Conference Paper

The Practice of Classroom Assessments during Curriculum Change: A Study of Inservice Teachers

Wiyaka*

English Education Department, Universitas PGRI Semarang, Indonesia

ORCID ID
Wiyaka: https://orcid.org/0000-0001-8073-2831

Abstract.

The education ministry has refreshed the past school curriculum to the Freedom Curriculum or *Kurikulum Merdeka*. Ideally, the change of curriculum is alignment with the methods of assessment. In terms of assessments, there are three types highlighted in the new curriculum: diagnostic, formative, and summative assessments. The question arises whether the teachers can really cope with the current issues in assessment. This research aimed to survey the teachers' perceived skills in assessment and to examine the practices of classroom assessments dealing with the latest issues, including the assessment they used most. A questionnaire was used to collect the data. The respondents were 276 teachers from various backgrounds. They were in-service teachers who joined the Teacher Certification Program managed by Universitas PGRI Semarang in the year 2022. Descriptive statistics was used to calculate the data. The results revealed that teachers reported being fairly skilled in doing assessments. However, several subskills in assessment need improvement. Among the three types of assessment, summative was mostly practised by teachers while teachers least frequently conducted the diagnostic. Based on the function, assessment of learning was mostly practised by teachers, while assessment as learning was the least frequently used. The in-service teacher training program curriculum must cover course materials in assessment to better equip teachers with the ability to execute various assessments.

Keywords: *Kurikulum Merdeka*; classroom assessment; diagnostic; formative; summative

Corresponding Author: Wiyaka; email: wiyaka@upgris.ac.id

1. Introduction

Indonesia has experienced several curriculum changes since its independence seventy years ago. A ten-year periodical change happened in the history of education in the country. The latest was when in 2020 the government released Freedom Curriculum or *Kurikulum Merdeka* as part of the Freedom of Learning policy which address the learning loss due to the outbreak of COVID-19 pandemic. While there are some constraints on the part of the teacher readiness, up to now schools are allowed to determine the curriculum to use, whether it is the 2013 Curriculum, the Emergency Curriculum, or the Merdeka

Curriculum. Merdeka Curriculum seems to be vastly introduced and adopted by lots of schools. Currently, the curriculum had been adopted by 140 thousand educational units all over the country (1).

Curriculum change usually takes a long process, and based on comprehensive evaluation of the previous document. It should not be based on a partial judgment from the authority. Ccurriculum design makes the connection between the research and theory of learning and the practice of designing lessons and courses (2). This confirms what (3) proposed: at least four fundamental things to be considered in the curriculum development: the purpose of education to be achieved, learning experience to achieve the goals, learning organizing experiences, and evaluation.

The present Indonesian curriculum seems to follow Top-Down Approach or Managerial Approach of curriculum design in that the initiatives comes from education officials and the administrators or the holders of the education policy such the ministry of education (4). The ministry of education has set the philosophy and the general goals of education. The curriculum was made to be more flexible, focusing on essential materials, and applications that provide various references for teachers to continue to develop teaching practices independently while also sharing good practices (5). Schools are allowed to make changes, enabling them to modify and choose simple materials.

Curriculum is not an exclusive domain the education system. There are two other factors that interconnectedly relate to curriculum: instruction and assessment (6). The relevance of the three cannot be over-emphasized because curriculum specifies the subjects and topics to be taught whereas, instructional activities determine the objectives of instruction in terms of skills, tasks and competencies expected of the learner as well as the methods, materials and strategies designed to accomplish them. Assessment measures the extent to which the curriculum and instructional objectives are accomplished (7).

A new paradigm of assessment which characterizes the Merdeka Kurikulum is the assessment approach. Despite the fact that there will be no more national examination as a requirement for student graduation, there are three issues highlighted in the classroom: diagnostic assessment, formative assessment, and summative assessment (8). Of the three, the first might be new to school teachers. The other two are not new issues in education, but it is still relevant to seek how teachers really understand and perceive about the assessments, and, this is most important, how they practice doing the assessments.

Theoretically, classroom assessment is commonly classified into three types: Assessment of Learning, Assessment as Learning and Assessment for Learning (9). Assessment

of learning (AoL) is recognized as summative assessment. Summative assessment measures students' achievement on certain learning standards with the use of evidences. It aims to measure, or summarize, what a student has grasped, and typically occurs at the end of a course or unit of instruction (10-11). It tends to be high-stakes testing as the results of the test is used for judgment, for example to put the students into ranks and to determine the success of a program. It is high stakes because the information collected from these assessments are used to make decisions such as selection and placement of students into higher educational levels.

Assessment for learning (AfL) is usually formative in nature. It occurs when teachers give constructive feedback to students based on inferences of student progress through observations, anecdotes, question-answer sessions, simple tests and so on. This motivates and encourages students to improve on their learning. Meanwhile, assessment as learning (AaL) encourages the students to reflect on their learning progress and monitor their progress and learning performance to achieve higher standards (12). It occurs when learners reflect on and regulate and monitor their learning progress. It comprises learner reflection and peer and self-assessment.

Apart from the above concept, the Indonesian Ministry of Education has stipulated the regulation No 21 Year 2021 on Standard of Evaluation for Kindergarten, Elementary, Junior, and Senior Schools levels, there are two forms of assessments that can be used by teachers, they are formative and summative assessment. Diagnostic assessment is embedded in the formative assessment as essentially, diagnostic purpose is basically formative.

There have been several studies on the classroom assessment practices. (13) did a research on the implementation od classroom practices. They reported that Malaysian teachers are generally unaware of the importance of feedback and effective questioning techniques to enhance learner's autonomy. Teachers lacked pedagogical skills such as giving appropriate feedback and questioning techniques to enhance students' learning. Similarly, (14) conducted a research on Classroom Assessment in Higher Education of Bangladesh. The findings indicate an overemphasis on 'assessment as measurement' and 'assessment as procedure' paradigms. Features of second generation of assessment are also identified in the current practice, type of learning is promoted through assessment in the higher education scene of Bangladesh. Teachers are trapped to use traditional classroom assessments. This finding is a stereotype to the research finding by (15) in which teachers had adopted an approach of learning assessment to make the assessment. Teachers used largely traditional assessment methods in classroom assessment practices.The above-mentioned research have not counted yet diagnostic

assessment as an approach to assessment. That's why, this research is expected to fill the gap in the research on assessment while expecting that the finding will contribute to the thorough description of classroom assessment.

This research was aimed to describe the teacher perceived skills and ability in classroom assessment, and to examine their practice of classroom assessment. Identifying the skills and practices will be valuable input for education office and institutions in efforts to enhance teachers' competence in dealing with assessment.

2. Method

This research applied a survey design to gather descriptive and comparative data in terms of the characteristics of several groups of subjects (16). Participants of this research were chosen through purposive and convenience sampling. They were teachers who were joining the in-service professional training program called Pendidikan Profesi Guru - Batch 1 dan Batch 2 - which were managed by Universitas PGRI Semarang Indonesia. They were joining online courses to complete the program. Their home basis are in various schools across different regions or provinces in the country such as Central Kalimantan, Nusa Tenggara, Central Java, West Java, and East Java. On the deadline of response submission, 376 respondents filled the online form.

An online questionnaire was used to collect the data. The questionnaire was adopted from Assessment Practices Inventory (17). The questionnaire contains 54 closed-ended items which are divided into three main sections. The first section asks teachers to provide background or demographic information, their educational background, length of teaching experience, teaching level, subjects they teach, and the expertise to join assessment training (5 items). The second section consists of items related to their perceived skills in conducting classroom assessments (25 items). The third section asks for information concerning teachers' practices in classroom assessments (24 items). A 5-point Likert Scale ranged from "Strongly Disagree" to "Strongly Agree" options.

Teachers' levels of skills with factors and the items loaded on each factor were determined using descriptive statistics in terms of means and standard deviations. Descriptive statistics in percentage were used for the types of classroom assessment teachers used most.

3. Result and Discussion

3.1. The Profile of Respondents

The study participants represented various teachers' backgrounds based on education, years of teaching experience, subject taught, school level, and training experience. The description of the samples can be seen in Table 1.

TABLE 1: Frequency Table of Demographic Variables.

Variable	Frequency	Percent
School level		
Kindergarten	28	7.4
Elementary School	124	33
Junior High School	106	28.1
Senior High School	118	31.5
Subjects		
English	78	20.8
Physics	22	5.9
Indonesia	72	19.1
Biology	33	8.8
Guidance and Counseling	47	12.5
Math	35	9.3
All subjects (for Elementary)	89	23.6
Work time		
<5 years	31	8.2
5 - 10 years	147	39
10 - 15 years	130	34.5
>15 years	69	18.3
Training Experience		
Never	63	16.7
Once	256	68.1
Often	54	14.4
Always	3	0.8

From the table it can be identified that majority of the participants are teachers at elementary schools (33%) as compared those at senior high (31.5%), junior high (28%), and kindergarten (7.4%). In terms of education background, most of them hold undergraduate degree (78%), master's degree (17%), and doctorate degree (5%). Their teaching experience ranges from less than 5 years to more than 15 years. Most of them have been teaching at the range of 5-10 years (34.5%), followed by teachers with working experience of 10-15 years (34.5%), more than 15 years (18.3%), and less than 5

years (8.2%). The surprising fact is that a great number of teachers got the opportunity to attend the program after working more than 15 years, while the program requires only five years minimum experience. The subjects they taught include all subjects (23.6%), English (20.8%), Bahasa Indonesia (19%), Guidance and Counselling (12.5%), Math (9.3%), and Physics (5.9%). They reported that teachers teaching all subjects are likely teachers in Kindergarten and elementary schools. In the Indonesian education system, elementary school teachers teach all basic subjects, including math, science, language and the other.

Since this study focuses on the assessment practices during the circular change, Figure 1 represents whether the participant's school has implemented *Kurikulum Merdeka* or not.
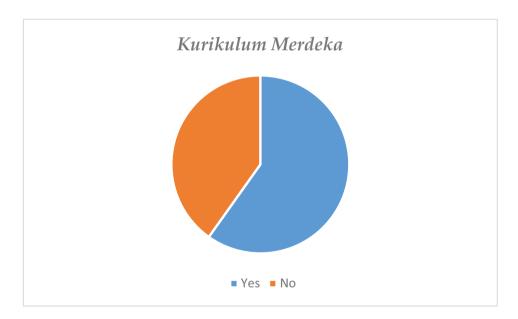

Figure 1: The implementation of *Kurikulum Merdeka* by in-service teachers.

Figure 1 shows that most in-service teachers have implemented *Kurikulum Merdeka* in their school (59.8%). Meanwhile some of them have not implemented *Kurikulum Merdeka* in their school (40.2%)

3.2. Teachers skills

Adopted from Setlhomo Koloi-Keaikitse (2012), teacher skills in classroom assessment are subdivided into six factors; Criterion-Referenced Testing, Grading Practices, Statistical Applications, Assessment Applications, Essay items, and Objective items). Descriptive statistics was calculated to assess teachers' levels of agreement with the factors

and items regarding their perceived skill in assessment practices. The result is shown in Table 2.

TABLE 2: Descriptive Statistics for Teachers Perceived Skills.

Factors/items	M		SD	
Criterion Reference Testing 7. Analysis of assessment results 8. Assessment result report 17. Align items to objectives 19. Test covers what taught 23. Giving fair value to students 24. Provide feedback in writing along with grades	3.10 3.11 3.49 3.43	3.02 3.46 2.12	0.93 0,87 0.93	0.93 0.87 0.99
Grading practices 13. Using portofolio assess 16. Using peer assessment 18. Develop a systematic assess procedure Statistics application 5. Create an assessment rubric 9. Calculates central tendency for tests 10. Conduct items analysis 11. Revise test based on item analysis 22. Use table of specification 25. Calculate variability	3.11 3 3.1 3.07 3 2.92 2.76	3.24 2.93 2.91 2.95	0.89 0.87 1.02 0.94 0.96	0.87 0.88 0.95 0.93
Assessment Application 6. Higher cognitive items 12. Assess problem solving skills 14. Use assessment results for decision making 15. Determine why student make mistakes 16. Use results to plan teaching 21. Assess answers honestly and consistent	3.17 2.99 323 3.48	2.84 3.27 3.24	0.93 0.87 0.87	0.90 0.87 0.82
Essay Items 3. Fair grading of essay items 4. Essay items	3.38 3.45	3.32	0.82	0.82
Objective Items 1.Multiple choice items 2. Write a statement items (STEM)	3.32 3.15	3.49	0.81	0.85

Table 2 tells that teachers reported to be mostly skilled in essay item practices. They reported to be skilful in constructing multiple choice test, but admitted that they are not trained in writing stems in multiple choice test. In criterion-referenced testing factor, teachers reported to be skilled in ensuring that the test adequately covers materials taught, and in assessing specific course objectives. In grading practices, they reported being less skilled in using portfolio assessment. Teachers reported to have a fair amount of skill in constructing essay items and applying assessment results. Teachers reported having low skill the statistic application, particularly in using table of specification, and calculate variability of the test items.

3.3. Teachers practice in different types of assessment

The practice of classroom assessment in this research was meant to examine how teachers practiced the types of assessments under six categories: Assessment of Learning, Assessment for learning, Assessment as learning, Diagnostic Assessment,

Formative Assessment, and Summative Assessments. Descriptive statistics was calculated to determine types of classroom assessment that teachers' use most (see Table 3).

TABLE 3: Descriptive Statistics of Teachers' Assessment Practices.

Factors/Items	M		SD	
Assessment as Learning 9. Understand the concept of assessment as learning 10. Assess is one of the learning techniques 11. Practice assess as a classroom leason 12. Using the result of assess as lesson for learning	3.61 3.78 3.92	3.92 2.83	0.54 0.71	0.73 0.53
Assessment of Learning 1. Understand the concept of assessment of learning 2. The purpose of assessment is measure the students' learning outcomes 3. Practice the assessment of learning outcome 4. Using the result to see students' achievement	3.65 3.85 3.06	3.80 3.91	0.61 0.74	0.61 0.65
Assessment for Learning 5. Understand the concept of assessment for learning 6. The purpose of assessment is to improve the quality of learning 7. Practice assessment of learning outcome 8. Using the results of the assessment to improve the quality of learning	3.24 2.81 4.13	3.02 3.02	0.67 0.83	0.57 0.69
Diagnostic 13. Teachers know the concept of diagnostic assessment 14. Diagnostic assessment to map the students potential 15. Teachers practice diagnostic assessment 16. Using the results of diagnostic test to planning for learning	3.41 3.95 3.70	3.66 2.36	0.67 0.83	0.57 0.69
Formative Assessment 17. Understand the concept of formative assessment 18. The purpose of assessment is to improve the quality 19. Teachers practice formative assessment 20. The result of formative assessment to improve learning 22. Formative assessment is used to measure the learning	3.68 3.99 3.95	3.84 2.81 3.86	0.55 0.74 0.63	0.49 0.53
Summative Assessment 21. Teachers understand the concept of summative assessment 23. Teachers practice summative assessment 24. Teachers using the results of summative assessment to see the students' achievement	3.51 2.85	3.83 3.85	0.59 0.62	0.75

Viewed from the functions, teachers reported using assessment practices most frequently was assessment of learning. Surprisingly, they claimed to practice more assessment as learning (AaL) rather than assessment for learning (AfL). They reported they understand the concept of assessment as learning, but at the same time they rarely practiced it in the classroom. In the assessment for learning factor, they claimed to have used the results of the assessment to improve the quality of learning. However, in the same factor they reported to doubt that the purpose of assessment is to improve the quality of learning.

Dealing with the purposes of assessment, teachers admitted that summative assessment was the assessment they used mostly. The formative assessment was more frequently practiced than the diagnostic assessment. Though, they reported having fairly good understanding in diagnostic assessment. They also believed that the results of diagnostic test is used to plan for future learning. It seems that the teachers' understanding on types of assessment is fairly good; however, the practices need improvement and support from educators and stakeholders.

3.4. Discussion

This research was aimed to examine classroom assessment practices to identify teachers' perceived skills and their practices in classroom assessment, and to provide suggestion for improving the quality of in-service teacher-training program. The discussion then focuses on the main findings of the research concerning the teacher perceived skill with its factors and their practice of classroom assessment, along with its factors.

Teacher's skills are subdivided into six "factors". In the Criterion Referenced Testing Factor, teachers reported mostly skilful in covering what they taught for a test. This indicates that teachers frequently practice summative assessments. Summative assessment measures the extent to which the materials have been given (12). They also reported having fairly good skill in Aligning Items to Objectives Factor. They were aware that constructing test items begins with looking at the teaching objectives. The assessments then map directly to the learning objectives. This in in line with the theory of backward design approach of course design by (18). This implies that teaching objectives mirrors what to be tested. What is questionable from the result is that teachers admitted less frequently giving fair value to students. This finding needs evidence, of course, by further exploration such as doing interview with the subjects; but, such a response indicated that so far the teachers have done the practice of giving unfair values to students.

In the Grading Practice Factor, teachers reported to use portfolio less frequently. This might be due to the fact that there is a lack of a standard for a portfolio. Unless there is a fixed rubric to use, using portfolio will lead to subjective individual judgement. Using portfolio presents several limitation as acknowledged by (19) that despite many potentials benefits, the use of portfolio assessment is limited by some issues such as lower reliability, high cost, high demand for well trained professional teachers and unsupported learning atmosphere in Indonesia. Learning atmosphere in the country is typical with traditional assessment, not alternative assessments.

The finding also revealed that teachers reported not skilful in Statistic Application Factor. They conducted item analysis and calculated variability less frequently than the other activities such us revising items and calculating central tendency. It seems that calculating central tendency became their routine when making report of the students learning. Calculating central tendency in terms of Mean is something common to be done by teachers, but finding item variability is rarely done by teachers. Assessing problem solving skill was also a problem for teachers. They were not accustomed to constructing tests on a problem-based learning. Problem-based learning is now popularized in the current curriculum as students have to be engaged with 21st century skills, including problem-solving skills.

From the Function of Test Factor, teachers admitted they were skilled in doing all types of assessment, including assessment of learning, assessment for learning, and assessment as leaning. the assessment of learning and assessment for learning. The result revealed that teachers had practiced the three types of assessment simultane-ously. They reported to have practiced assessment of learning more frequently than assessment for learning. This indicates that they used more summative tests during their teaching than formative assessment. Whereas teachers least often practiced the assessment as learning. It does not matter which must be practiced more than the others, but one fact to note is that teachers have attempted to use the three types of test or assessment because one is not to substitute the other. They should be integrated in the classroom as suggested by (20), which are not necessarily discrete approaches and may be used individually or together and formally or informal.

Viewed from the purposes of assessment which are subdivided by three types of assessment; summative, formative, and diagnostic, teachers reported to have practiced summative assessment most frequently. This confirms other research findings that teachers mostly adopted summative assessment (21-24). Teachers reported greater use of summative assessment than the others because of the fact that the students learning attainment is usually measured at the end of teaching. Formative assessment which is given during the learning process is usually given informally; thus, it is not considered as an assessment. Whereas diagnostic assessment was just highlighted in the new Kurikulum Merdeka, it was also given in the classroom. This form of assessment belongs to pre-assessment (25). This includes checks of prior knowledge and skill levels and surveys of interests or learning-style preferences. However, this form of assessment is basically informative in nature because diagnostic assessments are used for formative purposes. In Indonesian education system, diagnostic is seen as a strategy of identifying

learner's entry behaviour in order that teachers can serve differentiated leaning in the classroom.

4. Conclusion

Inservice teachers are the terminal academic professional training expected to produce skilled teachers who are only experts in teaching, but also capable of assessing students. Assessment is not only given to measure the learning outcomes, but also used to monitor the learning progress, to identify the students' problem with their learning and even to check the prior knowledge just before the new lesson is given.

The study revealed that the in-service teachers reported they were skilful all types of assessment. Some subskills need improvement on analysing the assessment results, giving values to students, using portfolio as alternative assessment, and using statistic application. Assessment *of, for*, and *as* learning have been done by teachers without any significant difference in the mean. Of the three types of test or assessment, diagnostic test was not frequently practice by teachers. This type of assessment is a new issue for teachers so that they are not accustomed to using such assessment in classroom practice. However, teachers reported they understood the principle of the assessment.

Teachers skills in assessment are affected by multifactor, among others are knowledge and the training that they afford to follow. As in-service teachers they got courses in several topics, but a particular topic for assessment is not yet offered to them. That is why the writer proposes that the in-service teacher training program curriculum must cover course materials in assessment to better equip teachers with the ability to execute various assessments based on the purposes and functions.

References

[1] Purwanti E. Preparing the implementation of *merdeka belajar–kampus merdeka* policy in higher education institutions. 4th International Conference on Sustainable Innovation 2020–Social, Humanity, and Education (ICoSIHESS 2020) 2021 Jan 21 (pp. 384–391). Atlantis Press; 2021.

[2] Danaye Tous M. A review and analysis of language curriculum design. Critical Studies in Texts & Programs of Human Sciences and Council for the Study of Humanities Texts and Books. 2021;21:61–85.

[3] Huth T. Classroom-based Conversation Analytic Research 2021. Cham: Springer; 2021. Conceptualizing interactional learning targets for the second language

curriculum. p. 359–381.

[4] Hidayah SN, Martani W, Supartono W. Accommodating stakeholders' voices in the curriculum development in an Indonesian higher education institution. Jurnal Educative: Journal of Educational Studies. 2021;6:18–31.

[5] Generoso JC, Arbon AM. Language needs analysis: An EAP curriculum design to develop foreign students' English skills. Journal of Asia TEFL. 2020;17:428.

[6] Khamying P, Chano J, Boonla W, Nithideechaiwarachok B. Thai language curriculum to enhance creativity thinking skills for primary school students. Journal of Education and Learning. 2022;11.

[7] Prastikawati EF. Technology-based formative assessments implemented by secondary school english teachers during remote learning. KnE Social Sciences. 2022:180–190.

[8] Krishnapatria K. Merdeka Belajar-Kampus Merdeka (MBKM) curriculum in English studies program: Challenges and opportunities. ELT in Focus. 2021;4:12–19.

[9] Pattalitan AP. The implications of learning theories to assessment and instructional scaffolding techniques. American Journal of Educational Research. 2016;4:695–700.

[10] Mahshanian A, Shoghi R, Bahrami M. Investigating the differential effects of formative and summative assessment on EFL learners' end-of-term achievement. Journal of Language Teaching and Research. 2019;10:1055–1066.

[11] Meletiadou E. Exploring the impact of peer assessment on EFL students' writing performance. IAFOR Journal of Education. 2021;9:77–95.

[12] Brown GT. Is assessment for learning really assessment? Frontiers in Education. 2019;4:64. Lausanne, Switzerland: Frontiers Media.

[13] Wilson DM, Narasuman S. Investigating teachers' implementation and strategies on higher order thinking skills in school based assessment instruments. Asian Journal of University Education. 2020;16:70–84.

[14] Azim DF. Learning promoted by classroom assessment in higher education of Bangladesh: A case study. International Journal of Social Sciences. 2014;26:165–171.

[15] Acar-Erdol T, Yildizli HÜ. Classroom assessment practices of teachers in Turkey. International Journal of Instruction. 2018;11.

[16] Radianti J, Majchrzak TA, Fromm J, Wohlgenannt I. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. Computers & Education. 2020;147:103778.

[17] Zhang Z, Burry-Stock JA. Classroom assessment practices and teachers' self-perceived assessment skills. Applied Measurement in Education. 2003;16:323–342.

[18] Wiggins G, McTighe J. Understanding by design. Alexandria, VA: Association for Supervision and Curriculum Development. Soc Stud Grade. 2005;6:31.

[19] Afrianto A. Challenges of using portfolio assessment as an alternative assessment method for teaching English in Indonesian schools. International Journal of Educational Best Practices. 2017;1:106–114.

[20] Black P, Wiliam D. Classroom assessment and pedagogy. Assessment in Education: Principles, Policy & Practice. 2018;25:551–575.

[21] Prastikawati EF, Wiyaka W, Lestari MY. Secondary school students' perception on Edmodo as online learning platform in English learning. Language Circle: Journal of Language and Literature. 2022;16:296–307.

[22] Firdaus MS, Prastikawati EF, Wiyaka W. Online formative assessments in English Teaching and Learning. SALEE: Study of Applied Linguistics and English Education. 2022;3:23–34.

[23] Qu W, Zhang C. The analysis of summative assessment and formative assessment and their roles in college English assessment system. Journal of Language Teaching & Research. 2013;4.

[24] Ishaq K, Rana AM, Zin NA. Exploring summative assessment and effects: Primary to higher education. Bulletin of Education and Research. 2020;42:23–50.

[25] Gan Z, Leung C. Illustrating formative assessment in task-based language teaching. ELT Journal. 2020;74:10–19.