

*Classroom Assessment Standards: Indonesian EFL Teachers' Assessment Practices
in the Amid of the Pandemic*

Romadhon, Universitas Pendidikan Indonesia, Indonesia
Didi Sukyadi, Universitas Pendidikan Indonesia, Indonesia
Pupung Purnawarman, Universitas Pendidikan Indonesia, Indonesia

The Asian Conference on Education 2022
Official Conference Proceedings

Abstract

This study investigated assessment standards conducted by EFL Indonesian secondary school teachers and how they coped with the emergency remote teaching during the pandemic. The Indonesian curriculum assessment standard and Brookhart's educational assessment standards were used as a research framework. An explanatory sequential research design was employed to capture the research problems holistically. First, using an online questionnaire based on Brookhart's educational assessment standards, 119 EFL teachers at junior and senior high schools who taught students in online settings during the pandemic were participated. Then, the obtained data were explored by conducting a semi-constructed interview to four selected teachers from different schools. In analyzing quantitative data, descriptive statistics was used to gain the frequency data of teachers' assessment. Sequentially, the interviews were transcribed, categorized, and interpreted following the Indonesian curriculum assessment standards as qualitative data. The findings show that EFL teachers performed most assessment practices in Brookhart's standards during the pandemic. However, through interviews, teachers admitted that they did not maximize classroom assessment mentioned in the curriculum assessment standards because the classroom situation was different and unpredictable in online learning during the pandemic. This study is theoretically and practically significant for implementing and evaluating classroom assessment practices amid the pandemic.

Keywords: Assessment Standards, Classroom Assessment, EFL Secondary Teachers, Online Learning Assessment, Assessment During the Pandemic

iafor

The International Academic Forum
www.iafor.org

1. Introduction

Almost two years since the Covid-19 pandemic was firstly spread. In education, countries worldwide try to adapt the pandemic situation by introducing emergency remote teaching as one of the responses. Amidst the global pandemic, the Indonesian Ministry of Education and Culture, on official letter No. 719/P/2020, launched an emergency curriculum during the pandemic. The emergency curriculum during the pandemic requires English teachers to be adaptable in teaching (Hapsari, 2020) and specifically in assessing students. Regardless of how the instructional teaching is simplified, the assessment standards and the essence do not change, still following the Ministry of education and culture No. 23/2016. Teachers must stick to the aims, principles, mechanisms, procedures, and classroom assessment instruments (MOE, 2016). Moreover, the pandemic leads to an educational shift from traditional to online learning or emergency remote teaching (ERT). According to Hodges et al. (2020), ERT refers to a sudden shift in education from face-to-face to an online learning environment responding to an emergency. Further, they state that ERT and online learning are significantly different in which online learning requires voluntary planning and design of virtual delivery, while ERT is used in an unpredicted situation. Regardless of the situation that makes those concepts different, they have the same underpinning framework, technology-based education.

However, based on the data revealed by the Indonesian Teachers Association (IGI), the implementation of distance or online learning in the last three months of 2020 showed only 60% because of teachers' inferior skills in utilizing technology (see Dewi & Wajdi, 2021). It is supported by several studies' findings that EFL teachers had a lack of readiness and inability to teach online because they still struggled using digital platforms. As a result, they transformed traditional learning activities into online learning without maximizing the technology (Atmojo & Nugroho, 2020; Ilmiani et al., 2021; Kusumastuti, 2020; Mardiah, 2020; Nugroho & Haghegh, 2021; Situmorang et al., 2020; Sumardi & Nugrahani, 2021; Supriadi et al., 2020; Syarifah & Zainil, 2021; Yundayani et al., 2021). Online learning impacts assessment and evaluation in which it has unique aspects compared to face-to-face learning (Dumford & Miller, 2018). For example, Serwatka (2002, in Dumford & Miller, 2018) stated that specific techniques in a traditional classroom do not always work in online learning. Although ERT forces teachers to utilize technology, pedagogical challenges reduce the function of technology for teaching and learning during the pandemic. Therefore, investigating the assessment practices of EFL teachers during the pandemic is pivotal.

This recent research investigated the frequent classroom assessment practices conducted by EFL Indonesian secondary school teachers and how they cope assessment standards during the pandemic with technology-based education mandated on the emergency curriculum with. The proposed research questions: What are frequent classroom assessment practices conducted by Indonesian secondary EFL teachers during the pandemic? How do Indonesian secondary EFL teachers practice their assessment standards regarding emergency remote teaching during the pandemic? This study is expected to contribute to some significance. First, theoretically, this study is significant as a framework for policymakers and educators in curriculum assessment and educational regulations during the pandemic. Secondly, this study is helpful practically to evaluate the implementation of curriculum assessment standards and educational regulations during the pandemic.

2. Literature Review

Assessing students is one of the integrated activities in the classroom that teachers must conduct. McMillan and Workman (1998) define classroom assessment as a process of collecting, interpreting, and using students' learning information to assist teachers in teaching decision-making. It is a fundamental aspect of learning instructions (Kearns, 2012). Then, Wang (2017) adds that classroom assessment assesses students' learning and achievement, covering eliciting-interpreting-using processes. Also, Cheng, Rogers, and Hu (2004) state that student achievements should be aligned with learning objectives for designing instructional, grading, and reporting practices. Karagül, Yüksel, and Altay (2017) mentioned that assessment and grading are essential aspects of effective teaching. Hence, it can be underlined that classroom assessment deals with teachers' instructional decisions for gaining students' learning information and achievement.

Classroom assessment in this current study cannot be separated from its standards. Assessment standards refer to a model of systematic guidelines for educational assessment. This current study used two standards as frameworks for collecting and analyzing data. The first assessment standard, taken from the 2013 Indonesian curriculum, is the national assessment standard (MOE, 2016). It is mentioned in the Ministry of Education and Culture policy (MOE, 2016) that the educational assessment standard refers to the criterion of scope, purpose, function, mechanism, procedure, and instrument for assessing student learning process and their achievement in three aspects: affective, cognitive, and psychomotor. It aligns with McMillan and Workman (1998) and Wang (2017) that assessment collects information about students' learning to evaluate their processes and achievements. Further, this assessment standard measures students' competency accomplishment in the daily and yearly reports and middle and final term. Teachers can give exams, observations, tasks, or other assessment forms following several procedures: a) setting aims of assessment referring to lesson plan, b) arranging question grid, c) creating instrument and its guidance, d) analyzing instrument quality, e) doing an assessment, f) collecting, analyzing, interpreting assessment results, g) reporting, and h) using assessment results. To capture student behavior, they observe during learning, then write it down in an observation sheet to be followed up and described. They have options to select a type of test, whether written, spoken, or tasks derived from selected competency for measuring student knowledge. At the same time, they give students practice, product, project, and portfolio for measuring students' skills. Teachers report the results in a range score from 0 to 100 with descriptions.

Another assessment standard for exploration and comparison, the researcher used Brookhart's educational assessment standard. Brookhart (2011:7) declares a model for educational assessment standards containing eleven points regarding classroom assessment practices. First, teachers should understand learning in the content area they teach. Second, they are expected to articulate clear learning objectives attainable and assessable by standards and curriculum. Third, they need to have a repertoire of strategies for communicating what achievement of a learning intention can be. Fourth is understanding the purposes and uses of the range of available assessment options and using them. Fifth, they can ascertain the specific knowledge and thinking skills required for students to do them. Sixth, they can provide practical, helpful feedback on student work. Seventh, they can construct scoring schemes that quantify student performance on classroom assessments into helpful information for decisions about students, classrooms, schools, and districts. These decisions should lead to improved student learning, growth, or development. Eighth, they are familiar with administering external assessments and interpreting their results to judge students,

classrooms, schools, and districts. Ninth, they are used to articulate their interpretations of assessment results and reasons about the educational decisions based on assessment results to the educational populations they serve (student and his/her family, class, school, community). The tenth is helping students use assessment information to make sound educational decisions. The last is understanding and carrying out their legal and ethical responsibilities in assessment as they conduct their work.

The researcher used those assessment standards as a theoretical framework for creating instruments in both quantitative and qualitative phases. The Indonesian educational assessment standards point to practices that English teachers applied before the pandemic with some principal adjustments in emergency remote teaching. This standard is suitable for acquiring assessment practices of English teachers during the pandemic. At the same time, having Brookhart's standard is an attempt to triangulate frameworks from the government policy to the perspectives of an expert as works of Brookhart are about educational assessment (Brookhart, 1993, 1997, 2011; Brookhart et al., 2016a).

3. Methods

3.1. Research Design

The researcher conducted this study using a mixed-method explanatory sequential design (Creswell & Guetterman, 2019) for several reasons. First, using quantitative and qualitative methods, the research questions can be understood and explored comprehensively. Furthermore, the researcher believes that this method is suitable for achieving a single study's research objectives (Creswell, 2012). This type of method is characterized by the quantitative data explored by qualitative data.

The researcher employed quantitative data collection using a questionnaire that was arranged and analyzed following Brookhart's educational standard (2011) to gain frequent classroom assessment practices by secondary EFL teachers during the pandemic. Then, the obtained data were followed up using a semi-constructed interview as a qualitative data collection. The interview questions were based on the Indonesian educational assessment standard (MOE, 2016) with some modifications mentioned in the ERT curriculum (MOE, 2020). The assessment standards used in this study: Brookhart's educational standard (2011) and the Indonesian educational assessment standard (MOE, 2016) have quite similar stages in assessment procedures as explained in the literature review. The data interview was employed to explore how they cope assessment practices in online learning with emergency remote teaching. Eventually, the interpretation of data was sequentially from all of those data.

3.2. Participants

This questionnaire was spread online targeting secondary teachers from junior and senior high in public and private schools under the Ministry of Education and Culture, and schools under the Ministry of Religion Affairs (Islamic junior and senior high in public and private schools). The researcher selected secondary schools to collect data because these levels of education significantly impacted the pedagogical shift compared with higher education. The specific criterion of EFL secondary school teachers is those who conducted online or distance learning during the pandemic. Due to the limited time of data collection, within a week, there were 119 respondents (see Table 1).

Table 1. Teaching experiences					
		Frequency	Per cent	Valid Percent	Cumulative Percent
Valid	< 1 year	7	5.9	5.9	5.9
	> 10 years	51	42.9	42.9	48.7
	1 - 2 years	23	19.3	19.3	68.1
	3 - 4 years	13	10.9	10.9	79.0
	5 - 10 years	25	21.0	21.0	100.0
	Total	119	100.0	100.0	

Table 1 above shows that English teachers with more than ten years of teaching experience are the most dominant respondents (N = 51) On the other hand, novice teachers who have experienced less than a year are only 5.9%.

Table 2. Education levels						
			Frequency	Per cent	Valid Percent	Cumulative Percent
Valid	Junior High School (SMP/MTs)		63	52.9	52.9	52.9
	Senior High School (SMA/SMK/MA)		56	47.1	47.1	100.0
	Total		119	100.0	100.0	

3.3. Data Collections

As mentioned above, the researcher conducted two phases in collecting the data. For the quantitative phase, the researcher distributed a questionnaire using Google Forms adopted four Likert scale responses: 4 = often, 3 = sometimes, 2 = rarely, and 1 = never. The questionnaire was arranged following eleven principles of Brookhart's educational standard (2011) to create 28 items as presented in Table 3.

Table 3. Questionnaire Items

Items	Item Numbers
I understand the steps how I am going to teach	Q1
I understand about what materials I am going to teach	Q2
The way I teach in the class is as same as what I have planned	Q3
I formulate learning objectives based on curriculum goals	Q4
I formulate learning objectives which are attainable and assessable.	Q5
I design classroom activities based on the learning objectives	Q6
I articulate learning objectives to student's beforehand classroom begins.	Q7
I understand the purposes of various assessment options that I will use to assess student' classroom performances.	Q8
I understand how I am going to use various assessment options to assess students' performances.	Q9
I am able to evaluate test items that I made.	Q10
I construct scoring schemes to assess students	Q11
I consider students' work into valuable scoring	Q12
I consider students' effort in learning into valuable scoring	Q13
I consider students' attendance into valuable scoring.	Q14
I consider students' test results into valuable scoring	Q15
I provide students time to do a follow-up of their feedback	Q16
I provide practical and valuable feedback on students' work	Q17
I use scoring results to evaluate learning activities.	Q18
I use scoring results to evaluate students' achievement.	Q19
I use scoring results to evaluate school assessment policies.	Q20
I use scoring results to evaluate regional assessment policies.	Q21
I report the scoring results and their reasons to students.	Q22
I report the scoring results and their reasons to students' parents.	Q23
I report the scoring results and their reasons to school committee.	Q24
I report the scoring results and their reasons to regional government.	Q25
I assist students to use the assessment result for further educational records.	Q26
I uphold rules in assessing students	Q27
I uphold ethics in assessing students	Q28

Since the questionnaire items were developed based on the theory without looking at previous related instruments, it was piloted to six EFL secondary schools who had the same characteristics with the targeted respondents. Then, it was evaluated using SPSS to measure the consistency of items. Table 4 below presents all items used in this study are reliable as

Shim (2009) stated that a Cronbach's alpha of 0.906 to 0,936 indicates high internal consistency of items.

Table 4. Reliability Statistics

Cronbach's Alpha	N of Items
.931	28

After that, systematically, the researcher employed a semi-structured interview. In the questionnaire initially shared, the interview consent was proposed to respondents. Among all of them, implementing random purposeful sampling (Collins et al., 2006), the researcher selected four participants from different school levels as each teaches at junior high school, Islamic junior high school, senior high school, and vocational school both a representative of private and public school and from the Ministry of Culture and Education and the Ministry Religious Affairs. The interviews were conducted in ten up to fifteen minutes consisting of eighteen questions. During the interview, the researcher recorded participant responses via Zoom meetings. The researcher used Indonesian to maximize responses (Filep, 2009).

3.4. Data Analysis

Using SPSS, the questionnaire data was calculated to measure the frequency (*Mean, Median, and Mode*) of teachers' classroom assessment practices as quantitative data. From eleven descriptions in Brookhart's standards, the researcher elaborated them into twenty-eight questions (see Table 3).

The researcher followed Lichtman's steps (2012) for analyzing the interview as qualitative data. First, the data from interviews were transcribed using google voice typing then the researcher verified its suitability. Then, from the data that were converted into texts, the researcher coded and categorized the data into several themes: a) assessment standards, b) lesson plan, c) learning objectives, d) assessment procedures, and e) feedback and follow-up. These themes are aligned with Indonesian educational assessment standards as procedures that teachers must follow (MOE, 2016) and also with practices mentioned in theoretical frameworks such as collecting, interpreting, and using the information (Mcmillan & Workman, 1998; Wang, 2017); designing instructional, grading, and reporting practices should be based on learning objectives (Cheng, Rogers, a& Hu, 2004); scoring, grading, and giving feedback (Bown & Abeywickrama, 2010); and eleven descriptions in Brookhart's standards. Data analysis and interpretation for both quantitative and qualitative, the researcher used the Indonesian educational assessment standards (MOE, 2016) and classroom assessment standards frameworks of Brookhart (2011). For triangulation data, previous studies that had been conducted were presented.

4. Results

4.1. Descriptive Statistics

English teachers at secondary schools during the pandemic often conducted most of the assessment practices in the questionnaire items coined by Brookhart (2011): Q1, Q2, Q4, Q5, Q6, Q7, Q8, Q9, Q10, Q11, Q12, Q13, Q14, Q15, Q17, Q18, Q21, Q23, Q27, and Q28 (see Table 3). They understood the materials, methods, and assessments used in the classroom because they had already prepared all of those instructions in the lesson plan. In the designing process, teachers valued what is mentioned in the curriculum. It was started from formulating

and articulating learning objectives that are attainable and assessable to students, evaluating test items, and constructing scoring schemes. Teachers accumulated students' tests and considered students' works, efforts, and attendances. Then, they used the results to grade students for evaluating learning activities and students' achievement. Also, they reported scoring results and their interpretations to students and the school committee. The data also shows that they obeyed rules and ethics in the curriculum for assessing students.

Besides the frequent practices, most teachers sometimes conduct assessment practices as mentioned in the questionnaire items: Q3, Q16, Q19, Q20, Q22, Q25, and Q26 (see Table 3). They did not precisely follow the lesson plan they have made. They also did not maximize feedback and follow-up on assessment practices. They were not familiar with using assessment results to evaluate school and regional assessment policies. Also, they did not prioritize reporting the results to students' parents and using students' assessment results for further education.

There is only one item that teachers had never done in classroom assessment that showed in item Q24 (see Table 3). Teachers did not report the assessment result and its reasons to the regional government even though Brookhart emphasized it in the framework.

4.2. Descriptive Qualitative

The researcher interviewed four selected teachers to follow up on the quantitative data. Several questions adhering to Indonesian educational assessment standards were proposed to deepen and explore items in the questionnaire. Since Brookhart's framework provides general classroom assessment practices, the Indonesian curriculum assessment standard was taken into account for the specification. As a result, based on those two frameworks, there are several highlighted themes to analyze the data: a) assessment standards, b) lesson plan, c) learning objectives, d) assessment procedures, and e) feedback and follow-up. The researcher arranged the explanations based on those themes in the data display below. Meanwhile, the participant codes describe as follow: participant 1 refers to an English teacher in a vocational school, and participant 2 is a junior high English teacher. Participant 3 comes from an Islamic senior high school, and participant 4 is an English teacher of an Islamic junior high school.

4.2.1. Assessment Standard

The participants designed lesson plans, syllabus, and academic calendar and formulated learning objectives following core and basic competencies. It includes assessing cognitive, psychomotor, and affective aspects based on the Indonesian educational assessment standards. It was stated by all of the participants in the interview. However, they admitted that adjustments in the curriculum assessment standard always happened due to school policy and consensus of MGMP (An organization of particular subject across the region). Moreover, one of the participants claimed that implementing what has been prepared on the lesson plan during the pandemic was challenging.

4.2.2. Lesson Plan

Specific to designing lesson plans, core and basic competencies are the navigators to arrange instructional teaching, including in assessing students. Most participants prepared the lesson plans and the other important documents such as yearly and semester programs and the academic calendar beforehand the teaching process. Although the participants completed

those administrations, in practice, sometimes they missed and changed them due to classroom situations during the pandemic.

4.2.3. Learning Objectives

Then, in more detailed descriptions, participants informed how they made learning objectives assessable in all domains: cognitive, psychomotor, and affective, by referring to basic competencies. Changes and adjustments also happened because of school policy, students' incapability, and classroom situation during the pandemic.

4.2.4. Assessment Procedures

Most participants employed similar assessment procedures from designing lesson plans to providing feedback after scoring since they followed the curriculum assessment standard as the primary resources. They used observations to assess students' affective. Besides, they conducted unit tests, middle and final exams, daily exercises to evaluate students' cognitive and skill performance. The difference occurs in a way that participants value the affective domain. Some of them emphasized the affective domain to evaluate students' achievement during the pandemic because they assumed that students' knowledge mastery was difficult to achieve.

Meanwhile, the rest of the participants emphasized students' middle and final exams because, based on the school regulation, students' affective score was integrated with all subjects, and it was challenging to record students' classroom activities during the pandemic. Teachers used rubrics, valuing scoring criteria in each skill 0-100. Eventually, teachers required students to get a minimum mastery criterion for the final grade. This grade is a collective score from cognitive, psychomotor, and affective domains.

4.2.5. Feedback and Follow-Up

In giving feedback and follow-up students' grades, approximately right away or in a week, the participants gave remedial to boost students' scores.

5. Discussion

Brookhart's educational assessment and assessment standards of the Indonesian 2013 revised curriculum were used as research frameworks in this study. As Rea-Dickins stated (2001, in Cheng & Wang, 2007) that a working framework is needed to analyze classroom assessment to comprehend how the teachers' practices conformed to the standards.

The questionnaire findings show that most EFL teachers at all secondary school levels performed classroom assessment standards following the educational assessment standards coined by Brookhart (2011), which is also aligned with the national assessment standards in the Indonesian 2013 revised curriculum during the pandemic. Teachers understood the instructional learning related to what materials, methods, and assessments they used. Teachers relied on learning objectives to design instructional, grading, and reporting practices, as suggested by Cheng, Rogers, and Hu (2004). Further, data from the questionnaire implies that the ways teachers formulated learning objectives were also aligned with the curriculum. Teachers were familiar with various instruments in assessing students, aspects they should consider, assessment criteria for plotting student scores, and reporting the

assessment results. The findings present the variety and complexity of classroom assessment practiced by teachers because classroom assessment involves various components, variables, and practices (Wang, 2017) and cross summative, formal to informal (Brown & Abeywickrama, 2010).

Following Brookhart's educational assessment, teachers did not frequently perform several practices even though the practices are needed in assessment practices. For example, teachers did not frequently report the assessment results to students' parents, give feedback and follow-up on student works, and assist students to use the result for further education. In terms of assessment evaluation, teachers did not report and involve in school and regional regulations because the curriculum assessment standards only obligate them to use the assessment results for evaluating: the achievement of students' competence, learning process, and student learning reports. Hence, reporting the assessment result should be an essential practice contributing to regulations from schools to regional policymakers since Brookhart (2011) describes reporting assessment results must be sequentially from teachers to regional policymakers. Referring to Brown and Abeywickrama (2010), this situation causes ineffective washback in which students did not find the assessments as learning experiences.

Based on major themes in qualitative data, the researcher concludes that English teachers did not maximize classroom assessment practices during the pandemic. According to four interviewed teachers, the classroom activities and situations were different and unpredicted during online or blended learning. They admitted that the implementation was different even though they followed curriculum assessment standards for designing lesson plans, learning objectives, assessment procedures, criteria, schema, instruments, and feedback and follow-up. Changes and adjustments cannot be neglected due to unpredicted classroom situations during the pandemic. This unpredicted situation leads to Emergency Remote Teaching explained by Hodges et al. (2020) in which English teachers did not have well preparation and voluntary planning in online learning during the pandemic. As a consequence, technology-based education that offers flexibility (Yin & Shi, 2021), efficiency, and convenience (Dumford & Miller, 2018) utilized in online learning during the pandemic could not be well-achieved.

The findings of this study support the data reported by the Indonesian Teachers Association (IGI) that the implementation of distance or online learning in the last three months of 2020 showed only 60% because of teachers' inferior skills in utilizing technology (see Dewi & Wajdi, 2021). It is also aligned with several studies' findings that EFL teachers had a lack of readiness and inability to teach online because they still struggled using digital platforms. As a result, they transformed traditional learning activities into online learning without maximizing the technology (Atmojo & Nugroho, 2020; Ilmiani et al., 2021; Kusumastuti, 2020; Mardiah, 2020; Nugroho & Haghegh, 2021; Situmorang et al., 2020; Sumardi & Nugrahani, 2021; Supriadi et al., 2020; Syarifah & Zainil, 2021; Yundayani et al., 2021). Further, assessment and evaluation in online learning have unique aspects (Serwatka, 2002, in Dumford & Miller, 2018) because specific techniques in a traditional classroom do not always work in online learning. Therefore, adjustments of assessment standards done by EFL teachers when assessing their students in this study are reasonable.

Scoring and grading practices done by English teachers are relevant with what was mentioned by Douglas (2011: 55-56) since they discussed how to score with other English teachers in MGMP (An organization of a particular subject across the region) to get consensus. Also, they upheld the curriculum assessment standard and adapted them with school regulations, as explained by Brown and Abeywickrama (2010). For instance, the

English teacher from the vocational school disclosed that she did not assess students' affective because that was responsible for the room teacher corresponding with all subject teachers. In addition, it seems that not all English teachers were aware of the emergency remote teaching with the simplification of instructional teaching. Explicitly, only the English teacher from vocational school made a lesson plan on a sheet of paper. It contradicts the findings of Hapsari (2020) shows that English teachers had a good perception of ERT.

Furthermore, there is a gap between teachers' knowledge and their practices. As admitted by the interviewed teachers, although they understood assessment standards and school regulations, they were still difficult to bear with the pandemic situation. Most participants argued that teaching in the pandemic through online or blended learning was more challenging than face-to-face learning because students lacked motivation, facilities, and understanding. Evaluating students' cognitive and psychomotor was challenging because online learning demotivates students' learning and decreases their knowledge mastery. Students' lack of understanding impacted their scores in any exercise, task, or test. Students also did not participate in the class actively. It became worst for junior high schools students who chose not to attend the class when it was online. It impacts their knowledge and skill mastery that were not well-accomplished. Therefore, English teachers in junior high schools emphasized students' affective. It is suitable with what is reported by Atmojo and Lase et al. (2021), Holisoh and Fitriani (2020), Nugroho (2020), and Nugroho et al. (2021) that English teachers in ERT struggled with different levels of mastery, low awareness on online learning, lack of students' motivation and engagement besides availability of internet connection and quotas. Hence, the interviewed teachers always compared online learning to face-to-face as students' engagement was less.

A dilemmatic perspective has risen in a way English teachers modified and adjusted classroom assessment standards during the pandemic as a priority of ERT and also suggested by the Ministry of Education and Culture that teachers must value simplicity, flexibility, and empathy since student security during the pandemic (Cahyadi et al., 2021). However, to make students pass, the teachers gave them minimum mastery criteria (KKM) by valuing their attendance and attitudes while learning. The Indonesian revised curriculum itself highlights students' attitudes as learning outcomes. As a result, teachers or school committees tolerated the assessment standards by modifying and adjusting the practice standards focusing on students' attitudes. These practices lead to grading inflation (Arrafii, 2020; Zulaiha et al., 2020). Grading inflation occurs when teachers consider academic and non-academic factors and internal or external factors (Arsyad Arrafii, 2020; Brookhart et al., 2016; Chen & Bonner, 2017; Cheng & Sun, 2015; Cox, 2011; Guskey & Link, 2019; Isnawati & Saukah, 2017; Karagül et al., 2017; Pulfrey, 2013; Randall & Engelhard, 2010; Widiastuti, 2018; Yesbeck, 2011; Zoeckler, 2007; Zulaiha, 2017).

6. Conclusion

Classroom assessment standards contain several practices that teachers should follow. Notably, the Indonesian curriculum regulates the practices of EFL teachers in the national assessment standards and generally uses Brookhart's educational assessment standards. Indonesian EFL teachers did not frequently report the assessment results to students' parents, give feedback and follow-up on student works, and assist students to use the result for further education. In terms of assessment evaluation, teachers could not report and be involved in school and regional policymakers. However, due to the pandemic, they did not maximize classroom assessment standards as online learning assessment differs from traditional ones.

Thus, adjustments and modifications in assessment practices happened to cope with the pandemic situation. This situation causes ineffective washback and grade inflation.

Recommendation

The finding of this study gives theoretical and practical significance for implementing and evaluating classroom assessment practices amid the pandemic. Further, since this study took place in a particular situation, it signifies some recommendations for the future studies such as presenting specific and comprehensive theoretical frameworks about online assessment and grading and investigating the differences of online and face-to-face assessment and grading whether during the pandemic or post pandemic situation.

Acknowledgements

The author disclosed receipt of the following financial support for the publication of this article: This work was supported by the Indonesia Endowment Fund for Education under the Ministry of Finance of the Republic of Indonesia.

References

- Andrade, H., & Cizek, G. J. (Eds.). (2010). *Handbook of formative assessment*. Routledge.
- Arrafii, M. A. (2020). Grades and grade inflation: exploring teachers' grading practices in Indonesian EFL secondary school classrooms. *Pedagogy, Culture and Society*, 28(3), 477–499. <https://doi.org/10.1080/14681366.2019.1663246>
- Atmojo, A. E. P., & Nugroho, A. (2020). EFL Classes Must Go Online! Teaching Activities and Challenges during COVID-19 Pandemic in Indonesia. *Register Journal*, 13(1), 49–76. <https://doi.org/10.18326/rjt.v13i1.49-76>
- Brookhart, S. M. (1993). *Teachers' Grading Practices: Meaning and Values*. 30(2), 123–142.
- Brookhart, S. M. (1997). A Theoretical Framework for the Role of Classroom Assessment in Motivating Student Effort and Achievement. *Applied Measurement in Education ISSN*, 10(2), 161–180. https://doi.org/https://doi.org/10.1207/s15324818ame1002_4
- Brookhart, S. M. (2011). Educational Assessment Knowledge and Skills for Teachers. *Educational Measurement: Issues and Practice*, 30(1), 3–12. <https://doi.org/https://doi.org/10.1111/j.1745-3992.2010.00195.x>
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016a). A Century of Grading Research: Meaning and Value in the Most Common Educational Measure. *Review of Educational Research*, 86(4), 803–848. <https://doi.org/10.3102/0034654316672069>
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016b). A Century of Grading Research: Meaning and Value in the Most Common Educational Measure. *Review of Educational Research*, 86(4), 803–848. <https://doi.org/10.3102/0034654316672069>
- Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: principles and classroom practices*. NY: Pearson Education.
- Cahyadi, A., Sri, H., & Suryani, W. (2021). COVID-19, emergency remote teaching evaluation: the case of Indonesia. *Education and Information Technologies*, 0123456789, 1–15. <https://doi.org/10.1007/s10639-021-10680-3>
- Can, I., & Silman, L. (2021). Insights into emergency remote teaching in EFL. *ELT Journal*, 76(1), 34–43. <https://doi.org/https://doi.org/10.1093/elt/ccab073>
- Chen, P. P., & Bonner, S. M. (2017). Teachers' Beliefs About Grading Practices and a Constructivist Approach to Teaching. *Educational Assessment*, 22(1), 18–34. <https://doi.org/10.1080/10627197.2016.1271703>
- Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods, and procedures. *Language Testing*, 21(3), 360–389. <https://doi.org/10.1191/0265532204lt288oa>

- Cheng, L., & Sun, Y. (2015). Teachers grading decision making: Multiple influencing factors and methods. *Language Assessment Quarterly*, 12(2), 213–233. <https://doi.org/10.1080/15434303.2015.1010726>
- Cheng, L., & Wang, X. (2007). Grading , feedback , and reporting in ESL / EFL classrooms. *Language Assessment Quarterly*, 4(1), 85–107.
- Collins, K. M. T., Onwuegbuzie, A. J., & Jiao, Q. G. (2006). Prevalence of Mixed-methods Sampling Designs in Social Science Research. *Evaluation & Research in Education*, 83–101. <https://doi.org/10.2167/eri421.0>
- Cox, K. B. (2011). Putting Classroom Grading on the Table : A Reform in Progress. *American Secondary Education*, 40(1), 67–87.
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative* (P. A. Smith (ed.); Fourth Edi). Pearson.
- Creswell, J. W., & Guetterman, T. C. (2019). *Educational research: planning, conducting, and evaluating quantitative and qualitative research* (Sixth edit). Pearson.
- Defianty, M., Hidayat, D. N., Kultsum, U., & Sufyan, A. (2020). Reframing Formative Assessment Practices in Emergency Remote Teaching Context of English Language Teaching in Higher Education institutions Indonesia. *ICIIIS*. <https://doi.org/10.4108/eai.20-10-2020.2305143>
- Dewi, M. P., & Wajdi, M. B. N. (2021). Distance Learning Policy During Pandemic Covid-19. *Journal of Education and Technology*, 4(3), 325–333.
- Douglas, D. (2014). *Understanding language testing* (B. Comrie & G. Corbett (eds.)). Routledge.
- Dumford, A. D., & Miller, A. L. (2018). Online learning in higher education: exploring advantages and disadvantages for engagement. *Journal of Computing in Higher Education*, 30(3), 452–465. <https://doi.org/10.1007/s12528-018-9179-z>
- Ferri, F., Grifoni, P., & Guzzo, T. (2020). Online Learning and Emergency Remote Teaching: Opportunities and Challenges in Emergency Situations. *Societies*, 10(86), 1–18. <https://doi.org/10.3390/soc10040086>
- Filep, B. (2009). Interview and translation strategies: coping with multilingual settings and data. *Social Geography*, 4(1), 59–70. <https://doi.org/www.soc-geogr.net/4/59/2009/>
- Grantor, D. (2021). *Teachers ' Perspectives on Emergency Remote English Language Teaching during the COVID-19 Pandemic*.
- Guskey, T. R., & Link, L. J. (2019). Exploring the factors teachers consider in determining students' grades. *Assessment in Education: Principles, Policy and Practice*, 26(3), 303–320. <https://doi.org/10.1080/0969594X.2018.1555515>

- Hampel, R. (2006). Rethinking task design for the digital age: A framework for language teaching and learning in a synchronous online environment. *ReCALL*, 18(1), 105–121. <https://doi.org/10.1017/S0958344006000711>
- Hapsari, F. I. (2020). The Implementation of Curriculum of English Teaching in Pandemic Era. *Journal of Applied Linguistics Indonesia*, 04(01), 32–37.
- Hasanah, E., Pendidikan, M., Pendidikan, F. I., & Dahlan, U. A. (2021). *Penerapan Kurikulum Darurat Pada Masa Covid-19 Di SMK Muhammadiyah Semin , Yogyakarta*. 7(2), 9–18.
- Hazaea, A., Bin-hady, W. R., & Toujani, M. M. (2021). Emergency Remote English Language Teaching in the Arab League Countries: Challenges and Remedies. *Computer-Assisted Language Learning Electronic Journal*, 22(1), 201–222.
- Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The Difference Between Emergency Remote Teaching and Online Learning. *Educause*, 1–12. <https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning>
- Holisoh, N., & Fitriani, H. (2020). English Learning Strategies of Using Application in Online Class: An Emergency Remote Teaching during Covid-19 Pandemic. *ELT in Focus*, 3(2), 42–45. <https://doi.org/10.35706/eltinf.v3i2.4626>
- Ilmiani, D., Muamaroh, & Nugroho, A. (2021). English Online Learning in the Age of Global Pandemic: How Do Teachers and Students Feel? Dwi. *Lingua*, 18(1), 62–81. <https://doi.org/10.30957/lingua.v18i1.682>
- Isnawati, I., & Saukah, A. (2017). Teachers' grading decision making. *TEFLIN Journal*, 28(2), 155–169. <https://doi.org/http://dx.doi.org/10.15639/teflinjournal.v28i2/155-169>
- Karagül, B. İ., Yüksel, D., & Altay, M. (2017). Assessment and Grading Practices of EFL Teachers in Turkey. *International Journal of Language Academy*, 5(September), 168–174.
- Kearns, L. (2012). Student Assessment in Online Learning: Challenges and Effective Practices. *MERLOT Journal of Online Learning and Teaching*, 8(3), 198–208. http://jolt.merlot.org/vol8no3/kearns_0912.htm
- Kusumastuti, W. P. (2020). *Emergency Remote Language Teaching During Pandemic*.
- Lase, D., Zaluchu, S. E., Daeli, D. O., & Ndraha, A. (2021). Parents' Perceptions of Distance Learning during Covid-19 Pandemic in Rural Indonesia. *Journal of Education and Learning*.
- Mardiah, H. (2020). The Use of E-Learning to Teach English in the Time of the Covid-19 Pandemic. *English Teaching and Linguistics Journal (ETLiJ)*, 1(2), 49–55. <https://doi.org/10.30596/etlij.v1i2.4894>

- McMillan, J. H., & Nash, S. (2000). Teacher classroom assessment and grading practices decision making. *Paper Presented at the Annual Meeting of the National Council on Measurement in Education*, 39.
- McMillan, J. H., & Workman, D. J. (1998). Classroom Assessment and Grading Practices: A Review of the Literature. *Metropolitan Educational Research Consortium*.
- MOE. 2016. Peraturan Kementrian Pendidikan dan Kebudayaan Republic Indonesia Nomor 23 Tahun 2016 tentang Standar Penilaian Pendidikan Indonesia (The Ministry of Education's Regulation No. 23, on Assessment Standards).
- Naqvi, S., & Zehra, I. (2020). Online EFL Emergency Remote Teaching during COVID 19, Challenges and Innovative Practices: A Case of Oman. *Arab World English Journal (AWEJ)*, 17–35. <https://doi.org/https://dx.doi.org/10.24093/awej/MEC2.2>
- Nugroho, A., & Haghegh, M. (2021). Emergency Remote Teaching amidst Global Pandemic: Voices of Indonesian EFL Teachers. *VELES Voices of English Language Education Society*, 5(1), 66–80. <https://doi.org/10.29408/veles.v5i1.3258>
- Pulfrey, C. (2013). *Autonomy and Task Performance : Explaining the Impact of Grades on Intrinsic Motivation*. 105(1), 39–57. <https://doi.org/10.1037/a0029376>
- Randall, J., & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education*, 26(7), 1372–1380. <https://doi.org/10.1016/j.tate.2010.03.008>
- Shim, 2009 – Shim, K.N. (2009). An Investigation into Teachers' Perceptions of Classroom-Based Assessment of English as a Foreign Language in Korean Primary Education. Doctoral Thesis. University of Exeter.
- Situmorang, K., Nugroho, D. Y., & Pramusita, S. M. (2020). English Teachers' Preparedness in Technology Enhanced Language Learning During Covid-19 Pandemic – Students' Voice. *Jo-ELT (Journal of English Language Teaching) Fakultas Pendidikan Bahasa & Seni Prodi Pendidikan Bahasa Inggris IKIP*, 7(2), 57. <https://doi.org/10.33394/jo-elt.v7i2.2973>
- Sumardi, S., & Nugrahani, D. (2021). Adaptation To Emergency Remote Teaching: Pedagogical Strategy For Pre-Service Language Teachers Amid COVID-19 Pandemic. *Turkish Online Journal of Distance Education*, 22(2), 81–93. <https://doi.org/10.17718/tojde.906553>
- Supriadi, Y., Nisa, A. A., & Wulandari, S. (2020). English Teachers ' Beliefs on Technology Enhanced Language Learning : A Rush Paradigmatic Shift during Covid-19 Pandemic. *Pancaran Pendidikan*, 9(2), 47–58. <https://doi.org/10.25037/pancaran.v9i2.290>
- Syarifah, I., & Zainil, Y. (2021). Teaching and Learning Challenges in EFL Classrooms During the Pandemic Era. *Proceedings of the Ninth International Conference on Language and Arts (ICLA 2020)*, 539(Icla 2020), 111–115. <https://doi.org/10.2991/assehr.k.210325.020>

- Wang, X. (2017). A Chinese EFL Teacher's Classroom Assessment Practices. *Language Assessment Quarterly*, 14(4), 312–327. <https://doi.org/10.1080/15434303.2017.1393819>
- Widiastuti, I. A. M. S. (2018). Teachers' classroom assessment and grading practices. *SHS Web of Conferences*, 42(4), 00052. <https://doi.org/10.1051/shsconf/20184200052>
- Widodo, H. P. (2016). Language Policy in Practice: Reframing the English Language Curriculum in the Indonesian Secondary Education Sector. *English Language Education Policy in Asia*, 11, 127–151. https://doi.org/10.1007/978-3-319-22464-0_6
- Yesbeck, D. M. (2011). Grading practices: Teachers' considerations of academic and non-academic factors. In *VCU Scholars Compass Theses*.
- Yin, H., & Shi, L. (2021). Which Type of Interpersonal Interaction Better Facilitates College Student Learning and Development in China : Face-to- Face or Online ? *Sage Publications*, 1–28. <https://doi.org/10.1177/20965311211010818>
- Yundayani, A., Abdullah, F., Tandiana, S. T., & Sutrisno, B. (2021). Students' cognitive engagement during emergency remote teaching : Evidence from the Indonesian EFL milieu. *Journal of Language and Linguistic Studies*, 17(December 2019), 17–33. <https://doi.org/10.52462/jlls.2>
- Zoeckler, L. G. (2007). Moral aspects of grading: A study of high school English teachers' perceptions. *American Secondary Education*, 35(2), 83–102.
- Zulaiha, S. (2017). Teachers' Grading Practices: In Search for Clear Grading Criteria. *Journal of ELT Research*, 2(1), 80–93. https://doi.org/10.22236/JER_Vol2Issue1pp15-23
- Zulaiha, S., Mulyono, H., & Ambarsari, L. (2020). An investigation into EFL teachers' assessment literacy: Indonesian teachers' perceptions and classroom practice. *European Journal of Contemporary Education*, 9(1), 189–201. <https://doi.org/10.13187/ejced.2020.1.189>