



## DEVELOPING AN OBSERVATION INSTRUMENT FOR ASSESSING THE EFFECTIVENESS OF ENGLISH TEACHING AT VOCATIONAL SECONDARY SCHOOLS

\*<sup>1</sup>Estu Widodo; <sup>2</sup>Franciscus Xaverius Sudarsono

<sup>1</sup>English Department, Universitas Muhammadiyah Malang, Jl. Bendungan Sutani No. 188A,  
Sumbersari, Lowokwaru, Sumbersari, Kota Malang, 65145, Jawa Timur, Indonesia

<sup>2</sup>Graduate School of Universitas Negeri Yogyakarta, Jl. Colombo No. 1, Karangmalang,  
Caturtunggal, Depok, Sleman, 55281, Yogyakarta, Indonesia

### Abstract

This study was aimed at: (1) developing an observation instrument for assessing English teaching at vocational secondary schools (VSSs), (2) examining the validity of the developed instrument, (3) examining the inter-rater reliability, and (4) assessing the effectiveness of the developed instrument. The subjects of the three-phase field-tests were English teachers and students of four VSSs. Observation sheets, questionnaires, and interviewes were employed to collect the data. The instruments were field-tested in three phases. Expert judgment, Kappa Coefficient, and Kane's (2006) argument-based validation and Bachman & Palmer's intepretive and use argument were used to determine the instrument validity and reliability. The field-test which involved stakeholders prove that the observation instrument could be implemented effectively to assess the teaching of English at VSSs.

**Keywords:** *assessment instruments, effective teaching, English, hotel accommodation*

### How to cite item:

Widodo, E., & Sudarsono, F. (2016). Developing an observation instrument for assessing the effectiveness of English teaching at vocational secondary schools. *Research and Evaluation in Education*, 2(2), 135-154. doi:<http://dx.doi.org/10.21831/reid.v2i2.8648>

\*Corresponding Author.

e-mail: [estu\\_wid@yahoo.com](mailto:estu_wid@yahoo.com)

## Introduction

As a key component of educational process, teaching quality constitutes educational quality in general. As far as educational quality is concerned, quality has been one of the big problems in the field of education in Indonesia since 1970s. According to some Indonesian experts, as noted by Kaluge (2003, pp.237-238), the quality problem has something to do with education process and outcome. The policy direction of educational development in this country advocating that our educational outcome must be able to fulfill what is needed by the worlds of business and industry in the form of, among others, student competency (Ministry of Education and Culture, 2015, p.75), is the right answer to the problem.

What is expressed at the policy direction of educational development should be the spirit which triggers all educational institutions, including vocational secondary schools (VSSs), to create innovations and get improved. As part of the educational system in Indonesia, VSSs play significant roles in developing the quality of human resources in general and workforce in particular. Therefore, Djojonegoro (1998, p.30) points out that VSS students and graduates must be equipped with strong basic skills, analytical skills, managerial skills, interpersonal skills, and a good command of English as an international language for communication.

The notion above lends to the credence that one of the primary aspects which determine the competitiveness of VSS graduates is English proficiency, more particularly the one related to the field. Globalization and Asean Economic Community have made up stricter competition to get a job, including for VSS graduates, and put good communication skills using English as a ticket to prospective jobs.

The high demand for VSS graduates with good English proficiency finally leads to the strategic role of English teaching at VSSs. That is, English teaching at VSSs must effective in the sense that it must be able to achieve the goal of English teaching at VSSs. In other words, English teaching at VSSs should provide students with sufficient opportunities to learn and practice English in

various communication activities so as to equips the students and graduates with good English proficiency required for further education or prospective jobs.

Unfortunately, English teaching has long been ineffective to foster VSS students to communicate in English. An effort to create effective teaching of English, therefore, must be encouraged in a systematic way. As stated by Kaluge and Hadiwijaya (2007, p.145), the principles of effective teaching are the basis of ongoing improvement. Moreover, as a foreign language, English is often deemed as one of the most frightening subjects; its effective teaching turns to be central to the teaching and learning processes of English not as a mother tongue (Curtis and Cheng, 2007, p.57).

The phenomenon of ineffective teaching of English actually is an iceberg of English teaching in secondary schools in Indonesia. The lack of competent teachers, teacher-centered learning, and inauthentic materials are some features sometimes found in Indonesian secondary schools. These, to some extent, have contributed to the failure of the implementation of Communicative Language Teaching in Indonesian secondary schools. It is no exaggeration, then, when Musthafa (2001, p.187) states that the most serious challenge faced by our English teachers is the lack of exposure to real-life English uses which engage students in real-life communication.

Some aforementioned facts related to English teaching at Indonesian secondary schools also worsen the quality of VSSs in general and in turn create a poor image of VSSs and VSS graduates. As stated by Maksum (in Burhanuddin, 2008, p.1), several problems encountered by vocational education in Indonesia are associated with several issues such as VSS graduates' competency and quality, the need to match VSS graduates to the work sector, the industry's concern about vocational education, and insufficient facilities to support teaching-learning processes. It is very likely that all of the aspects or issues are closely interrelated with the process and outcome of English teaching as described previously.

Based on the previously-explained description, many actions for improvement are needed. An initial effort to improve the quality of English teaching at VSSs could be made through the process of gathering data about the teaching process itself from which English teachers at VSSs will be informed about whether or not their teaching has been effective. With this in mind, an assessment of a teaching process will play an important role as the information gathered can reveal dimensions of teaching behaviors which need more emphasis or further improvement or enforcement.

Stemming from the tenet that the essence of assessment is to make improvement, innovative methods of assessment involving teachers and other stakeholders need to be promoted. Assessing teaching-learning processes involving teachers is also consistent with the Law No. 19 of 2005. Therein, it is mandatory for educational stakeholders to perform three kinds of assessments. The first is the assessment conducted by a teacher in a continuous way to monitor the process, progress, and improvement of learning. The second is the assessment which is undertaken by schools aimed at examining the accomplishment of graduate competency standard for all courses to indicate transparency, professionalism, and accountability. The third is the assessment carried out by the government in an attempt to find out the extent to which graduate competency has been accomplished. From the three kinds of assessments, assessing teacher behavior during the teaching and learning process is important to evaluate and to ensure that the teacher is teaching effectively and all activities in the classroom lead to the accomplishment of appropriate graduate competency.

In Indonesian context, assessing whether or not English teachers teach English effectively is absolutely required for two reasons. Firstly, VSSs are expected to bridge the big hope of many of Indonesian youths and their parents to get a job more easily so that VSS English teachers need to teach English relevant to VSS graduates' future jobs. Students's and graduates' competency to use English in real-life and workplace com-

munication, as a consequence, must be highlighted. An assessment of a teaching process is then required to monitor whether the teaching process is directed to the real objective of English teaching at VSSs and relevant to the hope of stakeholders, including those from related business and industries. Secondly, the state examination is still used as one of the measures of quality of English teaching in Indonesia, including at VSSs. The exam has pervasive effect on many aspects of teaching processes. Apart from the advantages of the state examination whose results provide the basis for the improvement of educational quality at the state level, the high-stake examination has instigated or incited Indonesian teachers to merely teach what to be tested in the examination. In such a situation, teachers must have a strong principle to use communicative language teaching (CLT) which prioritizes the communicative skills of students. Otherwise, as stated by McCaffrey, Koretz, Lockwood, and Hamilton (2003), a teacher is sometimes deemed to be effective when he or she just teaches skills to take tests most of his or he students pass a high-stake exam.

Drawn from the elaboration mentioned previously, with the intention to improve teaching processes and outcomes, a new set of assessment instruments is absolutely needed. The set of instruments is required to assess teaching processes through a classroom observation to examine what is taking place in the classroom, that is, whether it has been effective or not, whether the process is teacher-centered in which the students have limited opportunity to practice the target language use, or student-centered where students are actively engaged in meaningful communicative activities. The challenge is then to design a valid personnel assessment system to measure teacher performance which contributes to teacher professional development.

The dimensions of effective teaching is based on the findings of a lot of research with the hope that they are objective and research-based. With regard to effective teacher characteristics, for instance, Mitzel (1982, p.1895) notes that the subjective impressions have been the basis of decisions in the educational field. According to Borich (2007, p. 2), it is

due to the fact that many widespread definitions of affective teachers are made without clear and objective standards of teacher performance which could be consistently employed; they are not relevant to day-to-day classroom work so that they do not meaningfully contribute to the definition of a good teacher.

It is noteworthy, in line with the above elaboration, that Creemers and Kyriakides (2008, pp.104-117) with their grand theory or dynamic model of educational effectiveness assert that their model is a generic model in that the effectiveness factors are generic in nature. The generic factors of effective teaching are orientation, structuring, questioning techniques, teaching-modeling, application, classroom as a learning environment, management of time, and assessment. According to Kyriakides and Creemers (2009, p.82), the factors are considered generic because they have been found to be associated with student achievement in various subjects and at different phases of schooling.

In their attempt to seek the definition of effective teaching, Nougaret, Scruggs, and Mastropieri (2005, p.3), after conducting literary studies and reviewing some research, come to a conclusion that teacher effectiveness refers to 'alterable behavior exhibited by teachers that demonstrate a clear and consistent relationship to student achievement'. Having studied various literary works, they identify three important aspects of teacher effectiveness, namely planning and preparation, classroom environment, and instruction.

Many of the characteristics of effective teaching of English not as a mother tongue refer to the generic characteristics which are supported by rigorous research employing different approaches and result in several specific characteristics of effective English teaching. Song-Ae (2005, p.208), for instance, finds that an effective English teacher has to have cross-cultural understanding, more particularly when the teacher's culture is different from the students' to avoid students' resistance.

Apart from a wide variety of factors of effective teaching behaviors, it must be kept in mind that the factors should not treat a teaching and learning processes as a me-

chanical object. All of the factors must be interestingly orchestrated in a flexible way. The statement of Ornstein (1991, p.17) that teaching also involves 'artistic judgment about the best ways to teach' is worth noting.

In teaching English as a foreign language, an effective teaching is characterized by meaningful and contextual interactions which lead to communicative competence (Oxford, 1990, pp.8-9), teaching of learning strategies (Gower, Philips, & Walters, 1995, pp.87-88). The use of CLT which features all components of communicative competence (Brown, 2007, p.86) must enhance student proficiency to communicate (Littlewood, 2011, p.546), and use accuracy practices performed by students as the bridge to fluency activities. Above all, as noted by Brown (2007, p.98) further, CLT is manifested in task-based instructions which emphasize classroom interactions, student-centered activities, on classroom interaction, student-centered activities in the classroom, authenticity, and student's own experience deemed to contribute to learning.

An assessment process by directly observing classroom activities, especially an EFL teaching, is important to capture classroom dynamics as well as to ensure that some typical aspects of effective English teaching such as authentic materials, meaning-focused teaching, interactive methods, intercultural activities, enthusiasm, student high motivation and engagement, and teacher expertise appear in the classroom. Pertaining to the existing assessment types such as student assessments and portfolios, they highlight exclusively a certain aspect of teaching which is in reality very multidimensional in nature. An assessment by students proves to be unable to reach what is beyond the subjective area of teaching or student judgment. This kind of assessment, however, is also believed a valuable learners' point of view (Ackerman, Gross, & Vigneron, 2009; Carruth & Carruth, 2003) and has been found to be 'reliable, valid, relatively unbiased, and useful' (Cashin, 1990, p.1). On the other hand, given that the student assessment is largely dependent upon perspective, this kind of assessment has some weaknesses. In fact, the teacher's ability and performance to pro-

mote learning and improve instruction could not be captured (Emery, Kramer, & Tian, 2003). Employing other kinds of assessments along with a student assessment, therefore, is strongly encouraged.

Rating scales frequently employed in assessments by students turn out to merely emphasize student reaction rather than student learning (Arthur, Tubre, Paul, & Edens, 2003; Stark-Wroblewski, Ahlering, & Brill, 2007), and student perception (Arthur et al., 2003). The notion illustrates that an assessment involving students is within the subjective dimension of teaching and learning processes, whereas student learning is within the objective one. Student assessment of teaching is then often deemed to rely on student opinion or 'qualitative judgments' (Mohanty, Gretes, Flowers, Algozzine, & Spooner, 2005, p.139). The same thing is true for some other common types of assessment of effective teaching, such as peer visit (Ackerman et al, 2009).

## **Method**

### **Type of Study**

This study is research and development (Borg & Gall, 1989, p.782) employed to devise a set of instruments to assess an effective English teaching at VSSs. The instrument development was conducted at four VSSs which have the Hotel Accommodation Program in Malang, West Java, Indonesia. The subjects of this study were English teachers at four VSSs in Malang.

### **Settings and Subject of the Study**

This study was conducted at three VSSs in Malang, from January to April 2014. Three accredited VSSs with Hospitality Program from Malang were designated positively as the subjects of the field-testing.

### **Procedure**

#### *Preliminary Study*

This phase was intended to establish the theoretical construct about effective teaching, effective teacher, and teacher effectiveness. The theoretical constructs are absolutely required for instrument development. Literary

studies, and research findings analysis and synthesis were therefore carried out to achieve that purpose. A lot of findings of research on various subjects and levels conducted for several decades were investigated.

#### *Design and Development*

In this phase, particularly the design one, the results derived from the preliminary study were elaborated and analyzed in order to specify instructional objectives of English teaching at VSSs and then design or construct some instruments. The instruments were designed for assessing teaching processes. From the designing process, the dimensions and their items were identified. The items were also elaborated to make a comprehensive picture of the designed instrument. In this way, a pattern of relationship among dimensions, descriptors, and their indicators was finally established.

In the wake of the design process, the development phase was then carried out. The main purpose of this phase was to construct an instrument draft. Experts were heavily involved in devising, revising, and refining the instrument drafts. Importantly, to make the instrument drafts relevant to the field and appropriate for stakeholders and future users, this process also engaged people from various fields, including: four VSS English teachers, four subject specialists, three hotel staff, one hotel manager, and two VSS curriculum specialists. It is noteworthy that from the design and development process, an early prototype of instruments was finally made up. Also in this phase, early modification and refinement of the prototype design of instruments was initiated in light of the involvement of people from a wide variety of fields. All of the experts involved have got doctorate degree majoring in the field relevant to their expertise. They were expected to establish the content validity of the instrument through the Delphi technique and interview.

After the prototype design in the form of assessment instruments was made up, more experts were involved to provide appraisal and judgment about the content validity of the instruments. The instruments were sent to seven experts for validation.

### *Model Implementation*

Upon completion of the modification process, the refined and validated instruments were utilized to collect data. The resulting data were analyzed where unreliable items identified were dropped. The results of using the fixed design and its instruments at different school contexts were finally known and described.

### Data, Instruments, and Techniques for Collecting Data

#### *Data*

The data of this study are a collection of classroom activity records, observation notes, interview notes, expert opinions, interview results, observers' notes, and stakeholder's opinions about the instrument effectiveness.

#### *Instruments*

There were a number of instruments used in this study. The first was a survey, developed to examine the instrument readability from which revisions and modifications were made. A teacher survey was also devised to assess teacher's perception of school facilities in supporting English teaching. Principal and vice-principal's survey was constructed to understand their vision and mission so as to comprehend school context. The second was an observation protocol in the form of observation sheets constructed as an assessment instrument in the form of a rating instrument and thereby used to collect data on teaching and learning processes in the classroom which in turn allowed the writer to judge whether the English teaching was effective or not. To achieve the purpose, teaching behaviors (Alkin, 1992, p.1345; Ornstein & Lasley, 2004, p.196) and classroom interactions between students and teacher and among students were noted in the observation sheet with scoring rubric and narrative summary. A video equipment was also used with teacher's permission in order to capture the classroom dynamics including instructional sequences from the start to the close of the lesson, including teacher behaviors (Shulman, 1986, p.30) and teacher talk.

### *Techniques for Collecting Data*

A number of techniques for collecting data were employed. The first was direct observation conducted to collect data related to the process of English teaching at VSSs within classroom context including teacher behavior, students' activities in the classroom, interaction between teacher and students as well as among students, group interaction, and teacher talk.

The second technique was interview, used during the instrument development and during the measurement process to obtain ideas. The in-depth-interview required key informants to provide important information associated with their knowledge, experience, and perception.

The third technique of data collection was survey which was administered to participating teachers, school principals, and vice-principals. School principals and vice principals were also surveyed to get data on school vision and mission so that the context of English teaching at VSSs could be figured out. The next technique of data collection was Delphie technique, used to gather data from experts to develop the research instruments.

### Data Analysis Technique

Given that this study was sought to develop a set of instruments, the technique for analyzing data had to do with the reliability estimation and setting assessment criteria. A pilot-test was administered in this research because the instrument development is part of this R & D. The reliability value to be obtained refers to Alpha Cronbach's Coefficient  $\geq 0.7$  (Kaplan & Saccusso, 1982, p.90) which is common for most teacher-made tests, and they are useful for making instructional decisions; the correlation between each item should be above 0.3 (Field, 2009, p.98). Besides, pertaining to classroom observation, Cohen's Kappa (Elliott & Woodward, 2007, p.84) was used. In respect, the value is vital to ensure that the items are strongly related which, in turn, ensure that the items of the instrument are internally consistent, meaning that they assess similar and homogenous thing—teaching effectiveness of Hospitality English at VSSs.

Both qualitative and quantitative descriptive analyses were also employed in the light of the nature of the data obtained. Several techniques of analyzing data employed in this respect were percentage, mode, standard deviation, and mean.

### Findings and Discussion

Literature studies and research finding analysis were conducted to construct an appropriate observation instrument for assessing English teaching at VSSs. Thus, a variety of concepts of effective teaching and teaching effectiveness were discussed, analyzed and synthesized. Findings of research published in the forms of journal articles, thesis reports, and dissertation reports were analyzed.

Table 1. Synthesis of effective teaching behavior

<p>The results of synthesis of Effective teaching behavior</p> <p><i>Teacher Dimension</i></p> <ul style="list-style-type: none"> <li>-planning</li> <li>-implementing lesson plan</li> <li>-establishing orientation of learning</li> <li>-organizing and activating knowledge</li> <li>-teaching ways of learning</li> <li>-questioning with strategy</li> <li>-communicating appropriate expectation</li> <li>-using varied teaching</li> <li>-creating supportive climate</li> <li>-setting realistic context</li> <li>-facilitating negotiation of meaning or interaction</li> <li>-assessing communicative ability</li> <li>-presenting subject matter.</li> </ul> <p><i>Student Dimension</i></p> <ul style="list-style-type: none"> <li>-engagement and motivation</li> </ul>
---

### Observation Instrument

With regard to the observation instrument, there were many feedbacks from English teachers and experts in TEFL/TESOL. The feedbacks highlighted the concepts and readability. Appropriate wording, revision on overlapping dimensions, and more-friendly terms were suggested. With regard to dimensions, an evaluation expert suggested the writer to incorporate planning, which was initially considered as an input instead of a process, into the dimension of effective teach-

ing. Importantly, to be able to assess dynamics of teaching as suggested by another expert in evaluation, the implementation of lesson plans, which is sometimes categorized as an input (Perez & Mardapi, 2015, p.10), was also incorporated to complement the dimensions. Thus, there are two more dimensions in the assessment instruments: lesson planning and the implementation.

Based on teachers' feedback, a number of terms and phrases were modified and revised in order to be more user-friendly and unambiguous. *Facilitating negotiation* of meaning, for instance, was changed into *facilitating interaction*. Similarly, *giving contextualized explanation* was modified into *giving concrete explanation*. The same thing was done for several terms such as *setting high expectation*, *information gap activities*, *choices and errors*, *accuracy*, and *fluency*.

### The Results of the Field-Test

#### *Inter-rater Reliability*

During the second-phase field testing, the coefficients of reliability were very low, mostly under 0.7, because the raters found difficulties in using the instrument. The process of teaching was observed by two raters at each school. To calculate consistency of rating between the writer and raters or inter-rater reliability, the formula of IRR (Adamson, 2008, p.195) was used. The formula states that:  $IRR = \frac{\text{the number of agreements}}{\text{the number of agreements} + \text{the number of disagreements}} \times 100$ . Rater 1 is a writer who developed the instrument, while the other one is the English teacher, a peer teacher and an external observer upon the writer's request. There is only one English teacher to be observed in each school.

The observation process was conducted five times for each school. The true calculation of inter-rater reliability was carried out at the last three observations given that the first two observation processes were a learning process in using the instrument for the peer teacher in which anomaly and confusion came about. The reliability study yielded the inter-rater reliability scores of 0.77, 0.82, and 0.85 for School A, and the last 0.62, 0.74, 0.71 for School B.

### Validation of the Instruments

With regard to the instrument validation process, the findings are as follows. Firstly, in terms of procedural validation (Oliveri, Lawless, & Young, 2015), the procedures employed to devise the observation instrument under study were strongly based on the findings of various studies in many countries for several decades. Individual reviews by experts and consultation with experts were also conducted and documented. Besides, the set of instruments was field-tested and revised many times. Upon the field tests, it was found that several terms relatively difficult to understand or not ‘user-friendly’ and, consequently, revised or deleted. Secondly, in terms of content validation, evidence associated with content and comprehensiveness of the construct is available. Included in this evidence is expert judgment.

As part of interpretive argument, the involvement of stakeholders to design the set of assessment instruments in terms of content and procedure of communication processes which are real-world communication, the use of the instrument will assist both teachers and students to understand the real workplace. In terms of consequences as evidence (Kane, 2013, p.54) and also the Assessment Use Argument of Bachman and Palmer (2010, p.112), there are some kinds of evidence used to justify the use of the instrument. The information resulting from the use of the observation instrument is the basis for making teaching and learning processes more effective. Thus, the decisions made upon the results of the observational assessment benefit the teacher and students.

### The Use of the Instrument

#### Assessment of Teaching

Figure 1 shows that the English teacher at School A performed much better than the teacher at School B. The largest gap occurred at the aspects of student engagement and motivation, assessing student communication ability, and communicating high expectation. The lowest gap can be seen at the aspect of asking strategic questions.

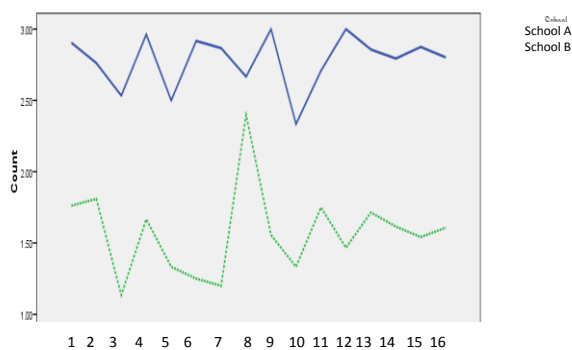


Figure 1. The comparison of process scores between School A and School B

Notes:

- 1 : Planning the lesson
- 2 : Implementing the lesson plan
- 3 : Establishing orientation of learning
- 4 : Organizing and activating knowledge
- 5 : Teaching strategies of learning
- 6 : Using varied instruction
- 7 : Communicating high appropriate expectation
- 8 : Asking strategic questions
- 9 : Facilitating interaction
- 10 : Using realistic context
- 11 : Creating supportive climate
- 12 : Assessing student communicative ability
- 13 : Presenting subject matter
- 14 : Teaching behavior
- 15 : Student motivation and engagement
- 16 : Process

*Planning the Lesson.* As displayed in Figure 2, it is obvious that the English teacher at School A was rated higher than that at School B by both raters. The writer as the first rater scored 280 for School B and 307 for School A.

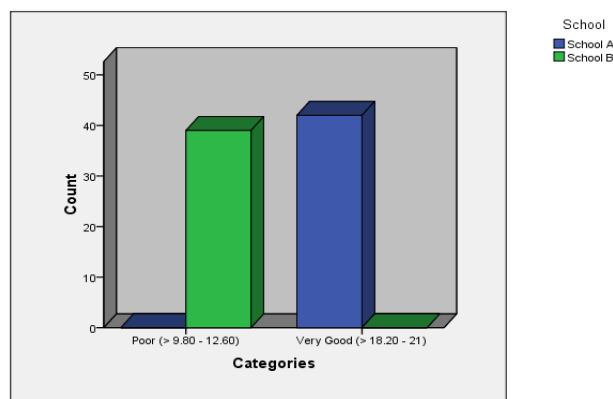


Figure 2. Planning the Lesson



School A: The lesson plan which was used by the teacher was developed in a comprehensive way. The teaching objective was set in the form of student behaviors. What kinds of communication tasks to be performed by the students were also clearly stated, including grouping students into a group of two or more. In order to achieve the objectives involving the groups, guided and independent practices were available.

School B: The teaching objective was stated in the short lesson plan by referring to the topic which was available in the textbook. No specific behavior to perform by students was explicitly stated. The teacher stated that she never used a lesson plan for teaching, except for the study upon the writer's request. No prerequisite knowledge for studying the teaching content was stated. What materials to be given were stated by stating the topic in the textbook and student worksheet or *Lembar Kerja Siswa* (LKS). The same thing was also done for assessment activities and practice or exercises.

*Implementing the Lesson Plan.* Among the three teachers at the three different schools, the teacher of School A was rated higher. It is normal, given that the lesson plans designed or written by the teacher were also more comprehensive and complete. These inevitably affected the scores of implementation.

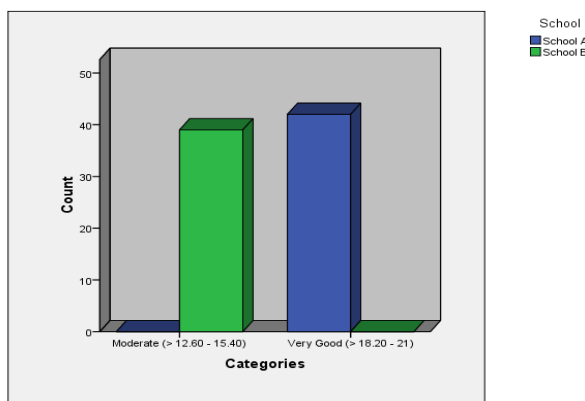


Figure 3. Implementing the Lesson Plan

*Establishing Orientation of Learning.* The orientation of learning is clearly presented in Figure 4. The orientation of learning of both School A and School B is elaborated as follows.

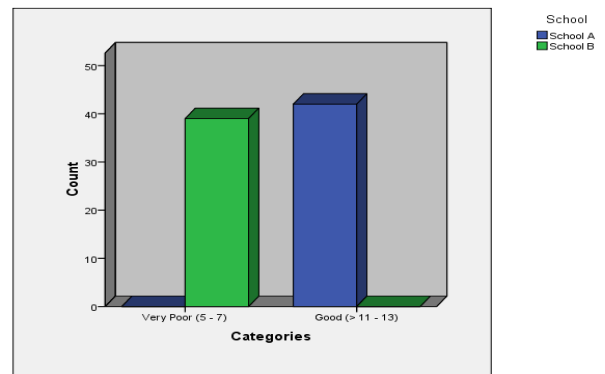


Figure 4. Establishing the Orientation of Learning

School A: The teacher had a problem with the computer for displaying the material for the first ten minutes so that what had been planned could not be implemented in a good fashion. In fact, pre-learning activities were limited to questions about students' opinion of the weather as the teaching activities that day was related to the ways of expressing opinion.

There was no opportunity to explain how the teaching material that day was linked to the previous. Importantly, however, the teacher still mentioned the teaching objective before something wrong happened to equipment. She tried to remind the objective of teaching again in the end part of the teaching through a reflection process. In the second transcript, the teacher began the lesson with pre-activities, that is, by reminding them of the previous lesson. How the new lesson was linked to the points of discussion already presented before was explicitly stated in the introductory statement. Likewise, the objective of the lesson and the reasons for studying the points of the lesson were explicitly stated.

School B: Pertaining to learning orientation, no pre-teaching activities were encouraged by the teacher. No relationship between what was going to be discussed and previous teaching, except for the homework given in the previous meeting.

The teaching objective and the importance of teaching materials were not communicated explicitly to the students. Similarly, in the end part of the teaching, no explanation or discussion about the teaching objective was

provided. In another meeting, the teacher talked about the previous meeting albeit very briefly. In her statement, she found that the students encountered difficulties in error analysis and incomplete sentences. What should be done in the session, therefore, was to continue the part of the unfinished exercise experienced in the previous meeting.

Those statements also imply that the teaching activities were analyzing incomplete sentences but without a clear teaching objective by mentioning student competency and the importance of studying the teaching materials of the day. Likewise, the teaching objective was not explicitly stated in the end of the teaching process.

*Organizing and Activating Knowledge.* The activity of organizing and activating knowledge is clearly presented in Figure 5. In addition, the explanation on the activity conducted in each school is described as follows.

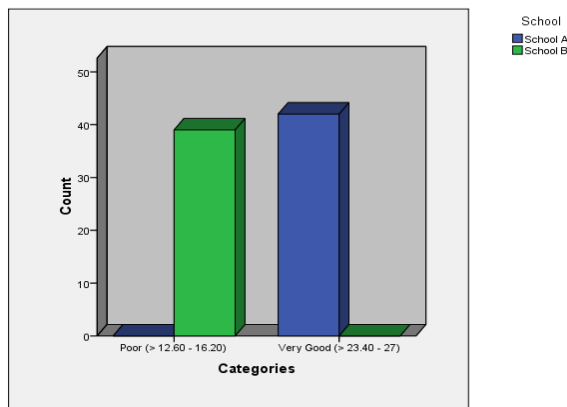


Figure 5. Organizing and Activating Knowledge at Three Schools

School A: With regard to teacher’s effort to organize and activate knowledge, the teacher prepared the students with Power Point teaching material to provide explanation and the worksheet to give students exercises to practice communication. The structure or brief description of the materials was also explicitly stated during the pre-activities and followed with the presentation of expressions usually used in job interviews.

Parts of the materials were presented in a logical fashion: it began with the examples of expressions commonly used in a job interview and then followed with group work to

practice using the expressions with the members of the groups. It is noteworthy that the teacher always used English language fluently and directly followed by Indonesian language.

School B: In terms of teacher’s effort to organize and activate students’ knowledge, the use of LKS (student worksheet) commercially available had been part of the teaching system at the school. Some of them, however, were not ready with their own LKS so that they had to share with their peers. Importantly, no explanation about the structure and description about the lesson except for asking the students to conclude the general topic drawn from items number five, six, and seven.

Repetition was carried out to emphasize important points especially in this teaching and learning process with regard to teacher questions as well as students’ correct answers. It means that the repetition also serves as corrective feedback.

In terms of managing transitions during teaching and learning activities, it is noteworthy that the teacher made smooth transitions except for the opening and the closing. In fact, the transition to the beginning of teaching activity came about very abruptly. Likewise, the classroom activities came to an abrupt end as the time was over; neither conclusion nor summary of key points were presented at the end.

*Teaching Ways of Learning.* The comparison of the teaching ways of both schools is clearly presented in Figure 6. Meanwhile, the details of the comparison are elaborated as follows.

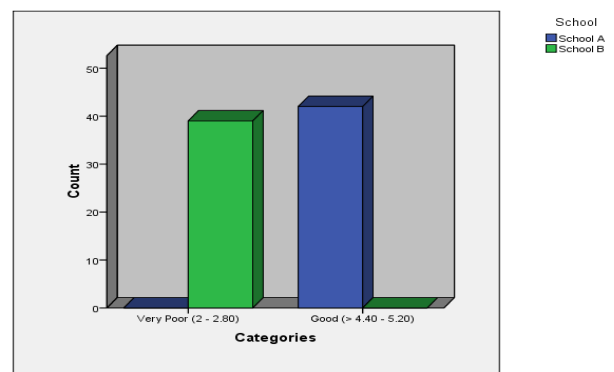


Figure 6. Teaching Ways of Learning

School A: Pertaining to teaching students about specific learning strategies to understand the subject or lesson, the teacher did it only to the low-performing group. She even sat with the special group more than six minutes providing assistance to do the task and then presented the result of their work in front of the teacher. It is important to note that the students in the class finished their interview by the world of hotel few days before. The teacher, therefore, encouraged them to take good opportunities offered such as working in the Front Office in which they were able to learn. The similar steps were also done by the teacher. In the end part of the class, the teacher reminded the students to practice outside what had been done during the session to enhance their English proficiency.

School B: As shown in the transcript, the teacher taught two learning strategies. The first refers to *budaya kamus* or dictionary culture by which every student consults a dictionary when finding difficulties pertaining to word definitions. It seems that the learning strategy had been long taught by the teacher. The second has to do with the strategy of dealing with incomplete sentences. The strategy which was taught by the teacher to deal with such a problem is recognizing or identifying the tense and the meaning of the sentence. During teaching-learning activities in the other session as shown in the Transcript, however, the teacher never taught or explained specific effective learning strategies to be used in the classroom. The teacher just had the students continue doing exercises as part of previous meeting exercises. Having provided explanations about hotel reservation as it was related to the exercise items, the teacher gave a formative test.

*Conducting Varied Teaching.* The condition of conducting varied teaching in both schools is presented in Figure 7. The elaboration of such a condition in both schools is as follows.

School A: Regarding the teaching variety, in the first transcript, various teaching strategies were employed. The strategies allowed the students to be actively involved in the

teaching and learning activities. The strategies are lecturing, imitating dialogues, pair-work, reading aloud, analyzing patterns, rearranging jumbled sentences in a paragraph, and question-answer most of which are teacher-centered.

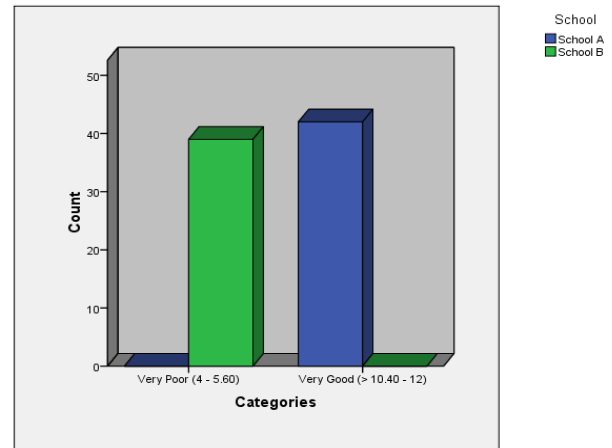


Figure 7. Conducting Varied Teaching

Similarly, some varieties were employed in the second transcript. They are asking questions, studying expressions displayed on a large screen, conducting conversation in the group work, and peer-assessment. Visual media in the form of Power Point presentation using a large LCD was used to explain expressions commonly used in job interviews. The visual material served as an authentic material as the students were offered real conversation in job interviews. Student’s work sheet also served as an authentic material since it allowed the students to have communicative activities.

School B: In an effort to create effective teaching, the teacher varied the teaching-learning activities. To achieve the purpose, she did different instructional activities. They are silent reading, doing exercises taken from *LKS*, brain-storming activities along with answering the exercise items pertaining to hotel reservation activities and hotel facilities, and a formative test. No multimedia equipment was used in this class to enhance the teaching effectiveness. Actually the teaching materials are authentic given that it is related to hotel reservation, meaning that the teaching material could have facilitated authentic interaction.

*Communicating High/ Appropriate Expectation.* The comparison of both schools' condition in terms of communicating high/ appropriate expectation is presented in Figure 8. The elaboration is also explained as follows.

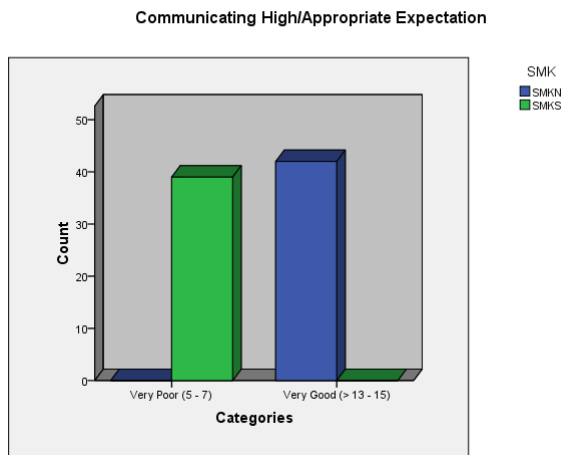


Figure 8. Communicating Appropriate Expectation

School A: To ensure the students that they were able to do the task, the teacher approached each group with enthusiasm that they were able to accomplish the difficult task. She kept staying with the group until the group members signaled confidence to do the task. Moreover, as discussed in the previous section, the teacher approached and stayed longer with a low performing group, provided assistance and advice to the group members, even asked them to make presentation in front of the teacher. With her enthusiasm and assistance, she indirectly ensured the students that they had the capability to master the subject. Likewise, the students were encouraged to express and communicate their ideas through group work. In the end of the session, the teacher also communicated her high expectation by saying that the students were actually capable of speaking English fluently.

In the first transcript, the teacher consistently did similar things, mostly at the end part of the session. The students were encouraged to ask questions. Importantly, although high expectations were not consistently communicated since the beginning of the lesson in an explicit fashion, the teacher's behavior which made students actively engaged implies that she had very high expectation.

School B: During the teaching-learning processes, no high expectation was orally communicated by the teacher particularly by having assignment, English mastery, and good English proficiency. In other words, the teacher did not say anything to convince the students that they can accomplish difficult tasks and have good command of English. However, the ways the teacher encouraged the students to answer the questions of the exercises correctly and express their ideas during question-answer activities and brainstorming indicated that the teacher held high expectation. The same thing was also shown in another session in which the teacher determined that the students had to be able to answer five items before moving to a more difficult topic.

It is important to note, however, that the teacher also showed her 'self-fulfilling prophecy' particularly to those deemed as low performing students. In fact, the teacher did not help the students to write English sentences when they tried to answer in English. As a result the students could not put forward their ideas completely.

*Asking Strategic Questions.* The different acts of asking strategic questions of both schools is presented in Figure 9. The details are elaborated in the following explanation.

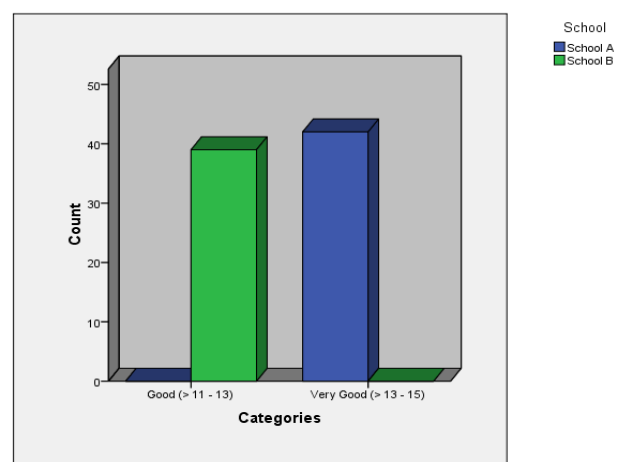


Figure 9. Asking Strategic Questions

School A: A lot of questions were addressed both to the whole class and more particularly to each group in her effort to make her students actively engaged in the

lesson. In the first transcript, ample times were provided to the students, particularly to those in group discussions which the teacher consistently approached during the session.

In the second transcript, a few questions were addressed to activate the whole class in the form of chorus answers. The product questions, however, tended to dominate. The teacher did not address probing questions when she should have done that and when the students' answer was very short.

Feedbacks were also given by the teacher every time a student or students answered her questions correctly or incorrectly. Expressions such as 'Good, OK', 'Yes, you are right', 'Right', 'Yea, asking about opinion...', 'Yes, there are too many people', 'No no no, not that one' were the most common feedbacks used by the teacher.

School B: As shown in the transcript, the teacher asked questions to activate the students during the teaching and learning processes. She tended to use product questions, instead of process questions. Actually she tried to use process questions, but the students always gave very short answers.

It is quite often that the teacher did not provide the students ample waiting time to think about the answers. The teacher also did not forward unanswered questions to other students. As a matter of fact, most of the teacher's questions were directed to the whole class.

*Facilitating Interactions.* Both schools have differences in facilitating interactions. Figure 10 presents the comparison and the explanation is elaborated as follows.

School A: As to interactive communication, the dominant activities of teaching and learning processes, as shown in the first transcript, were communicative activities which made the class student-centered. There were two kinds of communication performed in the class. The first is teacher-student communication where the teacher dominantly used English to explain and ask questions. This kind of communication succeeded in inviting the students to pay attention to the teacher's explanation and respond to the teacher's questions. The second one is student-student

communication as a pair-work or activity assigned by the teacher where they were required to make a group of two. Through the pair-work the students were asked to give their opinion and to give agreement or disagreement based on some statements provided.

Information gap activities and negotiation of meaning were implemented in the classroom. These all allowed the students to make choices and errors in a natural communication. Students' efforts to communicate were also highly appreciated through group discussion. Teacher's appreciation was apparent when she approached and stayed with each group during discussion.

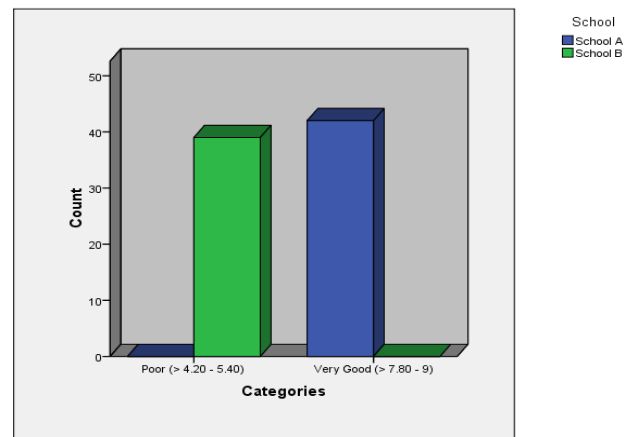


Figure 10. Facilitating Interactions

School B: The negotiation of meaning through information gap activities was limited to the interaction between the teacher and students, particularly when the teacher asked questions in English along with brainstorming activities. These information activities were very limited given that the teacher used similar patterns and the students gave answer in the forms of words and phrases instead of sentences which could have created meaningful teacher-student communication. In another session as shown in Transcript, there were no information-gap activities in English as the emphasis was on grammatical materials. Interestingly, two students actually tried to communicate in English that is to explain their reasons for choosing the answers, but their efforts were in vain. They got no help from the teacher when they found difficulties expressing their ideas.

*Creating Conducive Climate.* The comparison of the efforts in creating conducive climate in both schools is presented in Figure 11. The description of the condition in each school is also elaborated as follows.

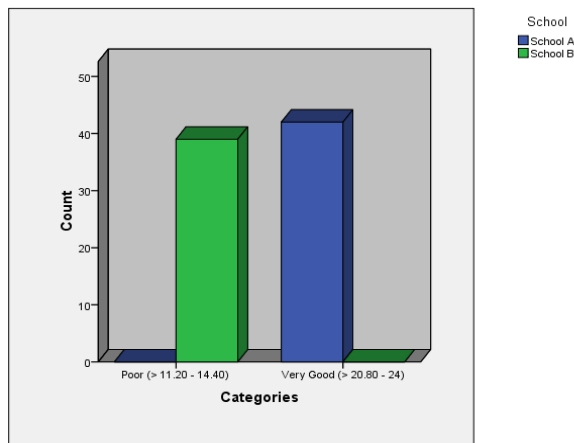


Figure 11. Creating Conducive Climate

School A: It is obvious that during the teaching and learning processes the teacher consistently communicated enthusiasm in front of the class through her voice, her facial expression, and body language, as consistently shown. She approached every group more than once to monitor their activities and provided help dealing with language used in the group discussion and interaction. Her friendliness was so obvious. The ways she handled the class also led to supportive learning atmosphere where the students seemed to feel free to express their ideas. She showed her ability in maintaining classroom discipline, even since the beginning of the class.

Interestingly, she also showed good use of humor to make students understand about how to use scoring rubric of peer-assessment. She expressively showed the way an incompetent speaker speaking English such as ‘e...e...e...’ after which several students repeated while laughing and let them know what to score such competency in the rubric ranging from 1 to 4.

In an effort to make her students understand better, she also came to a group and made special expressions and body language using her hands and face which made members of a group laugh. In other words, she was always expressive in explaining, with

very clear and loud voice. Eye contact, therefore, was consistently maintained during the teaching and learning processes.

As part of building a supportive climate, she appropriately gave special treatment to some students in response to their nonverbal cues indicating confusion. She did this particularly to a low performing group that seemed to be off the task and gave their confused look to the teacher. It must be noted, however, that she showed limited use of students’ ideas when having communication with the whole class, particularly during the reflection process. Actually she had a lot of opportunities to do that since the pre-activities with the whole class. With the groups, however, she seemed to use students’ ideas much.

School B: As shown in the recording, during the teaching and learning processes the teacher did not exhibit great enthusiasm in front of the class through her voice, her facial expression, and body language. It can be seen obviously that she never left her desk. Furthermore, she did not approach the students to monitor their activities. She did all things from a distance.

The ways she handled the class failed to create supportive learning atmosphere where ideas expressed or answers made by the students were difficult to catch, or in other words, were unheard because of the noise. The relatively large number of the students in the class on the one hand and the teacher’s attention and gaze limited to those sitting in front on the other contribute to the atmosphere. It stands to reason that she was confronted with classroom management problems most of the time. As a matter of fact, many students talked to each other, and some of them answered one after another without raising their hands.

It is noteworthy, however, she performed several positive actions which made the students actively engaged. First, she always used students’ ideas. As shown in the previous sections, she often repeated students’ answers. Second, she was sensitive to nonverbal gestures particularly confusion. Some good responses were shown by the teacher to deal with the problem.

*Using Realistic Context.* Figure 12 shows the comparison of both schools in using the realistic context. The elaboration is also presented as follows.

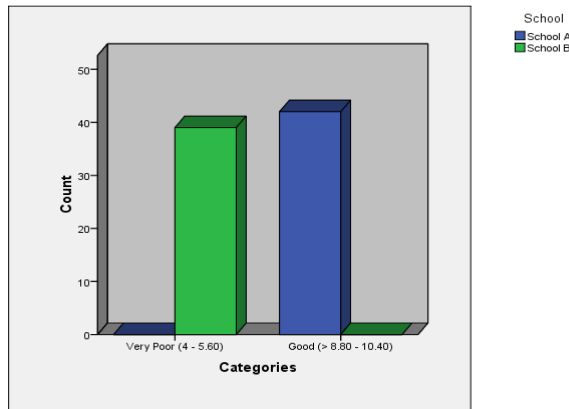


Figure 12. Using Realistic Context

School A: Regarding the realistic context, the interaction activities were not explicitly linked to the world of hotel. It is noteworthy, however, that good skills in expressing agreement or disagreement and giving opinion as taught and practiced during the teaching and learning activities are strongly needed in any field, more particularly in hospitality in which communicative skills are highly demanded. For that purpose, specific expressions were taught and explained. The use of visual aids in the form of power point presentation in which the expressions were shown interestingly with lovely pictures and displayed on a large screen in front of the class proved to make the classroom activities effective.

The second one also demonstrates that the interaction between the students in the form of job interview strongly led to hotel and real context given that such kind of interview is actually required to enter the hotel world, even for students applying for an internship program at star hotels. Before the students were involved in the job interview activities, the teacher gave and taught them particular idioms and expressions.

Visual aids and relevant pictures were also used and displayed to make the expressions and idioms more understandable. The visual aids were presented in the form of Power Point Display on which the expres-

sions along with pictures depicting the expressions were displayed on a large LCD screen on the wall.

School B: A limited hotel context was used during the interaction process between the teacher and student, particularly when the teacher asked the students about types of rooms and their facilities. In addition, some dialogues were presented in the previous section, another dialogue is as follows.

Relevant to hotel contexts, specific idioms and vocabulary items related to hospitality were taught. They are types of rooms (standard, deluxe, family), hotel facilities (single/double/triple bed, standing shower, living room, bathtub, mini-bar, international channel, and air conditioner), and specific verbs (to confirm reservation, to have air conditioner, and so on). In another meeting, however, completely different things came about.

*Assessing Communication Ability.* The comparison of the activity of both schools in assessing communication ability is shown in Figure 13. The explanation of what has been conducted in each school related to assessing communication ability is also presented as follows.

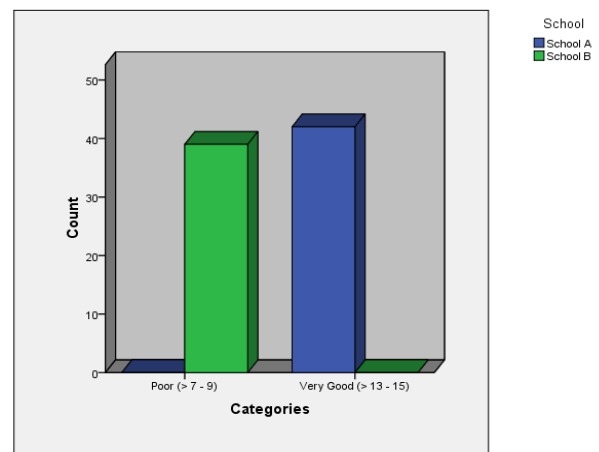


Figure 13. Assessing communication ability

School A: The teacher continuously came to each group, including a low-performing group, to monitor the students' activities and provide assistance to students in the group. She also provided answers to students asking her questions. It is important to note that the teacher never sat on her desk

during the teaching and learning processes, except when she prepared her teaching materials as soon as she entered the class.

Authentic assessment was implemented to her students. The assessment allowed her students to perform interactive communication in English with their partner or group members. The assessment for learning takes the forms of job interview conversation and expressing opinion and agreement or disagreement. Corrective feedbacks were always given, consistently, during teaching and learning activities.

School B: Monitoring student work and activities was carried out for the whole class in light of her position which was never away from her laptop. Assessment, however, was conducted by answering the questions based on a dialogue although they did not get opportunity to get engaged in active interaction using the hotel-real context. These information-gap activities were limited to teacher-student interaction, without student-student communication involving negotiation of meaning.

In the teacher-student interaction, corrective feedbacks were given, but not consistently, by the teacher: Corrective feedbacks were not given when students made mistakes in pronouncing several words such as service, separate, standard, superior, suit, *apotik*, and restaurant, including minor mistakes in pronouncing 'reservation, hotel, motel, method'. Corrective feedback in grammatical errors or mistakes did not take place because the students did not make any statements in English. Instead, they tended to speak in Indonesian language, except for words and phrases.

No authentic assessment was done; the focus was on form, that is, grammatical exercises. Grammatical corrective feedback would be given in the other day because the exercises were submitted. Corrective feedback focusing on pronunciation was not carried out. The teacher did not do anything to correct student's pronunciation of 'come' incorrectly. The dominant use of Indonesian and Javanese languages as a medium of instruction also contributed to this aspect.

School B: Giving various questions to students or the whole class is the only way the

teacher conducted to monitor student learning. The questions were put forward based upon the items in the LKS and as part of brainstorming processes to elicit students' understanding. She never approached students sitting away from her, more particularly those around the back of the classroom, due to her position which was always at her desk.

In terms of accuracy, no corrective feedback was given by the teacher. No English interaction between teacher and students and among students themselves is among the factors. The teacher made grammatical errors many times and these brought about students' confusion. The responsiveness of the teacher prompted her to revise her own errors right away. Many students look confused with the question by saying after the teacher.

Likewise, students' mistakes and errors dealing with fluency were not corrected either. In fact, they incorrectly pronounced very common words related to hotel industry such as restaurant, standard, email, and air conditioner. The mistakes came about when they answered the teacher's question during brainstorming and when they were reading the text. Usually the rest of the class laughed at a student making mistakes in pronunciation.

*Presenting Subject Matter.* Figure 14 presents the comparison of both schools in presenting subject matter. The details are elaborated as follows.

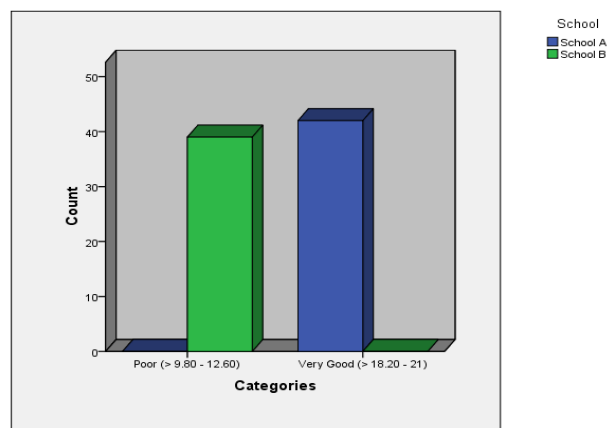


Figure 14. Presenting subject matter

School A: One of the strengths possessed by the teacher is that she was very articulate. In other words, she explained mate-



rials and expressed her ideas both in English and in Indonesian in a very clear voice and correct pronunciation, except for the intonation. She always came to students raising their hand for questions and then listened to their questions patiently and attentively. Demonstrating a good command of the teaching content and good proficiency of English allowed her to be able to answer questions put forward by her students in a very good way. Her appropriate body language also helped to make her teaching effective.

School B: Interestingly, as far as teaching content is concerned, the teacher's mastery of hospitality and hospitality English deserves special attention and appreciation. She knew well about different types of rooms at hotels, facilities, reservation process, and hotel amenities. Her good knowledge of hotel industry made up an interesting teaching and learning processes because the English teacher also played the role of a specialist.

*Student Engagement and Motivation.* The comparison of both schools' condition in terms of student engagement and motivation is clearly shown in Figure 15. The explanation is also presented as follows.

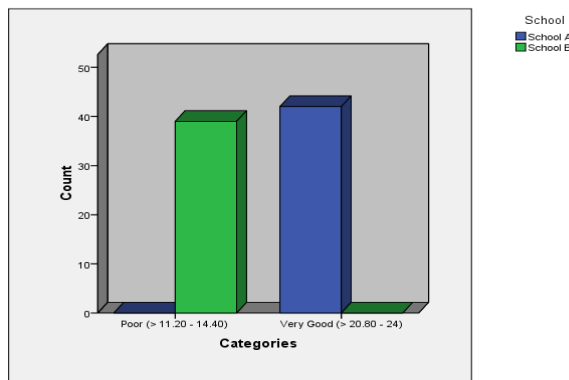


Figure 15. Student engagement and motivation

School A: The teaching and learning processes which were student-centered undoubtedly allowed most of students actively engaged and always on-task most of the time. They were very active in their group work or with their pairs in pair work. There was only one-low performing group. They seemed difficult to be on task. Teacher's good responsiveness, therefore, proved to be effective to

make them take part in activities along with their peers.

Several students were also active asking questions while they were working with their groups. Several times they asked about specific terms in English, how to state specific expressions in English, and comparison between several expressions pertaining to opinion, agreement, and disagreement.

School B: Looking at the teaching-learning process as whole, it is obvious that the teacher was very dominant. The teacher's position which was always at her desk and teacher talk that was so apparent during the session made the class teacher-centered. The voices of many students which deserved special attention were often unheard. It is quite often, therefore, the students got no teacher's feedback for their responses.

Neither interactive task nor group work was assigned and thereby no opportunity for expressing ideas in groups was offered to students. As it was teacher centered, the opportunity to express ideas was largely dependent upon teacher talk. Moreover, the seating arrangement, the number of students, and the teacher's gaze proved to make such opportunity harder to get. Students' participation, unsurprisingly, tended to come and go.

In the end part of the class, the students were not involved in reviewing the teaching objective. In fact, there was nothing under review, including summary or conclusion of the teaching content. The class came to end very abruptly as the bell rang.

## Conclusions and Suggestions

### Conclusions

The effective English teaching at VSSs has fourteen indicators: designing lesson plan, implementing lesson plan, organizing and activating knowledge, doing varied teaching, teaching strategies of learning, communicating appropriate/high expectation, using realistic context, asking strategic question, facilitating interaction, assessing student communication ability, presenting subject matter, student engagement and motivation.

From several observations which were conducted at the participating schools, the

observation instrument has acceptable inter-reliability coefficient. More training to use the instruments and make sense of key words and concepts on the instrument is likely to increase the coefficient.

Validation by arguments proved that the developed instrument was a valid representation of effective teaching. More evidence is likely to improve its validity. Based on the application of the assessment instruments and feedbacks from teachers and school managers as stakeholders, the instruments are considered effective or successful to achieve the desired results. The observation instrument which highlights relatively many aspects with a lot of items takes much time and concentration in its implementation.

#### Suggestions

Based on the findings of the study, some ideas are suggested. First, it is suggested that English teachers at VSSs conduct self-assessment to assess their teaching effectiveness by using the developed instrument in a regular fashion to make continuous improvement which leads to the enhancement of students' skills in hospitality English. Second, it is strongly suggested that evaluators and English teachers comprehend a wide variety of English teaching approaches and methods and keep abreast the development of English language teaching. Third, to all users of the instrument, it is very important to take into account the number of items in the observation instruments in the light of time and observer's concentration. Focusing on two or three aspects of teaching is strongly suggested. Fourth, it is suggested to use the observation instrument to get more evidence as a basis of argument for further validation.

#### References

- Ackerman, D., Gross, B.L., & Vigneron, F. (2009). Peer observation reports and student evaluations of teaching: Who are the experts?. *The Alberta Journal of Educational Research*, 55, 18-39.
- Adamson, L.S. (2008). *Development and evaluation of an instrument to assess data informed instructional practice* (Doctoral dissertation, John Hopkins University, 2008). Ann Arbor, MI: ProQuest/UMI. (Publication No. AAT 3309798).
- Alkin, M.C. (Ed). (1992). *Encyclopedia of educational research* (6<sup>th</sup> ed.). New York, NY: Macmillan.
- Arthur, W., Tubre, T., Paul, D.S., & Edens, P.S. (2003). Teaching effectiveness: The relationship between reaction and learning evaluation criteria. *Educational Psychology*, 23, 275-285.
- Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. & Palmer, A. (2010). *Language assessment in practice*. Oxford: Oxford University Press.
- Borich, G.D. (2007). *Effective teaching methods* (6<sup>th</sup> ed.). Upper Saddle River, NJ: Prentice Hall.
- Borg, W.R. & Gall, M.D. (1989). *Educational research: An introduction* (4<sup>th</sup> ed.). New York, NY: Longman.
- Brown, H.D. (2007). *Principles of language learning and teaching* (5<sup>th</sup> ed.). White Plains, NY: Pearson Education.
- Burhanuddin. (2008). Mengeliminasi kesulitan siswa menjawab soal ujian nasional bahasa Inggris melalui pembelajaran. *Pendidikan Network*. Retrieved from: <http://research-engines.com/burhanudin20708.html>.
- Carruth, P.J. & Carruth, A.K. (2003). Evaluation of college faculty: What do accounting students really think?. *Journal of College Teaching and Learning*, 9, 1-8.
- Cashin, W.E. (1990). *Student ratings of teaching: Recommendations for use. Center for Faculty Evaluation and Development*. Kansas, KS: Kansas State University.
- Creemers, B.P.M. & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. London: Routledge.
- Curtis, A. & Cheng, L. (2007). TAPPING into teaching effectiveness: A collaborative

- approach to performance appraisal of ESL teachers in a Canadian context. In Coombee, C., M. Al-Hamly, Davidson, P. et al. (Eds.). *Evaluating teacher effectiveness in ESL/EFL contexts*. Ann Arbor: The University of Michigan Press, 57-72.
- Djojonegoro, W. (1998). *Pengembangan sumber daya manusia melalui SMK*. Jakarta: Depdikbud.
- Elliott, A.C. & Woodward, W.A. (2007). *Statistical analysis with spss examples quick reference guidebook*. Thousand Oaks, CA: Sage.
- Emery, C.R., Kramer, T.R., & Tian, R.G. (2003). Return to academic standards: A critique of student evaluations of teaching effectiveness. *Quality Assurance in Education*, 11, 37-46.
- Field, A. (2009). *Discovering statistics using SPSS*. (3<sup>rd</sup> ed.). London: SAGE.
- Gower, R., Philips, D., & Walters, S. (1995) *Teaching practice handbook* (New ed.). Oxford: Macmillan Education.
- Kaluge, L. (2003). Inequity progress of pupil groups of upper primary schools. *Jurnal Ilmu Pendidikan*, 10 (3), 237-253. <http://journal.um.ac.id/index.php/jip/article/view/242>
- Kaluge, L. & Hadiwijaya, A.S. (2007). Good practices among current educational programs for basic education in Indonesia. *Jurnal Ilmu Pendidikan*, 14 (3), 144-153. Retrieved from <http://journal.um.ac.id/index.php/jip/article/view/34/310>
- Kane, M. (2006). Validation. In Brennan (Ed.). *Educational measurement* (4<sup>th</sup> ed.). Westport, CT: American Council on Education and Praeger, 17-64.
- Kane, M.T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50, 1-73.
- Kaplan, R.M. & Saccusso, D.P. (1982). *Psychological testing: Principles, application, and issues*. Monterey, CA: Brooks/Cole.
- Kyriakides, L. & Creemers, B.P.M. (2009). The effects of teacher factors on different outcomes: Two studies testing the validity of the dynamic model. *Effective Education*, 1 (1), 61-85.
- Littlewood, W. (2011). Communicative language teaching. In Hinkel, E. (Ed.). *Handbook of research in second language teaching and learning volume II* (pp.541-557). New York, NY: Routledge.
- McCaffrey, D.F., Koretz, D., Lockwood, J.R., & Hamilton, L.S. (2003). *Evaluating value-added models for teacher accountability*. Santa Monica, CA: RAND.
- Ministry of Education and Culture. (2015). *Indonesia education strategic plan 2015-2020*. Jakarta.
- Mitzel, H.E. (Ed). (1982). *Encyclopedia of educational research* (5<sup>th</sup> ed., Vol. 4). New York, NY: The Free Press.
- Mohanty, G., Gretes, J., Flowers, C., Algozzine, B., & Spooner, F. (2005). Multi-method evaluation of instruction in engineering classes. *Journal of Personnel Evaluation in Education*, 18, 139-151.
- Musthafa, B. (2001). Communicative language teaching in Indonesia: Issues of theoretical assumptions and challenges in the classroom practice. *TEFLIN Journal*, 12, 184-193.
- Nougaret, A.A., Scruggs, T.E., & Mastropieri, M.A. (2005). Does teacher education produce better special education teachers?. *Exceptional Children*, 71, 3, 217-229.
- Oliveri, M.E., Lawless, R., & Young, J.W. (2015). *A validity framework for the use and development of exported assessments*. Princeton, NJ: ETS.
- Ornstein, A.C. (1991). Teacher effectiveness research: Theoretical considerations. In Waxman, H.C. & Walberg, H.J. (Eds.). *Effective teaching*. Berkeley, CA: McCutchan, 63-80.
- Ornstein, A.C. & Lasley II, T.J. (2004). *Strategies for effective teaching*. (4<sup>th</sup> ed.). Boston, MA: McGraw-Hill.

- Oxford, R.L. (1990). *Language learning strategies what every teacher should know*. Boston, MA: Heinle and Heinle.
- Perez, B., & Mardapi, D. (2015). Evaluation of the bridging course offered at a university to foreign students: Batches of 2012 and 2013. *Research and Evaluation in Education*, 1(2), 146-157. doi:<http://dx.doi.org/10.21831/reid.v1i2.6667>
- Song-Ae, H. (2005). Good teachers know where to scratch when learners feel itchy: Korean learners' view of native speaking teachers of English. *Australian Journal of Education*, 49, 2.
- Stark-Wroblewski, K.S., Ahlering, R.F., & Brill, F.M. (2007). Toward a more comprehensive approach to evaluating teaching effectiveness: Supplementing student evaluations of teaching with pre-post learning measures. *Assessment and Evaluation in Higher Education*, 32, 403-415.