article/view/6213

Cassano, R., Costa, V., & Fornasari, T. (2019). An effective national evaluation system of schools for sustainable development: A comparative European analysis. *Sustainability*, *11*(1), 195. https://doi.org/10.3390/su11010195

Hartini, S., & Suryanti, H. H. S. (2019). Analisis sinkronitas rencana pembelajaran semester (RPS) dengan rencana tugas semester mahasiswa (RTM) dan rencana evaluasi pembelajaran (REP) dosen FKIP Unisri tahun 2018. *Research Fair Unisri*, *3*(1). Retrieved from http://ejurnal.unisri.ac.id/index.php/rsfu/article/view/2620

Hussey, T., & Smith, P. (2010). *The trouble with higher education: A critical examination of our universities*. https://doi.org/10.4324/9780203866344

Kaklauskas, A., Banaitis, A., Ferreira, F., Ferreira, J., Amaratunga, D., Lepkova, N., … Banaitienė, N. (2018). An evaluation system for university–industry partnership sustainability: Enhancing options for entrepreneurial universities. *Sustainability*, *10*(1), 119. https://doi.org/10.3390/su10010119

Kazimirov, A. N. (2018). Education at university and industry 4.0. *Proceedings of 2018 Global Smart Industry Conference (GloSIC 2018)*, 1–6. https://doi.org/10.1109/GloSIC.2018.8570076

Law No. 12 of 2012 of Republic of Indonesia on Higher Education. , (2012).

Luttenberger, S., Macher, D., Maidl, V., Rominger, C., Aydin, N., & Paechter, M. (2018). Different patterns of university students' integration of lecture podcasts, learning materials, and lecture attendance in a psychology course. *Education and Information Technologies*, *23*(1), 165–178. https://doi.org/10.1007/s10639-017-9592-3

National Accreditation Agency for Higher Education. (2008). *Buku II: Standar dan prosedur akreditasi program studi sarjana*. Jakarta: National Accreditation Agency for Higher Education, Ministry of National Education.

Presidential Regulation No. 8 of 2012 on the Indonesian National Qualification Framework. , (2012).

Pribadi, H. (2019). Implementasi kurikulum perguruan tinggi (KKNI dan SN-Dikti) pada program studi sarjana Bimbingan dan Konseling Universitas Borneo Tarakan. *Proceeding of Konvensi Nasional XXI Asosiasi Bimbingan Dan Konseling Indonesia*, 39–46. Retrieved from http://proceedings.upi.edu/index.php/konvensiabkinxxi/article/download/448/402/

Rahman, A. M., Mutiani, M., & Putra, M. A. H. (2019). Pengaruh kompetensi pedagogik dosen terhadap motivasi belajar mahasiswa Pendidikan IPS. *Jurnal Darussalam: Jurnal Pendidikan, Komunikasi Dan Pemikiran Hukum Islam*, *10*(2), 375–387. https://doi.org/10.30739/darussalam.v10i2.380

Rahmawati, D., & Anggraini, A. D. (2017). Evaluasi program kurikulum berdasarkan standar isi, standar proses, dan standar kompetensi lulusan di SDN Pisangan Timur 10 Pagi. *Jurnal Pendidikan Ekonomi Dan Bisnis (JPEB)*, *5*(1), 35–50. https://doi.org/10.21009/JPEB.005.1.3

Regulation of the Minister of Education and Culture No. 49 of 2014 on the National Standard of Higher Education. , (2014).

Regulation of the Minister of Research, Technology, and Higher Education No. 44 of 2015 on the National Standard of Higher Education. , (2015).

Rodríguez-Conde, M.-J., Olmos-Migueláñez, S., Gamazo, A., & O'Hara, J. (2018). Evaluation in education and guidance: A perspective from 2018. *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality - TEEM'18*, 129–133. https://doi.org/10.1145/3284179.3284203

Setyawati, R. (2015). Peran guru bahasa dan sastra Indonesia dalam pembelajaran yang inovatif. *Buana Bastra*, *2*(2), 165–174. Retrieved from http://jurnal.unipasby.ac.id/index.php/bastra/article/view/231

Sitepu, B. P., & Lestari, I. (2018). Pelaksanaan rencana pembelajaran semester dalam proses pembelajaran di perguruan tinggi. *Perspektif Ilmu Pendidikan*, *32*(1), 41–49. https://doi.org/10.21009/PIP.321.6

Tiwari, S. K. (2018). Evaluation of modern education system in India: Problem and opportunities. *International Journal of Advanced Research and Development*, *3*(2), 468–470.

Wijanto, S. (2009). *Pengelolaan perguruan tinggi secara efisien, efektif dan ekonomis (Untuk meningkatkan mutu penyelenggaraan Pendidikan dan mutu lulusan)*. Jakarta: Salemba Empat.