JOURNAL OF BALTIC SCIENCE EDUCATION

ISSN 1648-3898 /Print/
ISSN 2538-7138 /Online/

EFFECT OF PROBLEM-BASED LEARNING ON IMPROVEMENT PHYSICS ACHIEVEMENT AND CRITICAL THINKING OF SENIOR HIGH SCHOOL STUDENT

Mundilarto,
Helmiyanto Ismoyo

Abstract. *The aim of the research was to know the effect of implementation of problem-based learning model on students' physics achievement and critical thinking. The research was a quasi-experimental using a pre-test-post-test control group design. It was conducted at a state senior high school in Indonesia. The data were collected by a students' physics achievement test, a students' critical thinking test, and an observation sheet. The data were analysed by calculating normalized gain scores and hypothesis testing by using MANOVA with the significance level of .05 as criterion. The research results indicated that the average gain scores for students' physics achievement of the experimental and control groups were respectively .63 and .32, the average gain scores for students' critical thinking of the experimental and control groups were respectively .49 and .34, and a value of p is .0001 was obtained from the MANOVA calculation so that it is concluded that such learning as that in the experimental group is better in comparison with such learning as that in the control group. The results of this research are expected to increase insight and knowledge of teachers in implementing the learning process in the classroom so as to apply varied learning models.*

Keywords: *critical thinking, physics achievement, problem-based learning.*

Mundilarto, Helmiyanto Ismoyo
Yogyakarta State University, Indonesia

Introduction

Students' competence in comprehending the concepts in physics at school greatly determines their learning achievement. Their success in learning could be known from the results of their learning achievement evaluation, whose objective is to know the learning outcome that they obtain after the teaching and learning process is carried out. When they already comprehend physical concepts well, they would overcome problems related to the said concepts in daily life or overcome similarly complex problems. Council for the Advancement of Standards in Higher Education (CAS) in Komives (2012, p. 78) states that there are six domains of students' learning outcomes expected from learning, namely, those respectively of (1) knowledge acquisition, construction, integration, and application, (2) cognitive complexity, (3) intrapersonal development, (4) interpersonal competence, (5) humanitarianism and civic engagement, and (6) practical competence.

Learning outcomes could not fully assist students in comprehending something. It is in line with what is delivered by Hussey and Smith (Brooks, 2014, p. 722) by suggesting that learning outcomes can never fully capture the open, creative, and dynamic process of learning and, instead, they may lead to a narrowing of students' learning and tutors' delivery around the predetermined outcomes. The cognitive level of scientific models held by students, appeared in Chinese national curriculum standards and national examination requirements are equivalent, but there are also some differences among students' learning, curriculum standards and examination requirements. A specific classification of the models based on the comprehensive understanding systems is conducted, and the impact on learning among different schools and regions and inherent relationship on the cognitive level of physics models are founded by Wang, Zhang, and Shi, (2016). Since learning outcomes do not fully capture the open, creative, and dynamic process in learning, learning achievement is therefore not the one and only most

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

important aspect in learning. Besides students' learning outcomes, students' way of thinking could also determine their success. One of the important ways of thinking is thinking critically. Critical thinking is an important element in learning. It has even been applied since the times of Socrates. However, many still do not master the said ability and not all learning could be well received by students. It concerns a skill which greatly assists students in solving more complex problems. Siew and Mapeala (2016) state that problem-based learning with thinking maps effects on fifth graders' science critical thinking. The result indicated that students in the PBL-TM group significantly outperformed their counterparts in the PBL group who, in turn, significantly outperformed their counterparts in the CPS group in comparing and contrasting, sequencing, and identifying cause and effect. The findings suggest that thinking maps, which were explicitly infused into problem-based learning is effective in promoting critical thinking among fifth graders in physical science lessons.

The survey shows that students do not possess explicit cognitive activity; learning happens rather passively, without initiative; however, they have rather pronounced interest to explore and solve problems connected with the real life. Some implications for teachers on how to increase learners' cognitive interest are provided in the conclusion (Cēdere et al., (2015). Everyone could think but not everybody thinks deeply and the tendency is to think of many things so that no focus is given to one problem. It is in line with what is delivered by Ennis (1991, p. 7) as follows: "critical thinking ... means reasonable reflective thinking that is focused on deciding what to believe or do." According to Ennis, critical thinking is a reflective way of thinking which makes sense and is focused on what to believe or do. Elder and Paul (2007, p. 4) state that "critical thinking is the art of analysing and evaluating thinking with a view to improving it." Critical thinking is considered an art practiced by means of analysing and evaluating the manner of thinking itself so that we could improve our way of thinking. In critical thinking, there is a demand for our being able to train our analysis in order to further improve our way of thinking. Rudinow and Barry (2005, p. 12) explain that the term critical thinking in a way comes from the Greek word ***kritikos*** derives from the word "**critic**", which means the ability to evaluate or to discern. Critical thinking is to an extent indeed the ability to find faults and negative judgments but actually it is not limited to only that. To think critically, one requires a clear and rational mind and follows the rules of logic and scientific reasoning above all so that one could determine the right reasons in making decisions. As said by Lau (2011, p. 1), "Critical thinking is thinking clearly and rationally. It involves thinking precisely and systematically, and following the rules of logic and scientific reasoning, among other things."

There are various explanations concerning thinking critically but there is one similarity among them, namely, conclusions are drawn by using precise and systematic thoughts and following the rules of logic and scientific reasoning to improve the way of thinking. Thus, thinking critically is thinking not only of how one could answer a question but also of what is the way to get to that answer systematically and precisely. Moore and Parker (2009, p. 5) state that there are three basic building blocks of critical thinking, namely, the claim, the issue, and the argument. The claim or statement is a basic matter in critical thinking, with a statement not always being something true (or right or correct) and something false (or wrong or incorrect) being also able to be expressed in a statement. However, a statement need not be critically evaluated when its being true or false is already obvious. As for an issue or problem, its concept is very simple, namely, that it is just a question, meaning that after a statement is obtained, generally a question would appear to make sure whether the statement is true or false. The question appears based on what is already recalled in oneself. After a question appears like that, then our self makes a response to the statement and gives a reason for its being considered true or false. That is what is interpreted as argument by Moore and Parker (2009, p. 10). The critical thinking ability, according to Ennis (2013), is divided into several large categories, namely, those involving respectively (1) basic clarification, (2) bases for a decision, (3) inference, (4) advanced clarification, (5) supposition and integration, and (6) auxiliary or facilitative abilities. The categories have respective indicators and sub-indicators supporting the development of the critical thinking skill.

One of the efforts to improve learning achievement and the critical thinking skill is conducting varied learning. One of the learning models proposed by Ministry of Education and Culture in Indonesia for helping to improve students' abilities and skills is problem-based learning. Problem-based learning (PBL) is a learning model that presents various problems occurring in students' life so that it could stimulate them to learn. In PBL, students are to work in groups to seek solutions of existing problems. PBL is a learning model centred on students so that they are to actively seek information on their own and determine which information should be studied and learned to assist them in learning.

The PBL model is developed and designed to help students to build up expansive knowledge and to be able to apply the knowledge obtained on effective problem solving. It is in line with the statement by Oon-Seng Tan (2009, p. 9) that "in PBL, the problem is cast in a realistic context that the student might encounter in future."

The PBL model presents a problem which is realistic and might well confront students in the future. Like other learning models, according to Tan (2004, p. 8), PBL has its characteristics, namely, (1) the problem is presented initially in the learning conducted, (2) it should be a problem being faced and like a problem in daily life or, if the intention is presenting a problem in a simulation, the problem should be authentically possible to occur, (3) the problem presented should cover all the existing perspectives, (4) it should challenge the knowledge, attitude, and competence that students possess, (5) independent learning is the key to PBL, (6) the utilization of various sources of knowledge and the use and evaluation of information resources are important processes in PBL, (7) the PBL process itself is collaborative, communicative, and cooperative in nature, (8) the development of the investigation and the problem solving ability is equal in importance with the acquisition of knowledge in seeking the solution to a problem, (9) the closing part in the process of PBL is a synthesis and integration in the learning process, and (10) PBL also ends with an evaluation and review of the learning experience and learning process.

One of the important points in PBL is that the problem presented is not an initial question and knowledge is given by using a lecture and handout by the teacher. The said problem should be relevant with and actual in the students' life so that PBL could train students in getting a solution to the problem by themselves. The application of PBL in the improvement of the critical thinking skill is not something new. Take, for example, the research by Henderson (2014) on the relation between problem-based learning and the development of the critical thinking skill in higher education. In the research, Henderson obtained pre-test and post-test data from two classes respectively serving as PBL-using class and traditional class. Henderson made gain and MANOVA calculations to evaluate the difference among components of the critical thinking skill. The results obtained by Henderson (2010) indicate that there is no significant difference in students' critical thinking skill between a PBL-applying class and a traditional class. Henderson finds that the PBL class is higher in level of analysis, evaluation, and induction skills compared to the traditional one.

PBL-related research was also done by Anderson II. The research was conducted on two classes consisting of 110 persons in all with 56% of them consisting of women. One of the classes served as experimental group applying PBL with 67% of the persons in it consisting of women and the other class served as control group applying TGL (traditional group learning) with 46% of the persons in it consisting of women. In the research, a statistical difference between the two classes has been found. A comparison between the post-test mean score and the pre-test mean score in the PBL class has indicated a decrease of two points in magnitude while that in the TGL class has indicated a decrease of five points in magnitude. It, therefore, means that a greater decrease has occurred in the TGL class compared to that occurring in the PBL class. Consequently, the researcher concludes that such a PBL class is better compared with such a TGL class.

The third research to mention in this relation was done by Agdas (2013) on the effect of PBL application on the improvement of critical thinking and characteristics of students of the mechanical engineering field of study. The research used two classes with one consisting of twenty-three students serving as experimental group applying PBL and one consisting of twenty-two students serving as control group employing the lecture or traditional method. In the control group, there has been a decrease in mean score from 79.2 in the pre-test to 78.3 in the post-test but there has been an increase in standard deviation from 6.5 in the pre-test to 7.7 in the post-test. In the experimental group, there has been an increase in mean score from 78.1 in the pre-test to 78.4 in the post-test with a decrease in standard deviation from 8.4 in the pre-test to 6.2 in the post-test. This research by Agdas does not indicate any difference between the experimental group and the control group.

Tiwari et al. (2006) conducted research on a comparison in effect between PBL and lecture learning on the critical thinking skill. The research used two classes, namely, a PBL class of forty students and a class of thirty-nine students learning with the method of lecturing. Measurements were done by using the California Critical Thinking Disposition Inventory (CCTDI). The result indicates overall improvement of the PBL class in the CCTDI (with $p = .0048$). On the whole, there is a significant difference in the class using PBL compared with the class using the lecture method.

With the exposition above as basis, it was then considered of interest to know the effect of problem-based learning on improvement in students' learning achievement and critical thinking at a state senior high school in Indonesia. The research concerned here was then aimed at knowing the effect of applying a PBL model on students' physics learning outcomes and critical thinking skills. The research was focused on the cognitive domain and specifically on its respective levels of applying, analysing, evaluating, and creating. Those were the levels selected for the reason that students at senior high school should already be at those levels. The critical thinking skills put under research were giving basic clarification, building bases for a decision, making an inference, and making an

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

advanced clarification. These four critical thinking skills were selected because they were still related to physics learning so that it was worthy to know to what extent the PBL model could improve those skills.

Based on preliminary observations in a senior high school in Yogyakarta Indonesia, can be described some of realistic problems, as follows: (1) in physics learning process teachers still use traditional methods so as not to give sufficient time to students to participate in learning, (2) in physics learning process students tend to be passive so that they are just follow what is said by teacher, (3) in physics learning process teachers tend to rarely address issues related to the material to be taught, (4) in physics learning process rarely started with a realistic problem, (5) in physics learning process students are only able to use the existing ability just to solve the problems posed by teachers and less able to apply them in everyday life, (6) in physics learning process teachers also rarely train students to think critically in solving problems, and (7) in physics learning process teachers more often use classical methods than student center methods. Referring to some of the field problems, this research wants to know whether there is a positive effect of PBL model on physics learning outcomes and critical thinking skills of students.

The result of this research is expected to give practical contribution as one of alternative choices in improving learning effort, as follows: (1) For students, PBL model is expected to improve students' learning outcomes and critical thinking skills that it is very useful in solving problems. In addition, through the PBL is expected to provide experience to students learn together with a group. (2) For teachers, this research is expected to increase the competence of teachers in implementing the learning process in classroom. By applying a varied learning is expected to help educators in facilitating students to develop their skills. (3) For researchers, the results of this study provide a fact that PBL model effects on students' learning outcomes and critical thinking skills.

Methodology of Research

Research Type

The research was quantitative in approach and used a quasi-experimental research design. It involved two classes, namely, one serving as experimental group given treatment in the form of the application of a problem-based learning model with the method of experiment and one serving as control group given the same form of treatment but with the method generally used at school, namely, the method of demonstration as a comparison.

Research Location and Time

As previously implied, the research was conducted at a senior high school in Yogyakarta Indonesia. It was conducted on the students of Grade X in Semester 2 of the academic year 2015/2016. The decision to conduct the research during Semester 2 of the 2015/2016 academic year was for the purpose of making the physics class material fit the physics material used in the research. As previously mentioned, the learning material used was about dynamic electricity.

Research Population and Participants

The population of the research consisted of seven classes or 224 students of Grade X in Semester 2 at a senior high school in the 2015/1016 academic year. By means of purposive sampling, two classes, namely, Classes X4 and X7, were selected as participants. Class X4, which consisted of 32 students served as experimental group and Class X7, which consisted of 32 students, served as control group.

Procedure

The research used a quasi-experimental research design. The steps in the research were (1) giving a pre-test to both the experimental group and the control group, (2) giving treatment by applying a PBL model with the method of experiment on the experimental group and the method of demonstration on the control group, and (3) giving a post-test to both classes.

Table 1. Research design.

Experimental group	O_1	X_1	O_2
Control group	O_1	X_2	O_2

O_1 : Pre-test of the theme-related learning achievement and critical thinking skill.
O_2 : Post-test of the theme-related learning achievement and critical thinking skill.
X_1 : Learning treatment in the form of problem-based learning with the method of experiment.
X_2 : Learning treatment in the form of problem-based learning with the method of demonstration.

Data, Instruments, and Techniques of Data Collection

In the research, a PBL model with the method of experiment in the experimental group and the method of demonstration in the control group served as independent variable while learning achievement and the critical thinking skill served as a dependent variable. The data about learning achievement were obtained through multiple-choice test items and the data about the critical thinking skill were obtained through essay-type test items at the beginning and end of the learning process. The multiple-choice test items and the essay-type test items were for the purpose of obtaining respective data about the learning achievement and the critical thinking skill of the students in both the experimental and control groups.

The research used two data collection methods, namely, test and observation. Observations were made using observation sheets to see the suitability between lessons plan made with the implementation in the classroom. Before the data collection itself was conducted, the data source, then the data type, the technique of data collection, and the instrument(s) to be used were first determined one after another. In order to obtain data that supported the research in answering the research questions, the researcher had constructed some instruments, namely, the learning achievement test, the critical thinking skill test, and the learning activity accomplishment observation sheet. All this could be seen in Table 2.

Table 2. Technique of data collection.

Data Source	Data Type	Data Collection Technique	Instrument
Students	Learning Achievement	Pre-test and Post-test	Thirty Multiple-Choice Test Items
Students	Critical Thinking Skill	Pre-test and Post-test	Four Essay-Type Test Items
Teacher	Accomplishment of Learning Process Stages	Observation	Guide to the Observation of Teacher Activity During the Learning in Line With the Learning Execution Plan Developed

The students' learning achievement test consisted of thirty multiple-choice items. The learning achievement test material covered Ohm's Law(s) in seven items, series-parallel circuits in twelve items, the factors influencing the magnitude of resistivity in eight items, Kirchhoff's Law in one item, and the voltmeter and the amperemeter in two respective items. These thirty items were divided into four groups according to the levels of the cognitive domain dealt with, namely, the levels of, respectively, applying (C3), analysing (C4), evaluating (C5), and creating (C6). The items for C3 were eleven in number, those for C4 were eight in number, those for C5 were six in number, and those for C6 were five in number.

The critical thinking skill test was a written test consisting of four essay-type items. The test was to measure students' critical thinking skill before and after the learning.

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Table 3. Critical thinking skill item distribution.

Critical Thinking Category	Material	Number of Item
Basic Clarification	Electric instruments: amperemeter, voltmeter	1
Bases for a Decision	Ohm's Law: electric current, electric resistance, voltage	1
Inference	Series-parallel circuits	1
Advanced Clarification	Factors influencing the magnitude of electric resistivity	1

Technique of Data Analysis

The resulting research data were analysed by using gain factor and hypothesis analyses. To calculate improvement in students' conceptual understanding and critical thinking skill from those occurring before the learning to those occurring after it, the gain was calculated according to the following formula developed by Hake (2008):

$$\langle g\rangle = \frac{S\,post - S\,pre}{100\% - S\,pre} = \frac{S\,post - S\,pre}{100\% - S\,pre}$$

in which
$\langle g \rangle$ = gain factor
Spre = pre-test mean score (%)
Spost = post-test mean score (%)

The hypothesis testing used the multivariate test. The test was conducted on the significance level of the F values in Hotelling's Trace statistics. The criterion for such testing is that H_0 is rejected if $F_{obtained} \geq F(p,n_1+n_2\text{-}p\text{-}1; .05 = 2.66; .05)$ or the level of significance obtained is smaller than .05.

Before such a test otherwise known as the MANOVA test is conducted, testing of normality and homogeneity should be done. In the case here, the normality testing was intended to reveal the distribution of the data scores for students' learning achievement and critical thinking skill in the two classes. In the research, the testing of normality used the one-sample Kolmogorov-Smirnov test. The homogeneity testing was done to see whether the variances in the data for improvement in learning achievement and critical thinking of the experimental group and the control group were the same or not by using the Levene's test. In short, the homogeneity testing was intended to reveal whether there was equality of variances between the two classes. The testing of homogeneity could also use the SPSS program.

Results of Research

The Accomplishment of the PBL Model

Data of the accomplishment of the PBL model in class were obtained by means of observation. The observer in each session was the same person. In doing the observation, the observer was equipped with the observation sheet provided. The observation sheet filled in by the observer indicated to what extent the application of the PBL model was accomplished.

In percentage of accomplishment, the first session in the experimental group did not reach 100%. The percentage of accomplishment in Phase V, namely, the phase of discussing, analyzing, and evaluating the process of problem solving, was 75%. The percentage of accomplishment of the phases in the second session did not all reach 100%, either. In Phase I, namely, the phase of student orientation to the problem, it was 80%. In the third session, the percentage of accomplishment in all phases reach 100%, possibly because the researcher had held discussions with the observer about existing shortcomings in the preceding two sessions. See Table 4.

Table 4. Degree of accomplishment of the learning in the experimental group.

Session	Syntax of Model	Degree of Accomplishment (%)
First	Phase I	100
	Phase II	100
	Phase III	100
	Phase IV	100
	Phase V	75
Second	Phase I	80
	Phase II	100
	Phase III	100
	Phase IV	100
	Phase V	100
Third	Phase I	100
	Phase II	100
	Phase III	100
	Phase IV	100
	Phase V	100

In every session of the control group, the accomplishment of the phases did not all reach 100%. In the first session, the accomplishment of Phase IV, namely, the phase of developing and presenting the results of their work, was only 50%, which was caused by forgetfulness in the researcher's part to ask students to deliver their conclusions, and the accomplishment of Phase V, namely, the phase of problem-solving process analysis and evaluation, was only 75%. In the second session, the accomplishment of Phase III was also only 75%, possibly because the researcher only did the experiment himself and directly presented only the results to the student moving to the front of the class, and the accomplishment of Phase V was 75%, with the researcher forgetting to give assignments or tasks to students. In the third session, the accomplishment of all phases was 100% because the researcher had learned from the mistakes occurring in the preceding sessions. See Table 5.

Table 5. Degree of accomplishment of the learning in the control group.

Session	Syntax of Model	Degree of Accomplishment (%)
First	Phase I	100
	Phase II	100
	Phase III	100
	Phase IV	50
	Phase V	75
Second	Phase I	100
	Phase II	100
	Phase III	75
	Phase IV	100
	Phase V	75
Third	Phase I	100
	Phase II	100
	Phase III	100
	Phase IV	100
	Phase V	100

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

The Results of Pre-Test and Post-Test of the Experimental Group and the Control Group

As previously said, the students' learning achievement test used thirty multiple-choice test items. The learning achievement test was done twice, as pre-test before the learning and as post-test after the learning. The results of the pre-test were used to know the students' initial condition in particularly their understanding of the material about dynamic electricity because it had previously been studied in Grade IX. The results of the post-test were used to know to what extent the students could master the material after treatment. These pre-test and post-test activities were equally applied on both the experimental group and the control group. The following are the results of the pre-test and the post-test of the experimental group and the control group.

Table 6. Results of the pre-test and post-test of students' learning achievement.

Category	Experimental group		Control group	
	Pre-test	Post-test	Pre-test	Post-test
Mean Score	27.74	73.01	30.93	53.75
Highest Score	50.00	86.67	53.33	73.33
Lowest Score	3.33	53.33	13.33	33.33
Standard Deviation	9.71	9.75	10.88	10.99
Average Gain	.63		.32	

Table 6 makes clear that there is difference in pre-test mean score between the experimental group and the control group with the pre-test mean score of the former being 27.74 and that of the latter being 30.93. The highest pre-test score of the control group is 53.33, which is higher than that of the experimental group, which is 50.00. Likewise, the lowest pre-test score of the control group is higher than that of the experimental group, these scores being respectively 13.33 and 3.33. The abovementioned difference in mean score, highest score, and lowest score is possibly a result of difference in pre-test time because the pre-test of the control group was still in the morning or, more specifically, before 10.00 a.m. while that of the experimental group was in the afternoon so that the students in that class were already beginning to feel tired. Difference in previous school of origin may be something else that possibly had caused the aforesaid score difference.

As for the post-test results, they also show differences. The post-test mean score of the experimental group is higher compared to that of the control group, these scores being respectively 73.01 and 53.23. The situation of the highest and lowest scores is also reversed. As previously mentioned, the control group is higher in pre-test highest score but the experimental group is higher in post-test highest score, with 86.67 being the post-test highest score of the experimental group and 73.33 being that of the control group. The post-test lowest score of the experimental group is 53.33 while that of the control group is 30.00. The pre-test standard deviation of the experimental group is 9.71 and that of the control group is 11.22; the post-test standard deviation of the experimental group is 9.75 and that of the control group is 11.5.

The average numbers of items for the considered cognitive levels correctly answered by students of each class in the pre-test are as follows. The items for the C3 cognitive level correctly answered by the experimental group average four in number; so do those correctly answered by the control group. The items for the C4 cognitive level correctly answered by the experimental group average two in number while on the average only one such item is correctly answered by the control group. The items for the C5 cognitive level correctly answered by the experimental group average two in number; so do those correctly answered by the control group. On the average, only one item for the C6 cognitive level is correctly answered by both classes.

The average number of items for the considered cognitive level correctly answered by students of each class in the pre-test shown in Figure 1.

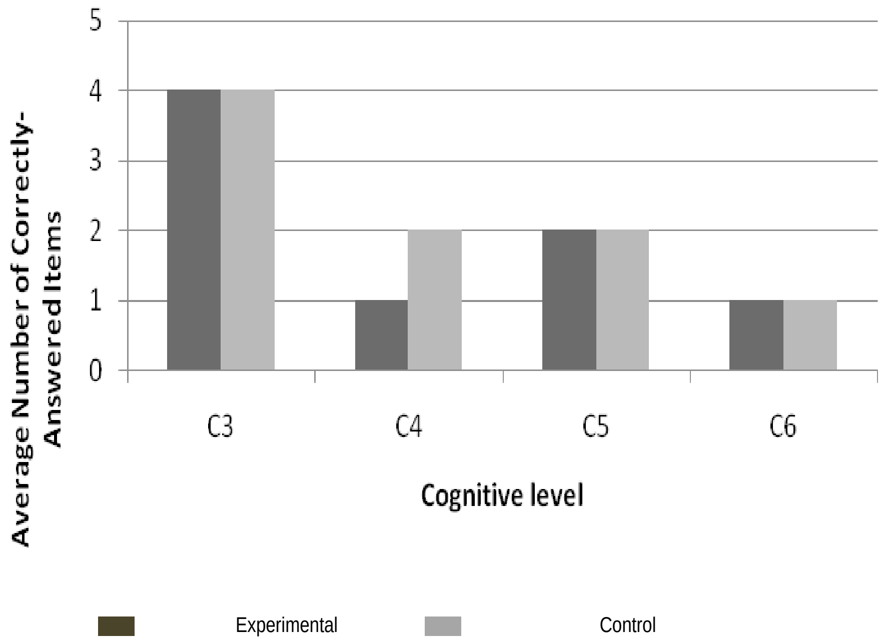

Figure 1. Diagram of the average number of correctly-answered items for each cognitive level in the pre-test.

The pre-test scores are distributed in the greatest number in the 16 – 30 score range and they are of 16 students in the experimental group and 15 students in the control group. Those distributed in the next greatest number are in the 31 – 45 score range and they are of 11 students in the experimental group and 13 students in the control group. The scores belonging to the 46 – 60 score range are of one student in the experimental group and three students in the control group. The scores belonging to the 0 – 15 score range are the smallest in number and they are of 3 students in the experimental group and one student in the control group.

The distribution of the pre-test scores of the experimental group and the control group shown in Figure 2.

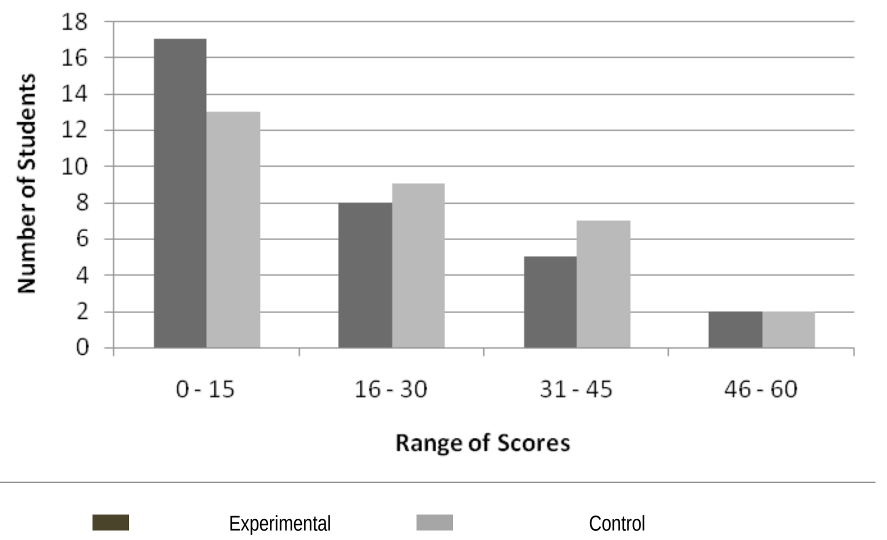

Figure 2. Distribution of the pre-test scores for students' learning achievement of the experimental group and the control group.

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

After the pre-test, the research continued with the learning stage by applying the method that had been planned. After the learning was conducted, post-test scores were obtained. The scores could be seen in Figure 3. The average number of items for the considered cognitive level correctly answered by students of each class in the post-test shown in Figure 3.

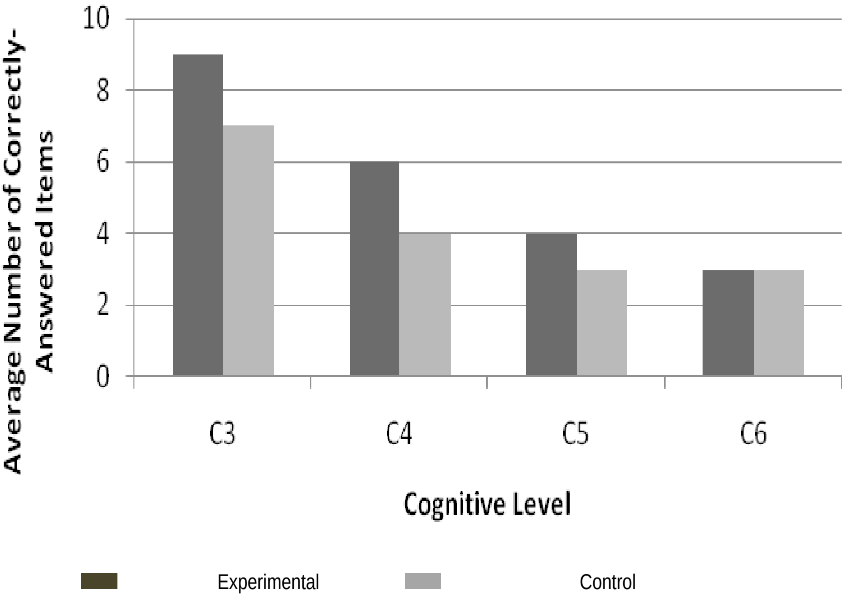

Figure 3. Diagram of the average number of correctly-answered items for each cognitive level in the post-test.

Figure 3 shows the average number of items for each cognitive level correctly answered by the two classes in the post-test. In this case, there is a difference in the situation compared with that in the pre-test. In the pre-test, the experimental group is lower in position than the control group in that matter but in the post-test the experimental group manages to be higher in average number of correctly-answered items for each cognitive level than the control group. The items for the C3 cognitive level correctly answered by the experimental group are 9 in average number while those correctly answered by the control group are 7 in average number. The items for the C4 cognitive level successfully answered by the experimental group are 6 in average number while those successfully answered by the control group are 4 in average number. The items for the C5 cognitive level correctly answered by the experimental group are 4 in average number while those correctly answered by the control group are 3 in average number. As for the items for the C6 cognitive level, those successfully answered by the experimental group are three in average number and so are those successfully answered by the control group. The distribution of the post-test scores of the experimental group and the control group could be seen in Figure 4.

The distribution of the post-test scores of the experimental group and the control group shown in Figure 4.

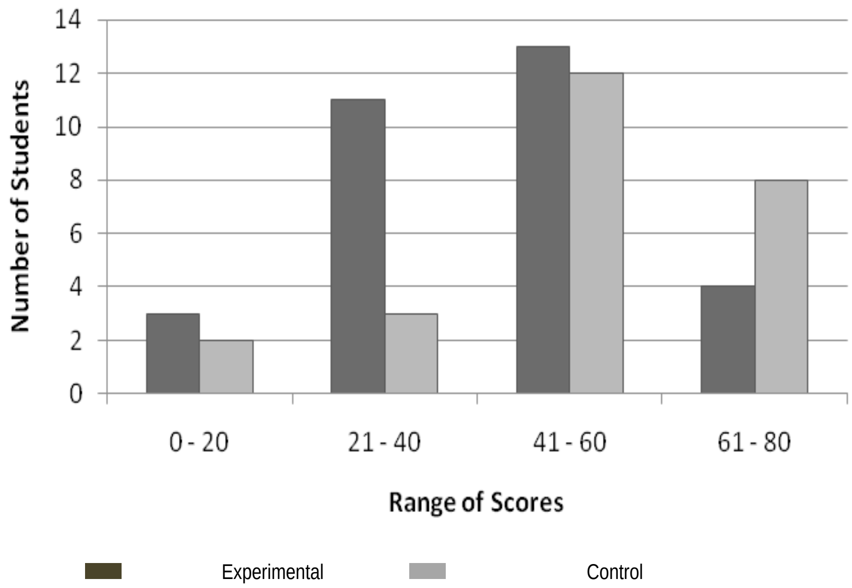

Figure 4. Distribution of the post-test scores for the learning achievement of students in the experimental group and the control group.

The distribution of the post-test scores of the experimental group and the control group appears quite different from that of their pre-test scores. While there are still 6 students in the control group with scores within the 31 – 45 score range, no student in the experimental group has a score within that score range. The 46 – 60 score range is the one receiving the greatest number of scores with the scores of 16 students in the control group included in that score range while there are only 5 students in the experimental group with scores within that score range. The students in the control group and the experimental group with scores within the 61 – 75 score range are respectively 10 and 13 in number. There are no students in the control group with scores within the 76 – 90 score range while there are 13 students in the experimental group with scores within that score range.

After viewing and recapitulating the pre-test results, the researcher's following activity before entering the calculation stage of using MANOVA was determining the homogeneity of the pre-test scores of the two classes.

Table 7. Results of the homogeneity test on the pre-test data.

Criteria	Result
Obtained Significance	.254
Criterion Significance	.05
Degree of Freedom	61
Variances	the same

The homogeneity testing is concerned with Ho stating that the two variances concerned are the same and Ha stating that the two variances differ. The testing of homogeneity used the Levene test with the level of significance of 5%. If the level of significance obtained > .05, then Ho would be accepted and conversely, if the level of significance obtained < .05, then Ho would be rejected and Ha would be accepted. From Table 7, it is seen that the obtained level of significance is .254, its being > .05 meaning that Ho is acceptable. It indicates that the pre-test scores of the two classes have the same variance value. Besides, determining the homogeneity by the calculation using the pre-test data, it was also done by analyzing the gain of the two classes, which was found to be .167 in magnitude. With the gain calculation in determining the homogeneity as basis, again the value obtained > .05 so that the use of the gain calculation also results in obtaining the same variance value for the data of both classes. In other words, the score data of the two classes are homogenous or could be said to be homogenous.

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

The normality testing is done for the purpose of knowing whether the data are normally distributed or not. The testing of normality was applied on the gain in students' learning achievement of the experimental group and the control group. The normality testing is concerned with H_o stating that the data are normally distributed and H_a stating that the data are not normally distributed. The results of the normality testing using the Kolgomorov-Smirnov test could be seen in Table 8.

Table 8. Results of the normality testing on learning achievement scores by use of the Kolgomorov-Smirnov test.

Group	Dependent Variable	Kolgomorov-Smirnov		
		Statistic	df	p
Experimental	Gain in Students' Learning Achievement	.115	31	.200
Control	Gain in Students' Learning Achievement	.138	32	.129

From Table 8 it is known that the obtained level of significance is .200 for the experimental group and it is .129 for the control group, in both cases > .05 so that Ho is accepted, indicating that the gain scores for the learning achievement of the two classes are normally distributed.

The testing of students' critical thinking used an essay-type test consisting of four test items. It was done twice, namely, once before the learning and once after the learning. The results of the test before the learning (or the pre-test) were used to know students' initial condition in understanding the material about dynamic electricity because that subject matter had previously been studied in Grade IX. The critical thinking test was also done after the learning was conducted for the purpose of knowing to what extent students could master the material after treatment. These activities were applied on both the experimental group and the control group. The following presents the pre-test and post-test results of the experimental group and the control group.

Table 9. Results of the critical thinking pre-test and post-test of the experimental and control groups.

Category	Experimental Group		Control Group	
	Pre-test	Post-test	Pre-test	Post-test
Mean	22.10	59.84	17.34	46.72
Highest Score	60.00	95.00	60.00	85.00
Lowest Score	.00	15.00	.00	10.00
Standard Deviation	16.37	21.00	18.84	16.78
Average Gain	.49		.34	

Table 9 shows that the pre-test mean scores of the experimental group and the control group are respectively 22.10 and 17.34 while their post-test mean scores are respectively 59.84 and 46.72. From these mean scores, it is seen that there is improvement in the two classes. Their highest pre-test scores are the same, both being 60.00 in magnitude, and so are their lowest pre-test scores, both being .00 in magnitude. In average gain score, the two classes differ, that of the experimental group being .49 with that of the control group being .34.

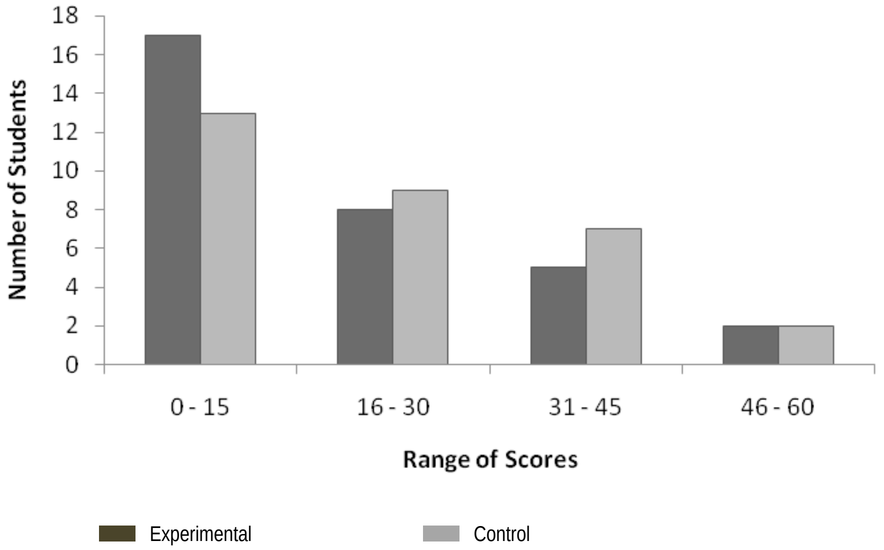

Figure 5. Distribution of the pre-test scores for critical thinking of the experimental group and the control group.

The students in the control group with scores within the 0 – 15 score range are 17 in number while those in the experimental group with such scores are 13 in number. Those in the control group with scores within the 16 – 30 score range are 8 in number while those in the experimental group with such scores are 9 in number. Those in the control group with scores within the 31 – 45 score range are 5 in number while those in the experimental group with such scores are 7 in number. Those in each class with scores within the 46 – 60 score range are the fewest, they being only two in number.

After viewing the pre-test results, we would view the post-test results and the data of the score distribution could be seen in Figure 6.

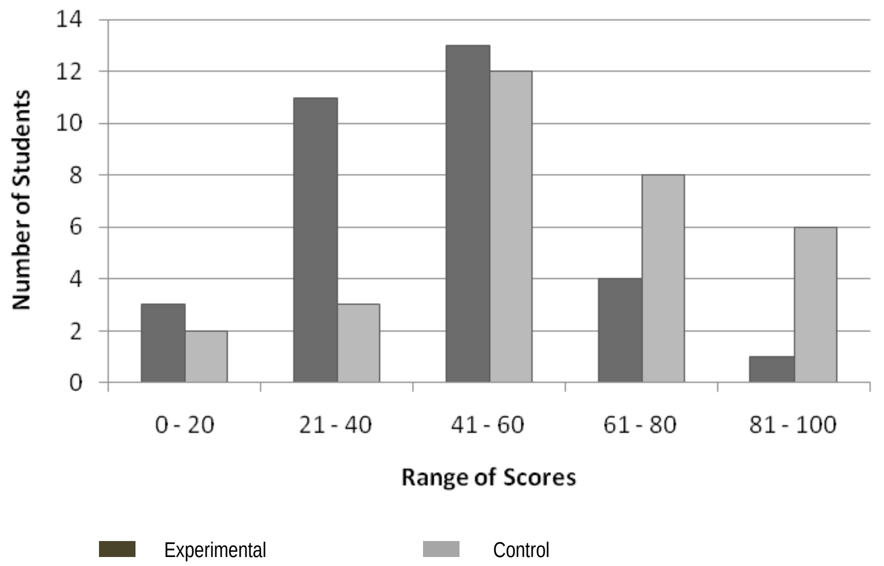

Figure 6. Distribution of the post-test scores for critical thinking of the experimental group and the control group.

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Figure 6 shows a sufficiently striking difference. The students in the control group with scores within the 0 – 20 score range are 3 in number while those in the experimental group with such scores are 2 in number. Far more students in the control group are with scores within the 21 – 40 score range compared to those in the experimental group with such scores, they being respectively 11 and 3 in number. The students in the control group and those in the experimental group with scores within the 41 – 60 score range do not differ far in number, they being respectively 13 and 12 in number. The students in the control group with scores within the 61 – 80 score range are 4 in number while those in the experimental group with such scores are 8 in number. There is only one student in the control group with a score within the 81 – 100 score range while there are far more students in the experimental group with such scores, they being six in number.

The testing of homogeneity was also conducted on the pre-test scores for critical thinking. Such testing concerns Ho stating that the two variances concerned are the same and Ha stating that those two variances differ. The homogeneity testing used the Levene test with the level of significance of .05 as criterion. If the level of significance obtained > .05, then Ho is accepted and, conversely, if the level of significance obtained < .05, then Ho is rejected and Ha is accepted. The results of the homogeneity testing of the pre-test data for critical thinking could be seen in Table 10.

Table 10. Homogeneity testing on the pre-test scores for critical thinking.

Criteria	Result
Obtained Significance	.281
Criterion Significance	.05
Degree of Freedom	61
Variances	the same

Table 10 shows that the level of significance obtained is .281. It being > .05 means that Ho is accepted, which means in turn that the two classes have the same variance value. Besides calculations using pre-test data, the homogeneity testing is also done by testing the gain scores. After the testing is done, the value obtained is .314. It being > .05 again means that the two classes have the same variance value and their scores are homogenous. After testing the homogeneity, it is also necessary to test the normality by using the MANOVA test before going to the stage of final completion.

The scores for critical thinking also need to be given a test of normality because the purpose is to know whether the score data are normally distributed or not. The testing of normality is done to the gain scores for the critical thinking of students in the experimental group and the control group. The normality testing is concerned with H_o stating that the data are normally distributed and H_a stating that the data are not normally distributed. The normality testing was done by using the Kolgomorov-Smirnov test and the results could be seen in Table 11.

Table 11. Results of the normality testing by using the Kolgomorov-Smirnov test on the scores for critical thinking.

Group	Dependent Variable	Kolgomorov-Smirnov		
		Statistic	df	p
Experimental	Gain in Students' Critical Thinking	.092	31	.200
Control	Gain in Students' Critical Thinking	.117	32	.200

According to Table 11, the obtained significance level for the experimental group is .200 in value and that for the control group is also .200 in value. Since .200 > .05, then Ho is accepted and it indicates that the gain scores for critical thinking of the two classes are normally distributed.

Such pre-test and post-test results could be used for hypothesis analysis when they already fulfill the requirements for homogeneity and normality. When the results of the data testing indicate that the two data batches are of the same variance and are normally distributed, then the hypothesis testing analysis would use parametric statistics.

Table 12. Testing of the variance and covariance assumption in MANOVA by using Box's M test.

Box's M	F	df	df	p
2.916	.937	3	6.931E5	.421

Table 12 shows the results of multivariate normality testing using Box's M test.
Such testing concerns the following pair of hypotheses:

Ho : the two dependent variables have the same variance and covariance

Ha : the two dependent variables have differing variance and covariance

The criteria for decision making in Box's M test are that if sig > .05, then Ho is accepted and if sig < .05, then Ho is rejected. From Table 12, it is seen that Box's M value is 2.916 with sig being .421. Because that value of sig > .05, then Ho is accepted. It means that the variance and covariance of variables for the experimental group and that of the control group are the same.

Table 13. Results of individual variable testing by using the Levene test.

	F	df	df	p
Learning Achievement	1.956	1	61	.167
Critical Thinking	1.029	1	61	.314

Table 13 shows the results of individual variable testing by using the Levene test. It is seen that the level of significance of students' learning achievement with F = 1.956 is .167 and that of students' critical thinking with F = 1.029 is .314. It is shown that the obtained significance level for either variable is greater than .05 so that Ho is accepted. It indicates that the two classes individually have the same variance.

Table 14. Results of the multivariate test.

Effect		F	p
Intercept	Pillai's Trace	4.120E2a	.0001
	Wilks' Lambda	4.120E2a	.0001
	Hotelling's Trace	4.120E2a	.0001
	Roy's Largest Root	4.120E2a	.0001
Method	Pillai's Trace	37.238a	.0001
	Wilks' Lambda	37.238a	.0001
	Hotelling's Trace	37.238a	.0001
	Roy's Largest Root	37.238a	.0001

The next results of analysis are of the multivariate test. The multivariate test itself is used to test the research hypothesis. The initial hypothesis of the research concerned here is that there is no significant difference between the two classes concerned in effect of problem-based learning on students' learning achievement and critical thinking in relation with the learning material concerning dynamic electricity. The criteria for decision making in the test are that if the significance level obtained > .05, then Ho is accepted and if the significance level obtained < .05, then Ho is rejected.

Table 14 shows that the significance level obtained by means of the respective procedures of Pillai's, Wilks', Hotelling's, and Roy's is .0001 in value. Since the obtained significance level < .05 in value, then Ho is rejected. It means that there is significant difference between the experimental group and the control group in effect of the PBL model on students' learning achievement and critical thinking.

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Table 15. Results of tests of between-subjects effect.

Source	Dependent Variable	df	F	p
Corrected Model	Learning Achievement	1	73.888	.0001
	Critical Thinking	1	7.046	.0100
Intercept	Learning Achievement	1	718.245	.0001
	Critical Thinking	1	222.183	.0001
Method	Learning Achievement	1	73.888	.0001
	Critical Thinking	1	7.046	.0100

Table 15 shows the same values for Corrected Model and Method, with the significance value of .000 for learning achievement and .0100 for critical thinking. The two figures for the significance value are less than .05 so that Ho is rejected. Therefore, it is concluded that the PBL model with the method of experiment as in the experimental group exerts significant effect on students' physics learning achievement and critical thinking skill compared with that with the method of demonstration as in the control group.

Discussion

The use of the PBL model employing the method of experiment and the method of demonstration hopefully would bring about a result conforming to an objective of physics learning, namely, making students show scientific behaviour and develop experience in using scientific methods. In addition, with the use of PBL students could develop the ability of reasoning in thinking because PBL puts emphasis on learning with a problem as basis. Tarhan and Acar-Sesen (2013) found that the mean scores of the students in the experimental group were instructed via PBL were significantly higher than those in the control group were instructed via teacher-centred approach. Problem Based Learning Assessment Scale results reflected that students' positive beliefs increased after each activity. Based on these results, it can be concluded that PBL instruction is effective in concept learning in chemistry education.

On the whole, the phases of PBL for the two classes in the research have been correctly applied. The pre-test and post-test results indicate that there is difference in the learning outcome of the two classes. The experimental group and the control group are the same in the items for the C3 cognitive level correctly answered by students in the pre-test being four in average number while in the post-test the items for the C3 cognitive level correctly answered by the experimental group and the control group are respectively nine and four in average number. It proves that the use of PBL brings more improvement to the cognitive level of C3, namely, the level of applying. In the average number of correctly- answered items for the C4 cognitive level in the pre-test, the experimental group is below the control group in position, correctly answering only one such item on the average while two such items on the average are correctly answered by the other class. There is significant improvement in the average number of items for the C4 cognitive level correctly answered by the experimental group, the items being six in average number while such items correctly answered by the control group are four in average number. It is therefore said that PBL could significantly affect improvement in learning achievement at the cognitive level of C4, namely, at the level of analysing.

So is the case with the C5 and C6 cognitive levels, the experimental group showing better improvement compared with the control group. In average gain score, the experimental group gets .63 while the control group gets .32. Those average gain scores of the two classes fall into the middle score range category. Based on self-determination theory (SDT) with a multilevel confirmatory factor analysis (MLCFA), it was hypothesised that the four factors of external regulation, introjected regulation, identified regulation, and intrinsic motivation account for the covariances among the items of the students' motivation to study physics in school (Byman et al, 2012). The results of the study were girls had a statistically higher mean score than did boys for all four factors. Otherwise only minor gender differences were found.

Such results show that the PBL model could improve students' learning achievement. The research indicates that using the method of experiment are more liable of showing better results compared to using the method of demonstration. It is in line with the objective of using PBL, namely, that the use of PBL would result in a learning that is meaningful for the students learning to solve problems by applying all the knowledge and ability in their possession or at least with the use of PBL students could try to obtain the knowledge that they require so that the learning process would become increasingly more meaningful when they are faced with a situation in which concepts are applied. Šorgo and Kocijančič (2011) found that the implementation of laboratory work in science teaching is not so positive. There are findings that many of the exercises are presented to the students as demonstrations and an expository style is preferred. The differences occur in the way laboratory work is performed. In the future, efforts should be made to transform expository labs into inquiry and problem-based laboratory work. Meanwhile, Hakkarainen and Ahtee (2010) find out how pupils will connect their observations and explanations from a demonstration. Both the observations and the explanations were improved significantly in the second demonstration even that no teaching was done between the demonstrations. The application of the method of experiment is more appropriate for moving to the objective favoured by the PBL model because students are taught to formulate a hypothesis, perform a process of experimenting, make an observation, retrieve data, and draw a conclusion from the experiment that they have conducted. Teacher participation in a sustained professional development intervention designed to improve the quantity and quality of guided inquiry-based instruction in middle school science classrooms and subsequent student academic growth. The results indicate that the students of participating teachers had significantly higher when compared to students of non-participants. This study supports prior research findings that inquiry-based instruction helps improve students' achievement relative to scientific practices and also provides evidence of increasing student conceptual knowledge (Marshall, Smart, and Alston, 2017).

The pre-test and post-test results in measuring the critical thinking skill could be seen from the mean score of 22.1 in the pre-test becoming 59.84 in the post-test in the case of the experimental group. The mean score for the critical thinking skill in the case of the control group also undergoes an increase from being 17.34 in the pre-test to being 46.72 in the post-test. Seen from the point of pre-test and post-test mean scores, both classes experience an improvement in critical thinking skill. The pre-test and post-test mean scores also show that on the average the gain score of the experimental group is .49 and that of the control group is .34. The average gain scores of the two classes are within a middle score range. From the gain scores it could be seen that the experimental group, by applying the method of experiment, is better in the learning compared with the control group applying the method of demonstration. The findings of the measurement of critical thinking overall suggest that the CT skills in electricity and magnetism (CTEM) test can be used to measure the acquisition of domain-specific CT skills in E & M, and a good basis for future empirical research that focuses on the integration of CT skills within specific subject matter instruction (Tiruneh et al., 2017).

From the MANOVA test used to compare the learning in the experimental group with that in the control group, the results obtained explain that the PBL model using the method of experiment is more influential compared with the PBL model using the method of demonstration. It is seen in the respective procedures of Pillai's, Wilks', Hotelling's, and Roy's all resulting with the value of .0001. That obtained significance value < .05 so that the related Ho is rejected. Besides, Table 23 shows that the aspects of Corrected Model and Method in the tests of between-subjects effect come to be the same in value. It could be concluded that the learning session using the PBL model with the method of experiment is significantly greater in effect compared with that using the PBL model with the method of demonstration. Therefore, it is suggested that that former kind of learning is the one to be used in every learning session instead of the latter.

With the above discussion as basis, it could be said that the application of PBL with the method of experiment could significantly improve achievement in learning physics at the respective cognitive levels of applying, analysing, evaluating, and creating. In addition, PBL with the method of experiment could also significantly improve students' critical thinking skill in the categories of respectively giving basic clarification, building bases for a decision, making an inference, and making advanced clarification. In line with the research by Tiwari et al. (2006) applying PBL to improve the thinking of nurses-to-be in Hongkong, PBL could also be applied on senior high school students to improve their physics learning achievement and their critical thinking skill. Yu, Fan, and Lin (2015) found that problem solving is often challenging for students because they do not understand the problem-solving process. The results indicate that context simulation is beneficial for cultivating students' abilities to establish and analyze questions and then select and develop solutions. In addition, the project design cultivated the students' ability to evaluate results and apply feedback. The findings of this study demonstrate that context-based learning may effectively enable students to establish and complete the problem-solving process.

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Conclusions and Suggestion

With all the preceding problem formulation, exposition, and calculation as basis, the conclusion could be drawn that, with the application of the PBL model in the learning, there is positive effect on students' learning achievement and critical thinking skill. In addition, it has been found that the average gain score for learning achievement of the experimental group and that of the control group have been respectively .63 and .32 and the average gain score for critical thinking of the experimental group and the control group have been respectively .49 and .34. Each of the MANOVA calculation procedures employed ends up with the value of .0001 and the tests of between-subjects effect shows the same value for the aspects of Corrected Model and Method. Therefore, the PBL model using the method of experiment as in the experimental group is significantly greater in effect compared with the PBL model using the method of demonstration as in the control group.

The PBL model is one of the learning models that could improve students' learning achievement and critical thinking skill but there is a matter that one needs to pay attention to in its application. It is showing the initial problem in the right way in order that students could understand what they should do. The problem presented should not be a problem stated as a question for it should be an actually occurring problem.

In order that the results of the research concerned here are improved, it is suggested that additional research be conducted to measure not only students' learning achievement and critical thinking skill but also other matters such as problem-solving abilities, other higher-level abilities, and science process skills or to measure students' collaboration abilities.

References

Agdas, S. (2013). *Effects of problem-based learning on development of critical thinking skills and disposition in engineering* (Doctoral dissertation, University of Florida, Florida). Proquest LLC, UMI 3586578.

Anderson II, J. C. (2007). *Effect problem-based learning on knowledge acquisition, knowledge retention, and critical thinking ability of agriculture students in urban schools* (Doctoral dissertation, University of Missouri, Columbia). Proquest LLC, UMI 3322674.

Anderson, L., & Kratwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objective*. New York: Longman.

Brooks, S., Dobbins, K., Scott, J. J. A., Rawlinson, M., & Norman, R. I. (2014). Learning about learning outcomes: The student perspective. *Journal Teaching in Higher Education*, *19* (6), 721-733.

Byman, R., Lavonen, J., Juuti, K., & Meisalo, V. (2012). Motivational orientations in physics learning: A self-determination theory approach. *Journal of Baltic Science Education*, *11* (4), 379-392.

Cēdere, D., Jurgena, I., Helmane, I., Tiltiņa-Kapele, I., & Praulīte, G. (2015). Cognitive interest: Problems and solutions in the acquisition of science and mathematics in schools of Latvia. *Journal of Baltic Science Education*, *14* (4), 424–434.

Elder, L., & Paul, R. (2007). *The miniature guide to critical thinking concepts and tools*. Retrieved from http://www.criticalthinking.org/files/SAM_Analytic_Think2007b.pdf.

Ennis, R. H. (1991, March). Critical thinking: A streamlined conception. *Teaching Philosophy*, *14* (1), 5-22.

Ennis, R. H. (2015). *The nature of critical thinking: Outlines of general critical thinking dispositions and abilities*. Retrieved from http://www.criticalthinking.net/longdefinition.html (2015, July16).

Hake, R. (2015). *Analyzing change/gain score*. Retrieved from http://www. Physics.indiana.edu/hake (2015, July 15).

Hakkarainen, O., & Ahtee, M. (2010). Pupils connecting observations and explanations in successive demonstrations. *Journal of Baltic Science Education, 9* (3), 167-178.

Henderson, G. (2014). *The relationship between problem-based learning and the development of critical thinking skill in higher education* (Doctoral dissertation, Texas A & M University, Texas). Proquest LLC, UMI 3636037

Komives, S. R., & Smedick, W. (2012). Using standard to develop student learning outcomes. *Journal New Directions for Student Services*, *2012* (140), 77-88. Doi: 10.1002/ss.20033.

Lau, J. Y. F. (2011). *An introduction to critical thinking and creativity: think more, think better*. New Jersey: John Willey & Sons Inc.

Marshall, J. C., Smart, J. B., & Alston, D. M., (2017). Inquiry-based instruction: a possible solution to improving student learning of both science concepts and scientific practices. *International Journal of Science and Mathematics Education*, *15* (5), 777–796. Retrieved from https://link.springer.com/article/10.1007/s10763-016-9718-x.

Moore, B. N., & Parker, R. (2009). *Critical thinking* (9th Ed.). New York: McGraw-Hill Companies.

Moore, D. T. (2007). *Critical thinking and intelligence analysis*. Washington: Diane Publishing Co, National Defence Intelligence College.

Oon-Seng Tan (2004). *Enhancing thinking through problem-based learning approaches*. Singapore: Cengage Learning.

Oon-Seng Tan (2009). *Problem-based learning and creativity*. Singapore: Cengage Learning.

Rudinow, J., & Barry, E. V. (2008). *Invitation to critical thinking* (6th Ed.). California: Thomson Wadsworth.

Siew, N. M., & Mapeala, R. (2016). The effects of problem-based learning with thinking maps on fifth graders' science critical thinking. *Journal of Baltic Science Education*, *15* (5), 602–616.

Šorgo, A., & Kocijančič, S. (2011). Presentation of laboratory sessions for science subjects in Slovenian upper secondary schools. *Journal of Baltic Science Education*, *10* (2), 98-113.
Tarhan, L., & Acar-Sesen, B. (2013). Problem based learning in acids and bases: Learning achievements and students' beliefs. *Journal of Baltic Science Education*, *12* (5), 565-578.
Tiruneh, D. T., Cock, M. D., Weldeslassie, A. G., Elen, J., & Janssen, R. (2017). Measuring critical thinking in physics: development and validation of a critical thinking test in electricity and magnetism. *International Journal of Science and Mathematics Education*, *15* (4), 663-682.
Tiwari, A., Lai, P., So, M., & Yuen, K. (2006). A comparison of the effects of problem-based learning and lecturing on the development of students' critical thinking. *Medical Education*, *40* (6), 547-554.
Wang, J., Zhang, Y., & Shi, W. (2016). Research on the cognitive level of students' perceptions of physics models and modeling mechanism in Chinese high schools. *Journal of Baltic Science Education*, *15* (2), 204–215.
Yu, C. K., Fan, S. C., & Lin, K. Y. (2015). Enhancing students' problem-solving skills through context-based learning. *International Journal of Science and Mathematics Education*, *13* (6), 1377-1401.

Received: *June 02, 2017* Accepted: *September 25, 2017*

Mundilarto	Professor of Physics Education Department Yogyakarta State University, Colombo Street No. 1, Karangmalang, Yogyakarta 55281, Indonesia. E-mail: mundilarto@uny.ac.id
Helmiyanto Ismoyo	Student of Post Graduate Yogyakarta State University, Colombo Street No. 1, Karangmalang, Yogyakarta 55281, Indonesia.