SELF-ASSESSMENT AND PEER-ASSESSMENT FOR STUDENTS’ ESSAY ASSESSMENT IN THE AUFSATZ II COURSE

Tamara Carolin Fitrisia and Primardiana Hermilia Wijayati
Universitas Negeri Malang, Jawa Timur, Indonesia
Jalan Semarang 5, Malang 65145, Indonesia
Corresponding Author: primardiana.hermilia.fs@um.ac.id

Abstract: Self-assessment and peer assessment is an essential alternative assessment in this 5.0 era and in this pandemic time. Thus, this study objectives are to examine differences between self-assessment and peer assessment, as well as to describe students’ perceptions of the application of self-assessment and peer assessment in the Aufsatz II course. This study uses a correlational research design. From the results of the study, it is known that the self-assessment and peer assessment of Aufsatz’s writing ability has some differences in the aspects of content and coherence, while self-assessment and peer assessments for the aspects of vocabulary, orthography, and grammar have no significant differences. Students still find some difficulties in making self-assessment and peer assessments, and they prefer teacher assessments to self-assessment and peer assessments.

Keywords: self-assessment, peer assessment, writing, essay

The most dominant characteristic of the learning process is interactive. So, to stimulate students interaction in teaching learning process during this pandemic is, one of which, by providing opportunities for students to assess their own work, through self-assessment and peer-assessment. The Self-assessment process is an internal evaluation process carried out by oneself and is aimed to identify strengths and weaknesses of oneself (Yan & Brown, 2017). The function of self-assessment is as a means of self-reflection on one’s own strengths and weaknesses in order to determine learning needs, to overcome weaknesses and to improve achievement (Qurbanov, 2016). This assessment method is used to settle on the attitude/ affective of an individual according to themselves (Betella & Verschure, 2016). In addition to self-assessment, there is also peer-assessment.

Peer assessments are usually carried out by people with the same status in work and education as the assessee and the assessor (Panadero, 2016). In addition, peer assessments are considered effective and efficient for students. In peer assessment, a person should reflect on the quality of the work of his peers based on criteria and use it to decide a standard by means of peer performance (Rotsaert et al., 2017) (Wuli et al.,
Self-assessment and peer assessment are important for learning because the purpose of learning is high student achievement (Wijayati et al., 2013).

The functions of self-assessment and peer assessment according to (Willey & Gardner, 2010) and (Wahyuni & Ibrahim, 2012) are regarding training student traits is to teach them as being honest, critical, independent, responsible, tolerant and self-confident; to increase motivation in learning; to save time and money and to teach subjectivity to the students because they reflect independently by clear criteria. However, this function can occur when self-assessment and peer assessments are applied repeatedly because these assessments are not instantaneous. This is due to students’ lack of language skills and their inexperience in assessing an essay. Therefore, students tend to judge subjectively because they are not used to it. In order for students to be familiar with this assessment, the one needs to be applied according to the correct procedure.

So as to the assessment to be applied properly and to minimize subjectivity, the assessment needs to be guided by clear assessment criteria (assessment rubric). Qurbanov (2016) states that the assessment rubric can be used to minimize subjectivity and to maximize objectivity. Based on the results of this study, it is important to conduct research that uses a clear assessment rubric both on self-assessment and peer assessment.

By using rubric, the assessment becomes more focused. Research on self-assessment and peer assessment can be applied to all fields, including the field of language. In German Literature Department, writing courses are set in Aufsatz II. One form of writing that is trained in the Aufsatz II course is writing essays. In this course, the performance of self-assessment and peer assessment can be carried out to assess the results of student essays which has a language level of B1 according to the CEFR standard.

German language assessment standards on writing skills according to Köster (2013) are content fulfillment (Erfüllung), text structure coherence (Kohärenz), vocabulary (Wortschatz) and grammatical structure or word building (Structural). In addition, the general assessment standards for foreign languages according to Marhaeni (2013) are content fulfillment (revising/content), language style (style), vocabulary (Wortschatz), technical and orthographic matters (mechanics/editing) and grammatical structure (grammar). In accordance with these assessment standards, Alfaki (2015) states several problems that are generally experienced by students when writing. A writer is not necessarily fluent in the writing process, sometimes he finds some problems in producing an essay. Problems in writing often include grammar, sentence coherence, word choice problems and cognitive problems consisting of problems with punctuation, capitalization, spelling, content/content and chronology.

Research on Self-assessment and peer assessment has been done a lot, such as in Marhaeni (2013) which has carried out self-assessment in English courses. In addition, research on peer assessment has also been carried out by Sriyati et al. (2016) in biology learning. However, the research on self-assessment and peer assessment in learning German has not been widely studied, including the assessment by using assessment rubric. Therefore, this study was conducted with a focus on the assessment of essays in the Aufsatz II course. The course Aufsatz II was chosen because there was not much research on self-assessment and peer assessment on the aspects of writing German essays which employed a clear assessment rubric.

**METHOD**

This study used correlational research design, with a sample of students of the German Literature Department, class of 2018, Universitas Negeri Malang who took the Aufsatz II course in the 2019/2020 academic year. The data collection technique in this research used three research instruments, i.e. self-assessment instruments, peer assessment instruments and questionnaire sheets. Self and peer assessment instruments were used to gather information about students’ writing skills in terms of content, coherence, vocabulary, orthography and grammar aspects. The questionnaire sheet is used to describe students’ perceptions of the two assessments in the Aufsatz II course.

The self-assessment instrument applied a combination of assessment standards from Marhaeni (2013) and Köster (2013), i.e. content, coherence, vocabulary, orthographic editing and grammatical structure. The content assessment aspect consists of identifying content, structure of
language functions and points and problems in terms of content, while the coherence aspect consists of analyzing language style, coherence in text and problems in terms of coherence. In the vocabulary aspect, the assessment consists of identifying the use of help, diversity of vocabulary and vocabulary problems. The aspect of orthographic editing consists of technical analysis of writing as well as problems in editing and finally the grammatical aspect which consists of identification of grammar and problems in terms of grammar.

The peer assessment instrument refers to Windarsih (2016) which consists of three parts. The first part is the essay analysis as a whole which has the same assessment as the self-assessment except for the problems at hand. In parts two and three, those are correcting essays and explaining the results of corrections which are different from self-assessments. This is because in this stage the feedback function on peer assessments is to correct the results of their friends’ essays. Then for the questionnaire sheet refers to Ernani (2012) which consists of self-assessment, peer assessment, characteristics, strengths, weaknesses, conclusions and opinions about self and peer assessments.

The results of the self and peer assessment were then tested using the Mann Whitney test to determine whether or not there was a significant difference between the two assessments. Of the five aspects of the assessment—content, coherence, vocabulary, orthography and grammar—are tested to prove the hypothesis. Then the results of the questionnaire on students’ perceptions on self and peer assessments were illustrated using a diagram and then the results were explained.

FINDINGS AND DISCUSSION

Findings

To answer the objectives of the first and second research, that is testing the differences between the results of self-assessment and peer evaluation on the results of the essays in the Aufsatz II course and describing students’ perceptions of these assessments, the Mann Whitney test was used. The following describes the Mann Whitney test result data from five aspects of the assessment, i.e. content, coherence, vocabulary, orthography and grammar.

The following are the results of the Mann Whitney test from content aspect.

Table 1. Mann Whitney Test for Content Aspects

<table>
<thead>
<tr>
<th>Essay Points</th>
<th>Mann-Whitney U</th>
<th>Wilcoxon W</th>
<th>Z</th>
<th>Asymp. Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>260.000</td>
<td>963.000</td>
<td>-4.734</td>
<td>.000</td>
</tr>
</tbody>
</table>

a. Grouping Variable: Assessment

In Table 1 it can be seen that Asymptotic Significance is 0.000, which means that there is a difference in results between self-assessment and peer assessment for the content aspect. If you look at the results of the self-assessment and peer assessment, the results of peer-to-peer assessments are better than the results of self-assessments. The content referred to in the iki aspect includes identifying the content, language function and points and structure of the text.

To find out the differences in self-assessment and peer coherence aspects, the data is presented as follows.

Table 2. Mann Whitney Test for Coherence Aspects

<table>
<thead>
<tr>
<th>Essay Points</th>
<th>Mann-Whitney U</th>
<th>Wilcoxon W</th>
<th>Z</th>
<th>Asymp. Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>443.500</td>
<td>1146.500</td>
<td>-2.816</td>
<td>.005</td>
</tr>
</tbody>
</table>

a. Grouping Variable: Assessment

In Table 2 it can be seen that the Asymp. Sig less than 0.05. This shows that on the self and peer assessment sheet, for the coherence element, the value is higher when assessed by peers than when assessed by themselves. This coherence aspect is in the form of language style analysis and coherence in the text written by students.

The third aspect that is assessed is the vocabulary aspect. The following shows the results of the Mann Whitney test.
Table 3. Mann Whitney Test for Vocabulary Aspects

<table>
<thead>
<tr>
<th>Essay Points</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>528.000</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>1231.000</td>
</tr>
<tr>
<td>Z</td>
<td>-1.763</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.078</td>
</tr>
</tbody>
</table>

a. Grouping Variable: Assessment

Different from the previous Mann Whitney test, the results of this test show the Asymp Sig. 0.078 or greater than 0.05, which means that the hypothesis is rejected. The results of the analysis in table 3 show that there is no difference between self and peer assessment for the vocabulary aspects of the essay. The things that students assess on the vocabulary aspect are identifying vocabulary such as the use of word classes or language selection (formal and informal).

Next is the presentation of the results and discussion of the Mann Whitney test on orthographic aspects in table 4 below.

Table 4. Mann Whitney Test for Orthographic Aspects

<table>
<thead>
<tr>
<th>Essay Points</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>605.000</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>1308.000</td>
</tr>
<tr>
<td>Z</td>
<td>-.955</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.340</td>
</tr>
</tbody>
</table>

a. Grouping Variable: Assessment

Considering the Asymp. Sig., then the Mann Whitney test hypothesis for orthographic aspects was rejected. It can be clearly seen that there is no difference between self and peer assessments on the orthographic aspects that have been assessed by students. The orthographic aspects that are assessed include the use of punctuation, capitalization and spelling.

To see the results of the fourth aspect, which is grammar, it is described in table 5 as follows.

Table 5. Mann Whitney Test for Grammatical Aspects

<table>
<thead>
<tr>
<th>Essay Points</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>538.000</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>1241.000</td>
</tr>
<tr>
<td>Z</td>
<td>-1.699</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.089</td>
</tr>
</tbody>
</table>

a. Grouping Variable: Assessment

Table 5 shows the results of the Mann Whitney test on the grammatical aspect. From the table, it is known that the significance value is greater than 0.05, which means that the hypothesis is rejected. This means that there is no significant difference between self-assessment and peer assessment in the grammatical aspect.

From the five results and discussion of the Mann Whitney test, it was concluded that there were significant differences in the content and coherence aspects of students' self and peer assessments; while for the aspects of vocabulary, orthography and grammar, there were no differences between self and peer assessments. The difference in the results of the assessment in the context and coherence aspects is caused by the subjectivity that occurs in peer assessments, namely the feeling of reluctance to judge their friends badly (Trisno, 2014).

Students are given an instrument in the form of a questionnaire sheet to find out how students' perceptions of the application of self and peer assessment. The following is an explanation of the results of the data that has been collected.

Student Perceptions of the Application of Self-Assessment and Peer Assessment in the Aufsatz II Course

Self and peer assessment generated many perceptions from students which have been summarized and their important points are presented in diagrams 1, 2 and 3 below. In each of these diagrams, data on the application of self-assessment is presented, as well as general self-assessment and peer assessment according to student perceptions.
It is known from the diagram 1 that not more than half of the students (40% of students) consider self-assessment easy to carry out. In addition, nearly 60% of students have difficulty carrying out self-assessments.

Furthermore, the results of students’ perceptions about peer assessments are described as follows.

The second diagram is a diagram of students’ perceptions on the implementation of peer assessments. It is known from the diagram that 56% of respondents considered peer assessments easy to carry out. While as many as 59% of students, however, had experienced difficulties in implementing peer assessments.

Then the data explanation of the results of student perception from the two assessments, i.e. self and peer assessments is illustrated in the following diagram.
Diagram 3 contains data from students' perceptions about the implementation of both self-assessment and peer assessment. It is obtained from the diagram, 76% of students are motivated to learn more and improve their writing skills. In addition, this assessment has also fulfilled the characteristics of self and peer assessment, that it is a student-centered because it is open to assess themselves and the peers. Furthermore, it also can be used as a benchmark in learning as stated by 75% of students. Most students (77% of students) felt the advantages of self and peer assessment, which are being able to practice honesty, to be critical, to be independence, to be tolerance and to be confidence, being objective and saving time. Nevertheless, on the other hand, as many as 63.4% of students still felt inadequate during the implementation of self and peer assessments, that is to say lack of knowledge and experience and tend to be subjective in assessing.

Regarding to the student’s opinion, there is tendency that they prefer the assessment from lecturers to self-assessment or peer assessment by the figure of 41% of students.

Discussion

Some of the factors that caused the low self-assessment for the content and coherence aspects were students' insecurity and the tendency to judge themselves more critically than others. This being modesty is a characteristic of Indonesian people who grew up with a culture of andhap asor (humble) and are reluctant to give bad pronounce other people (Jatirahayu, 2013). In addition, students read their friends’ essays at a glance so that they might ignore aspects of content and coherence. In assessing the context and coherence of the essay, accuracy and the ability to see the text comprehensively are needed. In the aspect of coherence, especially, the assessment was not done not only in students' text, but also in the ability to connect the meaning of the reading such as sequences and continuity (Lestari, 2019). Therefore, it can be concluded that the content and coherence aspects are deep or implicit in an essay.

On the aspects of vocabulary, orthography and grammar, there is no significant difference between the results of self and peer assessments of students. This is because the assessment of aspects of vocabulary, orthography and grammar is written and explicit. In addition, the rules for writing in German are very clearly stated in the Rat für deutsche Rechtschreibung (2017) or the German spelling rules that have been established by the German government. To regulate German grammar, it is clear that it is also listed in several books such as Duden 4: Die Grammatik. Errors in these three aspects can be recognized immediately by students. German Literature Department students are familiar with the procedures of writing (orthography), vocabulary and German grammar, because this topic is studied in depth in several subjects such as Structure und Wortschatz I, II, III, Aufsatz I and II.

Writing is the most complex language skill compared to other skills because it requires knowledge outside the language itself to compose a text (Fitri, 2017). Especially in writing essays in German, students should be sensitive to other issues besides academic issues. Students' insensitivity in looking for references beyond the language is because they are accustomed to only focusing on academic problems when they were guided by an academic supervisor (Kuntarto, 2020).

Based on this research, one solution for students to be able to think deeply and not only focus on one problem is to compile an essay framework based on five aspects of assessment, i.e. context, coherence, vocabulary, orthography and grammar. However, in its compilation it is regrouped into implicit and explicit ones.

Diamond (1987) argues that social science is more complicated than exact science. When connected with this research, students need to think abstractly according to their respective points of view in assessing aspects of content and coherence so that it is difficult and prone to subjectivity, while students do not need in-depth analysis to find out errors in the aspects of vocabulary, orthography and grammar.

Based on the results of data revealed to answer students' perceptions in conducting self and peer assessments, the perceived benefits are not immediately visible. Students find it difficult to carry out self-assessments compared to peer assessments. This is in line with Everhard's (2015) theory that peer assessment is a stepping stone to be able to assess peers.

On the other hand, students argue that by being given the opportunity to assess independently they can practice honesty,
criticality, tolerance and confidence when assessing their own essays and those of their peers. Students’ rate two different essays of the same material by deepening the material without repeating the material. Students know which parts are wrong and which parts are right according to the results of research from Wahyuni & Ibrahim (2012) that self and peer assessments can be used as independent reflections for students.

Students feel that self and peer assessment is difficult due to a lack of assessment skills and high subjectivity. To overcome this problem, it is necessary to replicate the implementation of self and peer assessments in a row with clear instructions and boundaries because replication or follow-up is the last step of the self and peer assessment procedure (Marhaeni, 2013); (Falchikov, 2001). If the replication is continuously conducted, the advantages of self and peer assessment can be seen. These benefits include increasing cognitive and metacognitive competence, serving as self-development and intellectual development, increasing social competence, self-confidence and positive influence (Everhard, 2015).

Assessment is conducted repeatedly because students still have not been able to assess well. This is due to their knowledge related to material and experience in assessing is still negligible. This inability also causes students not to be objective in assessing.

From the research results, it is obtained that students prefer peer assessment to self-assessment. This comes to pass because since in elementary education, students are accustomed to correcting the mistakes of others, as expressed by Wahyuni & Ibrahim (2012). However, when it is reevaluated between self, peer and teacher assessments, students tend to like teacher assessments based on distributed questionnaire data. Everhard (2015) also states that even though peer assessments are equivalent to teacher assessments, students still prefer teacher assessments. They feel uncomfortable if their work is judged by their friends or they have to judge their friends’ work. The influencing factor is because students feel that the lecturers ‘knowledge is much broader than their knowledge.

CONCLUSION

The results of self and peer assessment of student essays have significant differences in the aspects of content and coherence but do not have significant differences in the aspects of vocabulary, orthography and grammar. The difference in the aspects of content and coherence is caused by the extra abilities students must have in analyzing it. Students should read essays entirely and have to have the ability to reason about essays in order to assess them. As for the aspects of vocabulary, orthography and grammar, there are no differences in the assessment of both self and peer assessment because these three aspects are definite in the essay. Students do not need to read the entire text to assess it. According to the students, self-assessment and peer assessment need to be trained repeatedly so that they get used to it. Students prefer to teacher assessments and feel more comfortable to have the ones.

REFERENCES


Willet, K., & Gardner, A. (2010). Investigating The Capacity of Self and Peer Assessment Activities to Engage Students and Promote Learning. *Faculty of Engineering and IT, University of Technology*.

