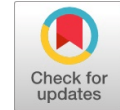# Investigating the Potentials of Classroom Assessment: A Critical Discussion

**Md Shaiful Islam, Mahani Bt Stapa**

*Abstract— Despite gaining currency in the contemporary assessment literature, the depth of the various dimensions of classroom assessment is yet to be explored. This systematic review aims to survey what the current literature reports on the evidence of success of classroom assessment in transforming students towards learning, and it also puts forth a number of implications. The review methodology we adopted includes inclusion/exclusion criteria, identification of the relevant literature, screening articles for the final selection and finally judgment used for the quality of the articles. The search for literature started by the keywords, such as, assessment, summative assessment, formative assessment, formative classroom assessment, and assessment literacy. The search engine and databases we used for the research articles entail Google Scholar, Web of Science, Scopus, JSTOR and ERIC. Considering the reality of the development of knowledge in continuum, we extended the period for literature search from 1989 to 2019 (thirty years). The major themes that surfaced incorporate formative assessment, self- and peer assessment, feedback, reliability and validity, alternative assessment, and assessment literacy. We then critically analyzed the themes and suggested implications.*

*Keywords: assessment, assessment literacy, classroom assessment, formative assessment, summative assessment*

## I. INTRODUCTION

Despite the tremendous importance of assessment in education, there is still a struggle to bring appropriate harmony among the terms used in assessment epistemology. Therefore, various types of dilemma exist among the instructors, educators, researchers, policy makers, and most importantly among students. Assessment in a generic sense refers to a wide variety of tools or methods used by educators to document students' progress in learning, skill acquisition, academic needs and preparedness. Assessment is an integral aspect of teaching and learning process as it provides guidelines to designing lessons, implementing them, and evaluating their success (Brown and Lee, 2015). Miller, Linn and Gronlund (2009) hold that assessment is a general term that includes the full range of procedures used to gain information about student learning (observations, ratings of performances or projects, paper-and-pencil tests) and the formation of value judgment concerning learning progress. According to Ovando, Combs and Collier (2006), "Assessment is the process of gathering data about what a student knows and what the student can do". Most often the learning/teaching process and assessment are conceived as two separate activities where assessment usually implies formal tests given on special occasions rather than collecting information about the students' performance during ordinary classroom activities.

The terms formative assessment, and summative assessment frequently appear in literature in their dichotomous relationship. Because of paradigmatic preferences, people choose either one, but their relationship has not been explored much. Most importantly, how they operate in the classroom has not been substantially investigated. Classroom plays a significant role when education is pursued in formal approach. Hence this paper attempts to critically review the influence of assessment in general, and classroom assessment (hereafter, CA) in particular. This paper also investigates the role stakeholders such as students, and teachers play to manipulate assessment for enhancing learning.

In this review, we critically analyzed the arguments of researchers, scholars, and practitioners on classroom assessment with a view to gaining more in-depth understanding, and insights of the impacts and potential of such type of assessment. Furthermore, various dimensions of classroom assessment were critically scrutinized and analyzed so that stakeholders can confidently make informed decisions regarding their approach, and application of classroom assessment in the real world situations.

## II. RESEARCH QUESTIONS

1. What are the factors influencing the shift from one assessment paradigm to another?

2. In what ways does the classroom formative assessment enhance students' learning?

3. What are the alternative methods of assessment currently in practice, and how do the assessment experts, teachers and the students perceive them?

4. How does assessment literacy influence students' learning?

## III. METHODOLOGY

The review methodology we adopted includes inclusion/exclusion criteria, identification of the relevant literature, screening articles for the final selection and finally judgment used for the quality of the articles. The keywords such as assessment, summative assessment, formative assessment, formative classroom assessment, alternative assessment, and assessment literacy were used for the initial search for literature.

The search engine, and databases we used to search for the research articles entail Google Scholar, Web of Science, Scopus, JSTOR and ERIC. Moreover, once we got a seminal paper, we used the method of citation search since the citations of that article are worth exploring. Finally, an approach familiar as 'snowball approach' (Black & Wiliam, 1998) was used to search for articles from the reference list. We extended the period for literature search from 1989 to 2019 (thirty years) so that we can identify the dynamics that influenced the innovations in assessment. For the selection of the articles, the titles and the abstracts were read, and articles which were close to the research questions were finally selected. With a view to gaining in-depth understanding, only 25 articles were finally short listed for critical analysis.

## IV   RESULTS AND DISCUSSION

### Paradigm Shift in Assessment

Black and Wiliam (1998) report a shift that has occurred during the recent years from the concentration on psychometric tests which are weakly linked to the learning experiences of the students towards a greater interest in the interactions between assessment and classroom learning. Coupled with this shift are the many expressions of hope that improvement in classroom assessment will make a strong contribution to the improvement of learning. It is evident that the top-down high-stake large-scale tests are built around instruction and outcomes; on the other hand, the classroom assessment integrates curriculum, instruction, assessment, and professional development (Shepard, Penuel & Pellegrino, 2018a). Rea-Dickins (2007) identifies that external tests exclude the lived curriculum which is collaboratively developed and used by the teachers and the students while they are engaged in classroom assessment. Furthermore, classroom assessment is used to inform language learning and teaching, and where assessment is seamlessly integrated into teaching and learning.

Classroom assessment integrates various formal and informal methods instructors use to make value judgment regarding student learning and performance (Volante, 2006). Observation checklists, anecdotal notes, reflective journals, group projects, portfolios, essay and interpretive exercises, open-ended tasks, concept maps, performance-based tasks, demonstrations, and oral questioning constitute the most popular CA methods and tools (El-Emam, 2006). According to Popham (2009), classroom assessment is, "… formal and informal procedures that teachers employ in an effort to make accurate inferences about what their students know and can do".  The dominant features of classroom assessment which are identified from the literature (Popham, 2009; Black & William, 2004; Garrison & Ehringhaus, 2011; Stiggins & Chappuis, 2005) include: 1) formative in nature with potential to improve students' learning, 2) internal to the classroom closely linking the classroom instruction, 3) implemented by the tests or assessment activities that are usually teacher-made, not standardized, 4) feedback-oriented, 5) informal in operation, 6) low stake in functions, and 7) not grade-oriented. Volante (2006) has identified seven interrelated principles which are supposed to operate in sequence. The first principle stresses student-centeredness in CA since the fundamental goal of CA is enhancing students' learning. Aligning explicitly expressed learning targets with the CA tasks comes next followed by equipping CA with numerous methods so that students with multiple learning styles feel comfortable to adopt their choice methods to perform in the assessment tasks. The fourth principle is the ability of the CA procedures to integrate a variety of student skills which inherently incorporates the fifth principle which is reducing bias in assessment. Reliability and validity constitute Volante's sixth principle which is often regarded as the most significant quality of a good assessment design. Finally, Volante believes that CA must be efficient since any method without efficient implementation proves to be useless at the end.

Assessment experts hold varied opinions about CA. Some think that CA, as a valid and reliable source of evidence regarding students' learning, can be graded and linked to accountability or summative assessment for a better inference about the quality of education. They believe that effective blending of these two assessments can yield benefits for the learners. The need for integrating formative classroom assessment and summative accountability assessment to enhance the reliability, validity, and utility of the accountability assessment data has been expressed by many assessment experts and educationists (Wilson & Carstensen, 2007; Banta, 2007; Wilson & Draney, 2004). Bennett (2011) also believes that the purposes of summative and formative assessments are not mutually exclusive. Therefore, they can coexist as primary and secondary purposes of the same assessment. However, guidelines on how appropriately the blending should be made are inadequate. Therefore, such blending may cause several pitfalls which may challenge the learners.

An opposing view regarding grading and blending CA with summative one has been presented by other researchers. For example, Black and Wiliam (2009) think that classroom assessment is effective in raising standards only when the purpose is to improve students' learning, rather than grading. Similarly, Shepard, Penuel, and Pellegrino (2018a) proposed that to have integrity in the intention of formative assessment culture and to motivate students in learning, grading policies should avoid using points and grades. Rather, assessment should create opportunities for students to use feedback to improve their work. Dunn & Mulvenon (2009) argue that although there is a plethora of literature arguing that classroom assessment in the form of formative assessment can raise educational outcomes of students, the definition of classroom assessment and formative assessment remain vague and excepting theoretical arguments, there is very little empirical research showing that real educational achievements result from formative classroom assessment. Another critique comments that the outcome of CA is measured through conventional paper pencil tests or summative tests which are not always a good measure of learning. Bennett (2011) argues that though widely acknowledged, the effectiveness of formative assessment claims is not always well grounded because of the lack of a uniform definition of the concept of formative assessment.

Therefore, more empirical research on classroom assessment practices would help this educational approach thrive.

*Classroom Assessment and Learning*

Opposed to summative assessment, such as semester final, year tests, or nationwide standardized tests, CA is integrated into the classroom instruction which makes it internal, and to a large extent informal in operation. Therefore, it is interactive in nature, and this interactive nature of CA provoked Hamp-Lyons and Tavaras (2011) to carry out a study on English as a Second Language (ESL). According to them, the term 'interactive assessment' (IA) refers to a conscious, deliberately sustained process through which the teacher monitors each individual's language production, progress and needs, and applies timely interventions that will encourage and scaffold the learners to produce an assessable performance. The teacher is aware of such performance because he has the experience of working with them over a consistent period of time. They also claim that 'interactive assessment' leads to the creation of learning and assessment opportunities that will stimulate and even challenge the learner to demonstrate a higher level performance than that being produced. They emphasized encouraging teachers to explore, in their own English language classrooms, ways of providing 'feedback' during the assessment process that move learners forward. Many researchers have described this as 'feed-forward' (Arbib, 1975; Cowie, 2005; Nicol & McFarlane-Dick, 2004). Furthermore, CA enhances the volume of responsibilities shared by the teachers, students, and peers (Black & Wiliam 1998b, 2009; Wiliam, 2011). CA elicits evidence of student achievement which is interpreted for use by teachers, learners, and peers for decision making regarding the next steps to modify instruction, or to change learning initiatives (Wiliam, 2011). However, we argue that teachers and students experience differently in various classroom contexts; therefore, a unitary approach of eliciting evidence will not work. Further intensive and context specific study is required to understand the local cultural dynamics.

The purposes of CA the researchers (Wiliam, 2011; Hattie, 2008; Carr, 2008; Perrenoud, 1991) have identified entail: 1) to provide the teachers with information necessary for adapting instruction to meet students' needs, 2) to help students develop knowledge, understanding, and metacognition, and 3) to make students aware of their responsibilities in learning. To achieve the desired results of CA, Wiliam (2011) maintains that teachers use a host of strategies to involve students. Teachers perform so by 1) illustrating the learning intentions and success criteria, 2) facilitating successful classroom discussions, activities, and learning tasks that extract evidence of learning, 3) supplying feedback which provokes students toward learning, 4) motivating learners to be instructional resources for others and 5) encouraging learners to take the ownership of their own learning. An example of formative CA can be an oral question-answer session after teaching a unit (Garrison & Ehringhaus, 2011). Wiliam (2011) described the formative role of classroom assessment as both 'Assessment for Learning' and 'Assessment as Learning' as opposed to summative, high-stakes, and accountability assessment which mainly plays the role of 'Assessment of Learning'.

While 'assessment of learning' refers to measurement by final scores, grades, certificates, or ranking which does not yield immediate improvement in students' learning, 'assessment for learning' is used to help students further enhance their learning. 'Assessment as Learning' which is a more sophisticated way of thinking about assessment means that the assessment process itself can be a way to help students learn. Self-assessment, peer assessment, and other alternative ways of assessment such as games, debates, projects, presentations used in classroom assessment are not only used for measuring learning but themselves constitute a rich learning process for the students. According to William (2011), the formative feature of classroom assessment is most powerful when it is used as a diagnostic tool. The goal of such tool is to identify what students have and have not achieved by using classroom assessment techniques such as teacher observation and classroom discussion, interviews and in-class brief writing assignments, class tests, and homework. Then, the teachers can make responsive changes in teaching and ultimately in students' learning. Since formative assessment is tightly linked with instructional practices, teachers must first consider how their classroom activities, assignments, and tests support learning aims and allow students to communicate what they know and then use this information to improve teaching and learning.

According to Black and Wiliam (1998), there are two actions that operate at the core of formative assessment: 1) perception by the learner of a gap between a desired goal and his or her present state (of knowledge, and/or understanding, and/ or skill), and 2) action taken by the learner to close that gap in order to attain the desired goal (Ramaprasad, 1983; Sadler, 1989). They find that involvement of students in formative assessment is determined by 1) factors which influence the reception of the message, and the personal decisions about how to respond to it, and 2) different ways in which positive action may be taken and the regimes and working contexts in which that action may be carried out. The focus here will be on study methods, study skills, collaboration with peers, and on the possibilities of peer and self-assessment. The major students-initiated assessment methods that emerge from the critical review of seminal papers on formative classroom assessment by Black and William (1998) include 1) assessment by students, 2) self-assessment, and 3) peer assessment. In formative assessment, any teacher has a choice between two options: 1) to help students develop capacity to recognize gaps that persist between the curriculum objectives and the students' current level of achievement and entrusting the students with the responsibility for planning and carrying out any remedial action that may be needed, or 2) to take responsibility by the teachers themselves for generating the stimulus information and directing the activity which follows. The two options often overlap. Thomas (1993) focused on self-directed learning which he finds is necessary for the students to perform practical works, develop study skills, and take responsibility for their learning. Self-evaluation is an intrinsic aspect of reflection on one's own learning.

1174

Elshout-Mohr (1994) in a review points out that students are often unwilling to give up misunderstandings; they feel to resolve their learning problems by themselves. These indicate that self-assessment is essential. Similarly, Hattie et al. (1996) argue that direct teaching of study skills to students without attention to students' reflective, and metacognitive development may well be pointless. One reason for the need to look for radical change is that students bring to their work models of learning which may be an obstacle to their own learning. Existence of such models, often culturally determined, is illustrated by a comparison of the approaches to learning by Australian and Japanese students (Purdie & Hattie, 1996). However, the findings suggest that the most able students in either country are more alike than their peers in having developed similar effective habits of learning, and therefore, such constraining traditions can be overcome.

*Alternative Methods of Assessment*

Because of its fluid and flexible nature, classroom assessment is enriched by a wide range of alternative methods. Of all the attempts of defining alternative assessment, the definition offered by Shepard (2000) seems more appropriate as he maintains that alternative assessment refers to an alternative approach that essentially embeds assessment into the teaching and learning processes, and in doing so it explicitly stresses students' performance. Therefore, some popular categories such as performance assessment, authentic assessment, dynamic assessment, or direct assessment also fall into this approach since all these methods intend to judge what students can do when they complete actual assessment tasks which are aligned with the goals of instruction. Although advocates of psychometric tests claim that such tests also support learning and they may do so, they are actually external in nature while alternative assessment methods are intricately connected with the teaching/learning processes. They also facilitate the view that the learners are actively engaged in the classroom to construct meaning, and their participation in assessment events reflects that alternative procedures promote self-learning and self-monitoring (Sadler, 1989).

In terms of processes involved in alternative assessment, there are two constructs: self-evaluation and feedback (Gipps & Stobbart, 2003). These constructs lead to the relationship assessment relation between the teacher and the learner. Through the active participation of the students into the assessment tasks and the support provided by the teachers, a classroom turns into a stage for dialogue and debate which stimulate both to construct innovative knowledge. Self-evaluation develops students' metacognition, i.e., thinking about thinking, which entails numerous self-awareness processes that help the students plan, monitor, orchestrate, and control their learning (Gipps & Stobbart, 2003). Learners' metacognition involves them in self-assessment, and self-evaluation. Feedback from the teachers to the students, on the other hand, is the process that integrates assessment in the teaching and learning cycle (Black & Wiliam, 1998; Crooks, 1988; Gipps, McCallum, & Hargreaves, 2000). A more detailed discussion on feedback has been presented in another section of this review.

When forms of alternative assessment are concerned, most popular in assessment literature are performance assessment, and portfolio. Performance assessment is a systematic attempt to gauge a learner's ability of using past knowledge to solve a new problem, and to do so, teachers develop a real life assessment task with a purpose to elicit authentic responses from the students which are observed and examined by the teachers (Stiggins & Bridgeford, 1982). This kind of assessment often involves problem-solving projects, or projects on oral presentation such as seminar, or public speaking on contemporary significant issues. To do such performances, students get involved to explore various dimensions of the problem, develop in-depth understanding, and positive transformation towards genuine learning occurs.

Portfolio assessment is another alternative to traditional paper and pencil tests. As Klenowski (2000) holds, a portfolio incorporates documenting achievements, monitoring self-development, reflecting on the artefacts and modifying, and analyzing learning experiences. Rationale for the support of portfolio assessment offered by the proponents of such method is built around a few arguments: 1) its power of building students' confidence, 2) its strong influence on enhancing students' learning, and 3) its unambiguous and accessible procedure of collecting, sorting, and annotating evidence (Gipps & Stobart, 2003; Broadfoot, 1998; Klenowski, 2000). This kind of assessment helps students develop new perspectives essential for positive transformation, be engaged in growth process, reflect on their learning initiatives, and become conscious of self-evaluation. To implement portfolio assessment effectively, Wolf and Koretz (1998) recommend some procedures and practices: 1) plenty of evidence, 2) detailed descriptive recording, 3) specific note-taking on teachers' feedback, and 4) sincere involvement of the students during the preparation of the portfolios.

*Reliability and Validity*

Classroom assessment offers larger scope for ensuring reliability, consistency in assessor's judgment about students' assessment performance, since students get multiple opportunities to exhibit their knowledge and understanding (Volante, 2006). Moreover, students can negotiate with the instructors if they have strong arguments and evidence in support of their responses to the assessment tasks. Volante (2006) also finds that validity in assessment is substantiated by classroom assessment as CA is authentic in nature, and it confirms more accuracy since it occurs in the real life contexts. Furthermore, Huerta-Macias (1995) reports that no significant change is necessary to implement classroom assessments since they are embedded within instruction. Therefore, such assessments are more valid in terms of their alignment with the curriculum, and more authentic since they are integral to the classroom teaching activities and learning processes. However, Gronlund (2003) makes us aware of the potential pitfalls that challenge the validity of classroom assessment. The pitfalls include 1) assessment tasks that supply insufficient sample of achievement, 2) unintended function of the procedures due to ambiguity, bias and difficulty, 3) improper administration, 4) impressionistic scoring, and 5) unclear task instructions.

1175

We find that post assessment dialogue and negotiations with the students occurring in the classroom may help avoid such problems, and can guarantee reliability and validity.

According to Huerta-Macias (1995), criteria of reliability and validity associated with traditional psychometric test are different from factors that determine reliability and validity in CA as held. She presents parallels between alternative assessment and qualitative research (e.g., Guba & Lincoln, 1994). She suggests that trustworthiness and triangulation of classroom assessment data determine quality in alternative assessment. However, she finds merit in what Wilde, Del Vecchio and Gustke (1995) suggest. According to them, trained judges, working with clear criteria may be consulted to ensure reliability in alternative. These experts monitor periodically to ensure that raters use criteria and standards in a consistent manner (Huerta-Marcias, 1995).

### Role of Teachers in Classroom Assessment

While reviewing relevant articles, Black and Wiliam (1998) identified choice of assessment tasks, classroom discourse, using questions, tests and feedback from tests as the common roles teachers play when the classroom assessment occurs. To begin with, it is obvious that formative assessment which guides learners towards valued learning goals can only be generated with tasks that are aligned with the goals. The diversity of classroom assessment procedures offers the teachers an opportunity to monitor how much the students have understood the lesson contents, and how successfully they have performed them (Garnett &Tobin, 1989). Ames (1992) highlights that the tasks should be novel and varied in interest and should offer reasonable challenge. However, too much challenging tasks may push the students to avoid the tasks with the risks that their confidence may fall (Blumenfeld, 1992). Similarly, tasks that are meaningful might be productive for learning.

Secondly, the quality of discourse exchanged between the teacher, and the students influences their interactions on various aspects of CA. The meaning encoded in any discourse depends also on the context, and the patterns of relationships between those involved in the production of discourse (Carlsen , 1991; Filer, 1995). Johnson & Johnson (1990) present a meta-analysis which illustrates that collaborative discourse can produce significant gains in learning. Going further, Roth and Roychoudhury (1994) recommend the use of concept maps as an aid for classroom discussions; such maps, drawn by the students, provide useful angles in clarifying the points under discussion and thus enable the teacher to engage the students in dynamic assessment.

Thirdly, Stiggins et al. (1989) value the quality of questions in CA. King (1990,1992a, b; 1994) promoted the generation of thought provoking questions by the students, and she interprets that such training helps students develop learner autonomy. Such questions also promote students' higher-order thinking, and more peer interaction. Therefore, teachers should encourage students in questioning about the various issues that occur in the class and also about the feedback they receive from teachers and peers.

Next comes the use of tests for CA. Bangert-Drowns, et al. (1991b) reviewed the effects of frequent class testing, and it indicated that several short tests were more effective than fewer longer ones. However, Iverson et al. (1994) did not find any significant correlation between frequent testing and students' improvement in performance even though the students in their experiment said that they would like to have such tests in other courses also. A similarly negative result was reported by Strawitz (1989), but a contrary, positive effect was found by Schloss et al. (1990) working with graduate students in teacher training for special education.

Another remarkable activity the teachers can exert through classroom assessment is sharing learning intentions and success criteria with the students (Cauley & McMillan, 2010; Heritage, 2007; Lysaght & O'Leary, 2013). Using these two, the teachers can monitor if the students are moving towards the right direction in their progress in learning. Besides, everyday practice of the classroom assessment methods helps the teachers elicit evidence about students' learning processes. The methods they can use include observations, classroom conversations, and homework assignments. The most important role of teachers in classroom assessment is providing feedback to the learners. Feedback is also embedded in classroom assessment for the teachers since they get information about the success or failure regarding their instructional methods and teaching materials. With this information, they can adapt instruction. Finally, the teachers can also engage the students in the classroom for self- and peer assessment.

### Role of Feedback in FCA

Provision of effective feedback is the key characteristic of classroom assessment (Black & William, 1998). Feedback is information about the gap between actual classroom performance and intended learning outcomes. According to Kluger and DeNisi (1996), feedback interventions are actions taken by an external agent to provide information regarding some aspects of one's task performance (p. 4). The main objective of using CA is to aid learning through the continuous provision of feedback. Unlike summative assessment, such as semester final, year tests, or nationwide standardized tests, CA is integrated into classroom instruction which gives immediate feedback to teachers and students on their teaching and learning so that teachers can improve their teaching and give feedback to students to improve their learning. Hamp-Lyons and Tavaras (2011) maintain that CA in the form of 'interactive assessment' encourages teachers to explore, in their own English language classrooms, innovative and interesting methods of providing 'feedback' during the assessment process that move learners forward. We hold the premise that need-based detailed feedback should be provided to the students only when they require to remove their misconceptions about the lesson contents, or to overcome any other kind of weaknesses they are facing. In addition, task-specific appreciation should be embedded in the feedback methods in order to keep the students' spirit of learning active; otherwise, feedback may appear to be counterproductive. Bangert-Drowns et al. (1991a) find that feedback is most effective when it is designed to stimulate correction of errors through a thoughtful approach.

However, use of feedback by students does not always yield uniform positive results. For example, students often ignore feedback if it is accompanied by grades since they value grades more that feedback (Buttler, 1988). In addition, there is no evidence that students unvaryingly respond to the feedback, rather differences do exist which may damage the purpose of providing feedback (Lofgren & Lofgren, 2017). Feedback with too many details (Brookhart, Andolina, Zuza, & Furman, 2004; Carr, 2008), or being treated too trivially (Brookhart et al., 2004; Wiliam, 2007) generate negative consequences on the learners. Wiliam (2011) reports that feedback must prioritize learning goals, stimulate learners' reflection, and promote self-learning. Feedback also motivates learners to be involved in learning (Perrenoud, 1991).

The reason for educators to spend a significant amount of time developing assessment tools for teachers to help them know their impact (Hattie, Brown, & Keegan, 2005). This is why teachers need to know how they can assist students to become assessment savvy so that they can diagnose their own weaknesses, can respond to the interventions and can evaluate their learning. The key to the maintenance of learning is feedback because it influences what happens after instruction.

Role of Assessment Literacy in Classroom Assessment

Classroom assessment will not yield desired results if it is not appropriately integrated into the teacher education programmes. The absence of solid assessment literacy is the reason of ineffective teaching (Eckhout, Davis, Mickelson, & Goodburn, 2005). Classroom assessment literacy is the essential knowledge and skill the teachers must possess to gather information about students' performance which they use to guide students in achieving the learning outcomes (Campbell, Murphy, & Holt, 2002; Popham, 2005). DeLuca and Bellara (2013) argue, "assessment literacy involves integrating assessment practices, theories, and philosophies to support teaching and learning within a standards-based framework of education".

Stiggins (1999) has identified seven competencies for teachers' assessment literacy: 1) connecting assessments to clear purposes, 2) clarifying achievement expectations, 3) applying proper assessment methods, 4) developing quality assessment exercises and scoring criteria and sampling, 5) avoiding bias in assessment, 6) communicating effectively about student achievement, and 7) using assessment as an instructional intervention.

All expectations generated by the potentials and promises of classroom assessment will not be realized if the main actors, the teachers, are not appropriately assessment literate. All teachers must have the competencies Stiggins has identified. Moreover, they should be familiar with the components of the framework Xu and Brown (2016) have recently proposed. The components include: 1) the knowledge base, 2) teacher conceptions of assessment, 3) institutional and socio-cultural contexts, 4) teacher assessment literacy in practice, 5) teacher learning, and 6) teacher identity.

## V  CONCLUSIONS AND IMPLICATIONS

Classroom assessment has tremendous potential to achieve educational goals regardless the academic subjects where it is applied. Because of its diverse nature, it can successfully integrate all paradigms of assessment. Currently popular concepts such as 'assessment for learning' (AfL) and 'assessment as learning' (AaL) are appropriately compatible with classroom assessment. However, it operates in a complex manner being influenced by a number of contextual factors. This review has critically discussed all major issues related to educational assessment.

Although the educators are excited about the paradigm shift, the use of novel assessment procedures may yield undesirable consequences if these are not implemented with adequate knowledge and skills. It is imperative for the stakeholders to gain substantial understanding of the purposes and implementation procedures of alternative assessment methods. Especially the students must be trained about the new methods of assessment. As the qualities of a good assessment scheme, reliability and validity are extremely crucial since the acceptability of the assessment heavily depends on them. Teachers' role to ensure these two qualities is crucial; therefore, teachers' awareness of them is expected.

In contrast to the past, learning is now viewed as an active process of knowledge construction, and making sense of the world. Numerous learning theories are in operation; assessment researchers often feel puzzled about how to connect one aspect of assessment with the other when they adopt one theory of learning for designing assessment tasks. Against this backdrop, development of teachers' assessment literacy is indispensible.

Considering the multidimensionality of classroom assessment as discussed above, a few implications are suggested. First of all, there should be more research about the connection between assessment and learning theories. Since students' learning is the ultimate goal, there should be more sophisticated study on this interconnectedness. Secondly, assessment methods must be compatible with the contemporary expectations of the real world. Hence, the policy makers must think of alternative forms of assessment suitable for the 21st century learning styles. Besides, emphasis on classroom-based assessment and its integration to the overall assessment scheme is essential as classroom

plays a crucial role in successful learning when formal education is concerned. Finally, 4th industrial revolution is instrumental in posing new challenges for all aspects of human life. So, educational approaches are assuming new shapes influenced by automation and digitization. Compatible assessment methods are necessary to be adaptable to the new educational realities.

# REFERENCES

1. Ames, C. (1992).Classrooms: goals, structures, and student motivation. Journal of Educational Psychology, 84,pp. 261-271.
2. Arbib, M.A. (1975). Artificial intelligence and brain theory. Annals of Biomedical Engineering, 3(3), 238-272.
3. Banta, T.W. (2007). Can assessment for accountability complement assessment for improvement?. Peer Review. Spring 2007, 9-12.
4. Bangert-Drowns, R.L., Kulick, J.A. & Morgan, M.T. (1991a).The instructional effect of feedback in test-like events. Review of Educational Research, 61, pp. 213-238.
5. Bennett, R.E. (2011). Formative assessment: a critical review. Assessment in Education: Principles, Policy & Practice, 18(1), 5-25, DOI: 10.1080/0969594X.2010.513678
6. Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education, 5(1), 71– 74. doi: 10.1080/0969595980050102
7. Black, P., & William, D. (2001). Inside the Black Box: Raising Standards Through Classroom Assessment. Phi, Delta, Kappan. 80(2). 139-148.
8. Black, P., & Wiliam, D. (2004). The formative purpose: assessment must first promote learning. In M. Wilson (Ed). Towards coherence between classroom assessment and accountability (pp. 20–49). Chicago: University of Chicago Press.
9. Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. Educational Assessment, Evaluation and Accountability, 21(1), pp.5-31.
10. Blumenfeld, P.C. (1992).Classroom learning and motivation: clarifying and expanding goal Theory. Journal of Educational Psychology, 84, pp. 272-281.
11. Broadfoot, P. (1998). Records of achievement and the learning society: A tale of two discourses. Assessment in education,S, 447-477.
12. Brookhart, S. M., Andolina, M., Zuza, M., & Furman, R. (2004). Minute math: An action research study of student self-assessment. Educational Studies in Mathematics, 57, 213–227.
13. Brown, H. D. & Lee, H. (2015). Teaching by Principles: An Interactive Approach to Language Pedagogy. New York: Pearson Education, Inc.
14. Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task involving and ego involving evaluation on interest and performance. The British Journal of Educational Psychology, 58(1), 1–14.
15. Campbell, C., Murphy, J. A., & Holt, J. K. (2002). Psychometric analysis of an assessment literacy instrument: Applicability to pre service teachers. Paper presented at the meeting of the Mid-Western Educational Research Association.
16. Carlsen, W.S. (1991) Questioning in classrooms--a sociolinguistic perspective. Review of Educational Research, 61, pp. 157-178.
17. Carr, S. C. (2008). Student and peer evaluation: Feedback for all learners. Teaching Exceptional Children, 40, 24–30.
18. Cowie, B. (2005). Pupil commentary on assessment for learning. The Curriculum Journal, 16(2), 137-151
19. Crooks, T.J. (1988). The impact of classroom evaluation practices on students. Review of Educational Research, 58, 438-481.
20. DeLuca, C., & Bellara, A. (2013). The Current State of Assessment Education: Aligning Policy, Standards, and Teacher Education Curriculum. Journal of Teacher Education, 356-372, DOI: 10.1177/0022487113488144
21. Dunn, K.E., & Mulvenon, S.W. (2009). A Critical Review of Research on Formative Assessment: The Limited Scientific Evidence of the Impact of Formative Assessment in Education. Practical Assessment, Research and Evaluation. 14(7). 1-11.
22. Eckhout, T., Davis, S., Mickelson, k., & Goodburn, A. (2005). A method for providing assessment training to in-service and pre-service teachers. Paper presented at the annual meeting of the Southwestern Educational Research Association.
23. El-Emam, Y. (2006). Rethinking classroom assessment: a case of paradigm shift. Mathematics Education Journal. doi: 10.13140/RG.2.2.32094.02880
24. Elshout-Mohr, M. (1994).Feedback in self-instruction. European Education, 26, pp. 58-73.
25. Filer, A. (1993).Contexts of assessment in a primary classroom. British Educational Research Journal, 19, pp.95-107.
26. Filer, A. (199 5) Teacher Assessment: social process and social products, Assessment in Education, 2, pp. 23-38.
27. Garrison, C., & Ehringhaus, M.(2011). Formative and Summative Assessments in the Classroom. Retrieved from http://ccti.colfinder.org/sites/default/files/formative_and_summative_assessment_ in_the_classroom.pdf
28. Garnett, P.J. & Tobin, K. (1989).Teaching for understanding: exemplary practice in high school Chemistry. Journal of Research in Science Teaching, 26, pp. 1-14.
29. Gipps, C., and G. Stobart. (2003). Alternative assessment. In International handbook of evaluation,ed. T. Kellaghan and D.L. Stufflebeam, 549–76. Dordrecht: Kluwer Academic Publishers.
30. Gipps, C., McCallum, B., & Hargreaves, E. (2000). What makes a good primary school teacher? Expert classroom strategies. London: Routledge Falmer.
31. Gronlund, N. E. (2003). Assessment of student achievement (7th ed.). Boston: Allyn and Bacon.
32. Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K.
33. Denzin, & Y. S. Lincoln (Eds.), Handbook of qualitative research (pp. 105-117). London: Sage.
34. Hattie, J., Biggs, J. & Purdie, N. (1996) Effects of learning skills interventions on student learning: a metaanalysis, Review of Educational Research, 66, pp. 99-136.
35. Hattie, J. (2008). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. London: New York: Routledge.
36. Hattie, J. (2013). Understanding Learning: Lessons for learning, teaching and research. Research Conference, 24-39.
37. Hattie, J. A. C., Brown, G. T., & Keegan, P. (2005). A national teacher-managed, curriculum- based assessment system: Assessment tools for teaching & learning. International Journal of Learning,10, 770–778.
38. Hamp-Lyons, L., and N. Tavares. Forthcoming 2011. Interactive assessment – a dialogic and collaborative approach to assessing learners' oral language. In Classroom-based language assessment, ed. D. Tsagari and I. Csepes. Frankfurt: Peter Lang.
39. Huerta-Macias, A. (1995). Alternative assessment: Responses to commonly asked questions. TESOL Journal, 5(1): 8–11.
40. Iverson, A.M., Iverson, G.L. & Llukin, L.E. (1994).Frequent, ungraded testing as an instructional Strategy. Journal of Experimental Education, 62, pp. 93-101.
41. Johnson, D.W. & Johnson, R.T. (1990).Co-operative learning and achievement, in: S. SHARAN (Ed.). Cooperative Learning: theory and research, pp. 23-27, New York: Praeger.
42. King, A. (1990).Enhancing peer interaction and learning in the classroom through reciprocal Questioning. American Educational Research Journal, 27, pp. 664-687.
43. King, A. (1992a).Comparison of self-questioning, summarizing, and note-taking review as strategies for learning from lectures. American Educational Research Journal, 29, pp. 303-323.
44. King, A. (1992b). Facilitating elaborative learning through guided student-generated Questioning. Educational Psychologist, 27, pp. 111-126.
45. King, A. (1994).Autonomy and question asking-the role of personal control in guided student- generated questioning. Learning and Individual Differences, 6, pp. 163-185.
46. Klenowski, Y. (2000). Portfolios: Promoting teaching. Assessment in Education, 7, 215-234.
47. Kluger, A.N. & Denisi, A. (1996).The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological Bulletin, 119, pp. 254-284.
48. Koretz, D. (1998). Large-scale portfolio assessments in the US: Evidence pertaining to the quality of measurement. Assessment in Education, 5, 309-334.
49. Löfgren, H., & Löfgren, R. (2017). Grades in the eyes of our parents: A narrative approach to
50. educational resilience in pupils' stories of getting their first grades. Nordic Journal of Studies in Education Policy, 3(2), https://doi.org/10.1080/20020317.2017. 1343624.
51. Miller, M. D.; Linn, L. R. & Gronlund, E. N. (2009). Measurement and Assessment in Teaching. New Jersey. Pearson Education, Inc.
52. Ovando, C; Combs, M & Collier V. (2006). Bilingual and ESL Classrooms: Teaching in Multicultural contexts (4th ed.): Chapter 8.Implications of High-Stake Testing 10.
53. Nicol, D. & McFarlane-Dick, D. (2004). Rethinking formative assessment in HE: A theoretical model and seven principles of good feedback practice. Higher Education Academy. Retrieved August 18, 2008. http://www.heacademy.ac.uk/assests/York/documents'/

54. Perrenoud, P. (1991). Towards a pragmatic approach to formative evaluation. In P. Weston (Ed.). Assessment of pupils' achievement: Motivation and school

55. Popham, W. J. (2005). Classroom assessment: What teachers need to know. Boston, MA: Allyn and Bacon.

56. Popham, W.J. (2009). Assessment literacy for teachers: Faddish or fundamental?. Theory into Practice, 48(1), 4-11, doi: 10.1080/00405840802577536.

57. Purdie, N. & Hattie, J. (1996).Cultural differences in the use of strategies for self-regulated learning. American Educational Research Journal, 33, pp. 845-871.

58. Rea-Dickins, P. (2007). Classroom-based assessment: Possibilities and pitfalls. In J. Cummins & C. Davison (Eds.). The international handbook of English language teaching, Vol. 1. (pp. 505–520). Norwell, MA: Springer.

59. Roth, W-M. & Roychoudhury, A. (1994).Science discourse through collaborative concept mapping: new perspectives for the teacher. International Journal of Science Education, 16, pp. 437-455.

60. Sadler, R. (1989).Formative assessment and the design of instructional systems. Instructional Science, 18, pp.119-144.

61. Shepard, L.A (1995). Using assessment to improve learning. Educational Leadership, 52(5), pp. 38-43.

62. Shepard, L.A. (2000). The role of classroom assessment in teaching and learning. In V. Richardson (Ed.), Handbook of research on teaching (4th ed., pp. 1066-1101). Washington, DC: American Educational Research Association.

63. Shepard, L. A., Penuel, W. R. and Pellegrino, J. W. (2018a) Classroom assessment principles to support learning and avoid the harms of testing, Education measurement: Issues and Practice, 37 (1), 52–57.

64. Schloss, P.J., Smith, M.A. & Posluzsny, M. (1990).The impact of formative and summative assessment upon test performance of science education majors. Teacher Education and Special Education Majors, 13, pp. 3-8.

65. Stiggins, R, & Bridgeford, N. (1982). The role of peiformance assessment in day to day classroom assessment and evaluation. Paper presented at annual meeting of the National Council for Measurement in Education, New York.

66. Stiggins, R., & Chappuis, J. (2005).Using Student-Involved Classroom Assessment to

67. Close Achievement Gaps. Theory Into Practice, 44(1), 11-18, DOI: 10.1207/ s15430421tip4401_3.

68. Stiggins, R.J., Griswold, M.M. & Wikelund, K.R. (1989) Measuring thinking skills through classroom assessment, Journal of Educational Measurement, 26, pp. 233-246.

69. Strawitz, B.M. (1989).The effect of testing on science process skill achievement. Journal of Research in Science Teaching, 26, pp. 659-664.

70. Thomas, J.W. (1993).Promoting independent learning in the middle grades-the role of instructional support practices. Elementary School Journal, 93, pp. 575-591.

71. Volante, L. (2006). Principles for effective classroom assessment. Brock Education Journal, 15(2), 134-147.

72. Wiliam, D. (2007). Keeping learning on track: Formative assessment and the regulation of learning. In F. K. LesterJr. (Ed.). Second handbook of mathematics teaching and learning Greenwich, CT: Information Age Publishing pp. 1053–1098.

73. Wiliam, D. (2011). What is assessment for learning? Studies in Educational Evaluation, 37, 2– 14.

74. Wilde, J. Del Vecchio, A. & Gustke, C. (1995). Alternative assessment for Latino students. In

75. M.Gonzalez, A. HuertaMacias, & J. Tinajero (eds). The Schooling of Latino Students: A guide to Quality Practice. Lancaster, PA: Technomic Publishing

76. Xu, Y. & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. Teaching and Teacher Education, 58, 149-162.

77. R. J. Vidmar. (1992, August). On the use of atmospheric plasmas as electromagnetic reflectors. IEEE Trans. Plasma Sci. [Online]. 21(3). pp. 876—880. Available:http://www.halcyon.com/pub/journals/21ps03-vidmar

ELT methods, education policy, and assessment in education.

**Mahani Stapa** is a Senior Leccturer at the Language Academy, Faculty of Social Sciences and Humanities, Universiti Teknologi Malaysia. She holds a Ph.D in Teaching English as a Second Language (TESL). Her areas of interest are Bilingualism/Multilingualism in Language Learning, Second Language Writing, Testing and Evaluation for Language Teachers and ELT Methodology.

## AUTHORS PROFILE

**Md Shaiful Islam** is Senior Lecturer, Department of Bangladesh (IUB). He obtained MA in Applied Linguistics and ELT from the University of Dhaka. Currently he is pursuing doctoral studies in TESL at Universiti Teknologi Malaysia (UTM). His areas of academic interest include curriculum, discourse analysis,