Submitted: 16/01/2021 - Accepted: 16/02/2021 - Published: 23/04/2021

The Use of Peer Assessment to Improve Students' Learning of Geometry

Dk Siti Nabilah Pengiran Omar[1]

Masitah Shahrill[2]

Masriatol Zuraifah Sajali[3]

[1]Sultan Hassanal Bolkiah Institute of Education, Universiti Brunei Darussalam, Brunei

Email: masitah.shahrill@ubd.edu.bn

DOI: 10.26417/ejser.v5i2.p203-222

Abstract

This mixed-method action research study aimed to examine the effect of the use peer of assessment in a Brunei Mathematics classroom in the learning of Geometry. This study offered insights into the use of a student-centred learning approach, which the participants held the role as an assessor of peer's work, and the use of peer feedback as a potential learning source in changing students' conception and understanding in the topic of Angle properties. The study revealed that the use of peer assessment had significance in improving students' performance in the learning of Geometry and there was evidence of knowledge retention as a result from the intervention as seen in the improved post-test performance on similar mathematical problem. In addition, the mathematical works in the post-test still showed evidence of misunderstandings and misconception in the concept of Angle. Despite the unsatisfactory quality of peer feedback given by the participants, the assessing activity and the student' role as the assessor had increased cognitive, metacognitive awareness and self-regulation in their learning. Overall, the participants showed positive perception and attitude towards the use of peer assessment as a learning tool in Mathematics and considered it as a means for knowledge sharing. There was still concern of

emotional sensitivity and anonymity despite the effort to maintain the anonymity of the students' work and identity as an assessor.

Keywords: Peer assessment, feedback, student learning, geometry

Introduction

The term assessment is often closely associated to tests, examinations and interviews. According to Brown (2004), assessment is defined as "any act of interpreting information about student performance, collected through multitude of means or practices" (p. 304). Ghaicha (2016) argued that the terms like testing, evaluation and measurement may not be synonymous with assessment, however, Black and William (1998) mentioned that there is no universally agreed upon term. Assessment is an approach to enhance the quality of instructions to suit the learners' different learning styles and needs by gathering information pertaining students' academic performance by means of assessments.

In Mathematics learning, giving assessments is already a common situation in classroom and serve as a way to practice the mathematical concept and the procedural steps that have been taught in the classroom. This is evident as according to the national survey of 555 teachers by Plake and Impara (1993, 1997) that reported that three-quarters of the participants gave minor classroom assignments at least once every week. Assessments do not just merely serve as an instrument to measure students' performance and behaviour but also for instructional evaluation to promote towards better teacher instructions by judging the effectiveness of instructions on students (Brown, 2004; Botty & Shahrill, 2015; Damit et al., 2015; Ghaicha, 2016; Kulm, 1994; Mohammad et al., 2017) and for students' accountability in their learning. In addition, Ghaicha mentioned that assessment is also used to categorise group of students for instructional purposes as also seen in the work of Othman et al. (2016) where a classroom assessment is administered beforehand to identify the learning styles of students in order to carry out the tiered assignments. Classroom assessment is a primary source for the process of differentiated learning.

Peer Assessment

Peer assessment is a form of formative assessment or assessment for learning that is used as a learning tool and no foreign in the field of teaching and learning. In educational context, the steps of peer assessment process are summarised in systematic order as follows: establish rapport, give out works to be assessed, distribute rubric and performance criteria, peer assessment debriefing, training using sample work, discussion and work revision (Black & William, 1998). The constructive feedback from the evaluation is the expected outcome that should serve helpful purpose towards the assessees' personal developments. Peer assessment comes in various formats in terms of its implementation and evaluation process. In learning

context, it ranges from just a simple marking of other's work for reducing teachers' load to assessing peers' performance and contribution towards group project.

In a traditional classroom setting, students are instructed to do a learning task and submit the completed task where a teacher will be in charge of making judgement and marking process. Once returned, there is likely students will not read the written feedback, as according to Thomas, Martin and Pleasants (2011) that once the work is off from the students' hands, they are no longer engaged with the work. Students are perhaps extrinsically motivated to complete or compelled to do learning tasks due to time-limit factor or fear of punishments. Subsequently, they become less reflective in doing the work and uninformed about what makes a quality work. In peer assessment, the students indirectly gain benefits for being reflective when judging or marking peer's work. The students are able to receive immediate feedback about their works.

In a study done by Adediwura (2015) on the use of peer assessment in Secondary Mathematics classroom, the findings revealed that the intervention had positive impact on the students' learning retention rate. In the study of 212 Secondary Mathematics students (Chukwuyenum & Adeleye, 2013) in Nigeria, the intervention had shown a significant improvement in the post-tests scores. In addition, students developed positive attitudes, behaviours and became more engaged in the learning process after being exposed to this learning approach (Kearney & Perkins, 2011; Topping, 2003). A study by Chan (2013) in an eighth grade Mathematics classroom in Macau had shown improvement in several aspects of students learning; mathematical reasoning skill, fluency in conceptual and procedural knowledge and positive growth of attitude towards Mathematics. These findings further support the benefits of peer assessment. According to Logan (2009), the improved academic achievement is resulted from the self-awareness and critical thinking that developed from the process.

Peer assessment to promote metacognitive thinking

According to Topping (1998), peer assessment has influences on the following domains namely "cognition and metacognition, affect, social and transferable skills" (p. 254). As shown in the study by Pantiwati and Husamah (2017) on university students undertaking Science courses, the use of peer assessment had influences on their metacognitive thinking. In addition, peer assessment is an active learning model, which helps students to develop collaborative and reflective skills through the result of metacognitive processes (Husamah, 2015). Hence, peer assessment encourages students to be accountable of their learning (Langan & Wheater, 2003; Vickerman, 2010) and leads to development of self-regulation, self-regulation and reflection (Egodawatte, 2010; Gielen, 2007; Langan & Wheater, 2003). Moreover, this could potentially improve learning as it involves a task requiring students to engage and encouraging them to reflect on the quality of work for improvisation (Chukwuyenum & Adeleye, 2013).

Peer assessment as a platform for peer tutoring

Peer assessment can be regarded as part of peer tutoring process (Chan, 2013; Donaldson & Topping, 1996). The characteristic of students interacting, supporting and learning from each other during the peer assessment reflects those of a peer tutoring activity (Topping, 2005). The students may not be interacting physically through verbal feedback but assessing their peers' works and giving feedback in the rubric are already considered as peer interaction. In this study, the students will be assuming the roles of both 'assessor' and 'assessee', which resembles reciprocal peer tutoring (Chan, 2013). According to Medcalf (1992), cooperative learning is defined as a learning approach that encourages the learning of peers or peer tutoring. This indirectly implies peer assessment is a form of cooperative learning, which shares similar role to peer tutoring. While the feedback aims to help peers to improve their quality of works and evaluate their learning, the act of giving feedback to peers incorporates sense of cooperation and also collaboration as it involves interaction between a minimum of two (Kollar & Fischer, 2010). As a consequence, it stimulates motivation among group members for peer tutoring and peer assessment, subsequently correction, which produce enhanced learning (Slavin, 1996).

Peer assessment as observational learning

A common feature of peer assessment is making judgement or evaluation of others' works. Not only this particular task provides opportunity to students to look into others' work, students get to monitor their current learning performance. This enables students to evaluate and make judgement by thinking critically and be reflective on the work, which becomes a learning opportunity for the learners. A study by Logan (2009) on implementing self and peer assessment on 11 higher education students revealed that majority of students found out they were able to learn more from looking at a variety of peer's works. Another study by Wood and Kurzel (2008) reported that student realised the value of doing comparison of their own and others' work. This provides opportunities for students to learn from the mistakes made and the criteria that make up an exemplary work through exposure of different peers' works. By exposing students to others' works and instructing them to make judgements, it gives them the opportunity to extend their knowledge and look at the different approaches made by their peers in the work. (Logan, 2009; Zevenbergen, 2001). A study by Tsivitanidou et al. (2018) also reported similar finding on the use of reciprocal peer assessment on Secondary school students in the learning of Physics.

Feedback

In peer assessment, peer feedback is the core component of the process meant for the peers (Topping, 1998). According to Gielen et al. (2010), the bidirectional nature of giving and receiving feedback potentially enhances students' learning as students could learn from different examples and approaches in the process. A classroom often consists of students coming from different academic ability and perhaps may have

different perceptions towards peer assessment due to prior experience. Hence, the quality of peer feedback generated from the assessor is likely to be affected by his or her domain knowledge (Patchan & Schunn, 2015; Van Zundert et al., 2012). This explains the concern on accuracy of feedback produced by peers reported in several studies (Alqassab et al., 2018; Falchikov, 2004; Liu & Carless, 2006; Strijbos & Wichmann, 2018). This is because for the feedback to be effective on the students' learning lies on the assessor's ability to critically link his or her prior knowledge with the work (Liu & Carless, 2006).

Use of rubrics in learning

A rubric is defined as a framework to assess students or mathematics learning task, which can be adopted or tailored by the teacher (Kulm, 1994). Rubric has been used to assess students' works across any disciplinary subject, often used to judge quality of performance or whether the criteria are met. Rubrics are used to help teachers analyse the information about students' understandings from performance-based task. Rubrics are categorised into different types according to its functions such as anaholistic rubric, process rubric, analytic rubric and anaholistic rubric (Kulm, 1994). Meanwhile, in a study by Idris et al. (2017) on the use of rubric in the History presentation of Year 10 student, it was found that students were motivated and anticipating the use of rubric which helped in completing the task by knowing teacher's expectations ahead. The explicitness of learning and task expectation set in the rubrics is an advantage for students in understanding clearly of the learning objectives to be reached (Andrade & Ying, 2005; Huba & Freed, 2000; Stiggins, 2001). Idris and colleagues stated that the use rubrics enabled teachers do systematic evaluation of their students' work and contribute to the students' learning process through constructive feedback whilst acting as the framework. On the other hand, Chong et al. (2017) used observational rubrics to monitor students' learning progress and to measure students' inquiry skill in the inquiry-based learning of conditional probability.

The Study

The aims of this present study are to investigate the impact of peer assessment on students' conception on the concept of angles, to investigate whether peer assessment affects students' learning retention and transferability skill after assessing their peers and reading feedbacks directed to themselves, to identify students' misconceptions in the topic of Angle Properties, and to inform educators and researchers particularly in Brunei of the use of peer assessment in the learning of Mathematics.

With the aforementioned aims the following research questions central to this study were posed:

How does peer assessment affect the students' performance in the learning of Geometry?

What are the students' perceptions on the application of peer assessment and its peer feedback in the learning of Mathematics?

Following the first research question stated above, two hypotheses are formulated:

H_0: There is no significant difference in students' performance after lesson intervention using peer assessment

H_1: There is significant difference in students' performance after lesson intervention using peer assessment

Methodology

This study adapted the Action Research design, which is known to employ a mix of qualitative and quantitative data collection approach. The rationale for choosing action research is because this particular research design aims to find solution to common problems encountered by schools and subsequently acts as a means of professional development (Mills, 2011). In accordance to the research questions stated above, imply the need of quantitative data type to identify any difference prior and post intervention. Collection of quantitative data is preferred for the first research question as Denscombe (2010) noted that quantitative data is analysed not based on intuition and others can validate it for authenticity. Whereas, for the second research question, qualitative data is required to explore further on participants' attitudes and perceptions towards any parts of the intervention process.

The participants that were involved in this study were 23 Year 11 students aged between 14 to 16 years old, from an all-female government secondary school in Brunei Darussalam. All participants were from a Year 11 class that consisted of mixed ability students. All participants studied Mathematics 'D' and were not taking Additional Mathematics subject. The school is government funded where the local Bruneian students are waived from paying tuition fee. The participants of this study were selected by convenience sampling. The school was chosen for this study on the basis that it was within the convenient commuting distance. A follow-up qualitative data collection procedure after the intervention, which is the focus group interviews, required a small percentage of study samples.

Consent to conduct the research study were sought from the Department of School, Ministry of Education, Brunei Darussalam and the university. According to Declaration of Helsinki (1946), children of age under 16 will be required parental consent to participate in any research study. Hence, parental permission letters along with participant information sheet were distributed to participants of the study. Any students' names mentioned here are pseudonyms.

Instrumentation

For quantitative data collection, the following instruments were employed:

The Pre-test was used to assess students' understandings and actual knowledge in Geometry at a point before the intervention is conducted. It assessed students' understanding on the concept of properties of angles. The pre-test allowed the researchers to tap into the students thinking by investigating the misconceptions and errors made by the students. The post-test assessed the change in the students' conception and understanding in Geometry after a period of intervention. In addition, the performance from the post-test served as an indicator whether the learning strategy, that is, the peer assessment approach, used in this study had effect on students' understanding. The contents of the post-test were an exact replicate of the pre-test instrument. The pre- and post-tests questions were pilot tested prior to the commencement of main study by test-retest reliability method. The pre- and post-tests were administered to Year 10 students at a different school setting. A one-week interval between the pre-test and post-test was ensured so there was no memory effect. Meanwhile the Pearson coefficient of correlation obtained from the Test-retest reliability (0.853) of the pre- and post-test instruments indicated a very high reliability.

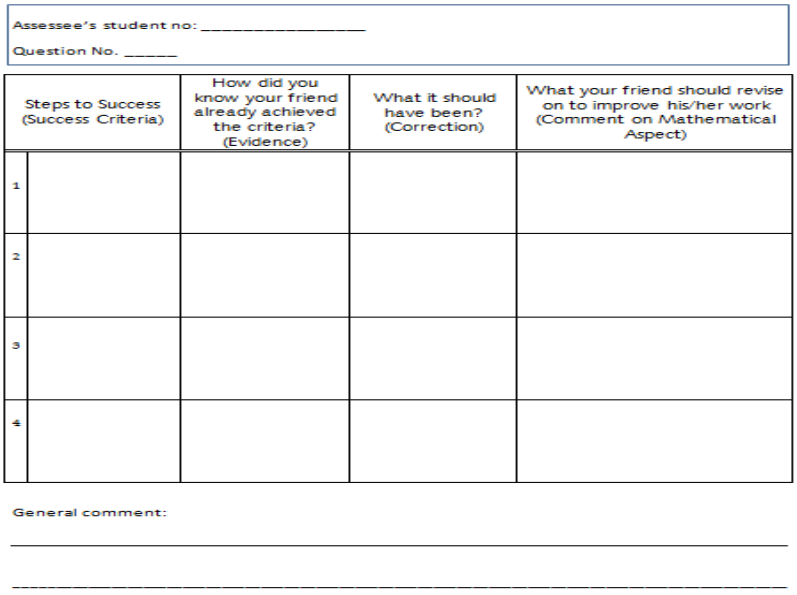
The classwork consisted of nine questions including its subquestions. It assessed students' knowledge on the concept of Angle Properties with some questions possessed similarities with the questions in the tests. This classwork would be used as the work to be evaluated by the assessor during the peer assessment.

The 5-points Likert Scale survey explored the students' perceptions and attitudes towards the intervention conducted in the classroom whether there was any relation with students' performance in the post-test. The survey was adapted from Chan (2013) and slightly modified for this present study with the addition of a few items to fit the design of the research study and the topic chosen for the intervention. This survey was particularly adapted and chosen for this study because this study shared resemblance with Chan's design of the intervention, that is, peer assessment, for her action research study. The use of survey allowed the researchers to collect wide range of response in a relatively short amount of time yet very informative towards study (Johnson, 2008). According to Revilla, Saris and Kronick (2014), the 5-points Likert Scale survey is preferred because higher odd-numbered point scales lead to lower quality data. In addition, the use of 5-point Likert scale instead of 7-point Likert scales helps to reduce respondents' frustration level, increase response rate and its quality as well (Babakus & Mangold, 1992).

For qualitative data collection, the following instruments were employed:

The rubric acted as a medium for the students to write qualitative feedback and to assess their peers' works against the criteria being set. The rubrics for the peer assessment were developed and designed by adapting the common structure in a

rubric developed by Chan (2013) for peer assessment in an eighth grade Mathematics lesson. The rubric was modified and designed to follow the content covered by the assessments given. The column under 'Steps to Success' was set blank for the purpose of student-teacher discussion of the required success criteria during later peer assessment. The modified rubric consisting of four columns is shown in Figure 1. Every student was required to fill in the success criteria they had come up for each question after discussing with peers and the teacher. If the assessee correctly answered a question, the assessor had to write down the evidence found in the assessee's work that met the success criteria into the column 'How did you know your friend already achieved the criteria?' However, if the assessee did not correctly answer the question, the column is not required to fill in and the column to its right, 'What it should have been?' had to be filled in by correcting the mistake found in the mathematical working. In addition, the assessors were recommended to give qualitative feedback on the mistake or misconception the student had made in the last column, 'What your friend should revise on to improve his/her work'. This included constructive feedback that focuses on correcting the mathematical understanding.



Assessee's student no: _____

Question No. _____

	Steps to Success (Success Criteria)	How did you know your friend already achieved the criteria? (Evidence)	What it should have been? (Correction)	What your friend should revise on to improve his/her work (Comment on Mathematical Aspect)
1				
2				
3				
4				

General comment:

_____

_____

Figure 1. Modified Rubric.

The focus group interviews allowed the researchers to elicit information from the participants as they were all sharing similar experience under the same intervention treatment. The group setting created dynamic interaction that means more varied responses and opinions could be obtained, and to gain insights on why they held such views (Denscombe, 2010). The interviews were in semi-structured format, and were carried out by asking verbally a series of open-ended questions with guided prompts

and follow-up questions. A pre-interview was conducted to determine whether the students had experienced peer assessment and their perceptions of similar processes, if they reported otherwise. The post-interview was to determine students' perception towards the use of peer assessment and its process as a learning tool in the learning of Mathematics.

Video recording of classroom observation was also employed to analyse the classroom behavior and interactions among the students. The information obtained helped to supplement what the interviews and surveys could not probe from the students, so that any important findings that caused concern would not be missed.

Data analysis

For the reliability of the test instruments, Pearson correlation coefficient or also known as the Pearson *R* test was used to compute the correlation coefficient. The tests instruments were piloted on a different school through test-retest reliability. The variables would be the students' test marks at Time 1 and Time 2 respectively. If the correlation coefficient was greater than 0.5 then, the test was reliable, whereas, if the correlation coefficient was between -1 and 0.5, then the test instrument was not reliable. The reliability analysis was done using a spreadsheet package namely Microsoft Excel.

The mixed method research design of this study allowed the triangulation of the findings from both qualitative and quantitative sources such as between focus groups and questionnaire. This approach helped to increase the validity of findings by means of their accuracy to gain more confidence (Denscombe, 2010). For this study, methodological triangulation was used where comparison between qualitative and quantitative data was made (Denscombe, 2010).

Quantitative data were collected for the first research question, the mean and standard deviation for both pre-test and post-test were computed. To accept or reject the hypothesis formulated earlier, the significance of the difference between the two pairs of the descriptive statistics aforementioned was checked. A repeated-measures *t*-test, or also known as, paired samples *t*-test, is an inferential statistics used to determine the statistical significant of the differences. If the *p*-value obtained is less than or equal to 0.05, then the null hypothesis is rejected. Whereas, if the *p*-value is larger than 0.05, the null hypothesis is accepted. This was done by statistical package namely SPSS (Statistical Package for the Social Sciences), which is known widely for its use in statistical analysis. Meanwhile, to analyse the quantitative data collected for the second research question, descriptive statistics were used to obtain a general view of students' perception towards the learning strategy implemented in this action study. Both video recordings of the classroom lessons and audio recording of the pupils' interviews were transcribed for the qualitative data analysis. Both focus group interviews were audio-recorded and transcribed verbatim. A word processing package, Microsoft Word, was used to create the interview transcripts. The

transcripts were analysed to form descriptions or codings that were further categorized into themes, which gave overviews of the students' perception towards peer assessment.

Intervention procedure

Before the intervention was carried out, a pre-test was administered to measure students' knowledge and to identify existing misconceptions on the topic of Angle Properties. The test papers were subsequently marked to record and measure the students' current level of understanding on the topic prior to the intervention. A short classwork was assigned to students to be attempted individually within 20 minutes. Each student was assigned a 'student number' consisting of alphanumeric characters. The students were not allowed to write their name on the paper nor did they were allowed to share their identification number with their peers. The purpose was to keep the anonymity when the works were exchanged among students later. The classwork was collected and followed by a short briefing. The briefing aimed to define peer assessment, its purpose in learning, examples of success criteria, and list of learning objectives. Students were reminded to give constructive feedback and avoid hurtful comments. Handouts of sample classwork and sample rubric on the same topic as the intervention were given to students. The explanation of how peer assessment is conducted was supplemented along with a handout on learning objective and a list of common success criteria expected in the topic. A training on peer assessment and setting of success criteria was given to students to familiarise themselves with the rubric and assessing peers' works against the criteria.

A one-hour Mathematics lesson was allocated for the actual peer assessment. Prior to the assessment, a rubric sheet was given to every student and a short time was allocated for the setting of success criteria in the rubric for each question. Each student was given an anonymous classwork of his or her peers. The assessors were not allowed to write their name in order to minimise the negative consequences of the interpersonal procedures (Panadero, 2016). Students were instructed to form into a group of four students. The groups were formed on the basis of teacher's selections so that to create a heterogeneous grouping. The classwork and pre-test marks were used to assist the teacher in doing the selection. They were encouraged to discuss among their group mates on the given works. The students were expected to discuss the correct procedures to solve the problem and write feedback by commenting on the mathematical aspect and suggest the correct solution for inaccurate or incorrect work. The teacher did not provide solution at all. At the end of the class, the rubrics were collected and the teacher marked the rubric to avoid the students from utilising inaccurate feedback for later work revision. The teacher refrained from giving Mathematical solutions in the rubric. During the subsequent lesson, the rubrics were returned and a similar blank classwork was distributed to students. The students were instructed to individually revise the classwork based on the feedbacks they received in the rubrics. The completed revised classwork was

returned and was evaluated by the teacher to check for its improvement and changes. A post-test was administered to students a week after the intervention. The post-tests were subsequently marked to compare the difference between their pre-tests and post-tests.

Results

The affect of peer assessment on the students' performance in the learning of geometry

To determine the significance of the effect of intervention on the participants' score mean, a parametric testing namely paired sample *t*-test was employed. Preliminary checks were made to ensure the data met the two assumptions for paired sample *t*-test. The two assumptions are the sampling distribution must be normally distributed and the data is measured at the interval level (Field, 2009).

The normality test was carried out using SPSS version 20 on the difference between the participants' scores in the pre- and post-tests. A histogram based on the difference between test scores was created and it could be seen that the sampling distribution is normally distributed (Figure 2). The normality of the distribution is further verified graphically as seen from the Q-Q plot diagram (Figure 3), which the plots are well aligned along the line. The absence of outliers in the boxplot diagram and both mean and median are in the centre, which forms a symmetrical-looking box indicating that the distribution is normal (refer to Figure 4). The dependent variable is an interval variable since meaningful mean scores could be obtained for both samples of data. Hence the two assumptions were satisfied and the hypothesis testing could proceed with paired sample *t*-test.

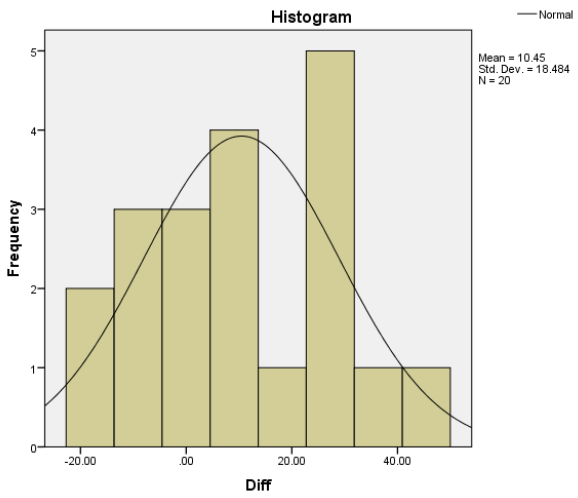


Figure 2. Histogram showing the distribution of the difference between the test scores.

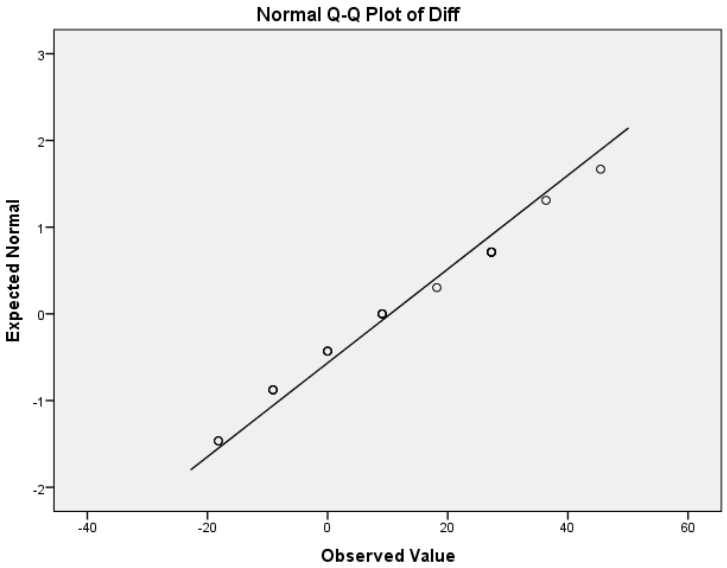


Figure 3. Normal Q-Q Plot of the mark difference between students' test scores

Figure 4. Boxplot of the mark difference between students' test scores

Table 1. Paired sample t-test, mean and standard deviation for mean differences between pre-test and post-test scores.

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Pre - Post	-10.45455	18.48433	4.13322	-19.10548	-1.80361	-2.529	19	.020

A paired sample *t*-test was conducted to evaluate the impact of the intervention, peer assessment, on students' test marks and to examine the differences in the scores of the sample group from pre-test to post-test (refer to Table 1). There was a statistically significant increase in test marks from pre-test (M = 25.0, SD = 15.3) to post-test (M = 35.5, SD = 21.2), $t(19) = -2.529$, $p < 0.05$ (two tailed). Thus, the null hypothesis (H_0) is rejected and there was significant difference in students' performance after the intervention. The mean increase in test marks was 10.5 with a 95% confidence interval ranging from -19.0 to -1.80. The eta-squared statistics (0.25) indicated a large effect size. This indicates that the difference is of practical significance. The raw scores for both pre-test and post-test were 11 marks; one mark for each subquestion. The students' overall marks were then converted into percentage. Overall, 60% of the participants in this study showed improvement in post-test scores after the intervention.

Figure 5 below shows the difference in the pre- and post-test scores. The largest improvement recorded was a positive change of 45.5% (5 marks increase), which the student initially did not manage to score any mark in the pre-test. The smallest improvement recorded was 9.1% change (1 mark increase). There were five students recorded with decreased performance in the post-test. The largest decrease in scores was only two marks decrease from her pre-test score. There were three students whose scores showed no changes in both pre-test and post-test. The student who previously scored the highest in the pre-test (10 out of 11, 90.9%) obtained an improved post-test score with an increase of one mark.

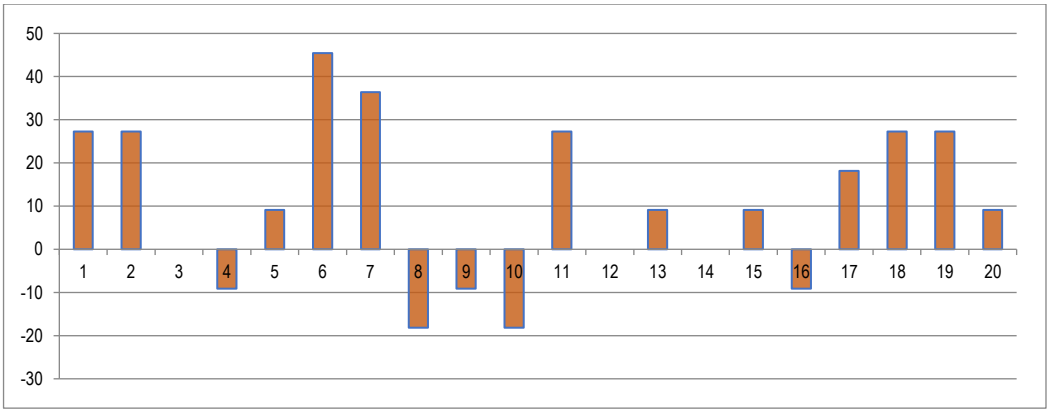


Figure 5. Difference in pre-test and post-test scores.

Table 2 shows the overall mean marks and standard deviations for both pre-test and post-test respectively. The scores for post-test are more spread out after the intervention compared to pre-test scores, as seen from the respective standard deviation values. Meanwhile, Table 3 shows the descriptive statistics of the students' scores in their pre-test and post-test following the intervention using peer assessment. There were only 20 students included in the analysis out of the actual 23 participants in this study due to absentees in either of the test administered. The comparison of test scores could then be made after excluding the absentees in order to ensure observable effect of the intervention. During the pre-test the highest score attained was 63.6%, while the lowest score was 0%. The mean mark for the pre-test was 25.0%, which is below the passing mark of 50%.

Table 2. Mean and standard deviation for pre-test and post-test.

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre	25.0000	20	15.29043	3.41904
	Post	35.4545	20	21.24854	4.75132

Table 3. *Descriptive Statistics for pre-test and post-test scores.*

	N		Mean	Std. Deviation	Variance	Range	Minimum	Maximum
	Valid	Missing						
Pre	20	0	25.0000	15.29043	233.797	63.64	.00	63.64
Post	20	0	35.4545	21.24854	451.501	81.82	9.09	90.91

Following the intervention, the highest mark attained in the post-test was 90.9%, while the lowest mark attained was 9.09%. This shows a slight improvement than the pre-test with every student seated for the post-test were now able to answer successfully at least a question. The overall mean mark for the post-test was 35.5%.

This is a noticeable improvement in the students' performance in post-test compared to in the pre-test. The change from the mean pre-test scores was 10.5%.

The rubrics were returned to respective students that were used in assisting them in making a revision of their classwork in individual setting. Both classwork were examined to identify any significant changes in the procedural steps or diagram work sketches before and after the intervention. The quantitative results (Figure 6) of the students' revised classwork show tremendous improvement in the quality and accuracy of their mathematical working of the problems. This shows that students utilised the feedback in the rubric in making revision of the work, although there were only few helpful qualitative feedbacks on improving mathematical understanding towards the topic.

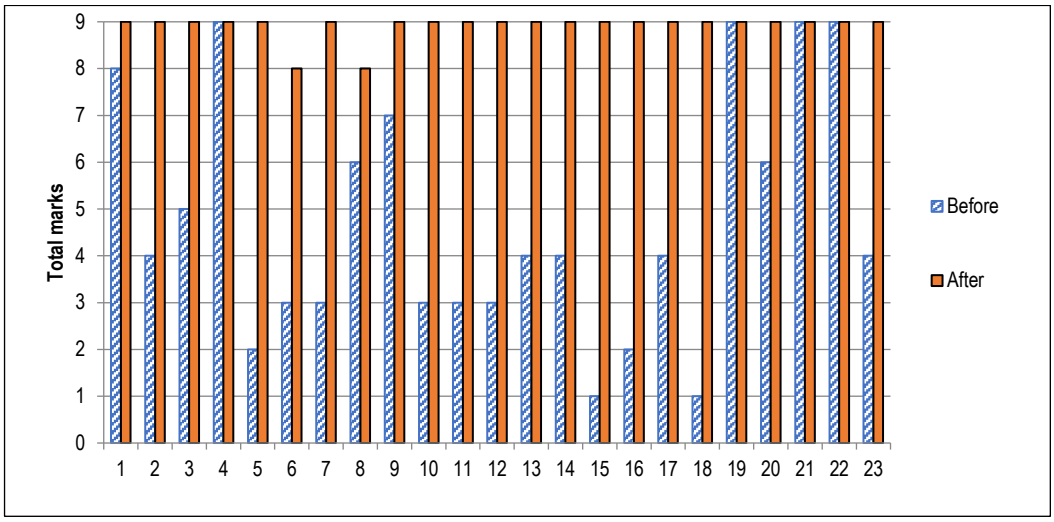


Figure 6. Classwork performance before and after revision of work using feedback from the returned rubric.

After the revision of their classwork, it could be seen that some students had shown improved mathematical reasoning in the work. Although the reasoning shown was often very brief and in short sentence, it was slightly accurate than before. As an example Farah previously stated 'parallel' to support her answer in Question 2b (Figure 7). After the revision of classwork, she wrote a more specific statement 'angle at alternate angle'. To state which angle property applied was not required in the question, however, this provided the opportunity to investigate the student' thinking and any possible misconceptions held.

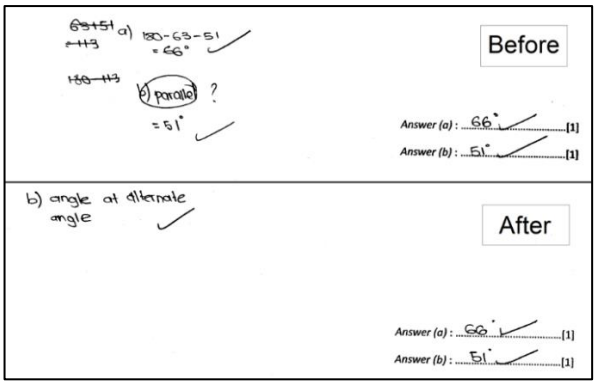


Figure 7. Sample work with improved reasoning.

Due to absenteeism during peer assessment, three students did not have an assessor for their work. Hence they received no feedback from peers. Despite the shortcoming, majority of these students were able to revise their classwork successfully without the aids of the written feedback. Perhaps the students' engagement in the discussion during peer assessment had memory effect on them.

The observation of rubrics revealed a few incorrect judgments and inaccurate mathematical explanations in the qualitative feedback. In the case of Wardah (a pseudonym), she successfully did the correct revision (Figure 9) for Question 3a despite the inaccuracy in feedback she received as seen in Figure 8. Perhaps Wardah's engagement in the discussion activity with peers or the assessing task, made her to reflect on her and peer's work, thereby allowed her to have a better understanding of the solution for the particular problem.

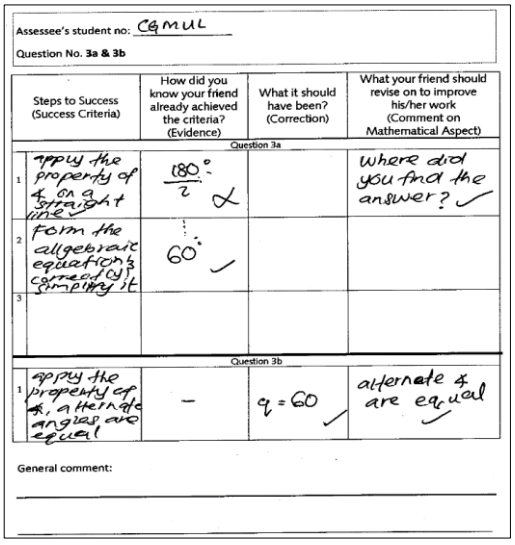


Figure 8. Inaccurate evaluation by the assessor for Wardah for Question 3a.

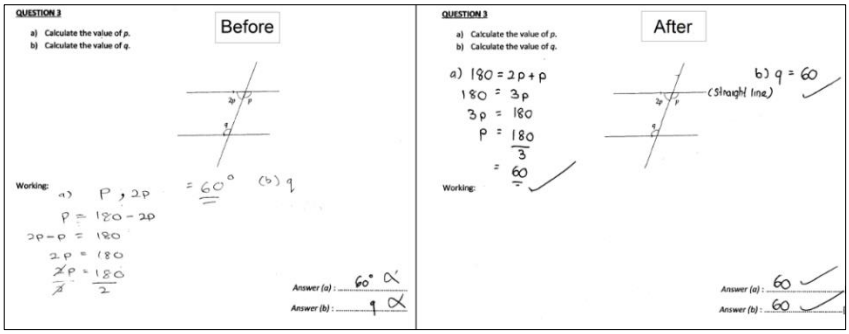


Figure 9. Wardah's changes in the classwork for Question 3a before and after revision given inaccurate feedback.

The correct mathematical solution provided by the assessor may have resulted in the drastic improvement in classwork revision. It was observed majority students followed and heavily depended on the solution given in the feedback. This includes copying a minor mistake made by the assessor. This case shows that it was not clearly evident whether students deeply reflected while revising their work with the feedback given in the rubric.

Students' perceptions on the application of peer assessment and its peer feedback in the learning of mathematics

A 25-item questionnaire was distributed to 23 students involved in this study. The questionnaire was distributed within the same week as the post-test. This multidimensional survey aimed to explore students' attitude and opinions on different aspects of peer assessment and towards the topic Angle Properties. The internal reliability of the questionnaire after reversing the score for negatively worded statements had a Cronbach's Alpha of 0.919. The response from each item was coded and assigned a score according to its scale: Strongly disagree (1), Disagree (2), Neutral (3), Agree (4) and Strongly Agree (5). The negatively worded Items 1.1, 5.3 and 6.5 were reverse coded.

Table 4 shows the results of the responses from the questionnaire. Generally there was neutral and positive tendency of the participants towards the setting of success criteria. Although there were two participants (8.7%, Item 4.1) disagreed of liking to be involved in the deciding of success criteria prior to both training of peer assessment and actual peer assessment, there were no participants (0%, Item 4.2 and Item 4.3) of being against the two statements about the helpfulness of making success criteria explicit towards their subsequent understanding of the classwork and their ability in assessing their peers' works. Five participants (21.7%, Item 4.2) strongly agreed that this task had contributed to their understanding on the classwork.

Table 4. Results of the questionnaire after reverse-coded on some items.

Item	Statement	SD =1	D =2	N =3	A =4	SA =5	Mean (*M*)	Std. Dev. (*SD*)
1.1	The most difficult part of Mathematics is angle properties	1	5	11	6	0	2.96	0.82
1.2	I am good at angle properties topic	0	9	11	3	0	2.74	0.69
1.3	I like the topic of angle properties	0	1	16	5	1	3.26	0.62
2.1	The peer assessment training is sufficient to prepare me to assess my peer	0	4	6	11	2	3.48	0.90
2.2	I understand what my role and what I must do during peer assessment	0	2	3	14	4	3.87	0.81
2.3	Teacher's feedback on my assessment performance is useful to improve my skill on assessing peers.	0	0	2	10	11	4.39	0.66
3.1	I am aware and understand of the purpose of using peer assessment in mathematics classroom.	0	0	8	12	3	3.78	0.67
4.1	I like to be involved in the deciding of success criteria required for the classwork	0	2	8	11	2	3.57	0.79
4.2	The discussion of success criteria is helpful for me in understanding the classwork questions	0	0	8	10	5	3.87	0.76
4.3	I am able to assess my peer's work well when involved in deciding the success criteria	0	0	12	9	2	3.57	0.66
5.1	The rubric is easy to use	0	4	3	11	5	3.74	1.01
5.2	After the setting of success criteria, I am able to assess my peer's work	0	0	9	11	3	3.74	0.69
5.3	I find it difficult when writing qualitative feedback	0	0	14	6	3	3.52	0.73

5.4	I can assess my peer's work with fair	0	0	11	8	4	3.70	0.76
5.5	I feel more comfortable writing feedback when assessing my peers than giving marks	1	5	6	8	3	3.30	1.11
5.6	Assessing peer's work makes me more aware of key mathematical concepts in the topic of learning	0	0	14	5	4	3.57	0.79
5.7	I always reflect back to my work while assessing my peer	0	2	7	11	3	3.65	0.83
6.1	I understand the feedback given to me	1	1	4	15	2	3.70	0.88
6.2	The feedback given to me is accurate	0	1	15	6	1	3.30	0.63
6.3	The feedback is fair	0	1	10	9	3	3.61	0.78
6.4	The feedback given to me is helpful for me to improvise my work	0	1	3	11	8	4.13	0.81
6.5	I am unsure how to revise my work after reading the feedback	1	5	13	4	0	2.87	0.76
6.6	I understand the topic better after reading the feedback and making revision of my work	0	1	8	12	2	3.65	0.71
7.1	I enjoy giving peer feedback	1	6	6	9	1	3.13	1.01

From Table, 4, it is found that 39.1% of the students agreed that they find it difficult in writing qualitative feedback for their peers. However 60.9% of the students were undecided about the difficulty of writing qualitative feedback. Item 5.5 had a wide range of response from the students. There were 6 out of 23 respondents who disagreed with the statement of Item 5.5 and preferred giving marks than writing feedback for their peers due to uneasiness feeling. For Item 5.7, two students admitted that they were not always reflective and thinking back to their works while assessing their peer's work. On the other hand, more than half of the respondents agreed that they thought about their work while assessing peer's work. About 65% of the participants were unsure of the accuracy of the feedback given by their peers. There were varieties of responses from the students when asked whether they enjoyed giving peer feedback (Item 7.1). It was also the second item in the questionnaire that found the highest number of participants disagreed with the statement besides Item 1.2 and also with lower overall mean. For Item 7.1, 30.4% of

students did not enjoy giving feedback to their peers. This result could reflect with the highly varying responses from Item 5.5, "I feel more comfortable writing feedback when assessing my peers than giving marks". Another factors that might contributed to the percentage of students disagreed with the statement were the easiness in using the rubric (Item 5.1) or perhaps the 'demanding' requirement to write feedback in the rubric for every step of success criteria that were not met by the assessee. On the other hand, Item 7.2 (M = 3.87) shows a higher mean than Item 7.1 (M = 3.13), indicating that students enjoyed receiving peer feedback more compared to giving feedback.

Overall the responses from the questionnaire show positive perception towards the use of peer assessment in the learning of Angle Properties with an overall mean of 3.53 and standard deviation of 0.81. The encouraging responses also include the reverse-coded of the negatively worded items.

Analysis of video recording

Analysis of the video recording of a classroom observation during peer assessment shows that the activity started off with minimal noise level. Interactive discussion among group members was evident and students were motivated to assess peer's work. In one case the teacher researcher allowed a student to seek assistance from other group to understand the approach to solving Question 3 in the classwork. Students cooperated well while working on the same task together. There were instances of students looking over each other rubrics with very minimal interaction. Off task and unrelated conversations were commonly heard as students made progressed further in the assessing task when the teacher researcher made a class round. The teacher researcher offered minimal assistance and only facilitated the task.

Result of the interviews

Each interview had a different set of six selected students, except for a student who happened to be selected to attend both interviews. The students were chosen for pre-interview on a responsiveness basis. Meanwhile, for the post-interview, the six students were chosen based on their difference of marks between pre-test and post-test which were coming from three categories; two from 'improved performance', two from 'no changes in performance' and two from 'decrease in performance'. The rationale for this was to have a varied responses and opinions coming from different abilities. Several themes emerged from the analysis of the interview codings, and the emerging themes from the pre-interview are as follows:

Use of peer assessment in classroom learning

Before the intervention, the students had never heard the term 'peer assessment' but have had experienced a peer assessment process informally through the marking of their peers' works. The student felt that assessing peer's work was helpful because

when they understood the mistakes made in the work they could teach their peers. In other words, they could assist their peers that had difficulty in understanding or solving particular Mathematics problem and help them to improve their understanding by tutoring them.

Usefulness and reliability of feedback

All students confirmed in the interview that they had experienced receiving feedback from peers. However, students admitted that they occassionally read the feedback given regardless from peers or teachers and hence in utilising the feedback in revising their work as well. The students showed positive attitude towards feedback by perceiving it as a step towards confidence, a comment to enhance skills and a motivation to improve quality of work. Afina displayed positive mindset for learning growth as "*does not mean we are wrong forever*."

Majority of the interviewee believed teacher's feedback were useful than peer's feedback. The reason for the preference from the teacher was the reliability and the concern of accuracy of the feedback given by their peers. A few students showed no preference over the other. Farah commented that teacher gave detailed feedback compared to their peers. Despite that, she believed that the peer feedback would help them to improve their work and work revision if the feedback commented on aspect they need to improve.

The need of explicit assessment criteria and its alternative mathematical solution

Students were sometimes unable to fully utilise feedback given because they were not able to comprehend the mathematical working given by the teacher. When the performance criteria of an assessment were not explicitly stated, students were unable to meet the teacher's expectation. As Nurul said, "*we don't understand [what] the teacher want*". They further added on that they prefer detailed explanation consisting of step-by-step mathematical working and also a simpler approach to solving a mathematics problem.

Exemplary learning

From assessing the students' work, students believed that they could learn how the mathematical formula was applied and the different approaches their peers used to solve the mathematical problems. A student said she preferred the mathematical feedback to be demonstrated rather than solely explained verbally by the teacher. She further specified that it had to be shown rather than just writing on the paper.

Meanwhile, the emerging themes from the post-interview are as follows:

Advantages of peer assessment

Use of rubric and explicitness of assessment criteria

Five out of six students gave positive responses in the interview, which implied the usefulness of the use of rubric in improving their understanding. The use of peer assessment also helped students to recall on the mathematical terms that might be necessary for their mathematical reasoning. For instance, peer assessment helped them to revisit the names given for each of the angle property such as 'z-shape', which were stated as success criteria.

A student responded that the use of rubric helped to improve the skills on remembering formula and on solving mathematical problems. The listing of success criteria in step-by-step characteristics in the rubric acted as a guide for the students to complete the work. In addition, another student said that the list of success criteria help to improve and clarify their understanding as it provided them examples to follow.

The use of rubric helped to supplement mathematics working with an explanation. A student stated that it helped to identify the mistakes they made in test. In addition the rubric provided clear expectation for the work as a student said "*it can helps us more understand the property by looking at the rubric, how to solve…*".

Learning from others and knowledge sharing

The peer assessment is also a way for students to learn from their peers the different approaches to solving mathematical problems such as, in this case, on the topic of Angle Properties. An interviewee added that they could follow the approach used by their peer and hence they could understand better. Peer assessment was perceived as a medium for knowledge sharing with peers, as evident from the following responses in the interview:

"*It [is] also we can share our understandings to any friends*" (Amy)

"*The way we try to understand the angle properties is different, isn't it. And their way is also different. And we can also give which one is easier for them; [they can choose] either their method or our method."* (Safi)

"*And then we can know . . . follow what they did. . . And then we can remember and we can understand.*" (Afiqah)

Assessing peers' works allowed students to discern the mistakes in the work. In one instance in the interview, the researcher asked the student how they could identify mistakes in the work they had completed, which the student responded by noting the difference between the submitted work and the peer's work they were assessing. All interviewees agreed that group discussion had helped them in understanding Mathematics better. An account from Amy implied that the discussion allowed them to explore different, simpler and easier mathematical strategy.

Disadvantages of peer assessment

Concern on anonymity and interpersonal sensitivity

Issue of anonymity was a concern for a student when assessing peer. An interviewee explained that their handwritings exposed their identity as an assessor and their weaknesses in the subject when their works were being assessed. Students disliked when their mistakes in the work became known. Issue of comfortability in giving qualitative feedback was also another concern, as it could possibly hurt their peer's feelings.

Reliability of peer feedback

There was an issue on reliability of feedback given by peers because of inaccuracy. A student was unspecified when asked regarding the helpfulness of peer feedback in making the revision of work because she had received an inaccurate feedback before. Other factor that affected the confidence in assessing was the different solutions offered by the students during the discussion that were often not in agreement with the teacher's solution. The reasons for the discrepancy was the criteria (and hence implying the solutions) suggested by the students were incorrect and hence not chosen as the criteria, or the question had alternative criteria which were also correct but not mentioned properly or written on the board by the teacher due to time constraint.

Issues pertaining success criteria

The interviewees were not confident in evaluating the work, particularly questions with few alternative approaches to solving. The responses from the interview suggested that majority were sometimes not confident in selecting their own criteria. Although the interviewees acknowledged there was an alternative approach or success criteria, the discrepancy between the teacher and students' decision in choosing criteria resulted in students' losing confidence. As Amy emphasised that it was difficult to be sure which of the answer was correct or not due to different saying from the teacher. The low confidence resulted from weak mathematical knowledge led to heavy dependent on teachers' final say. In addition, the interviewees were confused by the success criteria listed, as they were not able to comprehend what each criterion meant.

Discussion

The results from this study had provided insights on the effect of a student-centred learning approach through formative assessment. The quantitative result of paired sample *t*-test for difference in students' test marks showed that the intervention had significant effect on the students' performance in the learning of Angle Properties. Possible reasons that contributed to this improvement in test were the reflective cognitive activity that happened during any part of the peer assessment process. By exposure of different quality mathematical works to students through the assessing

activity in a group, it encouraged students to reflect on their works while making judgement of the peer's work. This indirectly led students to measure their current level of understanding against other and make an initiation for a change towards improvement. The reflective practice had contributed to enhance understanding through learning from exemplary work (Langan & Wheather, 2003).

The three out of five students that were recorded decreased performance were among the absentees during the actual peer assessment. Subsequently, they did not assess any of the peers' work. Perhaps these students benefitted less from the peer assessment process specifically being the role as an assessor. This is because the learning gains from being an assessor are more significant than just being an assessee, as the assessing activity makes them self-reflective while they are comparing between their own works and their peers' (Logan, 2009; Tsivitanidou & Constantinou, 2016). It is important to note this study had provided opportunity on every participant to become both an assessor and an assessee at the same time, if they were present.

The discussion that was incorporated in the peer assessment through grouping had helped students in understanding the approach to solving a particular mathematics problem. As an example, Question 4ai seemed to capture students' interest during the peer assessment activity and more discussion was put into this question. The discussion allowed students to maximise the benefits of peer assessment. The interactions with peers on a productive discussion enabled students to share their thoughts and promote critical thinking (Piaget, 1971). In addition, the discussion promoted peer tutoring and subsequently students' understanding on this particular question was embedded in their memory. Their memory retention was further enhanced with their roles as an assessor to write down the correction or feedback for this question to their peers. As Lin, Liu and Yuan (2001) stated that students are able to engage in important cognitive processes when writing feedback. Furthermore, Pugalee (2004) explained that writing assists students in critical or metacognitive thinking while engaging in the mathematical thoughts.

The use of rubric also indirectly acts as a medium for peer tutoring with others who were not assigned together in a group. This corresponded with a student's view towards peer assessment as a platform for knowledge or opinion sharing. This was consistent with the statements by Donaldson and Topping (1996) and Chan (2013) that peer assessment is a subset of peer tutoring activity.

The group discussion was lacking of explanation on the conceptual knowledge and often centred on the procedural knowledge. It did not contribute much in improving students' understanding of the topic, but the procedural steps to getting correct solution for the questions in the classwork. This was evident in Question 3 and 4bi of the post-test, which showed no improvement in marks, although there were similar questions in the classwork testing on the same concept knowledge. Although the students depended heavily on their peers as seen from the video observation and classroom observation, the assessing activity had shown to help in promoting self-

reflection and awareness of their understanding (Langan & Wheater, 2003; Logan, 2009; Wood & Kurzel, 2008). This was evident in the improved performance on some questions in the post-test. This showed that students retained the knowledge constructed during the peer assessment activity and consequently their ability to apply the knowledge (transferability) in similar context. This could be seen from the changes in approach to solving problems as evident from the large improvement in post-test for Question 2a, 4ai, 4aii that had similarities with Question 2b and 3 in the classwork.

Analysis of feedback in the rubric indicated that students did not provide sufficiently helpful feedback towards their peers and often lack of Mathematical content. Factors such as concern of emotional sensitivity, weak mathematical knowledge in Angle Properties and even time constraint, perhaps contributed to the low quality of feedback and students' willingness in giving feedback in the rubric. In addition, students' proficiency in English Language limits the students' ability to express their thoughts qualitatively. Writing in mathematics is perceived as a demanding task in the aspects of language skill and mathematical knowledge (Huang & Normandia, 2009), which perhaps resulted in some students preferred giving marks than feedback. This agrees with the result from the questionnaire that a proportion of participants did not enjoy giving feedback. This was further supported from a response in the interview, which the interviewee admitted of disliking writing feedback but numerical working only. Students' reluctance to participate in the peer assessment perhaps was also another factor, as this possibly contributes to the less reliable assessment (Liu & Carless, 2006), in addition to their weak fluency in the subject.

The students' domain knowledge influences the reliability and the style of feedback given (Alqassab, Strijbos & Ufer, 2018; Van Zundert et al., 2012). Hence, the students' strength of mathematical knowledge influences the quality and the reliability of the feedback given. It was observed that majority of the qualitative feedback given by the participants did not contribute much in correcting the students' misconception due to sometimes incorrect judgement and lack of accurate mathematical content. This was understood as the participants were coming from class of mixed ability. This corresponds to the findings from the large amount of respondents in the survey were undecided about the accuracy of the feedback they received and also one similar response from the interview. It was noted that students were dependent on their peers in the group to write feedback as evident in the strong similarity among several feedbacks.

After the intervention, it was observed that the improvement in the post-test was attributed to questions that required simple direct application of angle properties such as the angle property on a straight line, the alternate angle property formed between parallel lines, the vertically opposite angle property and those without

requiring formation of algebraic equation. This revealed the current geometric level of thinking of the students in this study.

Time factor such as one-hour lesson might have limited the students from using the opportunity in the peer assessment to reflect deeply on other's works. Peer distraction and motivation could influence the student's engagement in the reflective process while assessing the work. Despite the discussion of success criteria conducted earlier, it was expected that students would internalise the criteria and would be able to solve similar problems. However, the weak prerequisite knowledge they possessed unable to let them to understand some of the criteria listed that led to the correct solution. Consequently, the students were having difficulty in making sense of the flow of the mathematical working. With the poor understanding of the criteria, this also affected their role as the assessor and consequently their ability to give differentiated feedback (Sadler, 1998). Time factor also seemed to be a concern when implementing the intervention using peer assessment approach especially when Mathematics lesson is bounded by syllabus content to be completed, because peer assessment was a time consuming process. An interviewee acknowledged the concern as they would be sitting for a public examination for students studying in the final year of secondary education and they were constrained by extra classes. This is consistent with the reports from several studies regarding time issue in classroom implementation of peer assessment (Chan, 2013; Falchikov, 2001; Langan et al., 2005; Tsivitanidou et al., 2018).

Generally students found peer assessment and the feedback received as useful in their learning regardless from the teacher or their peers. They valued peer assessment as an opportunity to learn the different approaches and strategies of their peers in completing the work (Logan, 2009; Tsivitanidou *et* al., 2018; Zevenbergen, 2001), in this case, the Mathematical problems. In summary, the result of this present study on the effect of students' performance confirms the findings of the previous studies in the learning of Mathematics at Secondary Education level (Chan, 2013; Chukwuyenum & Adeleye, 2013).

Conclusions

The result from the paired sample *t*-test to compare the means of students' pre-test and post-test marks had shown that the use of peer assessment had resulted a significant difference in the students' performance between both tests and there was an improved performance in the post-test. In other words, the intervention had positive impact on the students' performance in the learning of Geometry particularly on the topic of Angle Properties. In addition, the analysis of students' revised classwork after the feedback was returned had shown drastic improvement in the quality of their mathematical working. However it is important to note the revision was done with reference to the feedback in the rubric and that the feedback received may not be fully engraved in their understanding as it was found that the performance of the revised classwork did not fully reflect in the post-test. Despite that, the

improved post-test scores indicated peer assessment had successfully made students becoming aware and reflective of their current understanding and subsequently their works.

The way in which the peer assessment was carried out indirectly plays a role in the students' learning. The discussion happened during the peer assessment had assisted students in understanding the approach to solve mathematical problems particularly the difficult questions from the classwork, as evidenced by the increased number of students in scoring similar questions in the post-test correctly. In summary, the intervention through peer discussion had positive impact on students' performance in the learning of Geometry, which promoted students' self-regulated learning and metacognition skills (Ahamad et al., 2018).

The findings from both interviews revealed that the students believed assessing peers' works allowed them to learn the mistakes made in the work and correct their peers by teaching. It was found that students occasionally utilise both teacher's and peer's feedback in revising their works. Despite that they preferred receiving feedback from the teacher and also showed positive attitude towards receiving feedback. The students perceived the peer assessment as a useful learning approach from which they could learn from their peer's works and the characteristics of accurate work through the criteria listed. Students showed positive attitude towards the use of rubric with success criteria listed. Negative perceptions only revolved around confidence in understanding and choosing the right success criteria, and those concerning anonymity in giving feedback. They were willing to have another session of peer assessment when there were no time constraints.

The quantitative results from the questionnaire showed an overall mean of above average. The outcome of the questionnaire was consistent with the responses in the interview. Analysis of video recording during the actual peer assessment showed that the students displayed positive behaviour through their interactions with their peers. The students were willing to discuss with their peers to assist them in the assessing task despite the difficulty they have. The group formation had contributed to this positive perception of the students as evident in the responses from students in the post-interview. To conclude, the multiple findings show generally encouraging responses that reflect students' positive attitude and perception towards the use of peer assessment as a platform for learning and knowledge sharing to improve their understanding in the topic of Geometry, that is, Angle Properties. Anonymity remained an issue despite names were being omitted from the works.

This study provides insights for the educators and researchers on the impact of peer assessment in the students learning. This study allows the readers to make evaluation how this approach could be implemented in the future classroom learning particularly in the learning of Mathematics. Peer Assessment promotes student-centred learning by making students accountable of their learning through making judgement of the quality of different works presented with less direct instruction

from the teacher. It is an approach that provides students the opportunity to reflect and link with their prior knowledge when assessing against the assessment criteria.

The finding of this study is consistent with other studies done previously in the learning of Mathematics. For the effective use of peer assessment would depend on the class ability; students' fluency in Mathematics knowledge in the area being assessed and the ability to express their Mathematical thoughts qualitatively. Students' confidence is also an important factor that has effect on how students would participate actively in the peer assessment task given. Students with weak knowledge in the topic of assessment are likely having lower confidence to assess their peers. More importantly, the effectiveness of peer assessment might result differently suppose it was done with higher degree of homogeneity in students' ability and even in a classroom of mixed gender. In addition, it is also important to note that this study investigated the impact of peer assessment on a closed-ended type of assessment, specifically, students' classwork. Different criteria and outcome of revised work might result suppose peer assessment is conducted on open-ended type of assessment such as Mathematics poster and presentation (Nor & Shahrill, 2014).

References

[1] Alqassab, M., Strijbos, J. W., & Ufer, S. (2018). Training peer-feedback skills on geometric construction tasks: Role of domain knowledge and peer-feedback level. *European Journal of Psychology Education, 33*(1), 11-30.

[2] Adediwura, A. A. (2015). Relationship between learning outcomes and peer assessment practice. *European Scientific Journal*, *11*(16), 353-368.

[3] Ahamad, S. N. S. H., Li, H-C., Shahrill, M., & Prahmana, R. C. I. (2018). Implementation of problem-based learning in geometry lessons. *Journal of Physics: Conference Series* (Vol. 943, No. 1, p. 012008). IOP Publishing.

[4] Andrade, H., & Ying, D. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation*, *10*(3), 1-11.

[5] Babakus, E., & Mangold, W. G. (1992). Adapting the SERVQUAL scale to hospital service: An empirical investigation. *Health Service Research*, *26*(6), 767-786.

[6] Black, P., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *80*(2), 139-144.

[7] Brown, G. T. L. (2004). Teacher's conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice*, *11*(3), 301-318.

[8] Botty, H. M. R. H., & Shahrill, M. (2015). A critical review of the school-based assessment in Brunei Darussalam. *International Journal of Education and Practice*, 3(1), 17-27.

[9] Chan, K. M. (2013). *Peer assessment in Mathematics lessons: An action research in an eight grade class in Macau* (Master's thesis, University of Hong Kong, Pokfulam, Hong Kong SAR). Retrieved from http://dx.doi.org/10.5353/th_b5210137

[10] Chong, J. S. Y, Chong, M. S. F, Shahrill, M., & Abdullah, N. A. (2017). Implementing inquiry-based learning and examining the effects in junior college Probability lessons*. Journal of Mathematics Education*, *8*(2), 157-164.
[11] Chukwuyenum, A. N., & Adeleye, B. A. (2013). Impact of peer assessment of performance in Mathematics among senior secondary school students in Delta state, Nigeria. *Journal of Emerging Trends in Educational Research and Policy Studies*, *4*(5), 719-725.
[12] Damit, A. H., Shahrill, M., & Roslan, R. M. (2015). Investigating the effectiveness of an assessment task through collaboration in a Bruneian classroom. *Mediterranean Journal of Social Science*, *6*(6 S1), 214-223.
[13] Denscombe, M. (2010). *Research guide for small-scale social research projects* (4th ed.). Maidenhead: McGraw-Hill Open University Press.
[14] Donaldson, A. J., & Topping, K. J. (1996). *The peer tutor training handbook for higher and further education*. Dundee: Centre of Paired Learning, University of Dundee.
[15] Egodawatte, G. (2010). A rubric to self-assess and peer-assess mathematical problem solving tasks of college students*. Acta Didactica Napocensia, 3*(1), 78.
[16] Falchikov, N. (2001*). Learning together: Peer tutoring in higher education*. London: Routledge Falmer.
[17] Falchikov, N. (2004). *Improving assessment through student involvement: Practical solutions for learning in higher and further education*. New York: Routledge Falmer.
[18] Field, A. P. (2009). *Discovering statistics using SPSS: And sex and drugs and rock 'n' roll* (3rd ed.). London: SAGE Publications Ltd.
[19] Ghaicha, A. (2016). Theoretical framework for educational assessment: A synoptic review*. Journal of Education and Practice*, *7*(24), 212-231.
[20] Gielen, S. (2007). Peer assessment as a tool for learning. (Unpublished doctoral dissertation). Catholic University Leuven, Belgium.
[21] Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyyen, K. (2010). Improving the effectiveness of peer feedback for learning*. Learning and Instruction, 20*, 304 - 315.
[22] Huang, J., & Normandia, B. (2009). Students' perceptions on communicating mathematically: A case study of a secondary Mathematics classroom. *The International Journal of Learning, 16*(5), 1-21.
[23] Huba, M. & Freed, J. (2000*). Learner-centered assessment on college campuses*. Boston: Allyn & Bacon.
[24] Husamah. (2015). Blended project based learning: Metacognitive awareness of Biology education new students. *Journal of Education and Learning, 9*(4), 274-281.
[25] Idris, S., Jawawi, R., Mahadi. M., Matzin, R., Shahrill, M., Jaidin, J. H., Petra, N., & Mundia, L. (2017). The use of rubrics in developing students' understanding in History. *Advanced Science Letters*, *23*(2), 901-904.

[26] Johnson, A. P. (2008). *A short guide to action research* (3rd ed.). Boston: Allyn & Bacon.
[27] Kearney, S. P., & Perkins, T. (2011, October). *Improving engagement: The use of 'authentic self and peer assessment for learning' to enhance the student learning experience.* Paper presented at The Academic and Business Research Institute Conference, Las Vegas, NV.
[28] Kollar, I. & Fischer, F. (2010). Peer assessment as collaborative learning: A cognitive perspective. *Learning and Instruction*, 344-348.
[29] Kulm, G. (1994). *Mathematics assessment: What works in the classroom*. San Francisco, CA: Jossey Bass Inc.
[30] Langan, M. & Wheater, P. (2003). Can students assess students effectively? Some insights into peer-assessment. *Learning and Teaching in Action*, *2*(1). Retrieved April 22, 2018 from https://www.celt.mmu.ac.uk/ltia/issue4/langanwheater.shtml.
[31] Langan, A. M., Wheater, C. P., Shaw, E. M., Haines, B. J., Cullen, W. R., Boyle, J., Penney, D., Oldekop, J., Ashcroft, C., Lockey, L., & Preziosi, R. F. (2005). Peer assessment of oral presentations: Effects of student gender, university affiliation and participation in the development of assessment criteria. *Assessment & Evaluation in Higher Education, 30*(1), 21 -34.
[32] Lin, S. S. J., Liu, E. Z. F., & Yuan, S. M. (2001). Web-based peer assessment: Feedback for students with various thinking styles. *Journal of Computer Assisted Learning, 17*, 420-432.
[33] Liu, N. F. & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education, 11*, 279-290.
[34] Logan, E. (2009). Self and peer assessment in action. *Practitioner Research in Higher Education*, *3*(1), 29-35.
[35] Medcalf, J. (1992*). Peer tutoring in reading*. Hastings: Medcalf Press.
[36] Mills, G. E. (2011). *Action research: A guide for the teacher researcher* (with MyEducationLab). (4th ed.). Upper Saddle River, NH: Pearson/Allyn & Bacon.
[37] Mohammad, H. H., Jawawi, R., Matzin, R., Shahrill, M., Jaidin, J. H., Mundia, L., & Shamsu, L. S. (2017). Engaging students' learning with elements of formative assessment. *Turkish Online Journal of Educational Technology*, December 2017, Special Issue for ITEC 2017, *12*(1), 281-290.
[38] Nor, H. N. H. M., & Shahrill, M. (2014). Incorporating the use of poster and oral presentations as an alternative assessment in the teaching of secondary mathematics. *Proceedings of the 2nd International Conference on Social Sciences Research* (pp. 369-378). Kota Kinabalu, Sabah, Malaysia: ICSSR 2014, WorldConferences.net.
[39] Othman, R., Shahrill, M., Mundia, L., Tan, A., & Huda, M. (2016). Investigating the relationship between the student's ability and learning preferences: Evidence from year 7 Mathematics students. *The New Educational Review*, *44*(2), 125-138.

[40] Panadero, E. (2016). Is it safe? Social, interpersonal, and human effects of peer assessment: A review and future directions. In G. T. L. Brown & L. R. Harris (Eds.), *Human factors and social conditions of assessment* (pp. 1-39). New York, NY: Routledge.
[41] Pantiwanti, Y. & Husamah. (2017). Self and peer assessments in active learning model to increase metacognitive awareness and cognitive abilities. *International Journal of Instruction*, *10*(4), 185-202.
[42] Patchan, M. M., & Schunn, C. D. (2015). Understanding the benefits of providing peer feeback: How students respond to peers' text of varying quality. *Instructional Science, 43*(5), 591-614.
[43] Piaget, J. (1971). *Science of education and the psychology of the child*. London: Longman.
[44] Plake, B. S., & Impara, J. C. (1997). Teacher assessment literacy: What do teachers know about assessment? In G. D. Phye (Ed.), *Handbook of classroom assessment* (pp. 55-68). New York: Academic Press.
[45] Plake, B. S., Impara, J. C., & Fager, J. J. (1993). Assessment competencies of teachers: A national survey. *Educational Measurement: Issues and Practice*, *12*(4), 10-12, 39.
[46] Pugalee, D. K. (2004). A comparison of verbal and written descriptions of students' problem solving. *Education Studies in Mathematics*, *55*(1), 27-47.
[47] Revilla, M. A., Saris, W. E. & Krosnick, J. A. (2014). Choosing the number of categories in agree-disagree scales. *Sociological Methods & Research*, *43*(1), 73-97.
[48] Slavin, R. E. (1996). Research for the future: Research on cooperative learning achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21,* 43-69.
[49] Stiggins, R. J. (2001). *Student-involved classroom assessment* (3rd ed.). Upper Saddle River, NJ: Merrill/Prentice-Hall.
[50] Strijbos, J. W., & Wichmann, A. (2018). Promoting learning by leveraging the collaborative nature of formative peer assessment with instructional scaffolds. *European Journal Psychology Education*, *33*, 1-9.
[51] Thomas, G., Martin, D., & Pleasants, K. (2011). Using self- and peer-assessment to enhance students' future-learning in higher education*. Journal of University Teaching & Learning Practice*, *8*(1), 1-17.
[52] Topping, K. J. (1996). The effectiveness of peer tutoring in further and higher education: A typology and review of the literature. *Higher Education, 32,* 321-345.
[53] Topping, K. J. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*(3), 249-276.
[54] Topping, K. J. (2003). Self- and peer- assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy, & E. Cacallaer (Eds), *Optimising new modes of assessment: In search of qualities and standards* (pp. 55-87). Dordrecht, The Netherlands: Kluwer Academic Publishers.

[55] Topping, K. J. (2005). Trends in peer learning. *Educational Psychology, 25*(16), 631-635.
[56] Tsivitanidou, O. E., & Constantinou, C. P. (2016). A study of students' heuristics and strategy patterns in web-based reciprocal peer assessment for science learning. *The Internet and Higher Education, 29,* 12-22.
[57] Tsivitanidou, O. E., Constantinou, C. P., Labudde, P., Ronnebeck, S., & Ropohl, M. (2018). Reciprocal peer assessment as a learning tool for secondary school students in modeling-based learning. *European Journal Psychology Education*, *33*, 51-75.
[58] Van Zundert, M. J., Könings, K. D., Sluijsmans, D. M. A., & Van Merriënboer, J. J. G. (2012). The differential effects of task complexity on domain-specific and peer assessment skills. *Educational Psychology: An International Journal of Experimental Educational Psychology, 32*(1), 127-145.
[59] Vickerman, P. (2010). Student perspectives on formative peer assessment: An attempt to deepen learning?. *Assessment & Evaluation in Higher Education, 34*(2), 221-230.
[60] Wheater, R., Durbin, B., McNamara, S., & Classick, R. (2016). *Is mathematics education in England working for everyone? NFER analysis of the PISA performance of disadvantages pupils*. Slough: NFER.
[61] Wood, D. & Kurzel, F. (2008). Engaging students in reflective practice through a process of formative peer review and peer assessment. In *ATN Assessment Conference 2008: Engaging students in assessment*.
[62] Zevenbergen, R. (2001). Peer assessment of student constructed posters: Assessment alternatives in preservice Mathematics education. *Journal of Mathematics Teacher Education*, *4*(2), 95-113.