


REVIEW

Open Access

English language assessment in Bangladesh today: principles, practices, and problems



Md Shaiful Islam¹, Md Kamrul Hasan², Shahin Sultana³, Abdul Karim^{4,5} and Mohammad Mosiur Rahman^{4,5*} 

* Correspondence: mosiurbhai2.0@gmail.com

⁴BRAC University, Dhaka, Bangladesh

⁵Universiti Sains Malaysia, Penang, Pulau Pinang, Malaysia

Full list of author information is available at the end of the article

Abstract

The achievement of curriculum goals and objectives, to a large extent, depends on how assessment methods are designed, implemented, monitored, and evaluated. English language learning in Bangladesh has miserably failed, and ineffective assessment methods may be largely attributed to this failure. This paper attempts to address various aspects and issues of English language assessment in Bangladesh in relation to English language learning as a curricular reform and the education policy of the country. The analysis revealed that there was always a gap between the principles of assessment embedded into the curriculum and the actual assessment practices. Furthermore, heavily hard hit by the high-stakes testing, the curriculum, the learners, and the instructors need to be liberated from this vicious policy. The review concluded with a recommendation that teachers need to develop assessment literacy through teacher education programs that are essential to helping teachers to acquire knowledge, skills, professionalism, and assessment expertise.

Keywords: Language assessment, Curriculum reform, Communicative language teaching, English language teaching, Bangladesh

Introduction

Two separate studies in early 1970s (Miller & Parlett, 1974; Snyder, 1971) concluded that it was assessment, not teaching, that influenced the students most in their learning. Assessment plays a critical role in the learning and teaching process in any domain of education. With regard to the achievement in English language proficiency, it may be stated that the development of proficiency happens gradually, and this gradual process, to a large extent, is effectuated by the adoption and implementation of appropriate assessment methods and their effective implementation. Assessment plans are guided by the language curriculum goals and objectives. Therefore, the establishment of an appropriate alignment between the design of assessment methods and the curriculum goals and objectives is essential. Most importantly, it is necessary for an education system to ensure if it has adequate infrastructure to implement the planned English language assessment scheme to achieve the goals stated in the English language curriculum.

Although “test” and “assessment” are in many cases interchangeably used in language measurement literature, distinctions exist between them (Brown & Lee, 2015). Brown and Lee (2015) define “test” as a carefully designed tool having rating scales that

systematically measures a person's knowledge or ability in a particular domain of learning demonstrated through performance. However, "assessment" as an overarching term refers to an ongoing process that involves various tools to identify if the students are progressing in learning (Islam & Bt Stapa, 2019). Assessment enhances learning through the provision of feedback, and subsequently, using the feedback assessment procedures facilitate constructive interactions between the students and the instructors as well as among students; such interactions engage students to be informed about their strengths, and weaknesses and also, help them adopt appropriate learning strategies, or adapt the existing ones (Islam, 2019).

The poor proficiency in English among Bangladeshis is also demonstrated in the EF English Proficiency Index 2019 (Education First, 2019). The index measures English proficiency in five levels: very high, high, moderate, low, and very low. Bangladesh is placed in the very low level with 71st position in the ranking while the Netherlands has secured the first position. Furthermore, such out of the nearly 1700 students who had sat for the Elective English paper admission test, only two were able to meet the minimum requirements set by the university's English departmental teachers who had devised the test to ensure that only students with the competence necessary to study English at the tertiary level were admitted to the undergraduate program. The tests revealed how poorly prepared the students were; 12 years of English education at the primary, secondary, and higher secondary level had apparently not prepared them for the kind of language skills they would require to study in the English department (Alam, 2018; Rahman & Pandian, 2018). According to Hamid, Jahan, and Islam (2013), the quality of overall English language education is so poor that even a Bengali-medium, Master's-degree-level student cannot speak decent English.

With regard to English language learning in Bangladesh, it is reported that English language learning has miserably failed, and ineffective assessment methods may be largely attributed to this failure (Ali, Hamid, & Hardy, 2018; Ali & Walker, 2014; Hamid & Baldauf, 2008; Rahman, Pandian, & Kaur, 2018a). Despite the immense importance of assessment in shaping students' learning of the English language, an insignificant volume of research has been conducted on English language assessment in Bangladesh. However, this review has searched these handfuls of published articles from the databases, such as Google Scholar, Scopus, and Web of Science. Articles were searched by using keywords and then, were short listed by skimming the abstracts and the introduction. Inclusion and exclusion criteria were used to finalize the most relevant and significant articles. Search engine Google was frequently used to search for articles identified in the references of the reviewed articles for further review.

Against this backdrop and information gathered from the abovementioned procedures, this review first presents a brief overview of the education system of Bangladesh followed by the discussion on the latest education policy of Bangladesh. Third, the unsettled status of English in Bangladesh has been analyzed chronologically. Fourth, manifold dimensions of English language assessment in Bangladesh, such as the issue of alignment between curriculum goals and assessment procedures, reforms of assessment methods, current assessment practices, washback effects, assessment literacy, and validity and reliability in assessment with reference to the contextual factors that influence the planning and decisions of assessment design have been critically examined. Finally,

the article concluded with the several context-based implication of language assessment and education of the country.

Education system in Bangladesh

Bangladesh follows a British model of education, and the mainstream education in Bangladesh is divided into three stages: primary, secondary, and tertiary (Ali et al., 2018). Grades 1-5 constitute primary education while secondary education has three levels: junior secondary (grades 6-8), secondary (grades 9-10), and higher secondary (grades 11-12). There is also a 2-year pre-primary education in Bangladesh. There is the National Curriculum and Textbook Board (NCTB) entrusted with the responsibility of developing curricula and textbooks for the students of these pre-university levels of education.

Asian Development Bank (2015) describes Bangladesh education system as a test-driven education system. Examinations are administered both internally and externally. The internal examinations are administered by schools and higher secondary colleges for the promotion of the students to the next grade while the external examinations are high-stakes tests and are administered by the educational boards. The examinations that are held nationally are the Primary Education Completion Examination (PECE) at the end of grade 5, the Junior Certificate Examination (JSC) at the end of grade 8, the Secondary School Certificate (SSC) examination at the end of grade 10, and the Higher Secondary Certificate (HSC) at the end of grade 12 (Ali et al., 2018). Nine education boards located in Dhaka, Chittagong, Comilla, Jessore, Mymensingh, Rajshahi, Barisal, Dinajpur, and Sylhet administer these examinations. There are also two alternative boards: Bangladesh Technical Education Board and Bangladesh Madrasah Education Board (MoE, 2020). Bangladesh Technical Education Board is a state regulatory board which develops the curriculum, and learning materials for technical and vocational education at the secondary level and grants affiliation to technical and vocational institutions that prepare the students for diploma/certificates. Bangladesh Madrasah Education Board, on the other hand, develops the curriculum on religious education and also integrates compulsory courses of general education. It awards degrees/certificates parallel to the degrees/certificates of general education such as Dakhil (SSC equivalent), Alim (HSC equivalent), Fazil (Bachelor equivalent), and Kamil (Masters equivalent). However, students here also sit for examinations for other purposes such as university admission, studying in the university, and also for job recruitment (Al Amin & Greenwood, 2018a, 2018b).

Bangladesh education policy 2010

Ever since Bangladesh emerged as an independent country in 1971 liberating itself from the former West Pakistan (now Pakistan) occupation, the education sector was operated by various Education Commissions until 2010 when the currently functional National Education Policy was formulated. This policy is comprehensive in terms of goals and objectives that emphasize the power of education which has been conceived as an active force to transform Bangladesh as a socially peaceful, economically sustainable, and politically stable state. Moreover, the Bangladesh Education Policy enacted in 2010 emphasizes that access to education is a fundamental right of all children and any educational initiative should be aligned with the broader educational goals of the country. The policy foci also include intricate integration of all levels of education (primary,

secondary, and tertiary), updating curriculum, extended use of ICT in all educational initiatives, improvement of teacher education, and unity among all stakeholders. The cornerstone of the policy is to ensure the quality enhancement of education at all levels. Most importantly, the Education Policy 2010 has put particular emphasis on educational assessment. It explicitly states that it is necessary to implement appropriate and effective examination systems or assessment methods, and it also mentions that successful assessment procedures should be embedded into the instructional methods.

Status of English in Bangladesh: a rollercoaster ride

Now English has emerged as the most influential language as a means of human capital development across the world; in fact, effective exposure to the global market seems to be impossible without minimum command of the English language (Chowdhury & Kabir, 2014). Popular in the contemporary industry discourse with regard to the twenty-first-century employability, communicative competence in English constitutes the main ingredient of the recipe. This kind of attitude to English is prevalent in other parts of the world as well. For example, Nunan (2003) reports the case of China where English was made a compulsory subject by government regulation at grade 3, and English was also chosen as the medium of instruction to teach business and engineering subjects at the colleges and universities. China is now yielding the benefits by ensuring its massive exposure to the entire world for trade and other purposes through English. Reference to the case of India may be relevant here too. During the British colonial expansion, and subsequently, colonial settlement, English permeated into the communication channel of the colonized Indian locale (Rahman et al., 2019). It began as a *lingua franca* which got gradually embedded into the social, political, economic, cultural, and esthetic domains of the colonized. Despite being linguistically one of the most diverse regions of the world, India could not resist the penetration of the English language into the socio-psychological behaviors of its people. This colonial hangover still persists, and India is yielding the benefits.

However, our country is yet in dilemma about appropriately positioning the status of English in our language policy. Failure to determine suitable status of English is affecting the people in general and the study of our students in particular. Rahman and Pandian (2018) find that a host of socio-political factors influence the English language education policy and planning in Bangladesh. Obaidul Hamid (2010) attributed the deficient English language teaching (ELT) to the erratic language policy and planning.

With a view to enhancing students' competence in a foreign language, English as an academic subject was recognized by the English Teaching Task Force commissioned by the Ministry of Education in 1976. The Commission explicitly recommended to introduce English in either class 3 or class 6, subject to the availability of teachers (Ministry of Education, 1976). However, English was first integrated in the curriculum as a compulsory subject in 1992 from class 1 when a competency-based curriculum was introduced at the primary level in that year (Ahmed, 2005). Hamid and Honan (2012) reported that English as a compulsory course was also introduced at the undergraduate level at the universities in Bangladesh in 2000. The National Education Commission of 2003 reviewed the status of English given in the past Education Commissions and was convinced that students at the primary level need exposure to this extremely necessary foreign language; therefore, the Commission declared that it should be integrated in

the curriculum (Chowdhury & Kabir, 2014). In addition, it stressed the importance of training of the trainers, from Primary Teachers Institute (PTI) and National Academy for Primary Education (NAPE), abroad. Finally, realizing the necessity as well as the presence of English in our day to day affairs, the government of Bangladesh integrated English as an academic subject in the curriculum at all levels of education in the latest National Education Policy published in 2010 (Rahman & Pandian, 2018). We believe the pre-university textbooks are aptly named “English for Today” as it is deemed that our every today is unimaginable without some use of English; in fact, English has entered into our natural communication in a spontaneous manner.

Principles of assessment in the NCTB curricula: the cornerstones of meaningful assessment

Curriculum, instruction, and assessment are intricately connected to each other (Pellegrino, 2010). The Primary English Curriculum 2012, a thoroughly revised one based on the National Education Policy, was introduced for the learners to acquire competence to communicate successfully at the local and global levels. The curriculum stressed importance on all four language skills: listening, speaking, reading, and writing. However, it is regrettable that the Primary English Curriculum has not integrated assessment policies into it. Without explicit assessment guidelines, the instructors at this level will naturally adopt any assessment procedures that they find convenient to them at the cost of students’ effective learning which is facilitated by assessment (Shepard, 2000). However, the National Curriculum and Textbook Board (NCTB), a unit of the Ministry of Education (MoE) in Bangladesh responsible for publishing and distributing textbooks, teachers’ guides, and other relevant materials, published a Teachers’ Guide in 2016 that incorporated a brief section on assessment which stressed continuous assessment (NCTB, 2016). The students at the end of grade 5 appear in the first public examination known as Primary Education Completion Examinations (PECE). The typical test items include matching, true/false, short answer questions, short composition, fill in the blanks with the given clues, letter writing, and rearranging words to make sentences, and fill in forms with information (Dainik Shiksha, 2020). These test items categorically demonstrate the traditional approach to assessment focusing on discrete test items. The test barely includes communicative tasks that can assess students’ ability to produce discourse. Because of the summative nature of this test, we believe that it is impossible to test a representative sample of students’ English language abilities. Moreover, the test severely suffers from validity since listening and speaking, two important English language skills are scrapped from the test although the curriculum has integrated all four language skills.

However, the junior secondary (grades 6-8) English Curriculum incorporates both formative and summative assessment methods (National Curriculum, 2012). The curriculum describes the formative assessment also as continuous assessment (CA) and finds that CA has tremendous potential to help students yield numerous benefits. For example, CA helps learners diagnose their strengths and weaknesses. Moreover, CA facilitates instructors to assess the students in a cost-effective manner at the minimum investment of time. Most importantly, CA helps instructors identify the psycho-social impacts of the assessment on the learners, and subsequently, the teachers may devise

aid strategies to help students overcome the negative side effects of assessment. The curriculum also states that through the CA methods, instructors can elicit information regarding the effectiveness of their instruction, and when flaws are detected, they can adopt ways to modify instruction. The summative assessment at this level involves half-yearly and final examinations while the students at the end of grade 8 sit for another public examination called Junior School Certificate (JSC). Although the curriculum integrates all four skills, the English test in the JSC Examination excludes listening and speaking tests; the test includes only reading and writing proficiency along with tests items on grammar and vocabulary.

The students of grades 9 and 10 are given the next level of the National English Curriculum which culminates in the most popular public examination known as Secondary School Certificate (SSC). This curriculum is implemented through two sets of contents labeled as paper 1 and paper 2. Paper 1 contents, such as learning outcomes, functions, and language points are delivered to the students through a textbook known as English for Today (Eft) which is developed based on various themes, such as home and family, our neighbors, pastimes and hobbies, adolescence, travel, and tourism while the contents of paper 2 entail grammar and composition skills (National Curriculum, 2012). The assessment methods stated in the curriculum include classwork, continuous assessment, terminal tests, and the public examinations. The curriculum also states the distribution of marks which demonstrates an unequal emphasis. Marks allocated for listening and speaking are 20 (10 marks for each skill) whereas 80 marks (40 for each skill) have been allocated for reading and writing skills. Listening and speaking skills, however, have been scrapped from the SSC.

The higher secondary English curriculum focuses on teaching and learning of English as a skill-based subject so that the learners develop competence in the language to successfully communicate in the real-life situations (National Curriculum, 2012). Consequently, learner-centered approaches to develop students' communicative competence were prioritized in the higher secondary English curriculum, and initiatives were taken to integrate all four English language skills: listening, speaking, reading, and writing as well as the other components of language: grammar and vocabulary (National Curriculum, 2012). This curriculum is also delivered to the students through two sets of contents labeled as paper 1 and paper 2. The contents of paper 1 have been developed as a textbook known as English for Today (Eft) which is based on a range of themes, such as human achievements in science and technology, myths and literature, traffic education, human rights, peace, and conflict. The contents of paper 2 are grammar and composition skills. The English assessment scheme at the higher secondary level includes continuous assessment in the classroom, internal examinations, and the public examinations known as Higher Secondary Certificate (HSC) (Ali et al., 2018).

Assessment and feedback at the tertiary level in Bangladesh are "parts and parcels of English language teaching" (Hasan, 2020, p.215). Assessment in the English foundation courses at the tertiary public sector has also been imprisoned in a particular structure (Hasan, 2020). There is also a fixed pattern of assessment of learners' proficiency in English like the pre-university practices. The practice, on the contrary, is slightly different in the private universities. As English language instructors in two pioneering private universities, we would say that English language assessment in the private universities in Bangladesh has largely been liberated and the instructors here enjoy greater

autonomy. Therefore, some formative classroom assessment methods are used by the teachers in private universities.

Reforms in English language assessment: the unsettled dilemmas

To help students acquire knowledge and develop skills, a new style of assessment known as School Base Assessment (SBA) was introduced in Bangladesh in 2005 from grades 6 to 10 with a view to reducing the sole dependence on high-stakes examination known as SSC (Bangladesh Examination Development Unit, 2016, p. 2). The assessment methods integrated into the SBA policy included class tests, class work, homework, assignments, oral presentations, and group work (Begum & Farooqi, 2008). However, it only continued till 2012 because NCTB (2012) reported that SBA was not successfully implemented. The report says that only 7% of the schools implemented SBA. The report further says: "The failure of the implementation of SBA is accounted for extra load of teachers, shortage of trained teachers, complex type recording system of SBA results, and lack of close monitoring and mentoring system" (Bangladesh Examination Development Unit, 2016, p. 2). When the new secondary curriculum appeared in 2012, another new style of assessment known as continuous assessment (CA) which was suggested in the curriculum was introduced in the secondary education of Bangladesh in 2013. The goal of CA is to treat assessment as an integral part of learning. Class tests, class works or practical works, and homework as assessment methods were recommended in the curriculum. Even CA implementation has raised questions since it is being implemented without training the teachers, and without developing teachers' manual which has been causing the lack of uniformity in application (Bangladesh Examination Development Unit, 2016). However, no such initiative was taken for the higher secondary level.

The report entitled "Status of Implementation of Continuous Assessment in the Secondary School in Bangladesh" presents mixed outcomes of this policy (Bangladesh Examination Development Unit, 2016). The study collected data from head teachers, other teachers, and students. Of head teachers, 51.7% are not clearly familiar with CA let alone their effort to encourage other teachers to implement it although they are the school level leaders, supervisors, and mentors (Bangladesh Examination Development Unit, 2016). The study says that the head teachers were not given training on CA implementation. However, the interesting fact is that according to the head teachers, 90% of students like CA while the teachers, in general, find CA a "nuisance" (Bangladesh Examination Development Unit, 2016, p. 16).

Of teachers who are the main actors for CA implementation, only 6.7% have clear understanding of CA while 50.3% partially understand what CA is, and CA is alien to 41.3% of teachers. The survey on teachers also demonstrates that students like CA practices. The teachers report that the students actively participate in the course works, take the class tests enthusiastically, attend classes regularly, and complete homework regularly. The main reasons why students favor CA, as reported by the teachers, include effectiveness of CA in improving students' learning; leverage of CA in obtaining good marks in the public examinations; coverage of curriculum facilitated by CA; more concentration on study throughout the year because of CA; and more motivation caused by CA. The teachers also shared that doing practical work, group presentation, peer work, class assignment, and oral questioning methods are used for continuous

assessment. The teachers also took the socio-psychological dimensions of the students into cognizance while conducting the CA. Punctuality, patriotism, leadership, honesty, discipline, perseverance, and cooperation to each other constituted the socio-psychological dimensions. The challenges the head teachers and the other teachers shared with regard to the implementation of CA include lack of classrooms; unwillingness of teachers; shortage of classroom teachers in the schools; high teaching load of the teachers; large class size; financial constraints; lack of time to check home works and tests; lack of CA manual; lack of monitoring; and lack of CA training.

With the exclusion of listening and speaking skills, partial assessment reform from grammar-translation oriented tests to skill-focused assessment occurred in English language assessment in Bangladesh ever since the communicative language teaching (CLT) approach was introduced in 1996 (Al Amin & Greenwood, 2018a, 2018b; Ali et al., 2018; Das, Shaheen, Shrestha, Rahman, & Khan, 2014; National Curriculum, 2012; Rahman & Pandian, 2018). Despite some remarkable achievements in the education section in Bangladesh, such as enhanced school enrolment and improved gender equity, required attention to the assessment methodology is still missing here (Ali et al., 2018). In fact, allocation of adequate fund in education and investing on educational research is secondary in this low-income country since poverty alleviation, and ensuring health and security have been prioritized in the agenda for development by the country (Ali et al., 2018). However, studies report that assessment reform is a complex process and it is often difficult to implement (Alderson, 2017; Cheng & Curtis, 2010). Although some studies on various dimensions of examination systems have been conducted in some Asian countries (Kwon et al., 2017; Qi, 2007; Qian and Cumming, 2017), there is little research in Bangladesh.

Although it was always expected that changes in the curricula will lead to the development and practice of reformed assessment, actual practices portray the reinvention of the wheels (Das et al., 2014). One impediment perceived as the potential barrier to assessment reform in Bangladesh is the lack of teachers' assessment literacy (Das et al., 2014). The crucial necessity of assessment literacy in bringing about meaningful improvement in learners' English learning effort is elucidated in another section of this article.

Another significant barrier to the assessment reform in Bangladesh is keeping the teachers, one of the principal stakeholders, away from the assessment policy formulation. A considerable amount of research investigated the capabilities of teachers in planning and implementing superior assessment methods, interpreting assessment evidence if the methods are yielding desired outcomes, and exploring ways how students can be engaged in assessment schemes (Looney, Cumming, van Der Kleij, & Harris, 2018). Moreover, the teachers have the lived experiences of judging the complexity of assessment tasks (Pryor & Crossouard, 2008, 2010). Nevertheless, the voice of the teachers who are the main players in implementing curriculum and conducting assessments are never heard in Bangladesh.

A noteworthy point, however, is that if it is SBA or CA or even the public examinations, none, to a large extent, are consistent with the assessment methods embedded into the CLT approach (Das et al., 2014). Therefore, we argue that the policymakers did not give adequate attention to the alignment between the assessment methods and the core competencies integrated into the curriculum. In fact, each reform effort lacks

vision as the efforts seem to be influenced by the external agencies, and the stakeholders involved in the assessment reforms were unable to consider the realities of the local context.

Current assessment practices: gap between intentions and implementations

The overall measurement approach in Bangladesh is test oriented since examinations are ritualistically implemented here (Al Amin & Greenwood, 2018a, 2018b; Ali et al., 2018). The few formative assessment practices reported in the literature are functionally testing (Das et al., 2014). We contend that assessment with its true purpose of facilitating learning by influencing pedagogical methodology, and materials used in teaching/learning are hard to implement in Bangladesh for various reasons which are discussed and explained below.

Ineffective assessment methods focusing mainly on the rote learning are prevailing in the testing and assessment culture in Bangladesh English language teaching and learning (National Education Policy, 2010). The Advisory Committee for the Development of English under the MoE (see Ministry of Education, 2010). also reported that the existing assessment methods are not aligned with the curriculum goals and objectives. Therefore, they risk the validity of assessment, an essential quality criterion of assessment. The validity issue has also been addressed by the researchers and educators since they find gap between what is “intended to be taught and what is measured” (Das et al., 2014, p. 330). Therefore, the assessment practices are unable to examine if the learners are genuinely acquiring communicative competence, a goal of learning English set in the curricula, to function effectively in the real-life contexts (Das et al., 2014).

Bangladesh adopted communicative language teaching (CLT) as a language teaching methodology at the primary and secondary level in 1996 (National Curriculum, 2012). It was expected that the assessment methods embedded into the CLT approach will be used by individuals and institutions. However, CLT assessment procedures had been facing resistance since the inception of CLT (Quader, 2001). Das et al. (2014) echo what Quader identified a decade and a half ago, and share that the current assessment practices are predominantly paper- and pencil-based high-stakes tests. These high-stakes large scale tests are developed on reduced English syllabi (Al Amin & Greenwood, 2018a, 2018b; Ali et al., 2018; Sultana, 2018), and they harm rather than help the students in learning the English language (Cheng, 1998; Ferman, 2004; Saif, 2006; Shohamy, 2014; Spann & Kaufman, 2015). Worst of all, such tests promote memorization of language structures, and premeditated pieces of language since the students are desperate to pass the tests (Rahman, Kabir, & Afroze, 2006). Since assessment impacts on the students’ approaches to learning, this practice of memorization is keeping students away from actual learning of English; rather, the tests are forcing the students to prepare themselves for the tests.

High-stakes tests exert negative impacts on learners and teachers (Cheng, 1998; Ferman, 2004; Saif, 2006; Shohamy, 2014; Spann & Kaufman, 2015). Smith and Rottenberg (1991) identified six negative effects of high-stakes tests: (1) reduced time for instruction; (2) minimizing curriculum during test construction; (3) assessing lower-order skills; (4) limiting students’ learning opportunities; (5) affecting teacher morale, and (6) torturing students. Many teachers have also informed the negative effects of the high-stakes test on students. According to Elliott (2000), teachers are concerned about the emotional effects

such as increased stress and anxiety which are caused by the high-stakes tests. The pressure can be especially difficult for low-performing students who might already have poor self-concepts and self-esteem. Gordon and Reese (1997) found: “Many of the teachers lamented that they had worked hard to build up at-risk students’ self-concepts and help them to achieve some measure of academic success, only to have the students’ progress wiped out by the [test] failure” (p. 357).

Measurement specialists have identified several weaknesses of the single high-stakes tests. One single high-stakes test is unable to help students, educators, or schools in making important educational decisions (Zernike, 2015). The students are being treated with a one-size-fits-all education approach which is causing unreliable test scores (Popham, 2015). Assessment experts oppose high-stakes testing because using a single indicator of competence to make important decisions about individuals or schools violates the professional standards of the measurement community (AERA, 2000). Other critics are concerned that the unintended effects of high-stakes testing lead to “perverse” (Ryan, 2004) and “corrupt” educational practices (Jones, Jones, & Hargrove, 2003; Nichols & Berlinger, 2005) while Pedulla et al. (2003) argue that the pressure of doing well on a test seriously compromises instructional practice. Where high-stakes test is in practice, the teachers put considerable importance on the contents that are tested, and the students’ achievement in the maximum test score because the test results are associated with the national test results (Harlen, 2007). As a result, most of the teaching time is devoted to preparing the students for the test or doing the testing. If the teacher spends a lot of school time preparing students for tests, the quality of teaching decreases.

Alarming is the fact these nationwide used English high-stakes tests exclude two essential language skills from the tests: listening and speaking (Al Amin & Greenwood, 2018a, 2018b; Das et al., 2014; Khan, 2010; Podder, 2013; Sultana, 2018). Even worse and unethical is the case that stakeholders concerned force teachers to cover only that part of the curriculum that will presumably be considered for the test construction in the public examinations (Choudhury, 2010; Maniruzzaman & Hoque, 2010).

At present, continuous assessment, internal examinations, and public examinations are used to assess students in Bangladesh at the primary and secondary levels (Ali et al., 2018; NCTB, 2020; Sultana, 2018). Continuous assessment using tests and homework is held in the classroom during the lessons while the individual institutions conduct the internal examinations and the public examinations are administered by the educational boards mentioned above (Ali et al., 2018). Although literature on assessment in the English foundation courses is not available, we as the instructors of this level report that assessment methods at this level include both summative and formative procedures which are, in fact, blended with the eventual summative purpose which is grading. Common assessment procedures include quizzes, tests, presentations, assignments, and interviews.

Amin and Greenwood et al. (2018) maintain that the current national examination system is not consistent with the curriculum goals and objectives. They believe that substantial change in students’ learning is just a daydream without establishing alignment between the assessment procedure and curriculum goals and objectives. They further comment that overhauling the assessment system is necessary if Bangladesh aims to produce effective communicators for the real-life situations, and the education

system should include not only the listening and speaking in the assessment but also opportunities for spontaneous, creative, and contextually appropriate use of language, rather than the rote learning that currently predominates.

Because of the test-oriented culture, feedback, mostly associated with formative assessment, has hardly received any adequate attention by researchers in Bangladesh. Feedback refers to the information regarding the gap between students' performance in the assessment task and the intended learning outcomes (Black & Wiliam, 1998). Need-based detailed feedback helps learners minimize their lack of content understanding and also boost up their motivation in learning (Islam & Bt Stapa, 2019). Hyland (2006) maintains that teachers may enjoy the rare opportunity of exercising significant task in the classroom by providing the students with feedback. Providing feedback by correcting students' language errors is a distinctive job of the language teachers, and it is highly expected by the students (Farjadnasab & Khodashenas, 2017). The study by Hasan (2020) on English language assessment at the tertiary level in Bangladesh reports conflicting findings. Students claim that they receive little feedback from their teachers while many teachers claim that they provide the students with feedback. Another study by Rahman, Babu, and Ashrafuzzaman (2011) find that the provision of feedback in the English language classes is irregular.

Washback: a trajectory of backward navigation

Washback in its simplistic causal operation refers to a test's influence on teaching and learning (Alderson & Wall, 1993; Bailey, 1996; Cheng & Curtis, 2004; Hughes, 1989; Saville, 2000) although a couple of other terminologies, such as "backwash" (Hughes, 1989) and "impact" (Wall, 1997) are also interchangeably used (Tsagari, 2007). While washback may affect teaching and learning in both negative and positive manners, negative washback is perceived as the outcome when the test's contents are based on the narrow definition of language ability, thus, constraining the teaching/learning context (Toylor, 2005). However, a test engenders positive washback when the testing procedure provokes effective instructional methods (Toylor, 2005). Considering the entwined link between washback and teaching/learning, it is essential to explore the impacts of washback caused by the English language assessment practices in Bangladesh.

Because of the test-dominant assessment culture in Bangladesh, the high-stakes testing generates dreadful washback effects which affect the English language instructions, and the learners in Bangladesh (Akther, 2016; Sultana, 2018). For one thing, the washback effects influence the learners to prepare for the test (Khan, 2010). This attitude of the learners is a threat to the achievement of the curriculum goals and objectives since the students will focus on only those areas of the curriculum contents that are integrated in the tests. In fact, the past practices of grammar tests are predominantly recycled in the name of communicative testing (Khan, 2010). Another serious problem is that the students will exclusively attempt to score high in the test ignoring the fundamental goal of the curriculum, i.e., acquiring communicative competence in the target language. And, to do so, the learners in Bangladesh get motivated to rote learning and memorization (Rahman et al., 2018a; Rahman, Singh, & Karim, 2018b; Rahman, Singh, & Pandian, 2018c) which eventually results in the surface learning by the students (Hamid et al., 2009). This attitude of the students to earn good grades at any cost drives

them toward alternative avenues, such as getting involved in private tuition for an intensive focus on test preparation, acquiring guidebooks which are written with suggested examination questions and answers, and going to coaching centers that train the students on how to perform well in the public examinations. Negative washback effects of tests are also explored by Kabir (2011) at the tertiary level in Bangladesh; Kabir concludes that students' learning is affected by the negative washback and stresses the importance of teacher training.

Beyond this linear simplistic causal relationship between washback and instruction, washback also has socio-political, economic, and psychological effects (Al Amin & Greenwood, 2018a, 2018b; Ali & Hamid, 2020). Examinations in Bangladesh have powerful impacts on teachers, students, parents, and the overall society (Al Amin & Greenwood, 2018a, 2018b). Perceiving the role of test scores as "gatekeeping mechanisms" (McNamara, 2000, p. 74), the students and their parents become desperate to score high in the English papers. They are also influenced by the social pressure since Bangladeshi society recognizes the achievements of the students in terms of their results in the public examinations (Al Amin & Greenwood, 2018a, 2018b; Hamid, Sussex, & Khan, 2009). From this evidence, we can figure out how much power tests hold in Bangladesh. Coupled with this, the washback effects are also aggravated by the political agenda of the governments since to fulfill their electoral commitment of improving education in the country, demonstrating enhanced passing rate through the standardized tests is a common practice (Ali & Hamid, 2020). This practice has implications in English language teaching in Bangladesh since the English teachers often feel obligated to prepare the students for the test. We argue at this point that if a conscientious pragmatic policy of assessment toward the meaningful learning of communicative competence in English is not adopted and implemented immediately, the negative washback will continue to exert its detrimental effects on English instruction and learning in Bangladesh.

Teacher education and language assessment literacy: a wonky state with a feeble nexus

Teachers' quality of teaching determines the quality of education (Cochran-Smith & Fries, 2005; Goodwyn, 1997; Hagger & McIntyre, 2006). Content knowledge, pedagogical knowledge, curriculum knowledge, pedagogical content knowledge, knowledge of the learners and their characteristics, knowledge of educational contexts, and knowledge of educational ends are all literacies that equip a teacher to be effective (Shulman, 1987). Teachers with deficiency in pedagogical knowledge are unable to deliver the contents well to the learners (Darling-Hammond, 2008; Evertson, Hawley, & Zlotnick, 1985). General pedagogical knowledge of the teachers entails knowledge of classroom management, knowledge of instructional methodology, assessment knowledge, knowledge about how students learn, and the ability to identify students' characteristics (Voss, Kunter, & Baumert, 2011). Kumaravadivelu (2012) went further and underscored the importance of a teacher education system that would enable teachers to form teaching/learning theories based on their classroom experiences and then they would put their theories into practice.

Recognizing the importance of the teacher education for improved teaching performance, the Bangladesh government has set up a number of teachers' training institutes

and centers to facilitate in-service teacher education (Karim, Shahed, Mohamed, Rahman, & Ismail, 2019) while graduate and postgraduate programs in TESOL, TEFL, TESL, and ELT constitute pre-service teacher education (Karim et al., 2019). Besides, various donor-aided training programs, such as English Language Teaching Improvement Project (ELTIP), English for Teaching, Teaching for English (ETTE), Teaching Quality Improvement in Secondary Education Project (TQI-SEP), Secondary Education Quality and Access Enhancement Project (SEQAEP), and English in Action (EiA) have also contributed to the capacity building of the teachers (Hamid, 2010). Although high-cost donor-funded projects claim success with regard to the improvement in teacher education, studies by Al Amin and Greenwood (2018a, 2018b), Anwaruddin (2016), Karim and Mohamed (2019), and Rahman et al. (2018a, 2018b, 2018c) oppose the claim. These studies have identified various challenges such as scanty training sessions, inadequate training opportunities for the rural teachers, lack of trainers, substandard training materials, and insufficient resources (Karim & Mohamed, 2019; Rahman et al., 2018a, 2018b, 2018c). The aggregated momentum of all these factors has precluded the effective implementation of CLT (Rahman et al., 2019) which ultimately affected the communicative English language assessment in Bangladesh.

Precisely, with reference to assessment literacy as a significant teacher education construct, it is regrettable to note that this vital assessment dimension has not been judiciously and professionally integrated into the teacher education programs (Sultana, 2019). Hence, the teachers here lack efficiency in assessing learners' proficiency in English. Although several English language education reform initiatives have been taken in the country, nothing significant has happened in English language assessment (Das et al., 2014). Das et al. (2014) also report that the failure may be attributed to the teachers' struggle to understand and to carry out communicative language assessment which requires advanced level English language proficiency of the teachers. The study by Sultana (2019) has revealed that most of the teachers have not received in-service training on testing or test design. She has also explored that the training programs are not appropriately developed, and the learning module on assessment literacy is almost missing in the programs. As a result, they are not familiar with the alternative classroom assessment procedures, such as portfolio assessment, self- or peer assessment, reflective journals, interview, observation, and conferences (Black & Wiliam, 1998; Brown & Lee, 2015; Sattar, 2006). Sattar (2006) has illustrated the distinctions between the alternative assessment methods and the traditional methods of assessment in his essay. Alternative assessment is learner-centered, formative, feedback-oriented, contextual, developmental, and all-encompassing while the traditional assessment is judgmental (about success, or failure), teacher-centric, anxiety-generating, and hardly motivating (Sattar, 2006).

Qualities of assessment: validity and reliability

Literature on measurement in education stresses validity as a prime principle for quality assessment (Brown & Abeywickrama, 2019). Validity refers to "the extent to which inferences made from assessment results are appropriate, meaningful, and useful in terms of the purpose of the assessment" (Gronlund, 1998, p. 226). Conceiving validity in a more technical manner, Messick (1989), a guru of validity, maintains "an integrated evaluative judgment of the degree to which empirical evidence and theoretical

rationales support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessment” (p. 11). A valid test measures what it is intended to measure, hinges on empirical evidence, integrates learners’ performance that models the test yardsticks, gleans worthwhile information regarding test-takers’ ability, and finally, is backed up by theoretical explanation (Brown & Abeywickrama, 2019).

Only a few studies have examined if the tests given to the students at the various levels of education are valid. The study conducted by Das et al. (2014) found the English tests in the Secondary School Certificate (SSC) examination invalid since they identified gap between what the test is supposed to measure and what it actually measures. The test is supposed to measure the achievement of students’ communicative competence in four English skills (listening, speaking, reading, and writing) as has been integrated in the curriculum, but the tests exclude listening and speaking skills from the assessment scheme. In their study on the validity of the English tests in the Higher Secondary Certificate (HSC) examination, the next tier of Education after SSC, Ali and Sultana (2016) concluded that the tests are not valid. They report and argue that the tests are built on the limited objectives of the curriculum, exclude listening and speaking skills, and are not representative of the breadth and depth of the construct communicative competence, a goal set in the English curricula. It should be noted that this legacy of ignoring validity in test construction has been continuing in Bangladesh for years, and the authorities concerned are least bothered about it. This attitude of indifference among the policymakers toward this fundamental assessment quality may be attributed to the debacle of English language learning in Bangladesh.

Next to validity, measurement experts would be interested to ensure reliability of a test. Reliability refers to the consistency of the test scores or the scores generated by other assessment methods (Miller, Linn, & Gronlund, 2009). In other words, reliability of a test will be ensured if the test generates consistent scores when it is administered from one occasion to another taken by the same candidate, and assessed by different examiners. When the reliability of the English tests of the public examinations in Bangladesh is concerned, they appear to be unreliable. While examining if the English tests in the Secondary School Certificate (SSC) examination are reliable, Billah (2018) identified that they are not reliable because they generated scores in two consecutive years with remarkable variations in one educational board, for instance. The fluctuations of the English test scores are a regular phenomenon (Sultana, 2018). Sultana (2018), and Ali and Hamid (2020) go even further raising doubt about the nature of recruitment of test developers, the level of their expertise, the status of training of the assessors on assessment quality, and so forth. Sultana (2015) in her study on the reliability of the English tests in the HSC examination has identified the lack of reliability. The researchers suggest that the consistency of the lack of reliability may be caused here for various reasons: absence of assessment criteria, lack of assessment literacy, and stakeholders’ indifference to quality assurance, to name a few (Billah, 2018; Sultana, 2015, 2018).

Only a few studies in the field of English language assessment in Bangladesh addressed validity and reliability; therefore, it is hard to present a generalized picture of these two vital constructs in assessment. However, the findings of the researchers discussed above may be considered as the representative scenario of validity and reliability

in Bangladesh since the assessment methods and the test items they have analyzed are essentially the replications of English testing and assessment at other levels of education.

Challenges in English language assessment: hurdles to cross

Numerous challenges have crippled the effective assessment implementation in Bangladesh. First, because of their inadequate understanding of the English curriculum, the teachers and the test developers are unable to confidently select appropriate assessment procedures, which maintain proper alignment with the curriculum goals and objectives (Das et al., 2014). They generally perceive curriculum as a document that contains the instructional contents. However, a curriculum is a set of objectives, courses of study, plans, documents, and experiences (Su, 2012). The breadth of a curriculum encompasses goals and objectives, contents, instructional methodology, and assessment methods. The stakeholders' understanding of these curriculum dimensions and how these dimensions are interconnected is crucial for formulating effective assessment policy.

Second, large class size prevents the teachers to engage the students in the implementation of varied assessment procedures (Ali, 2016). Ali (2016) explored that during the short duration of the class, it is hard to assess a large number of students and to provide them with feedback. In fact, short duration of the class has also been shared by the teachers as a barrier to the implementation of effective assessment. In addition, Ali (2016) identified the traditional set up of the classroom obstructing the teachers to reach the students for giving them individualized feedback.

Third, students' beliefs and attitudes to learning English go against effective assessment methods (Ali, 2016). The general perception of the students with regard to language learning is to acquire grammatical competence which leads them to memorize grammatical structures. As a result, they are not motivated to respond to the alternative assessment procedures. However, we argue that learning a language is not memorizing grammar rules in isolation; rather, learning a language is developing communicative competence in order to produce a stretch of language in a variety of situations. Four components of communicative competence involves (1) linguistic competence, (2) discourse competence, (3) sociolinguistic competence, and (4) strategic competence (Canale & Swain, 1980). When the teachers attempt to engage the students in innovative assessment methods based on communicative competence, the students invariably lose motivation (Ali, 2016).

Another challenge is the recruitment of teachers to teach English who do not have the required academic qualifications. The analysis of the demographic information of the study conducted by Sultana (2018) revealed that eight of ten participants of her study had Bachelor and Master degrees in English literature while one teacher had Bachelor and Master degrees in Economics, and only one teacher had a relevant degree which is MA in English Language Teaching (ELT). It is reasonable to assert that teachers without degrees in ELT will struggle to understand the complex nature of language teaching and learning. Despite having some generic pedagogic principles, every academic field has its own unique set of pedagogic approaches, strategies, and techniques. Applied disciplines, such as applied linguistics and ELT expose the potential practitioners to the discipline-specific theories, and the popular and effective practices.

Without this exposure, it is not only hard for the practitioners to function in the professional world that demands hands-on training facilitated by them but also they may cause damage to the recipients of the training, i.e., the students.

Worst of all, the assessment practices compromise with the social expectations of the parents and students, such as their expectation of high grades (Ali & Hamid, 2020). According to Ali and Hamid (2020), these social expectations influence the schools and the teachers, and therefore, they adopt assessment methods compatible with “teaching to the test” at the cost of “teaching to the students” (Volante, 2004). This kind of compromise poses threat to the ethics in educational assessment, a significant principle that establishes fairness in assessment (Green, Johnson, Kim, & Pope, 2007; Johnson, Green, Kim, & Pope, 2008; Pope, Green, Johnson, & Mitchell, 2009; Ryan, 1997). A framework of ethics in assessment is essential to guide the teachers, and the administrators to ensure equity (Gipps, 1994). It may be noted that any deviation from the ethical practices in assessment will give the students a false notion of self-efficacy (Bandura, 1977), a construct essential for students’ self-regulated growth as an independent learner.

Conclusion and implications

The critical examination and analysis of the various dimensions of English language assessment in this article has promises to improve the knowledge and understanding of the stakeholders concerned. A country with chaotic English language planning and policy (Rahman & Pandian, 2018) has been struggling to settle stable assessment policy consistent with the curriculum goals and objectives. Since assessment shapes learning, the progress of students in learning English has always been affected because of the unsettled dilemmas in the English language assessment policy. There was always a gap between the principles of assessment embedded into the curriculum and the actual assessment practices. Although a few studies could identify some formative assessment practices, the overall assessment practices are summative in Bangladesh.

Heavily hard hit by the high-stakes testing, the curriculum, the learners, and the instructors need to be liberated from this vicious policy. The literature reports that students’ achievement in the high-stakes tests is not what is expected of the curriculum. Assessment and the students’ processes of learning have intricate relationship. For example, the formative dimensions of assessment integrate learners’ processes of learning into consideration when the assessment methods are formulated. Assessment literature produced in Bangladesh has completely ignored this vital factor. While the low achievement in the English courses in general is persistent in Bangladesh, there is research-based evidence in other parts of the world that classroom-based assessment is yielding benefits. However, the success of continuous assessment which is similar to what we describe as classroom assessment is not optimistic. Hence, the continuous assessment should be strengthened by evidence gathered from more empirical studies. Moreover, alternative assessment methods, such as portfolio assessment, self- and peer assessment, and performance-based assessment may also be integrated into the assessment scheme. These assessment methods may respond to some of our challenges, such as large class size, short duration of the class, and the sole dependence on teachers for the provision of feedback. Feedback is generated by the learners themselves when they employ self-assessment while their peers provide them with feedback when peer

assessment is in operation. These feedback practices may take the feedback beyond the four walls of the classroom.

However, one important caveat: it is important to carefully examine the presence of a few alternative assessment methods being used in our assessment system. Researchers (Al Amin & Greenwood, 2018a, 2018b; Sultana, 2018, 2019) have observed that many instructors are imposed to integrate such assessment methods in their instructional designs although they are unsure about how to implement them. The implication is that we are feeding them with these innovative methods without adequately orienting them to the methods and also, without giving them literacy on how to integrate them into the instructional design. More alarming is the fact that we hardly accommodate any channel to listen to their voices, to gather ideas about their reactions to the innovations, to offer scaffolding so that they feel at home, and most importantly, to empower them as critical agents in the entire process of implementing innovations. One way to do that may be to strengthen the teacher education programs.

Teacher education programs are essential to help teachers gather knowledge, skills, professionalism, and expertise. Although teacher educators focused on transmitting knowledge and information to the teachers in the past (Johnson, 2006), the current paradigm for the teacher educators is to perform the role of transformative agents (Kumaravadivelu, 2012) since the ELT field is currently addressing various critical, socio-political, economic and cultural issues (Izadinia, 2012). Hence, the teacher educators will orient the teachers to the notion that teachers will not only help students improve communicative competence in English but also will create platforms for the students where they will use the language to interrogate and to challenge the current state of affairs (Safari & Pourhashemi, 2012). To be in line with this shifting process, the teacher education institutions and centers in Bangladesh should focus on time-to-time curriculum revision since patterns of teaching are constantly changing to cope with the twenty-first-century learning styles shaped by the contemporary sociopolitical, economic, and cultural realities. Without revision and innovation, the curriculum will not be able to train teachers to be effective for the audience coming from the next generations.

Negative washback effects are well documented in the literature as a harmful phenomenon that affects teaching and learning. However, no significant attention has been given by the authority concerned to reverse the consequences of this evil. Washback effects of assessment have been addressed by only a few studies which are limited to the high-stakes examinations. Therefore, more robust empirical studies on washback are required regarding continuous classroom assessment since teachers and students are more frequently affected by the washback during their day-to-day interactions in the English language classroom.

Furthermore, it may be noted that existing studies mostly focused on examining the limitations and flaws of the existing assessment practices, but they did not give enough attention to how a robust assessment culture can be developed. To respond to this, the Ministry of Education may develop an English language assessment manual with detailed discussion on the implementation processes for all levels of education with a view to creating awareness among stakeholders of assessment as a serious educational enterprise. It should be intended to let all think about how assessment operates, and the way students should respond to it. At least, the teachers will be able to make sense of their

current assessment performance and can plan remedial measures if anything goes wrong.

Last but not least, the underlying testing ideologies prevailing in Bangladesh testing culture warrants critical examination. Rahaman (2015) critiques that testing culture in Bangladesh is influenced by three dominant ideologies: (1) tests are coercive; (2) tests are mechanical; and (3) tests are dehumanizing. Tests are coercive since tests are used as political weapons to exert power on the students to learn English in a particular approach while tests are mechanical since the tests in Bangladesh promote utilitarian gains by the students through rewards and punishments. The dehumanizing nature of tests is evident since the test results socio-psychologically affect students and kill their morale. They often fail to provoke the creative energy of the students, an essential element of effective learning.

Abbreviations

CLT: Communicative language teaching; ELT: English language teaching; GTM: Grammar-translation method; SSC: Secondary School Certificate; HSC: Higher Secondary Certificate

Acknowledgements

Not applicable.

Authors' contributions

MSI made substantial contributions to the study conception and design and acquisition (literature review), analysis, and interpretation of data. MKH, SS, AK and MMR have contributed equally in the review of literature and analysis and involved in drafting the manuscript and revising it critically for important intellectual content. All authors read and approved the final manuscript.

Funding

Authors received no funding for the study.

Availability of data and materials

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Independent University Bangladesh, Dhaka, Bangladesh. ²United International University, Dhaka, Bangladesh. ³Goa University, Taleigão, India. ⁴BRAC University, Dhaka, Bangladesh. ⁵Universiti Sains Malaysia, Penang, Pulau Pinang, Malaysia.

Received: 29 July 2020 Accepted: 6 December 2020

Published online: 05 January 2021

References

- AERA (2000). *AERA position statement concerning high-stakes testing in PreK-12 education*. Retrieved April 11, 2002 from <http://aera.net/about/policy/stakes.htm>
- Ahmed, M. (2005). Teaching English in the primary school: Challenges and options. *Bangladesh Education Journal*, 4(1), 17–23.
- Akther, R. (2016). *Washback effects of Secondary School Certificate (SSC) English Paper-I test in Bangladesh (Unpublished Master's thesis)*. Chittagong: International Islamic University Chittagong.
- Al Amin, M., & Greenwood, J. (2018a). The examination system in Bangladesh and its impact: On curriculum, students, teachers and society. *Language Testing in Asia*, 8(1), 1–8 <https://doi.org/10.1186/s40468-018-0060-9>.
- Al Amin, M., & Greenwood, J. (2018b). The UN sustainable development goals and teacher development for effective English teaching in Bangladesh: A gap that needs bridging. *Journal of Teacher Education for Sustainability*, 20(2), 118–138 <https://doi.org/10.2478/jtes-2018-0019>.
- Alam, F. (2018). Revisioning English studies in Bangladesh in the age of globalisation and ELT. In *Engaging in educational research*, (pp. 241–261). Singapore: Springer.
- Alderson, J. C. (2017). *Foreword to the special issue "the common European framework of reference for languages (CEFR) for English language assessment in China" of language testing in Asia*. <https://doi.org/10.1186/s40468-017-0049-9>.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115–129.
- Ali, M. (2016). Assessment of English language skills at the secondary level in Bangladesh. *The Journal of EFL Education and Research*, 1(1), 1–7.
- Ali, M. M., & Hamid, M. O. (2020). Teaching English to the test: Why does negative washback exist within secondary education in Bangladesh? *Language Assessment Quarterly*, 17(2), 129–146 <https://doi.org/10.1080/15434303.2020.1717495>.
- Ali, M. M., Hamid, M. O., & Hardy, I. (2018). Ritualisation of testing: Problematising high-stakes English-language testing in Bangladesh. *Compare*, 00(00), 1–21 <https://doi.org/10.1080/03057925.2018.1535890>.

- Ali, M. M., & Walker, A. L. (2014). 'Bogged down' ELT in Bangladesh: Problems and policy. *English Today*, 30(2), 33–38 <https://doi.org/10.1017/S0266078414000108>.
- Ali, & Sultana (2016). A Study of the Validity of English Language Testing at the Higher Secondary Level in Bangladesh. *International Journal of Applied Linguistics & English Literature*, 5(6).
- Amin, M. A. & Greenwood, J. (2018). *The examination system in Bangladesh and its impact: on curriculum, students, teachers and society*. Language Testing in Asia.
- Anwaruddin, S. M. (2016). ICT and language teacher development in the global south: A new materialist discourse analysis. *Educational Studies*.
- Asian Development Bank [ADB]. (2015). *Policy reform in Bangladesh's secondary education (1993–2013) tracing causal processes and examining ADB's contribution*. Retrieved from: <https://www.adb.org/sites/default/files/evaluationdocument/177777/files/topical-ban-sec-educ.pdf>
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13(3), 257–279.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215.
- Bangladesh Examination Development Unit (2016). *Status of the implementation of continuous assessment in secondary schools in Bangladesh*. Dhaka: Directorate of Secondary and Higher Education.
- Begum, M., & Farooqui, S. (2008). School Based Assessment: Will It Really Change the Education Scenario in Bangladesh? *International Education Studies*, 1(2), 45–53.
- Billah, M. (2018). পাসরে হার ও জুপিপএ ৫-এর চসে পকিষার মান বৃজা বিপার [It is important to increase quality of education than the passing rate]. Banik Barta: Retrieved from <http://bonikbarta.net/bangla/>.
- Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, 5(1), 7–74.
- Brown, H. D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices*, (3rd ed.,). New Jersey: Pearson Education, Inc.
- Brown, H. D., & Lee, H. (2015). *Teaching by principles: An interactive approach to language pedagogy*, (4th ed.,). New York: Pearson.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47.
- Cheng, L. (1998). Impact of a public English examination change on students' perceptions and attitudes toward their English learning. *Studies in Educational Evaluation*, 24(3), 279–301.
- Cheng, L., & Curtis, A. (2004). Washback or backwash: A review of the impact of testing on teaching and learning. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods*, (pp. 3–17). Mahwah: Lawrence Erlbaum Associates, Inc..
- Cheng, L., & Curtis, A. (Eds.) (2010). *English language assessment and the Chinese learner*. New York: Routledge: Taylor's & Francis Group.
- Choudhury, R. K. (2010). *Appropriateness and relevancy of communicative language teaching (CLT) for Bangladesh: A perspective from Bangladeshi rural secondary school English teachers*, (PhD thesis). New York: Columbia University.
- Chowdhury, R., & Kabir, A. H. (2014). Language wars: English education policy and practice in Bangladesh. *Multilingual Education*, 4(1), 1–16 <https://doi.org/10.1186/s13616-014-0021-2>.
- Cochran-Smith, M., & Fries, K. (2005). The AERA panel on research and teacher education: Context and goals. In M. Cochran-Smith, & K. Zeichner (Eds.), *Studying teacher education: The report of the AERA panel on research and teacher education*, (pp. 37–68). Mahwah: Erlbaum.
- Dainik Shiksha. (2020). Prathomik Shomaponi Porikhar Proshno Kathamo o Nombor Banton. 2 June 2020. Retrieved from, www.dainikshiksha.com/প্রাথমিক-সমাপনী-পরীক্ষার-পরশ্ন-কাঠামো-নম্বর-বিস্তার/190156/
- Darling-Hammond, L. (2008). The case for university based teacher education. In M. Cochran-Smith, S. Feiman-Nemser, D. J. McIntyre, & K. E. Demers (Eds.), *Handbook of research on teacher education: Enduring questions and changing contexts*, (pp. 333–346). New York: Routledge.
- Das, S., Shaheen, R., Shrestha, P., Rahman, A., & Khan, R. (2014). Policy versus ground reality: Secondary English language assessment system in Bangladesh. *Curriculum Journal*, 25(3), 326–343 <https://doi.org/10.1080/09585176.2014.909323>.
- Elliott, J. (2000). NCAE: ABCs accountability statement. Retrieved March 3, 2003 from, <http://www.ncae.org/news/000717pr.shtml>
- Everton, C., Hawley, W., & Zlotnick, M. (1985). Making a difference in educational quality through teacher education. *Journal of Teacher Education*, 36(3), 2–12 <https://doi.org/10.1177/002248718503600302>.
- Farjadnasab, A. H., & Khodashenas, M. R. (2017). The effect of indirect written corrective feedback on students' writing accuracy. *International Journal of Research in English Education*. <https://doi.org/10.23971/jevl.v6i2.401>.
- Ferman, I. (2004). The washback of an EFL national oral matriculation. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods*, (pp. 191–210). Mahwah: Lawrence Erlbaum Associates.
- Gipps, C. V. (1994). *Beyond testing: Towards a theory of educational assessment*. Washington, DC: The Falmer Press.
- Goodwyn, A. (1997). *Developing English teachers: The role of mentorship in a reflective profession*. Philadelphia: Open University Press.
- Gordon, S. P., & Reese, M. (1997). High-stakes testing: Worth the price? *Journal of School Leadership*, 7, 345–368.
- Green, S., Johnson, R., Kim, D., & Pope, N. (2007). Ethics in classroom assessment practices: Issues and attitudes. *Teaching and Teacher Education*, 23, 999–1011.
- Gronlund, N. E. (1998). In 4th (Ed.), *Measurement and Evaluation in Teaching*, (p. 96). New York: Macmillan Publishing Co., Inc.
- Hagger, H., & McIntyre, D. (2006). *Learning teaching from teachers: Realizing the potential of school based teacher education*. New York: Open University Press.
- Hamid, M. O., & Baldauf, R. B. J. (2008). Will CLT bail out the bogged down ELT in Bangladesh? *English Today*, 24(3), 289–310 <https://doi.org/10.1017/S0266078408000254>.
- Hamid, M. O., & Honan, E. (2012). Communicative English in the primary classroom: Implications for English-in-education policy and practice in Bangladesh. *Language, Culture and Curriculum*, 25(2), 139–156.

- Hamid, M. O., Jahan, I., & Islam, M. M. (2013). Medium of instruction policies and language practices, ideologies and institutional divides: Voices of teachers and students in a private university in Bangladesh. *Current Issues in Language Planning*, 14(1), 144–163.
- Hamid, M. O., Sussex, R., & Khan, A. (2009). Private tutoring in English for secondary school students in Bangladesh. *Tesol Quarterly*, 43(2), 281–308.
- Harlen, W. (2007). *Assessment of Learning*. London: Sage Publications.
- Hasan, M. M. (2020). Using various assessment and feedback tools for teaching English at the undergraduate level in Bangladesh. *International Journal of Science and Business*, 4(1), 214–227 <https://doi.org/10.5281/zenodo.3703145>.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: CUP.
- Hyland, K. (2006). *English for academic purposes. An advanced resource book*. London: Routledge.
- Islam, M. S. (2019). *Assessment dilemmas in educational institutions*. Dhaka Tribune: Retrieved from <https://www.dhakatribune.com/opinion/op-ed/2019/07/25/assessment-dilemmas-in-educational-institutions>.
- Islam, M. S., & Bt Stapa, M. (2019). Investigating the potentials of classroom assessment: A critical discussion. *International Journal of Engineering and Advanced Technology*, 8(5), 1172–1179 <https://doi.org/10.35940/ijeat.E1165.0585C19>.
- Izadinia, M. (2012). Teacher educators as role models: A qualitative examination of student teachers' and teacher educators' views towards their roles. *The Qualitative Report*, 17(24), 1–15 <http://nsuworks.nova.edu/tqr/vol17/iss24/1/>.
- Johnson, K. E. (2006). The sociocultural turn and its challenges for second language teacher education. *TESOL Quarterly*, 40(1), 235–257 <http://www.jstor.org/stable/40264518>.
- Johnson, R., Green, S., Kim, D., & Pope, N. (2008). Educational leaders' perceptions about ethical assessment practices. *The American Journal of Evaluation*, 29, 520–530.
- Jones, M. G., Jones, B. D., & Hargrove, T. Y. (2003). *The unintended consequences of high-stakes testing*. Lanham: Rowman & Littlefield Publishers.
- Kabir, M. H. (2011). Backwash effects of examinations upon teaching and learning of English. *Horizon*, 4, 121–142.
- Karim, A., & Mohamed, A. R. (2019). Examining the impact of an English in action training program on secondary-school English teachers' classroom practice in Bangladesh examining the impact of an English in action training program on. *Qualitative Report*, 24(3), 441–469.
- Karim, A., Shahed, F. H., Mohamed, A. R., Rahman, M. M., & Ismail, S. A. M. M. (2019). Evaluation of the Teacher Education Programs in EFL Context: A Testimony of Student Teachers' Perspective. *International Journal of Instruction*, 12(1), 127–146.
- Khan, R. (2010). English language assessment in Bangladesh: Developments and challenges. In Y. Moon, & B. Spolsky (Eds.), *Language assessment in Asia: Local, regional or global?* (pp. 121–157). South Korea: Asia TEFL.
- Kumaravadivelu, B. (2012). *Language teacher education for a global society: A modular model for knowing, analysing, recognizing, doing, and seeing*. New York: Routledge.
- Kwon, S. K., Lee, M., & Shin, D. (2017). Educational assessment in the Republic of Korea: lights and shadows of high-stake exam-based education system. *Assessment in Education: Principles, Policy & Practice*, 24(1), 60–77.
- Looney, A., Cumming, J., van Der Kleij, F., & Harris, K. (2018). Reconceptualising the role of teachers as assessors: Teacher assessment identity. *Assessment in Education: Principles, Policy and Practice*, 25(5), 442–467 <https://doi.org/10.1080/0969594X.2016.1268090>.
- Maniruzzaman, M., & Hoque, M. E. (2010). How does washback work on the EFL syllabus and curriculum? A case study at the HSC level in Bangladesh. *Language in India*, 10(12), 49–88.
- McNamara, T. (2000). *Language testing*. Oxford: Oxford University Press.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement*, (3rd ed., pp. 13–104). New York: American Council on education and Macmillan.
- Miller, C. M. I., & Parlett, M. (1974). *Up to the mark: A study of the examination game*. Guildford: Society for Research into Higher Education.
- Miller, M. D., Linn, L. R., & Gronlund, E. N. (2009). *Measurement and assessment in teaching*. New Jersey: Pearson Education, Inc.
- Ministry of Education (1976). *Bangladesh Shiksha Commission Report*. Dhaka: Ministry of Education, People's Republic of Bangladesh.
- Ministry of Education (2010). *National education policy*. Dhaka: Ministry of Education, Government of Bangladesh http://old.moedu.gov.bd/index.php?option=com_content&task=view&id=338&Itemid=416.
- MoE. (2020). *Intermediate and Secondary Education Boards, Bangladesh*. Retrieved from: <http://www.educationboard.gov.bd/>
- National Curriculum (2012). *English: Classes I-XII*. Dhaka: National Curriculum and Textbook Board (NCTB) <http://www.portalnctb.gov.bd/>.
- NCTB (2012). *National Curriculum*. Dhaka: National Curriculum & Textbook Board.
- NCTB (2016). *Teachers' Edition*. Dhaka: National Curriculum & Textbook Board.
- NCTB (2020). *Teachers' Edition*. Dhaka: National Curriculum & Textbook Board Retrieved from: <https://drive.google.com/file/d/0B8L6VJQZe3EEaXNKdkhLZm40bWM/view>.
- Nichols, S., & Berlinger, D.C. (2005). *The inevitable corruption of indicators and educators through high-stakes testing*. *EPSL-0503-101-EPRU*. March 23, 2005. Retrieved from, <http://www.greatlakescenter.org/pdf/EPsl-0503-101-EPRU.pdf>.
- Nunan, D. (2003). The impact of English as a global language on educational policies and practices in the Asia-Pacific Region. *TESOL Quarterly*, 37(4), 589–613.
- Obaidul Hamid, M. (2010). Globalisation, English for everyone and English teacher capacity: Language policy discourses and realities in Bangladesh. *Current Issues in Language Planning*, 11(4), 289–310.
- Pedulla, J. J., Abrams, L. M., Madaus, G. F., Russell, M. K., Ramos, M. A., & Miao, J. (2003). *Perceived effects of state-mandated testing programs on teaching and learning: Findings from a national survey of teachers*. Boston: Boston College, National Board on Educational Testing and Public Policy Retrieved January 7, 2004 from <http://www.bc.edu/research/nbetpp/statements/nbr2.pdf>.
- Pellegrino, J. W. (2010). *The design of an assessment system for the race to the top: A learning sciences perspective on issues of growth and measurement*. Princeton: Educational Testing Service.

- Podder, R. (2013). English aural-oral skills assessment policy and practices in Bangladesh secondary school. In J. Greenwood, J. Everatt, A. H. Kabir, & S. Alam (Eds.), *Research and educational change in Bangladesh*, (pp. 53–68). Dhaka: University of Dhaka.
- Pope, N., Green, S., Johnson, R., & Mitchell, M. (2009). Examining teacher ethical dilemmas in classroom assessment. *Teaching and Teacher Education*, 25, 778–782.
- Popham, J. (2015). Why standardized tests don't measure educational quality. *Education Leadership*, 56, 8–16.
- Pryor, J., & Crossouard, B. (2008). A socio-cultural theorisation of formative assessment. *Oxford Review of Education*, 34, 1–20 <https://doi.org/10.1080/03054980701476386>.
- Pryor, J., & Crossouard, B. (2010). Challenging formative assessment: Disciplinary spaces and identities. *Assessment & Evaluation in Higher Education*, 35, 37–41 <https://doi.org/10.1080/02602930903512891>.
- Qi, L. (2007). Is testing an efficient agent for pedagogical change? Examining the intended washback of the writing task in a high-stakes English test in China. *Assessment in Education*, 14(1), 51–74.
- Qian, D. D., & Cumming, A. (2017). Researching English language assessment in China: Focusing on high-stakes testing. *Language Assessment Quarterly*, 14(2), 97–100.
- Quader, D. A. (2001). Reaction to innovation in a language teaching project. *Journal of the Institute of Modern Languages*, 5, 22.
- Rahaman, A. (2015). *Archaeology of research*. Dhaka: Muktochinta.
- Rahman, A., Kabir, M. M., & Afroze, R. (2006). *Effect of BRAC–PACE training on English language teachers of rural non-government secondary schools*. Dhaka: Research and Evaluation Division, BRAC Retrieved from http://www.bracresearch.org/srch_dtls.php?tid%4417.
- Rahman, F., Babu, R., & Ashrafuzzaman, M. (2011). Assessment and feedback practices in the English language classroom. *Journal of NELTA*, 16, 1–2.
- Rahman, M. M., Islam, M. S., Karim, A., Chowdhury, T. A., Rahman, M. M., Seraj, P. M. I., & Singh, M. K. M. (2019). English language teaching in Bangladesh today: Issues, outcomes and implications. *Language Testing in Asia*, 9(1), 9.
- Rahman, M. M., & Pandian, A. (2018). The Chaotic English language policy and planning in Bangladesh: Areas of apprehension. *Pertanika Journal of Social Sciences and Humanities*, 26(2), 893–908.
- Rahman, M. M., Pandian, A., & Kaur, M. (2018a). Factors affecting teachers' implementation of communicative language teaching curriculum in secondary schools in Bangladesh. *Qualitative Report*, 23(5), 1104–1126.
- Rahman, M. M., Singh, M. K. M., & Karim, A. (2018b). English medium instruction innovation in higher education: evidence from Asian contexts. *Journal of Asia TEFL*. <https://doi.org/10.18823/asiatefl.2018.15.4.20.1156>.
- Rahman, M. M., Singh, M. K. M., & Pandian, A. (2018c). Exploring ESL teacher beliefs and classroom practices of CLT: A case study. *International Journal of Instruction*, 11(1) <https://doi.org/10.12973/iji.2018.11121a>.
- Ryan, A. (1997). Professional obligation: A dimension of how teachers evaluate their students. *Journal of Curriculum and Supervision*, 12, 118–134.
- Ryan, J. E. (2004). The perverse incentives of the No Child Left Behind Act. *New York University Law Review*, 79, 932–989.
- Safari, P., & Pourhashemi, M. R. (2012). Toward an empowering pedagogy: Is there room for critical pedagogy in educational system of Iran? *Theory and Practice in Language Studies*, 2(12), 2548–2555 <https://doi.org/10.4304/tpls.2.12.2548-2555>.
- Saif, S. (2006). Aiming for positive washback: A case study of international teaching assistants. *Language Testing*, 23(1), 1–34.
- Sattar, S. (2006). The alternative modes of assessing students at tertiary level: The ELT perspective. *Stamford Journal of English*, 2, 178–187.
- Saville, N. (2000). Investigating the impact of international language examinations. *Research Notes*, 2, 2–7.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4–14.
- Shohamy, E. (2014). *The power of tests: A critical perspective on the uses of language tests*. London: Routledge.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57, 1–22 <https://doi.org/10.17763/haer.57.1.j463w79r56455411>.
- Smith, M. L., & Rottenberg, C. (1991). Unintended consequences of external testing in elementary schools. *Educational Measurement: Issues and Practice*, 10(4), 7–11.
- Snyder, B. R. (1971). *The hidden curriculum*. Cambridge: MIT Press.
- Spann, P., & Kaufman, D. (2015). *The negative effects of high-stakes testing*. Retrieved from: <http://www.luc.edu/media/lucedu/law/centers/childlaw/childed/pdfs/2015studentpapers/Spann.pdf>
- Su, S. (2012). The various concepts of curriculum and the factors involved in curricula-making. *Journal of Language Teaching and Research*, 3(1), 153–158.
- Sultana, N. (2018). Test review of the English public examination at the secondary level in Bangladesh. *Language Testing in Asia*, 8(1), 1–9 <https://doi.org/10.1186/s40468-018-0068-1>.
- Sultana, N. (2019). Language assessment literacy: An uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia*, 9(1), 1.
- Sultana, R. (2015). Reliability of the currently administered language tests in Bangladesh: A case study. *Journal of Literature, Languages and Linguistics*, 17, 76–85.
- Taylor, L. (2005). Washback and impact. *ELT Journal*, 59(2), 154–155 <https://doi.org/10.1093/eltj/ccj030>.
- Tsagari, D. (2007). *Review of washback in language testing: What has been done? What more needs doing?*. Online submission. Available online at <https://files.eric.ed.gov/fulltext/ED497709.pdf>. Accessed 1/06/2020.
- Volante, L. (2004). Teaching to the test: What every educator and policy-maker should know. *Canadian Journal of Educational Administration and Policy*, 35 Retrieved from <http://umanitoba.ca/publications/cjeap/articles/volante.html>.
- Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology*, 103(4), 952–969 <https://doi.org/10.1037/a0025125>.
- Wall, D. (1997). Impact and washback in language testing. In C. Clapham, & D. Corson (Eds.), *Encyclopedia of language and education: Vol. 7. Language testing and assessment*, (pp. 291–302). Dordrecht: Kluwer Academic.
- Zernike, K. (2015). *Obama administration calls for limits on testing in schools*. New York Times Retrieved from <http://www.nytimes.com/2015/10/25/us/obama-administration-calls-for-limits-on-testing-in-schools.html>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.