BIOLOGY LEARNING EVALUATION MODEL IN SENIOR HIGH SCHOOLS

Sri Utari [1]*, Djukri [2]
[1]8 State Senior High School, [2]Universitas Negeri Yogyakarta
[1]Jl. Sidobali No.1, Muja Muju, Umbulharjo, Yogyakarta, DIY 55165, Indonesia
[2]Jl. Colombo No. 1, Depok, Sleman 55281, Yogyakarta, Indonesia
* Corresponding Author. Email: utedelayota@gmail.com

Abstract

The study was to develop a Biology learning evaluation model in senior high schools that referred to the research and development model by Borg & Gall and the logic model. The evaluation model included the components of input, activities, output and outcomes. The developing procedures involved a preliminary study in the form of observation and theoretical review regarding the Biology learning evaluation in senior high schools. The product development was carried out by designing an evaluation model, designing an instrument, performing instrument experiment and performing implementation. The instrument experiment involved teachers and Students from Grade XII in senior high schools located in the City of Yogyakarta. For the data gathering technique and instrument, the researchers implemented observation sheet, questionnaire and test. The questionnaire was applied in order to attain information regarding teacher performance, learning performance, classroom atmosphere and scientific attitude; on the other hand, test was applied in order to attain information regarding Biology concept mastery. Then, for the analysis of instrument construct, the researchers performed confirmatory factor analysis by means of Lisrel 0.80 software and the results of this analysis showed that the evaluation instrument valid and reliable. The construct validity was between 0.43-0.79 while the reliability of measurement model was between 0.88-0.94. Last but not the least, the model feasibility test showed that the theoretical model had been supported by the empirical data.

Keywords: *evaluation model, biology teaching, senior high schools*

Introduction

This article reviews the Biology learning evaluation model in senior high schools. The evaluation model that will be developed consists of the following components: input, activities, output and outcomes; this model itself refers to the logic model evaluation. Then, the components of input cover students ' facility and initial states; the components of activities cover teachers' performance, students' performance and classroom's atmosphere; the components of output refer to the Biology concept mastery; and the components of outcome refer to the scientific attitude. The evaluation is focused toward the activities or the learning process.

The condition of and the needs toward the Biology learning evaluation model in senior high school lever has become the background of the problems that underlie this review. Evaluation is an important stage in measuring the success of a program. Evaluation toward learning program is necessary in order to measure the learning effectiveness within the efforts of improving the learning quality. Evaluation might encourage the students to be more motivated in their learning programs continuously and might encourage the teachers to improve the quality of their learning process. For schools, evaluation might encourage the improvement toward facilities and school management quality. Up to date, learning evaluation has been prioritizing more on the learning results rather than the learning process. The evaluation of learning results has been based on the results of daily test, mid-semester test, final-semester test, final examination and even national examination without considering the process; whereas, learning process is a black box that might uncover the teachers' and then students ' activities in achieving the learning objectives. The improvement toward the learning process becomes an important step in improving the educational quality. Therefore, there should be a learning program evaluation that might describe the learning process effectiveness.

Biology learning has object characteristics and their problems and, as a result, this learning demands a specific evaluation model. Reinburg (2009, p. 29) states that an individual who have completed his or her Biology learning should understand the main concept of Biology science, the impact of human activities toward biosphere, the inquiry process and the history of biological development. The statement implies that students should understand the Biology concepts, their implementation for solving real-world problems and the scientific investigation process. Finally, a Biology science-literate individual should be able to think creatively, to formulate problems regarding the nature, to have logical and critical reasons, to use efficient technology and to take ethical personal decisions in relation to the biological issues. Therefore, Biology learning should provide facilities in order to achieve that objective.

There are two bases of scientific learning namely process and product. Amien (1987, pp. 16-17) states that the effectiveness of each instructional approach depends on the desired product and process. The final target of learning experience is the implementation of the experience in the future. The learning product might be implemented in the future through the transfer of scientific concepts and attitudes.

Biology learning evaluation in senior high schools should be in accordance to the Biology learning characteristics namely involving the dimension of process and the dimension of results. The dimension of process relates to the activities of designing, implementing and reporting the investigation along with the results (Ludwig & Reynold, 1998, p. 1). Scientific process might be defined as an approach in learning science that has been based on the observation toward what an individual has committed (Rezba, 2007, pp. 4-5; Holt, Rinehart & Winston, 1989, pp. 18-22). The scientific process covers the activities of observing, measuring, classifying, predicting, hypothesizing, investigating, drawing conclusion and communicating results. Then, the dimension of results includes the mastery of scientific concept and attitude.

Theoretical reviews and studies should be conducted in order to provide an alternative of Biology learning evaluation model in senior high schools. The relevant theoretical reviews and studies that have supported this study will be elaborated further. Then, the theoretical review itself includes the Biology learning evaluation characteristics evaluation model and the Biology learning evaluation model.

Program Evaluation Models

There are several models that have been frequently applied in evaluating a program namely CIPP, Kirkpatrick, Stake and logic model. Model selection is mainly based on the evaluation objective. Then, the concept of CIPP (Context, Input, Process and Product) evaluation model was introduced by Stufflebeam in 1965 with his opinion that the important objective of an evaluation has been to improve instead to prove (Madaus, Scriven & Stufflebeam, 1993, p. 118). There are four dimension in this evaluation model, namely context, input, process and product that might be the targets of an evaluation. The CIPP evaluation model might be implemented in any sectors such as education, management and company.

The Kirkpatrick evaluation model is mainly implemented in evaluating the training program that aims to develop human resources. The evaluation model that had been developed by Kirkpatrick is widely known in *Evaluating Training Programs: The Four Levels*. This model covers four level of an evaluation namely reaction, learning, behavior and result (Kirkpatrick, 1988, p. 20).

Logic model is model tool that describes the theory of change that underlies an intervention, a product or a policy (Frechtling, 2007, pp. 21-22). Logic model might be turned into a tool or an approach in order to describe the important elements of a program and to identify the focus of an evaluation. Logic model might be implemented in order to optimize a program through logical and directed planning and evaluation. There are four basic compo-

nents in the logic model namely: (1) input, namely the human resources that a program has including the fund or the labor contribution; (2) activities, namely the actions that might be willingly conducted in order to achieve the desired objectives; (3) output, namely the direct results of an action that are stated in the form of number such as the number of service, event, document or participation; and (4) outcome, namely the aspect that displays the occurring change or achievement that leads to the final objective.

The review of this study applies the logic model evaluation because the focus of the evaluation is the learning process. Logic model is an evaluation model that has been based on performance and that described the causality logical path so that an individual might understand the components that influence the evaluation results. Logic model consists of four basic components that might be developed and be modified according to necessities.

Characteristics of Biology Learning

Biology has special characteristics that are different to other science in terms of object, problem and method. Biology has a clear scientific structure as having been developed by the Biological Science Curriculum Study (BSCS). According to the BSCS, Biology learning should not only be limited to the textual manner but also be followed by an observation that views any phenomena that might be objects or events. Carind & Sund (1989, pp. 4-5) state that scientific learning aims to train the children to master scientific methods in order to generate scientific products through investigating activities. Biology learning emphasizes scientific process or scientific method in order to generate scientific product.

Scientific process includes the activities of observation, hypothesis formulation, prediction, investigation, data interprettion, inferrence and result communication. Harlen (1992, pp. 83-93) elaborates the indicators of scientific process that might a basis for designing a research instrument.

The types of action that displays the activities of observation are as follows: (1) paying attaention to surrounding objects in detail; (2) identifying similarities and differences; (3) defining sequence of occuring events; and (4) operating assisting tools in order to learn objects in details.

Then, several types of action that might be categorized into the activities of hypothesis formulation are as follows: (1) displaying consistent explanation along with evidence; (2) displayiong consistent explanation along with scientific principles or concepts; (3) implementing knowledge that has been attained previously in providing consistent explanation; (4) realizing that there are more than one possible explanation that might be given to one phenomenon; and (5) realizing that consistent explanation is temporary.

Next, several types of action that might be categorized into the activities of prediction are as follows: (1) using evidence from previous observation or present observation in order to state what might happen; (2) using evidence from previous observation or present observation for extrapolation and interpolation; (3) stating what has happened in relation to the past evidence or experience; (4) recalling the implementation of a pattern that does not have evidence in stating assumption; and (5) differentiating prediction and guess.

Furthermore, the indicators of investigation skills include: (1) deciding independent and dependent variables; (2) manipulating variables in order to carry out appropriate investigation; (3) identifying variables that will be measured or that will be compared; (4) measuring and comparing dependent variables by using the right instrument; and (5) working under appropriate stages.

The indicators of data interpretation and inferrence skills are apparent in the following actions: (1) using information in order to state several meaningful statements; (2) finding certain patterns in observation or investigation results; (3) identifying the relationship between one variable and others; and (4) ensuring that a pattern has been in accordance to data.

Last but not the least, the behaviors of communication results are apparent in the following actions: (1) talking, listening or writing opinions in order to filter ideas or clarifying meanings; (2) taking notes on observation within the investigation process; (3) using graphics and tables in order to deliver information; (4) selecting appropriate communication media so that the informationn might be understood by other people; and (5) using secondary information source.

Biology learning activities that habituate students to implement scientific process will create scientific attitude. Scientific attitude is the attitude toward science that should be developed within the scientific learning (Harlen, 1992, p. 39). This attitude includes curiosity, critical reflection, creativity and discovery, environmental care and cooperation.

Definitions and indicators of scientific attitude has been elaborated by Harlen (1992, pp. 39-44). Curiosity is the attitude of being interested in all matters within the surrounding environment (Harlen, 1992, p. 41). The indicators of curiosity are as follows: (1) being interested in new matters; (2) displaying interest in performing detailed observations; (3) asking many questions in order to find answers; and (4) using multiple information sources2 in order to find new matters.

Respect for evidence is the attitude of not easily trusting opinions and/or information (Harlen, 1992, p.42). This attitude is apparent from the following actions: (1) providing reports according to facts although the facts might be different than expectations; (2) investigating evidence that is different than the pattern that has been found previously; and (3) using conclusions as part of further study.

Next, critical reflection is the analysis or the efforts of reviewing knowledge, understanding and belief that has been understood previously (Harlen, 1992, p. 43). The indicators of critical reflection are as follows:

(1) having desires to review what actions that have been taken; (2) considering alternative procedures that have been implemented; (3) identifying results that deny or support the results of previous studies; and (4) carrying out critical reflection on the results of previous studies in planning and implementing a study.

Learning objectives influence the way teachers teach Biology. The skills of scientific process and attitudes will not be attained if teachers teach Biology genetically. Multiple learning methods might be implemented by teachers in order to achieve expected competencies within the Biology learning process. The scientific learning experts emphasize the importance of scientific and inquiry method. Harlen (2007) recommends the inquiry method within the science learning. Learning by means of inquiry method does not only engage students into learning science but also into learning itself. Both of these aspects are the important results for future of science-literate society. The interaction between students and learning objects and between students and teachers are necessary in the inquiry-based learning process. In such learning,the skills of performing observation, raising questions, clarifying information sources, designing plan, performing experiment, gathering data, analyzing data, interpreting data, performing prediction and communicating results are necessary.

Biology Learning Evaluation in Senior High Schools

Learning program evaluation is an integral part of learning plan (William, 2012, p.2). As the integral part of a program, evaluation will answer the following questions: What will be evaluated? Who will use the evaluation results? What are the criteria that will be applied within the evaluation? Evaluation will help clarifying values, identifying needs and considering alternative manners in order to meet the needs of learning design conceptualization, to conduct the learning process and to improve self-evaluation. In learning evaluatio, there are two

most components namely teacher and learning participant; then, the third component is learning supports (librarian, laboratory operator and alike).

Biology learning process evaluation is a process of providing learning activities information in order to determine the learning effectiveness. Teaching, according to the context of standard educational process, is not a mere transfer of knowledge from educators to students . In general there are two concepts of learning process,namely learning as a process of delivering learning materials and learning as a process of regulating environment (Department of National Education, 2008, p.4). The first concept emphasizes learning materials mastery, while the second concept emphasizes process in which students might change their behaviors. This statement implies that teaching-learning process should be turned into the center of the activities. Therefore, teachers should be skillful in mastering learning materials, learning management, evaluation and conducive learning environment.

According to Doran (2009, pp.9-15), based on the focus, evaluation might be categorized into learning evaluation that focuses to students , to teachers, to classrooms and to curricula. The focus and the criteria that will be used in the evaluation influence the type of data that should be gathered. Recommendations from multiple society elements such as educators, parents, students and other interested parties are very necessary. Each group may have different emphasis. The method for attaining these recommendations might be various and might come from multiple questionnaires, checklists and written forms.

The Biology learning evaluation that will be developed in the study will emphasize the teacher performance, the learning participant performance and the classroom atmosphere in accordance to the Biology learning characteristics. The evaluation is focused on uncovering information on the inquiry-based Biology learning activities. The information source is attained from the

teachers and the students through the questionnaire and the interview.

Harlen (1992, p.130) explains the guidelines of teacher and learning participant performance according to the objective of scientific learning. The objective of scientific learning is to provide an opportunity for the students to: (1) carry out investigation; (2) develop scientific skills; (3) develop scientific concepts according to the curriculum; (4) develop scientific attitude; (5) develop interests in performing activities; (6) develop learning activities in the daily life; and (7) associate scientific learning to other lessons.

Students' activities in a scientific classroom includes: (1) completing assignment according to the learning objective; (2) enjoying the learning activities; (3) displaying curiosity, concept understanding and discipline; (4) associating the learning activities to the daily life; (5) discussing assignments with peers or teachers; (6) displaying shared cooperation and decision-making activities; (7) displaying critical attitude and open mind; (8) conducting study in order to find answers from the problems provided by teacher; (9) conducting study in order to develop interests; (10) predicting the results that will be attained; (11) observing and recording observation results systematically; (12) classifying and looking for the pattern of observation results; (13) drawing conclusions; (14) selecting accurate and appropriate measurement tools; (15) planning and designing experiment in order to display the concept understanding; (16) completing multiple aspects in a study independently in order to conduct the study thoroughly; (17) checking the results that have not been in accordance to the expectation and repeating the measurement; and (18) applying the results of the study in order to conduct another study so that more convincing results might be attained.

Flick & Lederman (2006, pp. 16-167) state a teacher's enormous role in supporting and developing the students ' thinking capacity. From the psychological perspective, teenage students are often inconsistent in performing their actions. Therefore, there should be an appropriate learning method that might support their metacognitive capacity. Metacognitive capacity is not instantly ready for use in the classroom or in the daily life; instead, this capacity should be demanded, be conducted and be trained. Teachers have a central role in developing this metacognitive capacity through the learning environment that supports the understanding and the implementation of scientific study. Teachers also have a role in creating opportunities for the students to develop the skills that they will use in conducting a study. From these statements, teachers have an important role in designing assignments, selecting learning methods and creating learning environments that support the development of students ' capacity.

Laboratory has another important role in the Biology learning. According to the Committee of National Research Council (Bybee et al., 2006), the learning objectives that should be met as the results of laboratory learning are as follows: (1) improving the mastery of learning materials; (2) developing scientific reasoning; (3) understanding the complexity of empirical works; (4) developing practical skills; (5) understanding the nature of science; (6) growing interest in science and in learning science; and (7) developing group cooperation capacity.

Laboratory learning is frequently conducted by Biology teachers. Several principles in conducting the laboratory learning are namely (Department of National Education , 2008, pp.34-35): (1) learning to do or to practice; (2) curiosity, being encouraged by students ' curiosity; (3) performing scientific thinking, developing the scientific thinking capacity. According to these principles, laboratory therefore is being used to perform experiment and demonstration.

Laboratory and laboratory learning become a general strategy among Biology teachers in order to develop the scientific capacity. Many laboratory activities that should be facilitated and be managed by the teachers in order that the students ' skills

may appear; the appropriate use of knowledge might trigger curiosity so that the students will be motivated to perform investigation. Flick & Lederman (2006, pp. 161-162) reviewed a study about laboratory and found several drawbacks. Most of the researchers paid less attention toward students' background, teachers' behaviors, learning environment within classrooms or interaction between teachers and students . As a result, these researchers did not attain a complete description with regards to what actually happened in the class or in the laboratory.

Studies regarding scientific learning process display a description that has been in accordance to the psychological teenage study. Students will be involved in a scientific inquiry if: (1) they are supported by appropriate learning process; (2) they are demanded to work with reflective- and critical-thinking people; and (3) their teachers have necessary knowledge for performing scientific inquiry. If the teachers have low knowledge, do not emphasize critical-thinking activities and propose lower cognitive demands then the students will be less focused toward their inquiry assignments. Learning science through inquiry manner places teachers as secondary information sources within the teacher-learning participant relationship. The teacher-learning participant interaction becomes important in assessing the learning effectiveness.

The success of learning process is influenced by many factors namely teachers, students, supporting facilities and classroom environment. Classroom environment refers to the classroom condition in relation to the learning process that has been marked by a pattern of inter-classroom member interaction or communication. Nasution (2003, pp. 119-120) states that there are three types of classroom situations based on the teachers' attitude. The first type is authoritarian situation where a teacher exerts his or her authority in order to achieve the learning objective and this includes the use of threat and penalty. The second type is permissive situation where a teacher lets the students

to develop under emotional freedom. The third type is real situation where the students have their freedom but they are still under control.

Levin & Nolan (1996, p. 147) mention that there are two most important variables that influence the classroom environment namely the physical environment and the classroom regulation, which has been influenced by the culture of the teacher and the students. Teacher has an important role in managing the learning activities so that these activities will be effective. Classroom and school are part of culture. If the school's culture and the students ' culture are appropriate from one to another, then both cultures might improve the positive behaviors and the positive relationship between the teacher and the students. However, if the teachers and the students have different values, norms, behaviors and expectations, then there will be misunderstanding, conflicts and disbelief improvement.

Based on the theoretical review and the results of discussion among the Biology teachers in the senior high schools located in the City of Yogyakarta, the researchers will design a Biology learning evaluation model for senior high schools. The evaluation model will emphasize the learning process that might be apparent from the teacher performance, the learning participant performance and the classroom atmosphere. The evaluation model that will be developed is displayed in the Figure 1.

Method

The study was a research and development research that had been adopted from the Borg & Gall (1983) model. The procedures of research and development activities in the study covered four stages namely: (1) information study and gathering; (2) product development; (3) product testing and model revision; and (4) implementation.
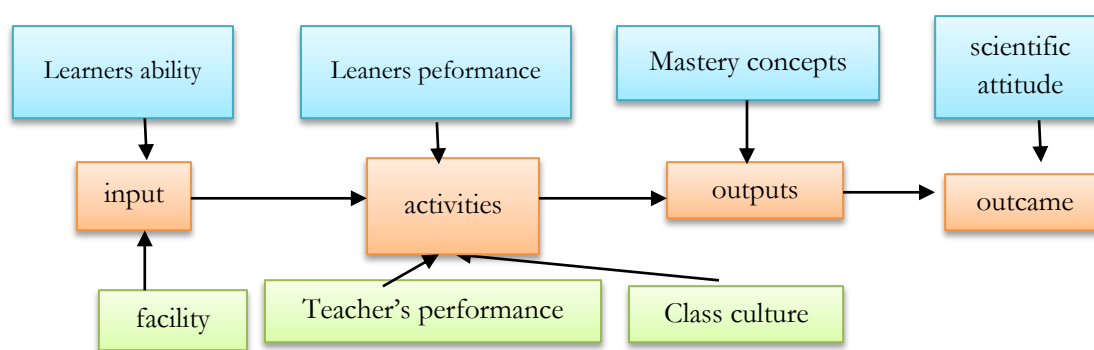
Figure 1.   Biology Learning Evaluation Model for Senior High Schools

The subjects in the study were the Biology teachers and the XII Grade students from the City of Yogyakarta. In conducting the instrument readability test, the researchers involved 12 teachers and 32 students. Then, in conducting the instrument experiment, the researchers involved 189 subjects who consisted of 183 students and 6 Biology teachers. In the implementation stage, the test respondents consisted of 16 Biology teachers and 250 students from 16 senior high schools located in the City of Yogyakarta. The selection of the schools as the sample in this study was based on the score of Biology National Examination from the previous year and the composition of these schools was as follows: 4 A-category schools, 4 B-category schools, 4 C-category schools and 4 D-category schools.

The instruments that would be used in the study were questionnaire, observation sheet and interview guideline. The questionnaire would the main instrument in the model development, whereas the observation sheet and the interview guideline would be the supporting instrument in order that the researchers might gather in-depth information regarding the Biology learning program evaluation.

The data would be analyzed by means of Exploratory Factor Analysis and Confirmatory Factor Analysis. The data that had been gathered in the instrument experiment would be analyzed by means of Exploratory Factor Analysis with the assistance of SPSS 17.0 software. The EFA analysis would be conducted in order to identify the presence of variable relationship and to reduce the data so that the researchers might attain new variables or factors. The data that had been attained in the implementation stage would be analyzed quantitatively and would be tested by means of Confirmatory Factor Analysis in order to attain information regarding instrument validity and instrument reliability as well as model fitness. The CFA analysis would be conducted with the assistance of LISREL 8.71 program (Wijayanto, 2008, p. 146).

Results and Discussions

The results of this studying an evaluation model complete with the instrument that might be applied in order to identify the learning process effectiveness. The elaboration of these results covered the evaluation instrument test and the evaluation process. The tests of instrument construct validity and reliability and the measurement model fitness had been an important of this study.

The results of the preliminary study showing that the evaluation that the teachers had conducted had only been in the learning results assessment level. The other type of evaluation was the teacher performance assessment that had been conducted by the principals or the learning supervision that had been conducted by the school supervisors. The program evaluation that

associated the learning process and the learning results through an integrated manner had not been implemented in senior high schools all over the City of Yogyakarta. The results of a review toward previous studies showed that there had not been any Biology learning evaluation models that considered the learning process and the learning results in the same place. On the other hand, the existing studies only partly reviewed the learning process and the learning results.

Evaluation Instrument

The instrument designing activities were initiated by theoretical review, review of previous studies and laws that regulated the learning process in senior high schools. These reviews resulted in an initial draft of the instrument in the form of teacher questionnaire and learning participant questionnaire. The teacher questionnaire entailed questions regarding the teacher performance, while the learning participant instrument entailed questions regarding the teacher performance, the learning participant performance, the classroom atmosphere and the scientific attitude.

After conducting the review, the researchers performed a validity test against the evaluation instrument. The evaluation instrument was discussed by the experts of Biology education, measurement and linguistics in order to attain the expert judgment. The experts who had been involved in the study consisted of two Biology learning experts, two measurement experts and one linguistics expert. This stage resulted the following changes: (1) the instrument should be clarified so that the instrument would be appropriate to the Biology learning in senior high schools; (2) the directions of instrument completion should be clarified; (3) the items on the learning participant performance should be added so that this instrument would be more specific; and (4) the language aspects should be improved in order that the instrument would be easier to understand.

The instrument readability test was conducted by the teachers and the students in order to gather feedback qualitatively. The instrument readability test by the teachers involved 11 respondents and this test gathered feedback of improvement on the aspects of language and teacher performance. On the other hand, the instrument readability test by students involved the students of Sports Classroom in the State 4 State Senior High School Yogyakarta and this test gathered feedback of improvement on the direction of questionnaire completion and the aspects of language. This readability test then became the basis of improving the writing mechanics. The revisions turned into a matter of reference in conducting the next stage namely the instrument experiment.

The instrument experiment was conducted in order to attain the valid and reliable instrument. The instrument experiment was conducted in three senior high schools that represented the high, the mauderate and the low senior high school in the City of Yogyakarta and these high schools were State 3 State Senior High School Yogyakarta, State 5 State Senior High School Yogyakarta and State 11 State Senior High School Yogyakarta. The number of returned learning participant questionnaire was 183 bundles but the number of learning participant questionnaire that might be processed was 120 bundles.

The data that had been attained were analyzed in order to identify the size of the reliability coefficient by using the Cronbach's Alpha formula. The Exploratory Factor Analysis (EFA) by means of SPSS 17.0 was conducted in order to explain the dimensions that had been measured. The criteria of reliability coefficient were as follows: if the reliability coefficient was closed to 1.000, then the coefficient would be better; if the reliability coefficient was lower than 0.600, then the coefficient would be inferior; if the reliability coefficient was around 0.700, then the coefficient would be acceptable; and if the reliability coefficient was higher than 0.800, then the coefficient would be good.

The reliability coefficient was apparent in the Table 1.

Table 1.   Instrument Reliability

No	Aspect	Reliability Coefficient
1	Teacher Peformance	0.908
2	Students peformance	0.903
3	Class Culture	0.856
4	Scientific attitude	0.805

Based on the variable reliability, the researchers found that the reliability coefficient for all of the variables that had been tested had been higher than 0.800. The reliability score of teacher performance had been equal to 0.908, the reliability score of learning participant performance had been equal to 0.903, the reliability score of classroom atmosphere had been equal to 0.856 and the reliability score of scientific attitude had been equal to 0.805. From these scores, the researchers found that the questionnaire that would be administered in the experiment had high reliability level.

Furthermore, the EFA analysis was conducted in order to identify the presence of inter-variable relationship and to reduce the data so that the researchers would attain new variables and factors which had been simpler. From the EFA analysis, the researchers would like to expect that they would attain the factors that had influenced the learning process and the learning results. The results of EFA analysis might be viewed in the Table 2.

Table 2.   Results of EFA Analysis

No	Aspect	EFA Results	
		MSA	Communalities
1	Teacher Performance	0.81 – 0.87	0.57 – 0.86
2	Learning Participant Performance	0.77 – 0.90	0.53 – 0.78
3	Classroom Atmosphere	0.72 – 0.90	0.52 – 0.73
4	Scientific Attitude	0.54 – 0.80	0.51 – 0.80

The MSA scores showed that the variables might still be predicted and might be analyzed further because the score of

each variable had been higher than 0.50. The further analysis was conducted in order to decide whether these variables might be grouped into one or several factors. The communalities score showed that the fixed factors might be determined because the average had been higher than 50.00%. From the total score of variance explained, component matrix and rotate component matrix, the researchers determined the categorization of variable input into certain factors based on the size of the correlation between the variables and the factors. The results of the product testing then were used in revising the product.

The EFA from the product experiment resulted in several changes on the number of factors and the items of the questionnaire. The determination on the number of the factors was based on the score of total variance explained, component matrix and rotate component matrix. The categorization of variable input into certain factors was based on the size of correlation between the factors and the variables.

After performing the instrument validation, the readability test and the product experiment, the researchers performed several revisions. The five factors of teacher performance variable now should be reduced into three factors namely the learning management capacity, the students ' characteristics understanding and the learning evaluation conducting capacity. The learning participant performance still had two factors namely the classroom performance and the laboratory performance. The classroom atmosphere was divided into two factors namely the class support and the self motivation. Last but not the least, the seven factors of scientific attitude now should be reduced into three factors namely the curiosity, the discovery/creativity and the sensitiveness toward surrounding environment.

The instrument that had been revised was administered toward 16 senior high schools located in the City of Yogyakarta. The results of CFA toward this evaluation instrument showed that this instrument that had been administered in the Biology learn-

ing evaluation model for senior high schools had good validity and reliability. The validity and the reliability score of this instrument was apparent from t-value score and CR (construct reliability) score in the following Table 3.

Table 3. The Instrument's Validity and Reliability

No	Variable	t-value	CR
1	Learning Participant Performance	4.82 – 7.09	0.88
2	Teacher Performance	5.83 – 10.53	0.94
3	Classroom Atmosphere	7.56 – 10.24	0.89
4	Scientific Attitude	4.91 – 9.08	0.92

The fitness index between the model and the data might be viewed from several GOF (Goodness of Fit) scores namely the score of Chi-Square, RMSEA, GFI and AGFI. These scores measured the fitness of measurement model that had been designed in the evaluation instrument. The objective was to attain empirical evidence regarding the existing factors and indicators in the measurement model of evaluation instrument. The model fitness index of this instrument can be be viewed from the Table 4 until the Table 7.

In several GOF measures that had been apparent from the Table 4 until the Table 7, the researchers found that the measurement model of teacher performance, learning participant performance, classroom atmosphere and scientific attitude had been apparent to the criteria of good fitness. The latent variables of teacher performance were designed by the learning conduct capacity (KG 1, first variable of teacher performance), the learning participant understanding (KG 2, second variable of teacher performance) and the evaluation conduct capacity (KG 3, third variable of teacher performance). The latent variables of learning participant performance were designed by the classroom performance (KPD 1, first variable of learning participant performance) and the laboratory performance (KPD 2, second variable of learning participant performance). The latent variables of class-

room atmosphere were designed by the classroom support (IK 1, the first variable of classroom atmosphere) and the self motivation (IK 2, the second variable of classroom atmosphere). The latent variables of scientific attitude were designed by the curiosity (SI 1, the first variable of scientific attitude), the discovery/creativity (SI 2, the second variable of scientific attitude) and the sensitiveness toward surrounding environment (SI 3, the third variable of scientific attitude).

Table 4. The Fitness Index of CFA for the Teacher Performance

	Limit	Fitness Index
Chi-Square	≥ 0.05	$x^2 = 227.70$; df=195; p-value = 0.054
RMSEA	≤ 0.08	0.026
GFI	≥ 0.90	0.920
AGFI	≥ 0.90	0.900

Table 5. The Fitness Index of CFA for the Learning Participant Performance

	Limit	Fitness Index
Chi-Square	$\geq 0,05$	$x^2 = 119.81$; df=98.00; p-value = 0.067
RMSEA	$\leq 0,08$	0.03
GFI	$\geq 0,90$	0.94
AGFI	$\geq 0,90$	0.92

Table 6. The Fitness Index of CFA for the Classroom Atmosphere

	Limit	Fitness Index
Chi-Square	$\geq 0,05$	$x^2 = 32.94$; df=23.00; p-value = 0.08
RMSEA	$\leq 0,08$	0.042
GFI	$\geq 0,90$	0.970
AGFI	$\geq 0,90$	0.940

Table 7. The Fitness Index of CFA for the Scientific Attitude

	Limit	Fitness Index
Chi-Square	$\geq 0,05$	$x^2 = 250.45$; df= 222.00; p-value = 0.09
RMSEA	$\leq 0,08$	0,023
GFI	$\geq 0,90$	0,92
AGFI	$\geq 0,90$	0,90

The results of Biology learning program evaluation model development for senior high schools results in an instrument that had met the requirements of validity and reliability. This instrument then was implemented in order to identify the level of learning effectiveness. The effective criteria were differentiated into very good (mean score > 4.20), good (mean score ranging around 3.40 – 4.20), moderate (mean score ranging around 2.60 – 3.40) and vey low (mean score < 1.80).

Evaluation Results

The evaluation model consisted of the following components: input, activities, output and outcome. The evaluation profile was presented in order according to the evaluation components namely the input evaluation, the activities evaluation, the output evaluation and the outcome evaluation. The graphic was presented in the sample group A, B, C and D. The input evaluation included students ' initial capacity and supporting learning facilities. The activities evaluation included teacher performance, learning participant performance and class-room atmosphere. The output evaluation included the national examination score and the outcome evaluation included the scien-tific attitude score.
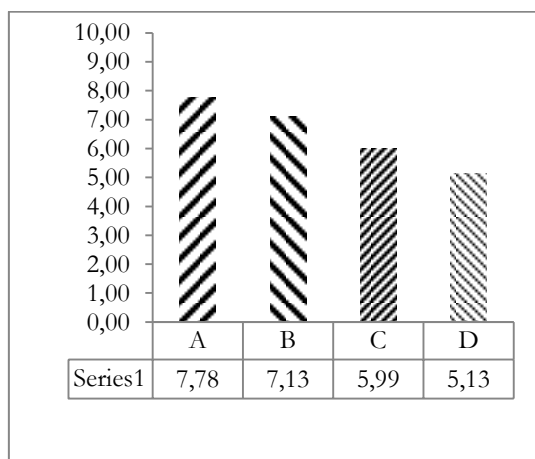
Input Evaluation

Figure 2. The Graphic of Biology National Examination Average Score in 2013/2014 Academic Year

The school classification based on the national examination scores in the City of Yogyakarta had been relatively stable over the years. The reason was that the senior high school Biology National Examination scores as the learning output indicator had become one of the students ' reference in selecting the school where they would like continue their study. The students with high national examination scores were inclined to select senior high schools that displayed high national examination score. In other words, the average national examination score of a senior high school might describe the students ' input. In the Figure 1, the researchers displayed the differences of Biology National Examination scores from the sampled schools, which had been ranging from 5.13 – 7.78. The average score gap between the A-classified school and D-classified school was 2.65, whereas the average score gap between the B-classified school and C-classified school was 1.14.
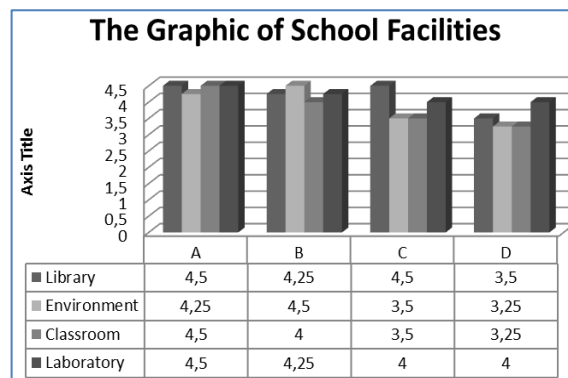

Figure 3. The Graphic of School Facilities

In terms of school facilities, all of the sampled schools had good library, school environment, classrooms and laboratory. On the other hand, in terms of completeness and quality, the A-classified school turned out to be the best.

Activities Evaluation

Learning activities had been the focus of this program evaluation. In the model logic evaluation, the activities were conducted in order to achieve the output. Output referred to the direct results that might be

measured from a program. Learning activities would be apparent from the teacher performance, the learning participant performance and the interaction between the teacher and the learning participant in order to create classroom atmosphere that might be helpful for the learning process.
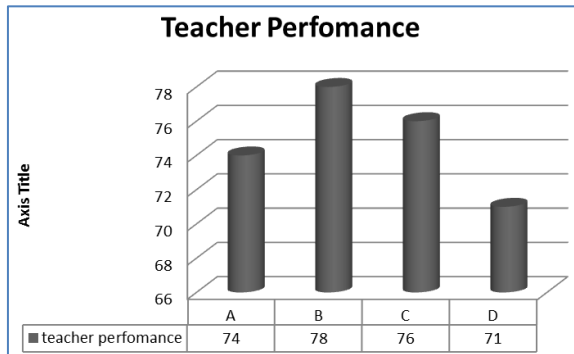
The learning activities were determined by the lesson plans that the teachers had prepared and were supported by the students ' motivation. The lesson plans prepared by the teachers determined the type of the activities, whereas the students ' motivation determined the quality.
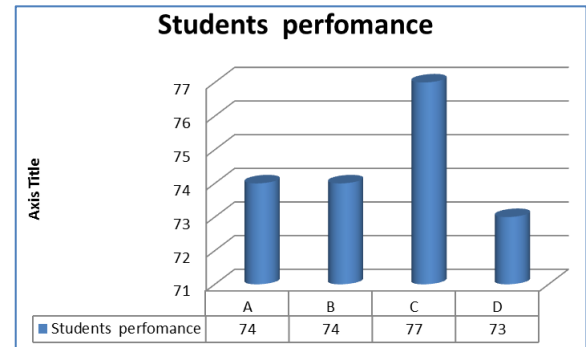

Figure 4.   The Graphic of Teacher Performance Achievement Level

Figure 5.   The Graphic of Students Performance Achievement Level

The teacher performance achievement level was ranging from 71.00% until 78.00%. The teachers assessed themselves higher than the students. The distribution of teacher quality was quite moderate; in terms average, the rank of teacher performance achievement level from the highest to the lowest was B-classified school, C-classified school, A-classified school and D-classified school. From these data, it was apparent that the highest teacher performance had been found in the B-classified school. Theoretically, teacher performance should influence learning results so that the researchers expected that the highest teacher performance might be found in the A-classified school. This difference showed that there had been other factors that influenced the learning results. The distribution of civil servant-status teachers in the City of Yogyakarta had been determined by the head of Education Office in order to meet the meeting the 24 teaching hours-obligation. Therefore, the schools could not choose the desired the teachers.

The learning participant performance achievement level was ranging from 73.00% until 77.00%; in other words, this achievement level belonged to the good category.
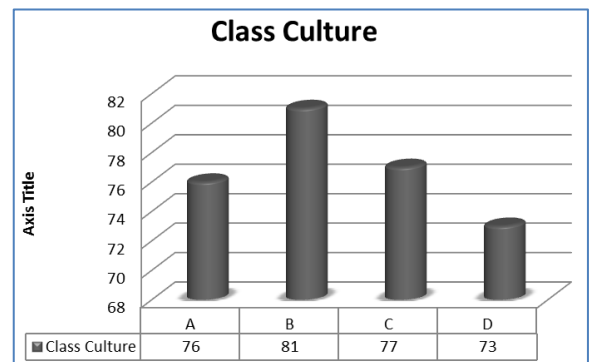

Figure 6.   The Graphic of Classroom Atmosphere Achievement Level

The classroom atmosphere effectiveness was ranging between 73.00% until 81.00%. The most effective classroom was found in the B-classified school, followed by the C-classified school, the A-classified school and the D-classified school. The classroom atmosphere was established by the components of self motivation and classroom support. The high self-motivation was supported by the good classroom culture and the good physical environment; in turn, the good classroom culture and the good physical environment established the classroom atmosphere that would be conducive for the learning process.
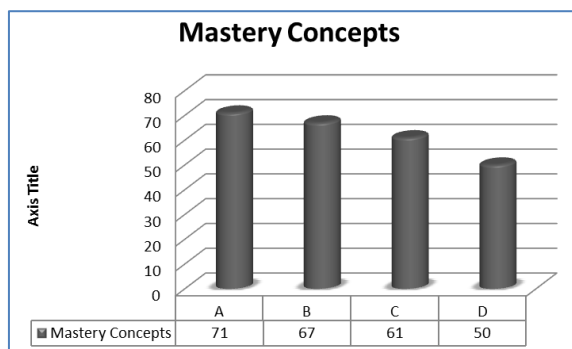
Output Evaluation

Figure 7. The Graphic of Students ' Concept Mastery Level

According to the essence of Biology learning, the Biology learning results were in the form of Biology concept mastery and scientific attitude. The Biology concept mastery as the direct result of the learning process might be viewed as the output in the logic model evaluation. The Biology concept mastery was quite various and was categorized into VG (very good), G (good), M(moderate) and P(poor). The categorization referred to the following criteria: 85.00 < SB ≤ 100.00, 70 < B ≤ 85.00, 55.00 < C ≤ 70.00 and 0 < K ≤ 55.00. The average national examination scores that described the students ' concept mastery might be viewed in the Table 7.

The students ' concept mastery was ranging from 50.00 until 71.00. These results were in line with the input that had been categorized based on the Biology National Examination scores in the previous year. The order of the concept mastery score did not change from the initial categorization namely the A-classified school that had the highest score and the D-classified school that had the lowest score. This matter implied that students ' input had a dominant role in determining the concept mastery output. The high input had a tendency to generate the high output, while the low input had a tendency to generate the low output.

Outcome Evaluation

Outcome referred to the indirect results of learning process that had been attained in the long term. Scientific attitude,

therefore, might be viewed as the outcome of Biology learning that had been established from a sequence of learning activities that benefitted scientific process. The score of students ' scientific attitude was ranging from 73.00% until 81.00%. There was also a tendency that the high concept mastery would be followed by the high scientific attitude.
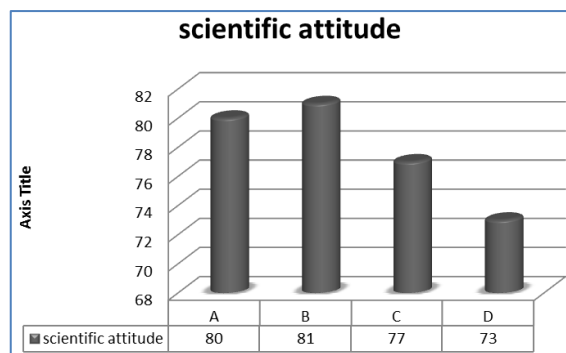

Figure 8. The Graphic of Students ' Scientific Attitude

The scientific attitude scores that had been almost similar in all schools were related to the learning activities. Flick & Lederman (2006, pp. 161-167) stated the enormous role of a teacher in supporting and in developing the students ' thinking capacity. Teachers have a central role in developing cognitive capacity through a learning environment that supported the students ' understanding and scientific study conduct. Several studies showed that teachers often decreased the cognitive demand and directed the students to the intended answers and, as a result, the students had decreasing motivation in accomplishing their investigation assignments.

Conclusions

The Biology learning evaluation model development for senior high schools has results in good evaluation model and instrument for evaluating the teacher performance, the students performance, the classroom atmosphere and the scientific attitude. The results of the significance test have implied that the classroom atmosphere, the teacher performance and the learning parti-

cipant performance are parts of the learning process. The profile of evaluation results for the Biology learning in the City of Yogyakarta shows that the learning process has been 100% good. On the other hand, the profile of these evaluation results also shows that 31.25% of the learning process has been good, 43.75% of the learning process has been moderate and 25.00% of the learning process has been low.

References

Amien, Moh. (1987). *Mengajarkan ilmu pengetahuan alam (IPA), dengan menggunakan metode "discovery" dan "Inquiry*. Jakarta: Departemen Pendidikan dan Kebudayaan Direktorat Jendral Pendidikan Tinggi.

Borg, W. R. & Gall, M. D. (1983). *Educational research*. New York: Pearson Education.

Carind, A. A., & Sund.R. B. (1989). *Teaching science through discovery*. London: Merril Publishing Company.

Department of National Education. Peraturan Menteri Pendidikan Nasional Republik Indonesia Nomor 3 Tahun 2008. (2008). Jakarta

Doran, R. L. (2009). *Basic measurement and evaluation of science instruction*. Retrieved February 18, 2013, from http://physicsed.buffalostate.edu/pubs/pdf.

Flick, L. B., & Lederman, N. G. (2006). *Scientific inquiry and nature of science.* Netherlands: Springer.

Frechtling, J. A. (2007). *Logic modeling methods in program evaluation*. USA: John Wiley & Sons.

Harlen, W. (1992). *The teaching of science*, London: David Fulton Publisher.

Herlen,W. (2007). *Assesment of learning* . Singapore: Sage.

Holt, et.al. (1989). *Modern Biology*, United State of America: Holt, Rinehart, and Winston Inc.

Kirkpatrick, D. L. (1998). *Evaluating training programs, The four levels*. (2nd ed.). San Fransisco: Barrett-Koehler Publisher, Inc.

Levin, J., & Nolan, J. F. (1996). *Principles of classroom management*. Boston: Allyn and Bacon.

Madaus, G. F., Scriven, M., & Stufflebeam, D. L. (1993). *Evaluation models, viewpoints on educational and human service evaluation*. Boston: Kluwer-Nijhoff Publishing.

Nasution. (2003). *Berbagai pendekatan dalam proses belajar & mengajar*. Jakarta: PT. Bumi Aksara.

Reinburg, C. (2009). *Theacher's handbook*. (4th ed.). Virginia: NSTA Press.

Rezba. (2007). *Learning and assessing Scieence Process Skills*. United States of America: Kendall/Hunt Publishing Company.

Wijayanto, S. H. (2008). *Sructural equation modeling dengan lisrel 8.8*. Jakarta: Graha Ilmu.