AN EVALUATIVE STUDY ON ABILITY GROUPING-
BASED ENGLISH LANGUAGE TEACHING AT
POLITEKNIK LP3I JAKARTA USING CIPP MODEL

1Herdiana, 2Diand Mardiana

1,2Politeknik LP3I Jakarta
Jl. Kramat Raya No.7-9, RT 4/RW.2, Kramat, Kec. Senen, Kota Jakarta Pusat,
Daerah Khusus Ibukota Jakarta 10450
Email: hair.de@gmail.com1, mardianadiand@gmail.com2

ABSTRACT

The objectives of the study were to investigate evaluatively the implementation of grouping-based English Language Teaching by Politeknik LP3I Jakarta, Pasar Minggu Campus. The study was held qualitatively with the participants of students and English lecturers at the Campus. The data were gained by observations, focus group discussion, and document analysis on the framework of CIPP (Context, Input, Process, Product). The data were analyzed through data collection, data display, data reduction, and the conclusion. The results reveal that the grouping is applicable although some challenges arise, particularly in teaching lower-level groups. The lower the level the more the efforts the lecturers need to adjust. This direct to conclude that the roles of lecturers in this grouping-based English language teaching are profound.

Keywords: Evaluative, Grouping, English Language Teaching, CIPP

1. INTRODUCTION

In English language teaching, especially at English language courses, grouping or placement or levelling is a common practice. English learning centers place the incoming students to class groups according to their scores of placement test. The practice is generally intended to accommodate the learners’ needs of suitable treatments from the learning institution according to their level of language skills. The practice of placing learners properly indicates the understanding that learners of language perform well when they are grouped into classes matching their proficiency level, immersed with their peers of equivalent competence or aptitude and that particular placement test is informative for the instructors. If students work on to engage in a class group too hard for them, they are likely to get frustrated and lose their learning enthusiasm; and vice versa, higher ability learners placed in a too-easy class group may become bored, which results in negative effect on learning motivation (Green, 2018).

Another point to consider related to ability grouping is efficiency. The case happens in Uni Arab Emirate (UAE), where English is spoken as a foreign language. Howling (2017) indicates that poor English competence among the newly accepted students at college has been a serious issue, although the students have been taught English for twelve years throughout their primary and secondary levels of education. As English is the language of instruction at public universities in the country, the incoming students are required to be mastering English. Yet, the problem happens when most of the admitted students are coming to the university without sufficient English. As the consequence, the federal government has to spend a third of the 2014 federal budget on English curative activities, such as English preparatory sessions (Swan and Salem as cited in Howling, 2017).

Howling (2017) further shows that there is a need to re-examine the placement method for English classes at universities in UAE. The current placement metod applied in the country is tracking, a system by which the students are placed based on their overall exam scores and academic performance. Due to this placement system, students are in the mixed-ability class groups when it comes to English learning. This varying ability class groups lead to negative students attitude and motivation (Howling, 2017). Then, after reviewing the advantages and disadvantages of three grouping methods for English class (tracking or streaming, ability grouping and within-class grouping) and examining research conducted at a state higher learning institution in UAE, concerning English learning experience of students taking parts in the English remedial foundational program, Howling concludes that current class grouping policy harms lower-achieving students and this policy needs to be reconsidered. As the recommendation, Howling (2017) suggests to revise the current student grouping in English classes to be student grouping in accordance to the students’ English language abilities.
In the context of secondary schools in Asia, ability grouping seems to be familiar enough among emerging Asian countries. Hwang (2014) indicated an example in South Korea, where the country’s authority is supporting the implementation of it. In the country, there has been a rising trend in implementing ability grouping particularly two lessons on which ability grouping was the most demanded (English and mathematics). Hwang (2014) highlighted that referring to reports from the government, roughly 77% of secondary schools were undergoing ability grouping by 2009 (Yonhap News as cited in Hwang, 2014). Accordingly, there have been growing needs on how ability grouping works or does not work. Related to this matter, Hwang (2014) noticed that there have been significant studies on ability grouping, although the studies are mainly on the roles of ability grouping towards academic achievement.

Contextually, the practice of grouping students based on their ability in the field of education is not an unfamiliar practice. According to Hwang (2014), ability grouping in the West has become a debated topic in education. The idea of placing students in one group with similar level of ability has been challenged by both its proponents and its opponents. The proponents underline that classes carried out in homogeneous class ought to beneficially fulfill learners’ specific needs and purposefully manage the class order, and also positively support the students to learn than class group consisting of learners having varied levels of ability (Hwang, 2014). On the contrary, the opponents of the idea notice that the convincing outcomes of competence-based grouping have not been adequately signified, and even if it is so, the positive outcomes only benefit higher-ability students. They also note that it is a drawback for lower-ability students in a means of academic actualization and self-esteem (Hwang, 2014).

As highlighted above, there are significant propositions whether grouping brings positive or negative effects in the EFL classes. At a nonselective college with students having mixed English proficiency, placing students correctly is crucial, because they have different needs which can be altered through appropriate supports (Barnett and Reddy, 2017). This leads to the thinking that an English language instructor cannot use the same approach, the same attitude and the same pace of teaching for all students who have different level of competence in the same class at the same timeline. On the contrary, arguments against grouping could not be put aside. There are study results indicating that English language teachers should reconsider grouping. The worry that adult students have their own rights to determine their decisions, even if they made failure; the reason that level placement or ability grouping only shows racial and economic background rather than academic capacities (Barnett and Reddy, 2017). Additionally, there is an indication that grouping may create unhealthy comparisons among learners that potentially harm the learning process. The proposition that ability grouping could affect academic self-concept as the students make comparison of their English competency (Sheppard et al., 2018) is among ideas that challenge grouping.

In Indonesia, one of the education institutions that have set and developed English Language Teaching (ELT) based on grouping is Politeknik LP3I Jakarta (Lembaga Pengembangan Pendidikan Profesi [Center for the Development of Professional Education]) Jakarta. Politeknik LP3I Jakarta (PLJ) is a Jakarta-based vocational higher learning institution offering Diploma Three (D3) programs. As English language skill has been an essential demand in the industry, it is a core competence to be possessed by PLJ students and graduates. In their Guideline Book for the implementation of ELT at LP3I, the vision of their ELT and learning program is to yield graduates who are able to communicate in English both verbally and in written forms according to the demand of workforce (Handbook for the Implementation of English Teaching and Learning at Politeknik LP3I, Politeknik LP3I Jakarta, 2015).

As the input of their education process, PLJ admits new students every year. The students come from different backgrounds of education and social status. As a private education institution that admits students without any standardized enrollment tests, the polytechnic could be categorized as a non-selective college among the stake holders, particularly senior high school and vocational high school students. Referring to that, as a nonselective college, their new students are typically heterogeneous especially in terms of academic qualification standards, included English language proficiency. Thus, to accommodate the heterogeneity, the institution carries out the English language teaching and learning process based on ability grouping for their first-year students.

In reference to the institution’s Guideline Book for the implementation of English language learning and teaching process:

**Effective language learning is one of which determined by the application of grouping in the learning system, namely classifications based on the level of competence. Classes with the same level of competency (among its students) can run more effectively than classes with heterogeneous abilities. This is because the teaching approach method carried out by the lecturer becomes more focused, the material provided can be more appropriate to the students’ abilities, interaction and communication that is built between students is more interwoven because they are in the same level. Therefore, leveling/grading (grouping) in English classes is important to do. Leveling (grouping) policy must be applied in all LP3I branches.**
In fact, since its first implementation, the Polytechnic has never carried out any study yet to investigate how and to what extent the ability grouping has been going and whether it brings positive or negative attainments and whether it is effective or ineffective on the ELT implementation. In connection to that, as an English lecturer teaching at the institution, the author has a sense of urgency to do an evaluative study on the program. This evaluation is significant for the improvement of ELT quality at the institution, as it can inform the stake holders (PLJ directors, academic staff, academic advisors, fellow lecturers, students and parents) about how and to what extent the policy is carried out and whether or not it affects the outcomes of the English language teaching and learning at the institution.

Concerning evaluation mentioned above, there are types of evaluation framework available in the field of educational research. From several existing evaluation models, the author applies Context, Input, Process and Product (CIPP) model developed by Daniel Stufflebeam (2007). The CIPP evaluation model has been widely used by evaluators, as the author considers this evaluation model applicable in comparison with other evaluation frameworks, particularly in the field of education. Besides, the model was originally devised for the evaluating Elementary and Secondary Education Act (ESEA) in the USA. CIPP model stretches evaluative areas of a program or policy in the aspects of context, input, process and product. These four evaluation aspects equip a decision-maker to answer four basic questions: What should we do? How should we do it? Are we doing it as planned? And, Did the programme work? The four words mentioned in the abbreviation of CIPP are the targets of evaluation, which are nothing but components of the process of an activity program. In other words, the CIPP model is an evaluation model that views programs that are evaluated as a system.

The main framework of CIPP model evaluation provides larger practical outlines in evaluating certain program or policy, as it emphasizes the comprehensive assessment in four areas (Context, Input, Process, Product) of the program or policy. Therefore, the author considers the model applicable to devise as the framework in evaluating grouping based English language teaching at PLJ. The grouping based ELT at the institution is evaluated through the four aspects of the CIPP evaluation model. Context evaluation takes the plans and the needs assessment: what the institution is supposed to achieve the planned targets and make sure whether or not the important needs were addressed. (Stufflebeam, 2007). The Input evaluation considers policy planning at the institution and spots comprehensively on whether or not the efforts by the institution or stakeholders team well guided (Stufflebeam, 2007). Process evaluation deals with the implementation of the the planned procedures and policy, looking at whether or not the implementation was designed and executed as stipulated by the institution (Stufflebeam, 2007). Product evaluation signifies the outcomes and impacts that the policy and program bring out. (Stufflebeam, 2007).

Based on the background presented above, the focus of this study is the implementation of ability grouping based ELT and learning at Politeknik LP3I Jakarta (PLJ) Pasar Minggu Campus, with the subfocus specified into: PLJ’s ELT and learning expected outcomes, teaching and learning methodology, teaching and learning results.

Refering to the background and the research framework highlighted above, the author intends to seek further understanding about ability grouping applied by PLJ Pasar Minggu Campus in ELT, with the specific questions as follow:

1. What are the expected outcomes of ability grouping-based ELT and learning at PLJ Pasar Minggu Campus?
2. How does PLJ Pasar Minggu Campus implement the ability grouping-based ELT and learning?
3. Is PLJ Pasar Minggu Campus doing the ability grouping-based ELT and learning as planned?
4. Did the ability grouping-based ELT and learning at PLJ Pasar Minggu Campus work as expected?

2. LITERATURE REVIEW

English Language Teaching (ELT)

“To teach is to learn twice.” (Joubert, [1754–1824]), these words catch the author’s attention when it comes to a discourse about teaching and learning. What the author could draw from this wise saying is that teaching process cannot ever be separated from learning process. In order to be a good teacher, one must be a good learner. Therefore, in the context of English teaching, somebody cannot be a good English teacher if he or she does not learn English well. A teacher is always supposed to learn. Learning is the key to teaching. The advancement of teachers in the subject they teach is contextually associated with better learning outcomes (Schleicher, 2016).

Accordingly, in regard to the importance of ELT and English language teachers, there are three aspects related to the topics that the author takes into consideration. The three aspects are the concept of
ELT, methodologies in ELT and assessments in ELT. These three aspects also contribute to one another in achieving the teaching and learning outcomes. The followings are the brief reviews of the three aspects:

As English has been a growing need globally and teaching is not merely transferring knowledge and skills, teaching English has been a significant concern in language education. Currently, what we call as English teaching could be depicted as TESOL (Teaching English as Other Language), TEFL (Teaching English as a Foreign Language), and TESL (Teaching English as a Second Language) (Lin, 2010). In general, the types of teaching English are the same. But what make them different in particular are their emphases, the backgrounds of the learners and also the area where English is spoken, namely English-speaking country and non-English-speaking country.

Typically, ESL is used in immigrants countries such as USA and Canada, in which many immigrants come to settle. Although English is not their native language, they still need to learn the language for the purpose of getting immersed with their local communities and get themselves integrated in the immigrants region (Lin, 2010). In the context of Asia, we can take Malaysia for an example, as English is one of their official languages and that they have historical links to British occupation, English is treated as their second language and the term for teaching it there is ESL. As for the TESOL, linguists usually use the term for giving emphasize to the methodologies of learning and teaching English as a profession. Lin (2010) suggests that we can most particularly perceive that TESOL is devised institutions for training teachers and programs that give education or instruction to English teachers, or certain communities that language learnig for learners of widening network. Meanwhile, TEFL is primarily focused on the use of English by learners and teachers whose mother tongue is not English. The emphasis of TEFL is that the learners and teachers position English as communication instrument in the context of diverse nations, instead of making use of it as an official language (Lin, 2010). Therefore, at this point we can see that context—the background of the learners and teachers and also the environment where English is used—more or less intersects with the terms or types of English teaching.

Still, despite the rapid development of ELT and learning and also the context of teaching EFL rather than ESL in Indonesia, the goals of English language teaching are relatively the same. According to Cook (2007), the goals could be classified to two, namely internal and external. Internal goals intersect with the learners’ self improvement as personalities, related to their being as parts of societies that understand mindfully the effects the second language possesses on them. While external goals deal with learners’ making use of language functions off the classroom: business or working activities, reading, traveling, even surviving in a new world or society. What is perceived as teaching language in traditional contexts oftentimes emphasizes the internal goals, for instance, studying Latin trains the brain; learning L2 literary works heightens up awareness of cultural and socioeconomical being, and so on and so forth.

In the implementation of ELT, at least there are four central areas to be concerned, namely management and interaction of classroom, methodology-method-post method, learners and the frameworks of social and institutional context of ELT (Hall, 2017). Regarding this proposition, in fact, from time to time, the ways English teachers teach and the learners learn have developed a lot. Nowadays, teaching and learning English is increasingly supported by the existence of technological innovations. Students do not just rely solely on classes but they can learn and even acquire English (through reading, listening, speaking and writing) anywhere and anytime using the gadgets available around them, such as smartphone and portable computer. The rapid development of teaching and learning contextually needs to be concerned well, especially those who teach at the junior and senior high school levels. The role of the teacher is no longer the only source of knowledge but rather directs the students to learn interactively in various media that are familiar to them.

Furthermore, it is also important to notice that the development of methodologies in ELT is a continuous process since the past, in line with human being cognitive advancement. Periodically, Lin (2010) divides the development of language teaching into three schools, known as three schools of language acquisitions. The three schools are the periods of 19th century, 1940s, to 1950s (known as First School: behaviorism and structuralism; periods of 1960s to 1970s (known as Second School: rationalism and cognitive psychology; and Third School, namely constructivism (from 1980s to 2000).

As digital technologies are growing rapidly nowadays, the methodologies of teaching and learning English keep up with the development. For instance, social media, that can influence its users in acquiring a foreign language. Social media also allows its users to interact with the speakers of the target language (English) in any topics of talk.

**Ability Grouping-Based English Language Teaching**

Ability grouping is not categorized a new practice in language teaching. Historically, placement at learning institutions has existed since early 20th Century. At that time, the policies of placement were accompanied by course pre-requrement, academic probation and progression policies, and other requirements related to entrance and graduation (Barnett and Reddy, 2017). However, later, in 1970s, the
popularity of such policies was decreasing as it got resistance from stakeholders arguing that as adults, college students should have the right to make their own decisions, even if it resulted in failing courses. Still, despite challenging debates on it, ability grouping in English language teaching is a significant issue to engage.

In English language teaching, especially at English language courses, grouping or placement or levelling is a common practice. English learning centers place the incoming students to class groups according their scores of placement test. The practice is generally intended to accommodate the learners’ needs of suitable treatments from the learning institution according to their level of language skills. The use of placing learners properly indicates the understanding that learners of language perform well when they are grouped into classes matching their proficiency level, immersed with their peers of equivalent competence or aptitude and that particular placement test is informative for the instructors. If students work on to engage in a class group too hard for them, they likely tend to get frustrated and lose their learning enthusiasm; and vice versa, and vice versa, higher ability learners placed in a too-easy lass group may become bored, which results in negative effect on learning motivation (Green, 2018). In addition, if not informed well about varied learner capabilities, teachers may find it difficult to plan and adjust the instruction.

In the practice of student grouping, there are three types, namely tracking, ability grouping or setting and homogenous within-class grouping. Tracking is placing students in class groups based on their scores of examination or any academic performance, such as National Examination (UN) in Indonesia or Common Educational Proficiency Assessment (CEPA) in UAE (Uni Arab Emirates). Ability grouping or setting is placing students based on their ability level (usually focusing on course-specific proficiency; so that students are in class groups in which they have similar or equivalent level of proficiency. Homogenous within-class grouping divides students in a class who have varying levels of competence into smaller class groups of students having equivalent or resemblant levels of ability (Lou et al as cited in Howling, 2017).

As a method in a field of language education, all types of grouping are certainly intended to equalize or accelerate students’ language skills. Still, it does not mean that each grouping type is without advantages and disadvantages. Howling (2017) points out those advantages and disadvantages as follow:

<table>
<thead>
<tr>
<th>Grouping Method</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tracking</td>
<td>● Simple to implement</td>
<td>● Tendency of benefiting students of advance ability level (Suknandan &amp; Lee, Gamoran, Callahan as cited in Howling, 2017).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Can give drawback mental effect on lower ability learners (Suknandan &amp; Lee; Ireson et al; Maclntyre &amp; Ireson as cited in Howling, 2017).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● There is a possibility of students being misplaced certain lesson-specific classes which are not in accordance to their level (Howling, 2017).</td>
</tr>
<tr>
<td>Ability Grouping/Setting</td>
<td>● Decrease the negative psychological impact of placement on students (Slavin as cited in Howling, 2017)</td>
<td>● Success depends on every school’s resources (Howling, 2017).</td>
</tr>
<tr>
<td></td>
<td>● Learners tend to be put in a class suitable with their level of proficiency (Gamoran as cited in Howling, 2017).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Possess a beneficial impact on the performance levels of learners of high medium and low levels of ability (Slavin as cited in Howling, 2017).</td>
<td></td>
</tr>
<tr>
<td>Homogenous Within-Class Grouping</td>
<td>● Tends to be the most beneficial for students of all levels (Suknandan &amp; Lee as cited in Howling, 2017)</td>
<td>● Class size establishes the efficacy of the groupings (Lou et al; Suknandan and Lee as cited in Howling, 2017).</td>
</tr>
</tbody>
</table>
|                          |                                                                          | ● Gives extra force on instructors, demands distinctive commands adjusted to every group needs (Lou et
In the context of ELT in Indonesia, ability grouping is not something new, especially at well-established English courses, such as English First, Wall Street English, LIA (Lembaga Indonesia Amerika), The British Institute, and the likes. While in higher learning ELT, ability grouping is not a common practice. Higher learning institutions in Indonesia generally set English language as a subject for pragmatical reason, namely to fulfill the requirement of curriculum credits. English instructors teach students directly according to their class of study programs, not according their level of ability or competency. Still, there are certain higher learning institutions that take into account ability grouping in their ELT, particularly vocational college like Politeknik LP3I Jakarta. Still, in spite of its well-known applications at schools, colleges or higher education institutions, the applicability and effectuality of ability-based grouping has not been scrutinized sufficiently (Sheppard et al., 2018).

Regarding ability grouping based ELT at Politeknik LP3I Jakarta Pasar Minggu Campus, in this thesis, the author does an evaluative study to reveal how and to what extent ability grouping based ELT has been implementing at the institution. In conducting this evaluative study, the author applies Context, Input, Process and Product evaluation model (CIPP). Hence, the following is a highlight about evaluation study and CIPP evaluation model.

Evaluation and CIPP Evaluation Model

In order for an organization or certain institution to achieve their goals, a particular program is set. For instance, a school applies an acceleration program to accommodate their bright students to graduate based on their pace of learning that is above the average of their peers. Then, after certain period of time, the school needs to find out whether or not the program affects the quality of their education process. At this point, a particular evaluative study seems to make sense to be carried out. Therefore, it is relevant to emphasize that evaluation means determining if a thing is worthy or not.

Viewed from the implementation timeline, evaluation can be formative and summative. Fedorchuk and Mykola (2016) suggest that formative evaluation is intended to view ongoing students’ learning to give information for the related stakeholders. Summative evaluation is supposed to view students’ learning performance at the end of a certain learning period by referring it to the set standard. Pragmatically, both formative and summative lead to the same objectives, to check the learning outcomes qualitatively and quantitatively pointing out to the expected standard to ensure the quality improvement.

The CIPP model is an evaluation framework for guiding project, program, product, institution, system and personnel evaluating efforts (Stufflebeam, 2007). The evaluation model can be used as a tool for sound and practical evaluations and making sure that the findings provide feedbacks for the accountability and improvement of a program. The overall CIPP Evaluation framework is addressed to assist evaluators and decision makers fulfill the expected standards. The following picture illustrates how CIPP Model maps out the study (Stufflebeam, 2014):

![The Flowchart of CIPP Model Evaluation](image)
The flowchart above illustrates how CIPP evaluation model is stretched out in the implementation. The implementation is in line with the formative and summative time scopes. The process of the evaluation starts with the context where the program is set to be implemented by dealing with the establishment of needs and objectives. Then, it goes on to evaluate the input by determining the most applicable strategies to reach recognized necessities. After that, the evaluation continues to dig in the process by examining the program implementation. Finally, it comes to the point of product evaluation by program outcomes assessment.

In connection to ability grouping based ELT and CIPP evaluation model, there are studies conducted previously the author considers relevant to the topic. These previous relevant studies add up some more insights and perspectives for the author in composing this thesis. Therefore, next sub-chapter deals with previous relevant studies and how these studies intersect with this research of the author’s.

**Previous Relevant Studies**

There are studies related to ability grouping in ELT (English Language Teaching). A study conducted in Japan (Sheppard et al., 2018) investigated the possibility of ability-based grouping being impactful to university students who had followed obligatory ESP (English for Specific Purposes) subjects. The findings show that ability grouping was beneficial for students of lower proficiency, but, by contrast, it was not impactful enough for higher level students. Still, the study also provides its limitations by stating that features of instruction and institution of the study program the learners were following and the possible impacts of ability-based grouping on learners’ intellectual concept of self are possible reasons for the findings. The author considers this research relevant in informing how ability grouping is evaluated in a tertiary education institution.

A study conducted in the USA also appears to be a case in ability grouping. This study (Gonzalves, 2017) dealt with the standardized language assessment typically taken by adult English as a Second Language (ESL) students after their enrollment at California’s adult schools. The study was conducted by participatory action research involving interviews. The findings of this study suggest that there are discrepancies between the state’s policies and actual pedagogical needs of the learners. Beside the findings, it also proposes ways to reconstruct how ESL assessment is carried out, such as making available a wider, more comprehensive base of assessments for schools to deploy, and offering an updated, common set of standards for use statewide.

Howling (2017) carried out a study entitled on the need of ability-based grouping for English classes at state secondary education institutions in UAE. The background of the study is the poor English performance of secondary school graduates in the United Arab Emirates. This has been an issue in the country as English is the instructive language of learning and teaching at federal universities. Poor English performance among school graduates who are just accepted at federal universities has made UAE government spend one-third of their university budget on English remedial courses.

Howling’s (2017) study weighed in research done in 2014-2015 on female students’ English learning experience in the foundational program at state tertiary education institution in UAE. The study suggests that there is a need to re-examine the placement method for English classes at universities in the country. The current placement method applied there is tracking, a system by which the students are placed based on their overall exam scores and academic performance. Due to this placement system, students are in the mixed-ability class groups when it comes to English learning. This varying ability class groups lead to negative students attitude and motivation (Howling, 2017). After reviewing the advantages and disadvantages of three grouping methods for English class (tracking or streaming, ability grouping and within-class grouping) and examining research conducted at a state higher learning institution in UAE, concerning English learning experience of students taking parts in the English remedial foundational program. Howling concludes that current class grouping policy harms lower-achieving students and this policy needs to be reconsidered. As the recommendation, Howling (2017) suggests to revise the current student grouping in English classes to be student grouping in accordance to the students’ English language abilities.

A CIPP approach was also used for evaluating grammar teaching programs at schools in Iran. Jafari and Shahrokhi (2016) applied CIPP model for evaluating the effectiveness of teaching second language grammar from the perspectives of students and instructors. The participants of the study were one hundred twenty students attending the second grade high-schools and 10 instructors teaching in the program. The data was gathered through a self-reported student questionnaire. An interview which was designed for the instructors was also used. While the data based on the questionnaire were analyzed through descriptive and inferential statistics, content analysis was carried out to analyze qualitative data. Results of the study indicated that the grammar program at a highschool served for its purpose. The findings revealed that
improvements in the objectives, teaching methods, and grammar curriculum are required to make the grammar program more effective.

Another similar study was carried out in Malaysia. Abdullah, Wahab, Noh, Abdullah and Ahmad (2016) co-wrote a study entitled The evaluation and effectiveness of school-based assessment among science teachers in Malaysia using CIPP Model. The purpose of this study was to identify the effectiveness of School Based Assessment (SBA) in science subject among secondary science teachers using CIPP evaluation model. The study also investigated the evaluation of science teachers’ regarding the advantages and disadvantages of SBA. A total of 112 science teachers from 21 secondary schools in Batang Padang District, Perak, Malaysia were involved in this survey. The data from the questionnaires and survey were analyzed using descriptive statistics and statistical inference. Frequency distribution with mean scores and percentages and Pearson correlation analysis were used. Results showed that the dimensions of context, input and product are at a moderate level while the level of process dimension is high. Hypothesis is also accepted, that there is no significant relationship between the level of dimensional context, input, process and product in the implementation of SBA with teaching experience. Analysis of the interviews showed that most respondents agree that there are advantages of SBA, thus this is suitable with the aim of Malaysia to create a world-class human capital. The author summarized that the research provided new insights to language instructors for the execution of SBA properly by support from the government, monitoring, courses and training.

In Indonesia, Firharmawan (2017) conducted a CIPP model evaluation study entitled CIPP-Based Evaluation on English for Sport Science at Sport Education Study Program of the University of Ma’Arif Nahdlatul Ulama (Umnu) Kebumen. The participants of the research were the students, the English teacher, and chief of the program at the university. The data was gained by in depth interviews with the research participants, analysis of the existing documents, and observation. The data were then analyzed through four steps by using interactive model as proposed by Miles and Huberman, (2014): collecting data, reducing data, displaying data and verifying. The technique used in data validation is by using source triangulation. The study pointed out that the program was ineffective. There were four main factors to cause the ineffectiveness: first, the teaching context was less supportive to good teaching practices; second, the inputs for the program were lack of quality, signaled by unprofessional teachers; third, the process of teaching and learning did not meet the stakeholders’ expectation; and the last, the product showed the students’ less competency on good communication skills required.

Another study using CIPP framework conducted in Indonesia is the one by Nyoman and Darma (2019) entitled Implementing the Context, Input, Process, Product (CIPP) Evaluation Model to Measure the Effectiveness of the Implementation of Teaching at Politeknik Negeri Bali (PNB). The study intended to find out the level of effectiveness of teaching implementation at PNB by viewing it from CIPP Components. The author collected data through questionnaire, guided interview, and related documents. The subjects of the study were students, lecturers, and the management staff of PNB. The collected data were the analyzed qualitatively and descriptively. To determine the effectiveness of the implementation of the teaching programs, the authors transformed the raw scores into the Z-scores and the T-scores, which were then verified into Glickman’s quadrant prototypes. The analysis results indicated that the effectiveness of the implementation of teaching program at PNB seen from the relatedness of context, input, process and product falls into moderately effective category. Beside the results, the authors also highlight the obstacles in the implementation of teaching programs at PNB, which are those in the components of context, input, process and product. The authors notice that the significant obstacle appears in the product variable.

In regard to the literature reviews with previous studies discussed above, the author could draw a lining on ability grouping based English language teaching and CIPP evaluation framework. Ability grouping in English language teaching at education institutions, especially at higher learning level, has been a significant topic to discuss, as many practitioners propose it and also not a few who are against it. Meanwhile, CIPP evaluation model has also played essential roles in educational evaluative research. The author views that ability grouping based ELT can be evaluated using CIPP evaluation design, as its CIPP evaluating components resemble the actualization of education process at an education institution: setting up the curriculum and the expected outcomes, admitting students, teaching and educating students and yielding competent graduates as the products. Thus, by taking into account the significance and also the pro and con of ability grouping based English teaching as well as the roles that CIPP evaluation model gives to the educational research, in this thesis the author applies the evaluation model in evaluating ability grouping based English language teaching.

3. RESEARCH METHOD

Research Design
The design of this research is evaluation method, with CIPP evaluation model as the framework. As there is a descriptive quantitative data analysis in the Product component of the CIPP applied in this thesis,
the author projects it as a mixed method. The process of education (ELT) at Politeknik LP3I Jakarta (PLJ) Pasar Minggu Campus contextually follows the cycle of yielding a product or impact (in this case: graduates who are able to communicate in English both verbally and in written forms according to the demand of workforce); therefore, the nature of the process is setting up the curriculum, receiving the input, processing the input and producing the outcomes.

In applying a certain method for research, not only does a researcher need to understand the approach, but also, as Creswell (2017) points out, need to take three framework elements into consideration: philosophical contentions on what generates knowledge projection, overall constructions of a study (inquiry strategies) and specified step by step of gathering data, analyzing, and composing (methods). Thus, in carrying out this research, the author takes these framework elements into consideration.

Plenty of evaluation designs and models have been used for evaluating the projects, programs or works of institutions. For the author, CIPP model is an evaluative framework that is relevant for evaluating a policy or program. The framework consists of four evaluative component: Context (C), Input (I), Process (P) and Product (P) (Stufflebeam, 2007).

Description of Setting and Participants

This research evaluates the implementation ability grouping based ELT at PLJ Pasar Minggu Campus of Academic Year 2017-2018 and 2018-2019, with specific description of setting and participants as follow:
1. Classroom teaching and learning activities in six ability-based class groups of Academic Year 2017-2018 and 2018-2019;
2. Students of six ability grouping-based class groups of Academic Year 2017-2018 and 2018-2019 (total participants: approximately 180 students);
3. Lecturers and academic staff as the stake holders in English learning and teaching activities; and
4. Related documents resulted from teaching-learning activities (scoring transcripts of summative tests [of General English 1 and General English 2 Courses] and TOEIC tests) and the documents of rationale and plans of ability grouping based ELT at the institution.

Research Instruments

Instruments of research are tools built up by researchers for reaching the projected purposes when conducting research. Practically, the instruments are designed devices assist in collecting data to be analyzed. Miles and Huberman (2014) describes research instrumentation as specific means which is focused to gather qualitative or quantitative information and it can be closely or openly structured. In this study, the author applies the instruments of Focus Group Discussion (FGD) guide, observation sheet and records from the institution.

1. Focus Group Discussion (FGD) Guide
The author designs this FGD guide based on the the framework of CIPP evaluation model projected to the implementation of ability grouping based ELT at PLJ. This guide is focused on directing the author to collect the data to be analysed.

Table 2
Focus Group Discussion Guide

<table>
<thead>
<tr>
<th>No.</th>
<th>CIPP Component</th>
<th>Points to Consider</th>
<th>Question to Answer</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Context</td>
<td>- Rational grounding for determining the targets</td>
<td>*Curriculum guidelines?</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Relevant environment</td>
<td>*How is the teaching and learning environment?</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Expected and real conditions of environment</td>
<td>*What is the ideal condition for ELT and learning at PLJ?</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Unfulfilled needs</td>
<td>*What is still left to be implemented?</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Missed opportunities</td>
<td>*What still can be done?</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Input</td>
<td>- how the resources are used</td>
<td>*How are the resources used by the lecturers and students?</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- cost/material effectiveness to meet objective and achieve goals</td>
<td>*What material and how much cost needed?</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- alternative strategies and procedures</td>
<td>*What strategies and procedures used as alternatives?</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- linkage between the means &amp; the desired ends</td>
<td>*How are the means and the targets linked?</td>
<td></td>
</tr>
</tbody>
</table>
3. Observation Sheet
In this instrument, the author designs a form to note any significant information in the process of ability grouping-based ELT at the institution based on the framework of CIPP evaluation model which focuses on the Process component: detect/predict defects in procedures or in implementation, gather information for programming decisions and record of procedures as it occurs.

3. Records
The plans or rationale and the desired outcomes of the grouping-based ELT of PLJ are available in the documents at the institution. In addition, the development or progress of learning in the phases of pre grouping, during grouping and post grouping is recorded in forms of test scores’ documents. These documents inform the author about the background and targets, as well as the impact or product of ability grouping based ELT implemented by the institution.

All of the instruments highlighted above are projected to gather comprehensive information and data according to the CIPP framework.

Data Collection
The sources of data in this study are the main stakeholders directly and indirectly involved in the implementation of ability grouping based ELT at PLJ (lecturers and academic staff, that can provide information, experience, insights and data) and the documents available at the institution related to the implementation of ability grouping: the guidance book as standard operating procedure and records of students’ test scores as the indicator of the attainments that have been achieved so far. The data sources are in line with the framework of CIPP. Stufflebeam (2007) indicates that the evaluator should obtain information both quantitatively and qualitatively.

Creswell (2017) states that data collection process goes like a “circle” of interrelated activities which cover up, but go beyond gathering data. In this case, the interrelated circle is the circle of CIPP framework. Each of these framework components explores how they are different from each of the approaches of inquiry, and, provides a data collection table. More specifically, in this study, the author collects the data through the research instruments with the process as follows:

1. Focus Group Discussion extracts information, experience, insights and data in general regarding the ability grouping based ELT at PLJ (See Table 3.1 for the points and questions to be concerned in the FGD). In this FGD, the author is dynamically immersed with the main stakeholders (lecturers and academic staff) to extract any significant information, experience, insights and data directly or indirectly related to the implementation of ability grouping based ELT.
2. Non-participant observation and participant observation add any significant information in the process of ability grouping-based ELT at PLJ. This could potentially strengthen what have been found from the FGD.
3. Records related to the implementation of ability grouping based ELT at PLJ are available at the institution. Through these records or documents, the author reveals the rationale and plans for implementing the ability grouping based ELT. Besides, the author could find out the impact or product of ability grouping based ELT implemented by the institution in forms of test scores as the standard indicators that have been gained so far.

Data Analysis
Following Miles and Huberman (2014) the data analysis method in this research underlines the interactive model. The model indicates that analysis flows concurrently in three activities: reduction of the data, displaying the data and synthesizing the data for a conclusion or for a verification. After the mass of data were collected, they were then reduced in order to be meaningful. Miles and Huberman (2014) cite it reduction of data. Reduction of data constitutes selecting, focusing, simplifying, abstracting and transforming the data in forms of written notes or transcriptional records.

<table>
<thead>
<tr>
<th>3.</th>
<th>Process</th>
</tr>
</thead>
</table>
| - detect/predict defects in procedures or in implementation 
- information for programming decisions 
- record of procedures as it occurs | *Is there anything wrong or confusing? 
*Any information to take into consideration? |
| 4. | Product | 
| - attainments to measure 
- attainments interpretation 
- evaluation as often as necessary during program | *What attainments are available? 
*How and do the attainments indicate? 
*How often is evaluation necessary? |
In this study framework, the author sets CIPP model for allowing formative and summative questions on all aspects of the ability grouping-based English language teaching and learning at the institution. The following is the framework of CIPP evaluation model that the author deploys in evaluating the program:

1. **Context Evaluation**
   What the author takes into consideration in this context evaluation of ability grouping based ELT at PLJ, specifically Pasar Minggu Campus, are the plans and the needs assessment: what the institution is supposed to achieve the planned targets and make sure whether or not the important needs were addressed (Stufflebeam, 2007). Therefore, to elaborate those considerations, the author presents the rationale of the institution in carrying out the ability grouping for the determination of the objectives, gives significant insights on the relevant environment, explains the expected and actual conditions of the environment, tries to spot the unfulfilled needs and finds out if there are any opportunities left unused.

2. **Input Evaluation**
   In this aspect of evaluation, the author considers policy planning of the ability grouping based ELT at the institution and spots comprehensively on whether or not the efforts by the institution or stakeholding team well guided. By comprehending this consideration, the author determines how and to what extent the institution uses resources; examines resource-effectiveness in meeting the expected outcomes and achieving the targets; weigh in any alternative strategies and procedures that the institution puts into consideration; and makes sure the significance between the means taken and the objectives aspired to achieve. (Stufflebeam, 2007).

3. **Process Evaluation**
   The key points of process component of CIPP evaluation is implementation of the planned procedures and policy, looking at whether or not the implementation was designed and executed as stipulated by the institution (Stufflebeam, 2007). And in doing so, the author checks or forecasts the defect in program and policy design or its implementation. Furthermore, the author provides information for programming decisions and maintains record of the procedure or steps as it occurs.

4. **Product/ Impact Evaluation**
   This aspect of CIPP evaluation model signifies the outcomes and impacts that the policy and program bring out. In other words, this aspect tries to figure out whether or not to what extent the policy and program worked and whether or not the efforts carried out by the stakeholders succeeded (Stufflebeam, 2007). Refering to those significances, the author digs in to measure the attainments that the institution achieved from the ability grouping based ELT, then interpret the attainments, so that necessary information could be extracted for the main stakeholders of the institution. In order for the author to gain more understanding of the impact of ability grouping based ELT at the institution has been going, a quantitative measurement is going to be added in the part of Impact or Product Evaluation in this thesis.

As for the phases of analyzing the data, it goes as follow:
   a. Description phase. In this step, the author describes the information seen from the aspects of CIPP refering to the data collection already acquired.
   b. Reduction phase. At this point, the author reduces the unnecessary data.
   c. In-depth analysis phase. At this stage, the author analyses results of the investigation, by projecting it to the theoretical framework and literature review of the evaluative study concerning ability grouping-based ELT at PLJ.
   d. Conclusion and recommendation phase. At this phase, the researcher concludes the analysis and recommends the stakeholders in accordance to the significances of the research which is carried out under the actual circumstances of the research object.

4. **RESEARCH FINDINGS AND DISCUSSION**

In this chapter, the author presents research findings and discussions. The research findings are classified into four components of CIPP evaluation model, namely Context evaluation, Input evaluation, Process evaluation and Product evaluation. Later, significant and potentially additional points from these research findings are discussed in the Discussions section.
Findings

1. Context

As the first point of context evaluation, the author highlights the rationale or the plans of Politeknik LP3I Jakarta (PLJ) in implementing ability grouping based ELT. As a vocational college, which is oriented to fulfil the needs of competent manpowers in industry, PLJ sets its English curriculum and programs specifically to meet the demands of industry. Therefore, the implementation of English language teaching at PLJ refers to its vision, namely to yield graduates who are capable of speaking English verbally and in written according to the needs of the workforce (LP3I, 2015). Meanwhile, to achieve the vision, the institution sets specific goals as follow:

1. Yield graduates who are able to communicate in offices, businesses and every day public conversations in English well;
2. Yield graduates who are able to do job interviews and presentations in English properly;
3. Yield graduates who have expertise in writing English correspondence or e-mail properly;
4. Yield graduates with a minimum TOEIC (Test of English for International Communication) score of 450.

As for the curriculum streamline, PLJ distributes the subjects of the English language teaching according to a credit system in four (4) semesters out of six (6) semesters of Diploma Three (D3) study. The distribution is as follows:

<table>
<thead>
<tr>
<th>Semester</th>
<th>Subject</th>
<th>SKS (Credits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>General English 1</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>General English 2</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>Business English 1</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>Business English 2 (ESP)</td>
<td>2</td>
</tr>
</tbody>
</table>

From the table above, the total volume of English language teaching is fourteen (14) credits, four in the first semester, four in the second semester, four in the third semester and two in the fourth semester. In the first and second semesters, students learn General English 1 (GE 1) and General English 2 (GE 2); while in the third and fourth semester, they are taught Business English 1 (BE 1) and Business English 2 (BE 2 or more specifically called English for Specific Purpose [ESP]). The implementation of GE 1 and GE 2 is the main focus of this study as the teaching for the two courses are based on ability grouping, which is the topic of this thesis.

Beside the curriculum streamline above, the author also finds the timelines of projected targets to be achieved by English language team at the institution per semester (from the first until the fourth semester). The timeline is as follows:
As the measurement of its teaching and learning development, there is a standard competence test applied by the institution. Brown (as cited in Lin, 2010) indicates that a test is a means to measure a learner’s learning development, knowledge and skills in a certain field of study. By looking at the test results, the main stakeholders or decision makers can recognize what appropriate teaching methods and teaching and learning policies are. For this purpose, PLJ selects TOEIC for the standard of English competence to be achieved by the students. TOEIC is used by the institution as the standard competence because it is a type of proficiency test that is typically intended for public or commercial or industrial English settings, which is similar to the orientation of PLJ in graduating the students.

The target of TOEIC score of 450 is in general and this target was made when TOEIC score had not been stipulated as a requirement for completing study at the institution. But, starting from 2017 academic year, TOEIC score has been stated as the compulsory requirement for finishing the study, in which the minimum score that has to be achieved is 350 for students of 2017 academic year and 450 for those of 2018 academic year.

On table 4.2, there is also a target of vocabulary mastery, namely 2500 to 4000 starting from the first until the fourth semester. Due to huge number of English lexical units, particularly the ones of English-Latin combination, it is understandable why English learners commonly need to spend fairly long time acquiring English (Lin, 2010). For the purpose of accumulating adequate words and build up vocabulary proficiency,
English language scholar have suggested the pedagogies of teaching and learning in connection to vocabulary mastery. In the case of PLJ, students are scheduled to do exercises on a vocabulary mastery system installed in a computer. Then, the results of their work are reported for evaluation, counted as a component of students’ final score for the subjects of GE 1 and 2 and BE 1 and 2.

While for the target of students communication skills, it is in line with the target of TOEIC score (300, 350, 400, 450). Here, 450 TOEIC score is in average considered as intermediate level of proficiency. Accordingly, the target of the amount of students who achieve intermediate level is 25 % in the first semester, 50 % in the second, 75 % in the third and 90 % in the fourth. In the academic system of the institution, the attainments of TOEIC is part of the Key Point Indicator (KPI), which is audited periodically by the management of the institution.

Certain significant insights on the relevant environment in which ability grouping based ELT is carried out are presented in the following. This starts with the beginning of the ELT at the institution in the first semester. Before the teaching and learning starts in the first semester, the students are supposed to do TOEIC test. Their TOEIC scores are then used to group them into classes in which they have equivalent scores or the same level of English competency. For the purpose of effectiveness in teaching, especially speaking parts, every class group consists of not more than sixteen (16) students. More details about this grouping process findings are elaborated in the next sub-chapter (Input).

In the academic year of 2018-2019 and 2017-2018, there are six (6) classes of grouping students. These classes are taught GE 1 and GE 2 in the first and second semester by four English lecturers according the schedules allocated (4 credits in one semester). The lecturers teach them using students talking time-oriented approach, based on the reference book entitled Speaking Practice for General English 1 and Speaking Practice for General English 2. This type of approach appears to be the combination of student-centered and interactive approach, as Lin (2010) points out that in student-centered learning, students are the centre of any learning interaction in the class and the teacher acts as a facilitator; while interactive approach conditions the learners for an active class atmosphere by getting them to participate with peers. Although the approach emphasizes student talking time rather than teacher talking time, the three other skills of English language (listening, writing and reading) are still covered proportionally.

In the first and second semester, students are taught foundational English language lessons through the subjects of GE 1 and GE 2. In these first two semesters, PLJ students’ English proficiency is projected in the level of Pre-Intermediate or B1-B2 in the standard of CEFR (Common European Framework of Reference for Languages). In these foundational learning stages, the learning and teaching materials in the text books are supplemented with guided-speaking practice, grammar for communication, pronunciation practices and substitution drilling methods. Lin (2010) concludes that in an interactive and immersing approach of speaking teaching, English instructors can aid their learners non-verbal ways of communicating, like body languages and gestures in the absence of sufficient target language understanding.

The following is the highlight of the desired and actual condition of the environment where this ability grouping based ELT is applied. Green (2018) suggests that learners can learn language best when they are grouped in classes with their peers who have similar ability or aptitude. Accordingly, ability grouping is a form of institutional policy for the students and the lecturers aimed to accelerate the ELT at the institution. Still, as cited earlier, PLJ is a non-selective college, not a college which admits the incoming students through a competitive or strict test. In fact, PLJ applies an unmeasured acceptance to the incoming students. Therefore, due to the unmeasured acceptance, the challenges faced by the English lecturers are more specific compared to other colleges especially competitive colleges. For example, among the incoming students there is sometimes one or two who are still not familiar enough with certain basic understanding of English language, for example: numbers (cardinal numbers and ordinal numbers), alphabet spelling and pronunciation, names of days, colors, and the likes. These challenges are more or less affect the ways the lecturers teach students, especially students with ‘special needs’.

2. Input

The framework of Input evaluation component considers the policy planning by viewing the resources available for the policy and how the resources are used in meeting the expected outcomes (Stufflebeam, 2007). In general, based on space and time, there are two types of teaching and learning activities, namely inside the classroom and outside the classroom. Inside the classroom are teaching and learning according to the fixed schedules with 4 (four credits) each, as long as fourteen meetings plus mid semester test and final semester test every semester, where in the first and second semester, the students are grouped based on their level of English ability (ability grouping). While, for outside classroom, there are English Conversation Journal, English Camp and Outing classes, all of which are counted in into assignment element of scoring.

The teaching and learning activities are supported by resources available at the institution, namely the English lecturers (four [4] English lecturers to teach six [6] class groups), English language curriculum,
programs and syllabus, learning environment (classrooms, laboratories and library) and academic staff. The first resource that the author emphasizes is English lecturers.

As it is explained in its Guidebook for the Implementation of Teaching and Learning Processes (LP3I, 2015), “effective language learning, one of which is determined by the application of leveling in the learning system, namely the grouping of classes based on the level of competence. Classes with the same level of competence can run more effectively than classes with heterogeneous abilities. That is because the teaching approach method used by the lecturer becomes more focused, the material provided can be more appropriate according to students’ abilities, the interactions and communication that are built between students are more interwined because they are at the same level. Therefore, leveling or grading in English classes is important to do”.

The ability grouping policy is supposed be applied at all PLJ campus branches. The steps for implementing the grouping are as follow:

a. Test material for grouping is a matter of TOEIC
b. Campus/ branch can do additional leveling test by interview/ oral test. But it is not permitted only by interview/ oral test because the assessment will be less objective and less standardized.
c. Class grouping is determined from the TOEIC test results. Or merging the TOEIC and Oral Test scores if an oral test is performed.
d. Classes are grouped with approximately 25 students.
e. Student grouping is done by sorting the list of new student names from the highest grade to the student with the lowest grade. The first 25 students with the highest grades are grouped in one class, followed by a list of names of the next 25 students who are included in the next class, and so on.
f. This classifications apply to all English courses
g. The class code to indicate the highest level to the lowest can be done by: (Example for EC courses): EC 1-1, EC 1-2, EC 1-3, etc. (Example BE courses): BE 1-1, BE 1-2, BE 1-3, etc. Information number 1-1: the first number 1 shows the semester, the second number 1 indicates the highest class level and so on.
h. To be able to implement the application of this ability grouping, the English course is scheduled on the same day for all existing programs. (This is in line with the policy of compaction of the English language curriculum, which is 8 SKS (4 Sessions) in Semester 1 and 8 SKS (4 Sessions of Credit Point System) in semester 2.

With this ability grouping the incoming students of three (3) study concentrations (Office Administration, Computerized Accounting and Computer Informatics) were divided into six class groups. By doing so, there are about eleven (11) until sixteen (16) students in every class group. These six class groups are taught by four English lecturers. Then, looking at this composition, the ratio of teacher : learner is considered ideal for a language class.

Regarding the lecturer – students ratio or class size in ELT, American Council on the Teaching of Foreign Language (2006) states that concerning the learning outcomes of a language program which is based on a standard, there is supposed to be chances for intensive and significant teacher-to-student and student-to-teacher interaction, structured practice, and customized feedback during instructional phase. This justifies the special attention to a class size that appears to be as small as possible.

For further understanding, the author reveals any alternative strategies used by the institution in achieving the expected outcomes. Brown (as cited in Lin, 2010) proposes eight principal factors for teaching English speaking well: conversational discourse, question about intelligibility, interaction effect, pronunciation teaching, accuracy and fluency, affective matter, genres of spoken language and the development of spoken corpus. Regarding this view, the author finds out that in trying to achieve the targets and realizing the vision, the institution also sets a compulsory activity for all students called English Conversation Journal (ECJ). The ECJ presentation or conversation could be carried out anywhere around campus and anytime at the range of time set as the schedules. The ECJ is the only component contributed to the score of assignment which is counted for the final score (of GE 1, GE 2, BE 1 and BE 2).

Based on the guidelines book of the institution (LP3I, 2015), English Conversation Journal is weekly conversation report book which is used as a monitoring of weekly activities in training students’ English communication skills. The journal contains some topics related to the lessons taught in the classroom (see Figure 4.1). Every student is supposed to be able to present and answer questions regarding topics that have been determined in the Conversation Journal. The development of communication skills is monitored and guided by one main contact person (usually the lecturer who teaches in their English class). In one semester, students must be able to complete a minimum of 7 (seven) topics. Related to the ability grouping, ECJ is a more personalized teaching-learning activity, as the student and the lecturer are face to face involved in a question and answer activity (conversation), so that the lecturer can closely observe the current English skills and the development of the skills.
The ECJ has certain purposes as stated in the following:

a. Familiarize students to practice describing something or certain topic and answer questions regarding the topic presented.
b. Familiarize students to practice their English outside of class hours.
c. Give additional time to students to study outside class hours
d. Give special attention to students who still do not meet the standards
e. Monitor the development of student communication skills on a weekly basis
f. Students are able to give opinions about topics contained in the English Conversation Journal in writing and verbally. These are the topics for ECJ used: *myself, me and my family, my ideal job, public transportation, Favorite TV program, my daily routines, places around my town, my idol, my holiday experience, my friend and her/his job, business activities, making an appointment, a company and its history, my favorite thing(s), what I did last week, what I am going to do/ what I will do, food and drinks that I have in my kitchen, places that I have ever been to and places that I have never been to, technologies that I know, my strengths and weaknesses, things I like and things I do not like, What is CV and how to make it, what is cover letter and how to make it, job interview, unforgettable experience, interesting experience, me and my hobby, my future career, a trip, my future plans, food, music, my dream company, how I improve my English*, and the likes (according to the creativity of every campus branch).

The ECJ is scored and counted as the assignment, which is contributed (as much as 25%) to the final score of every English class that a student takes. The score is the accumulative scores of all topics in the ECJ.

The standard reference for scoring the journals are as follow:

<table>
<thead>
<tr>
<th>Score</th>
<th>Grammar</th>
<th>Vocabulary</th>
<th>Pronunciation</th>
<th>Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>86-90</td>
<td>The response demonstrates effective use of grammar. It exhibits a fairly high degree of automaticity with good control of basic and complex structures (as appropriate). Some minor (or systematic) errors are noticeable but do not obscure meaning.</td>
<td>Rich, precise and impressive usage of vocabulary words learned in and beyond of class.</td>
<td>Pronunciation was very clear and easy to understand.</td>
<td>Student was able to comprehend and respond to all of the questions and the topics that were being discussed with ease</td>
</tr>
<tr>
<td>76-85</td>
<td>Response may exhibit some imprecise or inaccurate use of grammatical structures used. Student was able to express their ideas and responses fairly well but makes mistakes with their tenses.</td>
<td>Student utilized the words, in an accurate manner for the situation given.</td>
<td>Pronunciation was good and did not interfere with communication</td>
<td>Student was able to comprehend and respond to most of the questions and topics that were being discussed.</td>
</tr>
<tr>
<td>70-75</td>
<td>Student was able to express their ideas and responses adequately but often displayed inconsistencies with their sentence structure and tenses.</td>
<td>Student was able to use broad vocabulary words but was lacking, making him/her repetitve and cannot expand on his/her ideas.</td>
<td>Student was slightly unclear with pronunciation at times, but generally is fair.</td>
<td>Student fairly grasped some of the questions and topics that were being discussed.</td>
</tr>
</tbody>
</table>
The standard scoring reference above is also used for speaking section of the Mid Semester Test (UTS) and Final Semester Test (UAS).

Another alternative strategy is the standard competency for English lecturers using TOEIC score, which is 650 minimum. Considering learning is the key to teaching, the advancement of teachers in the subject they teach is contextually linked to better learning outcomes (Schleicher, 2016). In fact, from the four lecturers teaching at the institution, 75% fulfil the standard. In addition, the institution also trains the lecturers regularly in the semester break, although in general, it is only the full-time lecturer who makes time for the lecturer development training. Although there is lecturer sharing session, the information shared is often not shared well as each of the lecturer has their own limited time spent at campus, due to tight teaching schedules at other campuses. As the results, there is some discrepancy at certain points in executing the curriculum and syllabus. The example of the discrepancy is in the English teaching thinking frameworks. As the teaching approach at the institution has emphasized more on the communicative approach, in which student talking time is more than talking time, traditional teaching approach has often still been dominant to be carried out rather than student-centered learning.

Furthermore, regarding the textbook used during ability grouping based English classes (GE 1 and GE 2), the institution formulated its own text books and the approach of teaching based on the books, namely student talking time oriented approach. The books are Speaking Practice for General English 1 and Speaking Practice for General English 2. To meet the desired learning outcomes, the books are enriched by guided-speaking practice, grammar for communication, pronunciation practices and substitution drill. And, as TOEIC has been set a standard competency to be achieved by students, the books are also supplemented by TOEIC-type questions in the listening, reading and grammar exercises.

Related to the resource-effectiveness and the significance between the means and the objective of ability grouping based ELT applied by the institution, the author looks at the guideline book of ELT (LP3I, 2015), in which the vision of their English language teaching and learning program is to yield graduates who are able to communicate in English both verbally and in written forms according to the demand of workforce. This vision is then broken down into four specific goals as mentioned in the previous sub-chapter (1. Produce students who are able to communicate in offices, businesses and everyday public conversations in English well; 2. Yield students who are able to do job interviews and presentations in English properly; 3. Yield students who have expertise in writing correspondence or e-mail properly; 4. Yield students with a minimum TOEIC score of 450; Related to the vision and specific goals, the resources available and how the resources are used at the institution as highlighted above and the expected outcomes are still in line. Still, related to the significance between the means and the objectives, one thing to note is that the institution admits the new students with heterogenous background of intellectual abilities, as it does not consider any pass-fail tests in the admission process. Ultimately, this could be a significant challenge for the English lecturers in treating them according to the curriculum and syllabus and the desired results.

### 3. Process

The general framework in the process evaluation is the implementation of the planned procedures (Stufflebeam, 2007). The process of grouping-based ELT at the institution starts just before the incoming students commence their study at the institution. The incoming students are given TOEIC test simultaneously on the same day. The results of test (the scores) are then used to map out and group them into class groups. The average of TOEIC score of the incoming students of 2017 academic year was 271, with the average of its lower level class group was 136 and the average of higher level class group was 503. While the students of 2018 academic year, the average was 238, with the average of lower level class was 91 and the higher level class group was 458.

Contextually, the measurement of whether the ability grouping based ELT at the institution is successful or not can be seen from their official TOEIC score (from authorized TOEIC test holder in Indonesia, namely ITC [International Test Center]). It is to note that the official TOEIC score used as the standard of competence is the TOEIC score the students get from ITC, which is usually held in the fifth semester; while the TOEIC test given to them every semester from the first semester until the fourth semester is only counted for the purpose of evaluation and enrichment.
As for the students of 2017 and 2018 academic years, there were six (6) class groups as the results of placement (TOEIC test). Then, these six class groups are taught by four English lecturers. Each class group consisted of ten (10) to sixteen (16) students. These amount ranges of student in each class are considered adequate for teaching and learning environment in the classroom. It corresponds with Lin (2010) who indicates that by having frequent interactions between students and students and teacher and students, an English class might appear to be effective.

In the process evaluation, Stufflebeam (2007) implies that the researcher needs to check or forecast any possible defects in the implementation of the program being evaluated. Concerning this, the author looks on how the expected target of students competence and the way PLJ admits its students. PLJ sets the competence standard of TOEIC minimum score of 350 (for 2017 academic year students) and 450 (for 2018 academic year students) and GE and BE scores of 75. At this point, the author sees a possible defect in the implementation ability grouping based ELT. First, in admitting the new students, the institution does not require any academic standard (pass/ fail admission test or academic assessment) but in the implementation, it directly sets an academic standard to them, namely TOEIC score of 350-450 and English subjects score of 75, which are in form of pass/ fail. Second, the increase of TOEIC score standard from 350 to 450. The author finds no rationale grounding or need analysis-based formula why it is 350, why it is 450, what makes it need to increase from 350 to 450. In other words, if the students do not need any academic standards for their entry to the institution, why they are required to meet certain academic standards to graduate from the institution.

Regarding information and any programming decisions from the implementation of grouping based ELT at the institution, the author views how to what extent the four lecturers (included the author) make sense of what they have been experiencing in their teaching activities. To formulate and extract any information and programming decisions, the author applies a focus group discussion (FGD) involving the lecturers (see table 3.1. for the FGD formula). The information from these lecturers could be significant points, as the lecturers are at the frontiers in ability grouping based ELT at the institution being concerned.

For the ethical matter, the author presents the three lecturers’ perspectives and insights by identifying themselves as L1 (Lecturer 1), L2 (Lecturer 2) and L3 (Lecturer 3). It starts from L1. According to L1, the idea of ability grouping based ELT at the campus is overall acceptable. Some points that A takes into consideration are mainly the day-to-day experience in the class. A notices that when doing a pair work in the class, especially for speaking drills, a lower-level student paired with a lower level student could be a case. They both have the same ability in trying to understand the instruction from the lecturer; so that sometimes the pair drilling does not move anywhere as they both do not have the clues on what to do purposely. Related to this, Yoshida’s study (2008) (as cited in Lin, 2010) suggests students’ satisfaction with their pair-work interactions in class and constructive feedback from the instructor is important in teaching speaking. Thus, the roles of a teacher in finding the understandable clues or key sentence are vital.

Furthermore, L1 indicates that independent tasks for lower level students need to be adjusted by the lecturer, as, for instance, there is a potential that they give unclear or out of context responses in the drills. This implies that the lecturer is conditioned to put extra efforts in directing lower level students. Lin (2010) argues that many English teaching specialists have proposed the significance of role play in interactive formats (Brash and Warnecke, 2009; Carlson, 2009; Powell, 2009; Sung and Hsu, 2009). For intrsucting and learning of speaking, writing and vocabulary, directing the learners to interact in pairs is a means as it may hint the interacting members to dig in the contents of interactive activities and recognize the partners intended meanings of vocabulary more productively.

Beside what are highlighted above, L1 also notified that there is ‘secure mental stability’ in the class. This is because the students feel that they are in the same level of English skills, so that the are not worried of being labelled as weak - strong or lower – higher in the class. This condition leads to the state of fair self confident among students in that class group. Slavin (as cited in Howling, 2017) indicates that (ability) grouping decrease the negative psychological effect of placement on students. Howling (2017) also underlines that grouping learners in a different way can assist in increasing learners’ motivation and development.

Another lecturer, L2, has some insights in the teaching practices at the institution. The lecturer notes that, overall, the ability grouping is relevant, in the sense that the lecturer knows or has a map about the incoming students, so that the lecturer could alter the materials and the teaching plans to be applied. As for the equal ability among the students in the class group, L2 sees that it may boost students’ confidence in the class. This is confirmed in Hwang’s study (2014) that points out students who take views of others into consideration and expect to get ahead of others are more likely to lose confidence in themselves and their academic ability when they are compared with stronger peers in the upper level.

L2 also cites that the teaching materials in the textbook already cover the four skills of English language, although L2 still thinks there is still room for improvement, which is about the sequencnes of the drills and exercises for every chapter or topic in the textbook. The usual sequence of teaching stages in
teaching are relatively presentation, practice and production. L2 perceived that the teaching material sequences in the book are disordered, in the parts of practice, still there is grammar explanation (which should be part of presentation), for instance. In addition, L2 conceived that the listening materials are adequate but reading materials are still insufficient. L2 concluded that in general the ability grouping has been running well, with a note in the roles of lecturer as a key factor, especially for lower-level class groups.

The lecturer, marked as L3, finds no serious issue regarding the implementation of ability grouping. Still, L3 has several insights to take into consideration. First concern to note from L3 is about the textbook. The concern of a textbook is confirmed in Lin (2010) that indicates that vocabulary pre-teaching and then reading the load of the textbook are common means how English instructor introduces or familiarizes new words. According to L3, if the implementation of ability grouping must be in line with the contents of the textbook used, or in other words, if the students are classified or grouped according to their ability, the textbook also needs to be adjusted based on students’ level of ability. Therefore, referring to L3’s point of view, the higher level class group must use higher level textbook as well; and vice versa for lower level class grouped. L3 emphasizes that if the textbook is the same for all levels of class groups, it is better not to apply the ability grouping; because some students whose abilities are below the average, will likely feel difficult in understanding the materials, and for those whose abilities are above the average, the materials are potentially too easy or boring. Yet, regarding the quality of the textbook used by the institution (for GE 1 and GE 2), based on L3’s teaching experience, the textbook is generally representative for middle and basic level students, but for the advance or higher level students, the textbook is too simple or not challenging enough if the lecturer does not do any improvisation or creativity. River (1987, 2001 as cited in Lin, 2010) mentions that development of language control goes on by creativity, being maintained by participatory and interactive activities.

4. Product

In this element of evaluation, the author measures and interprets the attainments gained from the implementation of ability grouping based ELT at the institution. At PLJ, to measure the quality of English teaching, every semester there are UTS (Ujian Tengah Semester [Mid Semester Test]) and UAS (Ujian Akhir Semester [Final Semester Test]). For the first and second semester students, the subject that are tested are General English One (GE 1) and General English Two (GE 2). In both mid semester test and final semester test, there are two types of test, namely oral (spoken) English test and written English test. Therefore, all skills of English language (speaking, reading, listening and writing) are covered in the tests. Accordingly, the final scores gained by the students are the combination of oral test scores and written test scores. The combination means the oral test scores plus written test scores, the it is divided into two, resulting in the final score of a student.

Beside GE 1 and GE 2 in the first and second semester, there are also TOEIC (Institutional TOEIC Prediction test) tests given every end of semesters (first semester, second semester, third semester and fourth semester. The minimum score as the standard of General English 1 and General English 2 is 75 (Seventy-five), while for TOEIC, the minimum standard scores are different for 2017 students and for 2018 students: for 2017 students, it is 350 and for 2018 students, it is 450. Students who fail to get the minimum scores are supposed to take remedial classes and students who fail to get minimum scores for official TOEIC test (held by ITC [International Test Center] in the fifth semester are considered not eligible to take part in Final Assignment Hearing (Sidang Tugas Akhir) as the main requirement to graduate from the institution. Practically, the scores of both subjects (General English and TOEIC) are the quantitative indicators of the success of ability grouping based ELT at the institution. Accordingly, in the following (Table 4.4 and Table 4.5), the author presents the findings of attainments that have already been gained so far by the year 2017 and year 2018 students on both General English (1 and 2) and TOEIC (Prediction test).
Table 6 shows what have been the product or outcome of the ability grouping based ELT at PLJ Pasar Minggu Campus. For the students of 2017 academic year, the average of TOEIC score when they start the English classes (Pre-Grouping) was 271. Then, the last time they were given TOEIC test was when they were in fourth semester (in 2019) and their average score was 368. It means that there was an increase of 36%. Meanwhile, for the higher level class group, their average pre-grouping TOEIC score was 503, and at the last time they were given TOEIC test (in the fourth semester, 2019), their average score was 615. Here, the author sees an increase of average TOEIC score of higher level class group of 2017 academic year student as much as 22%. On the other side, the author sees an interesting finding from the lower level class group. Their average score in the pre-grouping was 136, and after they were given TOEIC test in the fourth semester (in 2019), their average was 277. At this point, the author sees a significant increase, namely 104%. Apparently, what could be the point to be concerned here is that the lower level class group made a bigger percentage of increase than their peers did in the higher level class group.

As for the 2018 academic year students, their average TOEIC test score was 238. Then, after they were given the test in the second semester (2019), their average score was 251, which means there is 5.4% rise. Meanwhile, for the higher level class group, their average pre-grouping TOEIC score was 458, and at the last time they were given TOEIC test (in the second semester, 2019), their average score was 446. Here, the author sees a decrease of average TOEIC score of higher level class group of 2018 academic year student as much as -2.6%. On the other side, the author sees a standing out finding from the lower level class group. Their average score in the pre-grouping was 91, and after they were given TOEIC test in the second semester (in 2019), their average reached 158. At this point, the author found an obvious increase, as much as 74%. Again, what could be the point to take into consideration here is that the lower level class group made a larger percentage of increase than their counterparts in the higher level class group, who even made a minus (-).
Scores of GE 1 and GE 2 Courses (1st Semester and 2nd Semester) of 2017 and 2018 Academic Year Students

The table of General English 1 and 2 attainments above points out what have been the impact or expected product of the ability grouping based ELT at Campus. The scores are what have been achieved by the students of 2017 academic year and 2018 academic year during and the end of ability grouping based ELT, which is the first and the second semesters. It is to notice that the GE 1 and GE 2 scores are derived from the same components of scoring, namely formative (20 %), assignment (30 %) mid semester test (20 %) and final semester test (30 %). The final score of every student is the combination of written test and speaking oral test (see Table 4.3 for the standard scoring reference for speaking or oral test).

For the students of 2017 academic year, the average of GE 1 score in their first semester was 74,64. Then, when they were in the second semester (in 2018), their average score of GE 2 was 71,62. It indicates that there was an decrease of -4,04 %. Meanwhile, for the higher level class group, their average score of GE 1 was 86,63, and their average score of GE 2 (second semester, 2018) was 83,90. Here, the author finds an increase of average score of GE 1 to GE 2, namely 1,25 %. Obviously, what could be the point to be concerned here is that, different from TOEIC attainments, higher level class group of 2017 academic year students made a bigger percentage of increase than their peers did in the lower level class group.

Going further, for the 2018 academic year students, their average GE 1 score was 75,71. Then, for their GE 2 (in the second semester, 2019), their average score was 68,39, meaning there is -9,66 % decrease. Meanwhile, for the higher level class group, their average GE 1 score was 87,00, and in the second semester, 2019, their GE 2 average score was 80,07. Here, the author sees a decrease of average GE 1 to GE 2 score of higher level class group of 2018 academic year student as much as -7,96 %. On the other hand, the author finds a contrasting figures from the lower level class group. Their average GE 1 score (in the first semester) was 68,8, and then in the second semester (in 2019), their average GE 2 score reached 73,9. At this point, the author sees an apparent increase of GE 1 to GEE 2 average score, as much as 7,4 %. Looking at these findings, the average score of GE 1 and GE 2 of 2018 academic year students in general and in specific— higher level class group and lower level class group, what could be the point to be concerned is that the lower level class group made a larger percentage of increase than their fellows in the higher level class group, who showed a minus (-): from 87,00 in GE 1 to 80,07 in GE 2.

Discussion

Based on the findings, there are certain points to be concerned. First of all, as a private polytechnic with the label of vocational, a tagline of prospective jobs after graduation to attract new students is something normative, because students are the main resources for the continuation of education service by the institution. But, it does not mean that it has nothing to do with the expected outcomes and the strategies applied. Ability grouping based ELT as one the strategies could not make sense if not being supported by other strategies and policies set by the institution. For instance, the author finds that the standard score of TOEIC of 350 for 2017 academic year students and 450 for 2018 academic year students and also the
standard score for GE 1 and GE 2 of 75 still lacks of rationale grounding. As the lecturers deal with teaching, learning and evaluating, the figure from 350 to 450 is a significant elevation that needs rationale calculation by the stakeholding parties in the institution. Lin (2010) emphasizes that more systematic measurements of English skills ought not to merely be built up as a statistical device for recognizing learners’ and teachers’ competencies of learning and teaching, but also be set up as a functional arrangement that asserts the effectiveness of teacher and students.

When the author seeks to find out the rational ground why the expected scores were 350 and (TOEIC) and 75 (GE 1, GE 2), the stakeholding parties deliver rationales that if their scores are below 450, they show no better English communication, which is not selling enough in the industry, and that if their scores in English subject are below 75, it means their English is still poor. What the author finds missing at this point is the empirical analysis on why the numbers are 350, 450 and 75, referring to the fact that the incoming students are admitted without any academic measurement (an admission test or the likes). This could lead to a question about the readiness of the students to achieve the targeted scores is still questionable due to the absence of admission test (pass/ fail test) for the incoming students. Cohen (as cited in Lin, 2010) suggests that assessments or tests are naturally developmental and and non-harming, enabling the test takers sufficiently show what they understand and do not understand, and giving constructive feedback for both teachers and learners. Hence, what could be a case is that there is still a gaping whole between the background of incoming students with the standard scores targeted.

In term of students readiness to meet the targeted outcomes from the ELT, in fact, there is an anticipation made by the institution, namely foundation program or matriculation or pre-lecture for the already registered students before they start their study. In Howling’s study (2017), it is recognized that once grouped in a class fit with their proficiency levels and received compatible instruction and assistance in the foundational program, the students gained confidence and began to enjoy learning English. But, in case of PLJ, the foundation program itself does not go as it is expected. Because the program is not categorized as compulsory for the incoming students, the students come to the program classes based on how comfortable they feel to come, instead of being a need to prepare themselves for improving their English. In addition, the institution cannot spot which students need English preparation and which students do not, as there is no mechanism or a test to detect the readiness of incoming students for standardized English English competency they need to achieve in their study. As the result, the preparation program is not focused to tackle students proficiency needs and tend to be random, even appears to be merely promotional introduction of the institution.

The next concern to discuss is related to the textbook used. This appears to be a major concern of this ability grouping based ELT because it is the major sources of teaching and learning activities for the courses of GE 1, GE 2 and relatively TOEIC as well. As explained by the lecturers who teach and use the books, there are still some loopholes with the the textbook, typos of spelling of words, the missing of audio files for certain listening materials, a disordered sequence of drilling and target language explanation, and so on so forth. But despite these discrepancies, overall, the lecturers still consider the book usable and teachable.

Still related to the textbook, there are certain pair work guidances for speaking which works for higher level students but do not for lower level students. For this matter, the role of a lecturer in using more understandable language in explaining the instructions is an advantage. From some information that the author collects through focused group discussion with the lecturers, there is a proposition that if the students are group based on their level of competence, the textbook applied also must be classified based on their level of competence. This proposition is interesting, but the implications also need to be re-assessed by the institution, especially in terms of budgets. Yet, one thing that can be a red lining from this textbook concern, is the lower the level of students ability, the higher the creativity and effort needed from the lecturer, and vice versa.

Then, moving on to the attainments that have been highlighted in the Product evaluation. There is an increase of average score for both TOEIC and GE 1 and GE 2. For TOEIC 2017 academic year students, pre-grouping: 271 and post grouping: 368 (36 % increase) and for 2018 academic year students, pre-grouping: 238 and post-grouping: 251 (5,4 % increase). But, it does not directly mean that the ability grouping goes as expected. From the two indicators (TOEIC score and GE scores), it is only TOEIC that can be made sure to notice. TOEIC, although in this case, it is a prediction test, is a standardized test that can be used to show an English language competency. As for GE test, it is a non-standardized test that is made by the English language team at the institution based on the guidelines from the syllabus and refering to the materials that have been taught to the students. And the scoring reference for GE is not based on a standardized formula, especially for scoring the writing or composition of the students. While, for TOEIC, all of the questions are in form of multiple choices and there is formalized standard reference for calculating the score. Therefore, the issue of subjectivity in scoring tends to be in GE test, rather than in TOEIC.

Moreover, it is not only the different reference standard of scoring between GE and TOEIC that can be an issue. The case of remedial also needs attention here. Starting in academic year 2015-2016, students
who fail to get minimum 75 for an English subjects has to take a remedial. The remedial costs 150.000 IDR per credit (SKS), so if the English subject is 4 credits, a student has to pay as much as 600.000 IDR. For the majority of PLJ students who come from middle and lower level family, that amount of money is a burden. This condition at certain points affect the attention of the lecturers who teach them. Then, pragmatically, instead of burdening the students and interfering the administration affairs of academic team, the students tend to do it ‘safe’ by giving satisfying scores (of minimum 75) to students, especially the ones who are in need of ‘intervention’. This concern of remedial, especially related to budget, is also amplified by Howling (2017), that to deal with the students’ low performance, a third of the 2014 (UAE) federal budget for higher education was allocated on remedial efforts, such as the English foundational year programs.

Another point to discuss is the characteristic and motivational issues among students, apticularly students in the lower level class group. Looking at the attainments of both TOEIC and GE of 2017 and 2018 academic year students (see Table 4.4 and 4.5), it is intriguing to see the fact that lower level group students achieve bigger percentage increase that their peers in higher level class group. For the increase of TOEIC 2017 academic year students from pre-grouping to post-grouping: higher level group = 36 % and lower level group = 104 % and of 2018 academic year students: higher level class group = 5,4 % and lower level class group = 74 %. Meanwhile, for the increase of GE 2017 academic year students from GE 1 to GE 2: higher level group = - 4,04 % (minus) and lower level group = 1,25 % and of 2018 academic year students: higher level class group = - 9,66 % (minus) and lower level class group = 7,41 %. These findings are in contrast with Howling’s study (2017) that concludes the on going class grouping policy at the secondary school level does not benefit lower-performing learners.

The advantages of percentage increase of lower level class group over higher level class group could raise a question in terms of learning motivation, or a motivation to elevate to the highest possible scores they could achieve, given they are labelled as a lower level students. Whether or not their advantage over higher level students is caused by their high motivation because they do not want to be labelled as lower level students, but for competitive students especially competitive students who happened to be placed at lower level class group, the label as lower level students could be a huge energy to prove that they are not actually lower. In other words, it is normative to see competitive students do more than what other people expect from them. This finding is in contrast with what Hwang (2014) concludes that placement students into lower stage tends to yield labeling effect, and this such distinct atmosphere appears to structurally position lower performing learners at harm.

Seeing from the fact that lower level students could show progress more than their counterparts did in the higher level class group could show, it cannot be put aside from the roles of the lecturers who teach them. As the students who are labelled as lower level learners, they are reluctant to feel desperate or to feel hopeless being left behind. Practically, it is their lecturer who is the closest to them in keeping or their learning motivation on the track or even boosting up. This finding intersects with what Howling (2017) concludes that ability grouping would possibly boost lower-performing students’ motivation and confidence, making them build more academic development and yield with satisfactorily proficiency level.

6. CONCLUSION AND RECOMMENDATION

Conclusion

Based on the findings in the evaluation components of CIPP and also in the Discussion, the author draws the following conclusion:

First, how PLJ perceives its institution as a vocational higher education more or less intersects with how the institution admits and educates the students. As a private polytechnic with the label of vocational, a tagline of prospective jobs after graduation to attract new students is something normative. But, it does not mean that it has nothing to do with the implementation of ELT, the expected outcomes and the strategies of teaching applied. The heterogenous level of English competency as the results of unmeasured admission leads to a need of accurate mapping of the incoming students’ English proficiency in order for the institution to accelerate the implementation of ELT. This ability grouping based ELT is one of the strategies applied to successfully achieve the expected learning outcomes as stated in the vision of English language teaching and learning of the institution.

Second, in setting up the competence standard, it is important not only to recognize the demand of the industry but also a rationale calculation, which relates to the academic or intelectual background of the incoming students, the resources and human resources available at the institution, and so on so forth. Practically, there is a need to conduct need analysis, especially for the expected standard score of TOEIC to be achieved by students, included the rationale grounding of setting up the targets of 350 for 2017 academic year students and 450 for 2018 academic year students. There should be a quantative measurement on why the target for 2018 academic year students is 28 % (100 points) higher than 2017 academic year students.
Third, overall, the ability grouping policy works or at least it does not go backward seen from the two indicators: scores of TOEIC and scores GE 1 and GE 2 that the author finds in from the attainments presented the Product evaluation (see Table 4.4 and 4.5). From the attainments of the ability grouping students, it is apparently seen that students of lower level class groups show bigger percentage of progress than their counterparts in the higher level class groups. The finding is interesting as it could lead to an indication that ability grouping may benefit more for lower level student group.

Still, at this point, the author as one the English language lecturers teaching at the institution, finds the difference of objectivity in the scoring of TOEIC and GE 1 - GE 2. TOEIC is a standardized English proficiency test although in its implementation at the institution, the TOEIC given each semester is a TOEIC compiled by an English language team from certain reliable resources (commonly name prediction test). While GE, it is a non-standardized test composed by English language teaching team based on the materials that have been taught in day-to-day teaching. Hence, in terms of scoring accuracy, TOEIC test is more measured and more systematic than GE, as TOEIC test has a standardized formula in calculating the scores, while GE test is scored subjectively by the lecturers particularly for the composition or writing part.

This subjectivity issue is also more or less intersected with the policy of the institution on remedial for students who fail to get score of more than 75 for GE. The students who is supposed to do the remedial has to pay 150.000 IDR for every 2 (Two) credits (2 SKS), which is more likely burdening the majority of students who come from low economic background, leading to a subjective consideration by the lecturers in giving scores (as discussed in the Discussion section of Chapter 4).

But related to the improvement of students, as a lecturer who directly teaches and interacts with the students, the author sees the impact of students of lower level group being labelled as marginalised students. For the competitive and highly motivated students, being labelled as marginalised students may, in fact, increase their confidence to prove that they can achieve what other students doubt that they can achieve. This proposition is on the contrary with what Hwang (2014) concludes that placing students into lower class groups possibly tends to yield a labeling effect or stereotype, and such atmosphere likely appears to structurally harm them. But, the contrast may likely occur because of the difference of the object of the study, where Hwang (2014) studies ability grouping impacts on Korean middle school learners, while in this study, the author deals with ability grouping at a tertiary level of education in Indonesia. Regarding this, the author perceives that the roles and the capacity of lecturers are significantly vital in stimulating and maintaining students’ learning motivation, particularly the students of lower level class groups.

Implications and Recommendation

Implications from the approach of CIPP model use in this study are related to its strengths and weaknesses. Its strengths are that it is practical to be applied for varied evaluation context, its implementation of the research is flexible, applicable for individual or teamwork or collaborative research, and so on so forth. Meanwhile, its weaknesses are that there is a bluring understanding between evaluation and other processes of investigative inquiry such as analysis or assessment of need; and that this evaluation framework is relatively less devised than other investigative or evaluative research frameworks (Stufflebeam, 2007).

Then, as the recommendation in regard to the results of this research, the author suggests more specific study, such as need analysis, especially for the expected standard score of TOEIC and GE. The rationale grounding and need based analysis of setting up the targets of 350 – 450 of TOEIC are important to give more focused guidelines for the lecturers. Suggestion: take the standard of TOEIC from the previous academic year graduates’ ITC official TOEIC scores.

Furthermore, based on the findings and discussion, there should be more focused study how and to what extent ability grouping affects learners’ English skills, particularly speaking. As part of the teaching and learning, how and to what extent English conversational Journal (ECJ) affects students’ English skills, especially speaking. While related to the textbook used, the author suggests there should be periodical improvement of the the textbooks based on the feedbacks from the main stakeholders, namely the lecturers and the students. Ultimately, last but not least, based on the findings, this ability grouping based ELT fundamentally—beside competence and reliability of the lecturers—requires more efforts from the lecturers, especially creativity and deep care on students’ motivation in both lower level class group and higher level class group, as lecturers are the ones who have more space and mental bonding with them.
REFERENCES


Gunung, I. N., & Darma, I. K. (2019). Implementing the context, input, process, product (CIPP) evaluation model to measure the effectiveness of the implementation of teaching at Politeknik Negeri Bali (PNB). International Journal of Environmental and Science Education.


Howling, C (2017). The need for ability grouping in English classes in public schools in the UAE: Policy Paper.

https://coerll.utexas.edu/methods/modules/classroom/02/howmany.php

https://www.bartleby.com/354/18.html


