Abstract: This study investigated teachers’ grading decision making, focusing on their beliefs underlying their grading decision making, their grading practices and assessment types, and factors they considered in grading decision making. Two teachers from two junior high schools applying different curriculum policies in grade reporting in Indonesian educational context were interviewed in depth to reveal their grading decision making. The results show that the teachers believe that assigning grades is not only for measuring the students’ ability, but also for making them active users of the language, giving them life skills and experience and motivating them. In addition, these teachers use various grading practices involving not only formal assessment but also informal one. Remedial tests, giving more tasks and grade adjustment are also done to reach the criteria of passing grade. The factors considered in grading decision making are achievement factors in the forms of scores from formal and informal assessment and non-achievement factors involving students’ effort and behavior, curriculum and school policy.

Keywords: grading decision making, grading practices, assessment

DOI: http://dx.doi.org/10.15639/teflinjournal.v2i2/155-169

Grading, a process within the practice of classroom assessment where teachers calculate students’ grades for standardized report cards (Tierney, Simon, &
Charland, 2011), is an important teachers’ professional responsibility. Grades, which result from grading, also play an important role in students’ life. They determine students learning paths, scholarships, post-secondary opportunities, and career choice (Tierney, 2015). Grades, therefore, should provide information about students’ achievement for stakeholders (Cheng & Sun, 2015) and rightly represent the students’ ability in learning.

Despite the importance of grading, a lot of teachers consider it a difficult task (Allen, 2005). Allen (2005) illustrates some of the difficulties, which includes handling non-academic factors, such as, students’ efforts which should not be part of a grade that represents academic achievement. Further, he posits that teachers also seem confused with the communication function of grades and they try to communicate multiple pieces of information about students that cannot possibly be contained within a single academic mark. As a result, many teachers assign invalid grades.

In fact, when valid and fair grading is required, teachers should follow some principles of grading as synthesized by Tierney et al. (2011) from several assessment guidelines. They are as follows:

1. When the purpose of grading is to report on student achievement, grades should be referenced to the curriculum objectives or learning expectations (criterion referenced); 2. A grade should be an accurate representation of achievement, so non-achievement factors should be reported separately to permit valid interpretation by stakeholders; 3. Results from multiple assessments should be combined carefully, with weighting that reflects the learning expectations, to ensure that the grade accurately summarizes achievement; 4. Information about grading should be clearly communicated so that grades are justified and their meaning is understood by students, parents, and other teachers (Tierney et al., 2011, p. 212).

Although the involvement of non-academic factors in grading decision making is not recommended, a lot of teachers still consider these factors in grading their students. Previous studies on teachers’ grading decision making revealed that both academic and non-academic factors were taken into account by the teachers and they affected teachers’ practices of grading decision making (Cheng & Sun, 2015; Cheng & Wang, 2007; Duncan & Noonan, 2007; Liu & Yan, 2015; McMillan, 2001; Prøitz, 2013; Tierney, 2015; Yesbeck, 2011; Zoeckler, 2005). The factors such as the learners’ effort, homework, and study habits were reported to have greater weight in grading decision making (Cheng
Subject matter was also found to affect teachers’ grading decision making since teachers from different subject areas have different considerations in using non-academic factors as contributing elements to their grading practices as well as different tendencies in choosing assessment strategies in their grading practices (Duncan & Noonan, 2007).

However, this involvement of non-academic factors by teachers in grading is not without reasons and support. It is argued by Zoeckler (2005) that teachers need to consider such factors as the students’ individual good personal characteristics, the teachers’ willingness to motivate, to encourage and to appreciate students’ greater learning effort and achievement for the sake of students’ improvement and further development, the accountability with the larger community, the school organization and parents, and the teachers’ own integrity. In line with that, Tierney (2015) revealed that the reasons behind what so-called grade alteration, where teachers do not report the students’ ability accurately by deliberately increasing or decreasing a grade for a reason that relates to achievement, were teachers’ empathy for students, and their wish to give students opportunity, and teach them about life where everyone gets his/her due, good or bad.

Further, grading decision making is not only determined by the students’ academic and non-academic factors, but also other factors. Cheng and Sun (2015) posited that grading decision making is influenced by both internal factors such as teacher’s values and beliefs about learning and external factors such as parental influences and state accountability testing pressures. Teacher’s belief is considered a major component in grading (Brookhart, et al., 2016; Tierney, 2015). However, review on grading research done by Brookhart et al. (2016) indicates that most studies on grading practices over the past 20 years (e.g. Randall & Engelhard, 2009, 2010; Sun & Cheng, 2014; Tierney, et al., 2011) focus more on teachers’ use of cognitive and non-cognitive aspects in grading practices. It implies a need for conducting more in-depth studies on teachers’ grading practices, especially dealing with teachers’ beliefs underlying their grading decision making. Further, contextual factors also have a strong influence on teachers’ grading decision making. One study by Cheng and Sun (2015), for example, found that teachers of English in the Chinese context considered achievement and non-achievement factors in grading and placed greater weight on non-achievement factors, such as effort, homework, and study habits. For the types of assessment methods, major examinations were most frequently used by teachers because of their dominant influence in China.
In sum, previous research has indicated that teachers’ beliefs of grading purposes, their grading practice and use of assessment types, as well as their consideration of factors in grading certainly make up important aspects of teachers’ grading decision making. It is crucial to do more studies on grading decision making in different educational contexts because each context is unique and has different characteristics affecting teachers’ grading decision making. One unique educational context is Indonesia where curriculum transition is in progress. A new curriculum called 2013 Curriculum is replacing the old one, a School-based Curriculum (Kurikulum Tingkat Satuan Pendidikan/KTSP). This change also causes a change in the assessment approach. The 2013 Curriculum emphasizes authentic assessment which requires teachers to assess students’ cognitive, affective and psychomotor aspects (Indonesian Ministerial Regulation No. 66, 2013). As a consequence of this assessment reform, 2013 Curriculum requires students’ report card format to include students’ achievement in those three aspects. This is obviously different from KTSP report card which only requires teachers to report students’ achievement in a single grade. This present study is, therefore, intended to investigate teachers’ grading decision making in Indonesian educational context where the curriculum transition is happening as well as to compare the teachers’ grading decision making between two professional teachers in two schools with different curriculum policy. The study should provide essential information about teachers’ grading decision making from different perspectives and broaden the theory of grading decision making in general.

METHOD

In this study, the grading decision making of two English teachers from two different junior high schools in Indonesia was examined. One of the teachers was from a private junior high school, henceforth identified as Teacher 1. The other was from a public junior high school (Teacher 2). The two teachers had taught English for more than five years and had been certified as professional teachers. Both of them had a bachelor degree in English Language Education. During their teaching, they also had joined some teacher professional trainings especially related to curriculum implementation. In a week, both teachers had to teach at least 24 hours as the requirement of teacher certification. They taught six classes, each of which has a four-hour meeting in a week.
These two teachers were from two different junior high schools which applied different curriculums. The private junior high school still applied the old curriculum which was the school-based curriculum, while the public one applied the new 2013 Curriculum. The two curriculums are pretty different in terms of assessment method where the new curriculum gives more emphasis on authentic assessment. The students are expected to be assessed not only on their cognitive aspect but also on psychomotor, affective, and religious aspects. Consequently, the form of students’ report card of 2013 curriculum is also different from that of the old curriculum which only requires a single grade in the form of number for each subject. The report card of 2013 Curriculum, on the other hand, requires teachers to give students grades for cognitive, psychomotor, affective and religious aspects with descriptive reports of students’ learning.

In order to reveal the two English teachers’ grading decision making in depth, an open-ended interview guide was used. The interview guide consisted of three sections following research questions; they are (1) teachers’ beliefs in grading practices, (2) grading practice and assessment types, and (3) factors considered by teachers in grading practices. The interview was retrospective in nature where the two teachers were asked some questions and required to recall and reflect on what they had done in grading the students. The interviews were conducted face-to-face and lasted 45 to 60 minutes. All of the interviews with these two teachers were recorded and transcribed in summary form. The teachers’ responses to the interview questions were then coded, classified, analyzed, and interpreted in order to answer the research questions.

In addition to the interviews, some documents showing teachers’ grading decision making practices were collected from the teachers in order to complete the data as well as for triangulation. The teachers’ scoring formula, the students’ raw scores, and the students’ report card were among samples documented in this study.

FINDINGS AND DISCUSSION

Findings

The results of this study are presented based on the three research questions; they are the beliefs underlying teachers’ grading decision making, teacher-
Teachers’ grading practices and their use of assessment types, and factors considered by teachers in grading decision making.

**Teachers’ Beliefs Underlying Grading Decision Making**

When asked about teachers’ beliefs in making decisions about the students’ grades, both teachers had similar answers. Teacher 1 said that so far she assessed the students to measure the students’ ability. Besides, she said that in assessing the students, she expected to motivate them to be better learners. She also decided to have some types of assessment which can give the students real and practical experience and get the scores from those assessments. This decision was made because she believed that giving them real experience was more meaningful for the students.

Teacher 2 also thought that the assessments she did so far were aimed at measuring the students’ ability as fairly as possible. In case she had to change some students’ scores because of some factors, she also gave some additional scores for the other students even though the addition was not the same. Concerning grade alteration, she agreed that grades can give the students motivation in learning. Therefore, although sometimes she gave higher grades for the students than their real ability, she thought that it was fine for it can motivate them to be better language users. Further, in grading practices, she used various forms of assessment. She believed that the more frequently the students were assessed, the better they were in using the language as a means of communication.

From the teachers’ responses, it can be seen that teachers in this study share similar beliefs in grading the students although they come from different schools applying different curriculums and reporting policies. They believed that assessment and grading practices are not only for the sake of assigning final grades to the students. Grades can serve more useful functions for the students’ learning and future life.

**Teachers’ Grading Practices and Their Use of Assessment Types**

Based on the results of the interview with two English teachers about their grading practices and their use of assessment types, it was found that both teachers had done various grading practices and used different assessment types in order to arrive at the students’ final grades in the report cards. They used both written and oral assessments during the teaching and learning pro-
The assessments were for individuals or groups depending on the tasks. The time for assessment was during and after the teaching and learning process in a semester. Formal assessment involving tests and in the forms of tasks and projects were also administered by the teachers to collect scores to assign grades to the students.

For formal assessment, Teacher 1 and Teacher 2 gave the students some kinds of written tests, i.e. daily tests, a midterm test, and a final test. The daily tests were done after the students finished some basic lesson units. They were usually done three to four times in a semester. Mid-term and final tests were done once a semester in the middle and at the end of the term respectively.

The scores obtained from these tests had to reach the Minimum Mastery Criteria (Kriteria Ketuntasan Minimal/KKM) set by the schools. In this case, the two junior high schools had the same Minimum Mastery Criteria of 75. The students who could not reach the criteria had to take another test to get better scores. The tests taken for the second time by the students were called remedial tests. Teacher 1 usually gave opportunity to have remedial tests for the students who failed to reach 75 in mid-term tests and final tests. In case the students could not reach the minimum criteria after having remedial tests a number of times, the teacher would deliberately increase the students’ score. For the daily tests, however, the students were given additional assignment when they could not reach 75. On the other hand, Teacher 2 only gave the students opportunity to take remedial tests when the students failed in the daily tests. There were no remedial tests after mid-term and final tests. Interestingly, Teacher 2 gave the students different questions from what was given in the regular tests. For some low-achieving students who always failed in reaching the minimum mastery criteria, both teachers deliberately increased the students’ score to reach the criteria.

Dealing with informal assessment, Teacher 1 and 2 gave the students quizzes, homework, and individual/group tasks to get the scores. Teacher 1 usually preferred giving the students story-telling tasks in written and oral forms. For quizzes, she commonly gave impromptu vocabulary quizzes by asking the students to translate the meanings of some words without telling them before about the quiz. Homework was given after a certain topic was discussed and used for giving the students more exercises at home. Teacher 2 also had similar activities for informal assessment. However, she gave more various tasks to the students, such as writing descriptive texts completed with pictures, drama projects, and some other projects.
In scoring the students’ homework, tasks and projects, both teachers usually considered some criteria although they did not prepare them strictly. They only announced orally some aspects of scoring such as the generic structures and language features of a certain writing text before the students worked on writing projects. For such activities, the teachers commonly gave higher scores than the Minimum Mastery Criteria. This was to appreciate the students’ effort in doing the tasks. Besides, the teachers did so to anticipate poor scores in the written tests.

Peer-assessment was also occasionally considered by Teacher 2 in deciding the quality of students’ task. For example, she asked the students whether or not their friends’ performance was good. Sometimes, she asked the students to exchange their work and check each other’s work. Similarly, Teacher 1 sometimes asked the students to check each other’s homework. Although such practices were not the ideal practices of peer assessment in which some clear and well-prepared criteria are necessary for guiding the assessment, they can be considered peer assessment practices since students did some evaluation on their friends’ work or performance.

Upon the end of the semester, the scores obtained from various activities were calculated for the final grade in the report card. Teacher 1 summed up the scores of homework, quizzes, tasks and daily tests, and averaged them to get the average of daily scores. The average of daily scores, the scores of the mid-term test, and the scores of the final test were then summed up and divided by three in order to get the final grades in the report cards. Teacher 2 had a different way in calculating the scores for the final grades. She summed up the scores of homework, quizzes and daily tests and got their average. The average of daily score was then summed up with the scores of the mid-term test and the final test which had been multiplied by two. The total scores were then divided by five to get the final grades for cognitive aspect in the report cards. For the scores of tasks, Teacher 2 averaged them and considered the average score as the grade for psychomotor aspect in the report cards.

It can be summarized that teachers in this study used various grading practices ranging from written to oral tests, individual to group work, and also formative to summative assessment in arriving at the final grades in the students’ report cards. The assessment types used by the teachers were both formal, involving written tests, and informal assessment which comprised homework, tasks, and projects. However, compared to teacher 1, teacher 2 seemed to do more efforts in assessing and grading students using informal assessment.
Factors Considered in Grading Decision Making

For both teachers, the process of deciding final grades for the students involved some factors. Teacher 1 and Teacher 2 said that students’ achievement is the main factor in assigning grades for the students. However, they had different opinions in considering the students’ achievement. According to Teacher 1, in assigning the students’ grades, the process of learning is the main consideration. For example, when the students get good scores in the test while they are not really good in the process of learning, the scores in the report card will be adjusted. On the contrary, Teacher 2 relied more on the result of written tests if there were different scores between tests and tasks or homework.

Besides the students’ achievement, both teachers also considered the students’ learning effort, study habits, and behavior in assigning grades to the students. They usually gave additional scores to the students who learnt hard and had good study habit although the results in the tests were not very satisfying. This extra score was given to appreciate their learning and motivate them to do better in the future.

The students’ active participation in learning was another consideration in grading decision making. However, according to Teacher 1, students who were active in classroom activities were usually good students and their scores in English had been satisfying even though the teacher did not give additional scores for their active participation. But still, she gave these active students better scores for the sake of fairness. Teacher 2, on the other hand, always encouraged every student to participate actively in the classroom activities. Therefore, according to her, it was fair to give her students extra scores in addition to their original scores.

School policy was also said to influence the teachers in assigning the grade for the students. Although both teachers were from two different schools with different conditions, the schools generally required the teachers to give the students grades equal or higher than the Minimum Mastery Criteria. When there were more poor grades in the report cards, the students were at a risk of not being able to go to the next level of education or even at a risk of dropping out of school. According to Teacher 1, in the private school where she was teaching, it was almost impossible to give the students grades lower than the Minimum Mastery Criteria although the students performed very poorly. The private school got the advantages of having more students. The more students they have, the more funding they get for doing better teaching learning process.
in the school. So, giving lower grades means discouraging the students to study at the school and there was a risk to have the students drop out of school. Although the case was not as serious as in the private school, the same policy was also applied in the public school where Teacher 2 was teaching. To accommodate such policy and give the students fair grades at the same time, both teachers usually gave more tasks to those poor students before improving their final grades in the report cards.

In addition, grading decision making cannot be separated from the government educational policy. Both teachers experienced how hard the students passed the school graduation criteria. According to them, the criteria for school graduation involved some aspects such as National Examination score and grades in their school report cards. Since a lot of students failed in National Examination, while the school demanded the students to graduate for many reasons, the teacher usually anticipated this by giving the students better grades in the report cards although the grades did not really represent the students’ ability. Good grades in the report cards helped students meet the graduation criteria. Such grading practice was considered common among the teachers.

Finally, grade reporting format based on curriculums also gave a big influence in teachers’ grading decision making. Since the two teachers were from different schools implementing different curriculums, the grade reporting formats required were also different. According to Teacher 1, she just had to provide a single grade in students’ report cards. The grade was in the form of number ranging from 0 to 100. To assign the grade, she considered the students’ scores from tests and some tasks. Non-academic aspects were the final consideration in grading decision making. A slightly different grading practice was done by Teacher 2 because of the different form of report card. Based on the 2013 Curriculum implemented in her school, teachers were required to provide numerical and letter grades as well as description for students’ knowledge and skill mastery. In addition, the social and spiritual attitude must be given letter grade with its description. Teacher 2 said that she had to collect as many scores as possible from the students’ activities in the classroom. She got the scores not only from tests but also from observing the students’ performance and activities. As a consequence, she tried to give the students more various tasks and projects in order to get more information about the students.

In short, in grading decision making, the teachers in this study considered not only the students’ achievement but also other factors such as students’ efforts, their study habit, and their active participation. The school policy, gov-
government policy, and curriculum were also inseparable considerations in their grading decision making.

Discussion

The results of the current study show that in terms of teachers’ beliefs underlying their grading practice and grading decision making, teachers in this study believe that assessment and grading practice are not only for measuring the students’ ability, but also making them active users of language, giving them life skills/experience and motivating the students. Such beliefs dealing with not only academic factor but also moral factors in grading practice are supported by O’Malley and Pierce (1996) who state that assigning grades has two basic purposes, that is, for reflecting on students’ learning accomplishment and for motivating students. Further, Tierney (2015) reveals his study about moral reasons underlying teachers’ altered grading practices such as the need for compassion; the desire to provide students with opportunity; and the intent to teach life lessons.

Dealing with grading and assessment practice, teachers have undergone relatively various and complex processes of grading practices by using formal and informal assessment types in order to obtain the students’ scores before they can decide their students’ final grade in the report card. This is in line with Tierney’s (2015) statement that in grading practice, teachers have to go through processes of scoring students’ various work and calculating the final grade from those scores. However, although the teachers in this study have used multiple assessment methods consisting of formal and informal assessment as recommended by experts of educational assessment (Stiggin, Frisbie & Griswold, 1989), they need to be more creative and well-prepared in implementing various kinds of informal assessment or authentic assessment such as observation, peer assessment, self-assessment, portfolio, and journal in their classroom as it is the recommended assessment procedures in 2013 curriculum in Indonesia. In addition, it is believed that grades can be useful if they are based on authentic assessment because they can provide a comprehensive picture of student growth and achievement especially when combined with illustrative samples of student work and informative scoring rubrics (O’Malley & Pierce, 1996).

The grading process that teachers have to go through is sometimes useless when teachers deal with low-achieving students who cannot achieve the Minimum Mastery Criteria. This happens because teachers have to manipulate those
students’ final grades to meet school and curriculum demands as well as government educational policy on criteria for graduation. Such practices actually put teachers in a dilemma between doing a fair assessment and fulfilling their school demands. As a result, the teachers in this study decided to give students more tasks to ‘legalize’ their practice of altering students’ grade. This is considered wise although this practice is against measurement theory and thus should be avoided (Yesbeck, 2011).

While it is clear from this study that teachers consider achievement factors involving scores obtained from formal and informal assessment in their grading decision making, their considerations of non-achievement factors, such as, students’ effort and behavior are also apparent. This contradicts the recommendation of measurement theory experts that grading should consider students’ achievement only (see e.g., Yesbeck, 2011). However, a number of studies report similar grading practices involving non-achievement factors (Bowers, 2011; Cheng and Sun, 2015; Cheng & Wang, 2007, Liu & Yan, 2015; Duncan & Noonan, 2007). Such grading practice might also be partly triggered by teachers’ empathy for their students so that they assign grades higher than what students can do (Brown, 2004: 290; Tierney, 2015). Involving moral dimension is considered fair in grading decision making since moral development influences students’ ways of study and thus, indirectly affects their grades (Zoekler, 2005). With the changing curriculum in Indonesia, there has been an effort to separate, in students’ report card, the achievement factors, that is, the students’ cognitive and psychomotor factors, and the non-achievement ones, that is, the affective factors. This might minimize the grading practice involving non-achievement factors.

CONCLUSIONS AND SUGGESTIONS

It can be concluded that while both teachers conduct a variety of grading practices and types of assessment in the classroom, non-achievement factors also play an important role in their decisions for the students’ final grades. Under their strong beliefs that assigning grades is not only for measuring the students’ ability, but also for making them active users of language and giving them life skills/experience and motivating the students, both factors, achievement and non-achievement factors interplay flexibly in the process of grading decision making. Remedial tests, giving more tasks and grade adjustment are also done to reach the criteria of passing grade in order to meet the demands of the
school, curriculum and government educational policies. In addition, although the differences in the teachers’ grading practices in this study are not very strong, the curriculum applied in the school is likely considered by teachers in their grading decision making.

If the ideal grading decision making based on assessment theory is sought, the teachers should be given more authority to assign an objective grade to their students without being restricted by the external factors of school, curriculum, and government educational policy demands. More training on assessment and grading should also be given to the teachers in order that they can be more skillful in assessment and grading since assessment is a key element in students’ learning.

Future studies involving more teachers with different characteristics are needed to reveal practices of grading decision making among teachers in different educational contexts. Involving teachers with a different range of teaching experience may also become an interesting focus of the next research studies.

REFERENCES


