



A Journal of Culture, English Language, Teaching & Literature

ISSN 1414-3320 (Print), ISSN 2502-4914 (Online)

Vol. 21 No.1; June 2021

Copyright © Soegijapranata Catholic University, Indonesia

A Systematic Literature Review on the Impact of EFL Classroom-Based Language Assessment

¹Ima Fitriyah and ²Utami Widiati

¹IAIN Kediri, Kediri, East Java, Indonesia

²English Language Education, Faculty of Letters, Universitas Negeri Malang,
Malang, Indonesia

email: ¹imafitria@iainkediri.ac.id, ²utami.widiati.fs@um.ac.id

Received: 22-01-2021

Accepted: 22-06-2021

Published: 29-06-2021

A Systematic Literature Review on the Impact of EFL Classroom-Based Language Assessment

¹Ima Fitriyah and ²Utami Widiati

¹imafitria@iainkediri.ac.id, ²utami.widiati.fs@um.ac.id

¹IAIN Kediri, Kediri, East Java, Indonesia

²English Language Education, Faculty of Letters,
Universitas Negeri Malang, Malang, Indonesia

Abstract: The trend in investigating how teachers apply assessment in EFL classroom has been at the momentum. Using systematic literature review and PRISMA protocol, as many as 44 research articles during this past ten years were found. Their findings are beneficial for EFL learning. Research investigating students' perceptions of Classroom-Based Language Assessment (CBLA) shows that students tend to have positive perceptions of CBLA. On another topic, students' motivation to implement CBLA depends on how the test itself is carried out. Few studies states that CBLA contributes greatly to students' achievement, what is there is that they learn a lot from CBLA, and they sufficiently understand CBLA's function for learning. From the EFL teachers' perspective, assessment literacy is crucial for them because at least 75% of research is found to focus on the application of assessment literacy, and the practice of assessment is one-way and balanced. This article discusses on how researchers have been using CBLA benefited for EFL teaching and learning.

Key words: literature review, classroom-based language assessment, assessment for learning

Abstrak: Tren penelitian yang berfokus pada bagaimana guru menerapkan penilaian bahasa Inggris di kelas semakin tinggi. Ditemukan sebanyak 44 artikel penelitian selama sepuluh tahun terakhir. Temuan-temuan ini tentu bermanfaat untuk peningkatan pembelajaran bahasa Inggris. Penelitian yang menyelidiki persepsi siswa tentang

Penilaian Bahasa Berbasis Kelas (CBLA) menunjukkan bahwa siswa cenderung memiliki persepsi positif tentang CBLA. Pada topik lain disebutkan bahwa motivasi siswa dalam melaksanakan CBLA sangat bergantung pada cara pelaksanaan ujian itu sendiri. Beberapa penelitian menyatakan bahwa CBLA memberikan kontribusi yang besar terhadap prestasi siswa, mereka banyak belajar dari CBLA, dan mereka cukup memahami fungsi CBLA untuk pembelajaran. Dari sisi guru bahasa Inggris di perguruan tinggi, literasi penilaian sangat penting untuk dimiliki oleh guru. Setidaknya 75% penelitian berfokus pada penerapan literasi penilaian, dan praktik penilaian bersifat satu arah dan seimbang. Ringkasnya, artikel ini mendiskusikan bagaimana penelitian-penelitian tersebut menggunakan CBLA untuk kepentingan pengajaran dan pembelajaran Bahasa Inggris.

Kata kunci: tinjauan pustaka, penilaian bahasa berbasis kelas, penilaian untuk pembelajaran

INTRODUCTION

Classroom-based assessment has played a critical role in the pedagogical practices and learning processes of both ESL and EFL educators and learners. When evaluating classroom-based assessment, the link between teaching, learning, and the actual assessment is critical. Teachers aim to collect data that could guide their decision-making in order to integrate a variety of assessment methods at crucial points during the class (Perrone & College, 2011). As assessment methods play a critical role in how language points are understood by students and their own competencies are strengthened, these aspects of the classroom have been referred to as testable moments. In short, the discussion on the effect of language assessments and test have long been of wide interest in general education due to its causal relationship with teaching and learning.

This relationship is clearly seen in the term of washback. Washback relates to the impact of the assessments on instruction in terms of how learners train for the exam in a large-scale evaluation. The information that "washes back" to students in the form of helpful diagnoses of strengths and deficiencies is another form of washback that happens mostly in classroom evaluation. According to Brown (2004), washback also includes the effect of an assessment on teaching and preparing prior to the evaluation itself, that is, on evaluation readiness (p.29). Washback can also have beneficial and detrimental effects (Ali & Hamid, 2020). Positive washback refers to test outcomes expected. For example, an exam may help students to learn more or

may promote a link between expectations and teaching. Negative washback refers to the unintended, negative consequences of an examination. For instance, at the detriment of other tasks, teaching can rely too heavily on test preparation. Thus, teachers should minimize the existence of negative washback in order to gain more positive one.

Educational preparation that connects teaching and research is one way to guarantee meaningful washback. Teachers will more effectively match testing with curriculum by choosing a measure that represents the instructional and program objectives. However, empirical evidence of the phenomenon only began to flourish in language education in the 1990s, especially after Alderson & Wall (1993) posed their famous question in their article's title, "Does washback exist?". Since then, a large number of studies in language education have been done to seek empirical proof for the common assumption that assessments have influence on teaching and learning. This phenomenon is acknowledged by many studies on it recently; i.e., Cheng (2005); Wall (2000); Lam (2016); Perrone & College (2011); Stoyloff (2012). Moreover, the literatures has also focused on the impact of assessment on teaching and learning, such as Barnes (2017); Furaidah, Ali & Utami (2015); Galikyan et al. (2019); Johnson & Shaw (2019); McKinley & Thompson (2018); Xu & Liu (2018).

On the other side, language assessment has also gained popularity in online education as a result of the advancement of increasingly complex learning technology. Further, since the assessment should be conducted online, the effect of CBLA in online learning is also a significant study topic (Ali, 2015; Abbad et al., 2009; Cai, 2012; Daniels et al., 2019; Doculan, 2016; Wagner, 2020; Wihastyanang, 2018; Mohamadi, 2018). And thus, the effect may occur differently between online tests and non-online assessment. Regardless of the context on how language assessment is carried out, the objective of the assessment can be identified on several criteria known as an assessment for learning (AfL), assessment of learning (AoL), and assessment as learning (AaL) (Gan et al., 2017; Lam, 2016; Lee, 2007). Studies confirm that a good assessment makes both students and teachers better even though many believe that tests are carried out to see student learning outcomes (AoL). Therefore, it is interesting when research on the effects of language assessment can be narrowed down to look at how the impact of the evaluation is on improving the quality of teaching and learning (Ridhwan, 2017).

AfL in formative assessments, distinguished from summative assessment (AoL) (Hill & McNamara, 2011; Ridhwan, 2017). Learning is an aim

throughout the AfL, and assessment is the means of completing the objective (Lee, 2007). Teachers need to combine instruction, studying, and assessment in carrying out AfL rather than abiding on how to measure the success of students themselves. The introduction of AfL continues with an appreciation of the relationship between teaching and learning, and how assessments can be utilized to advise teaching and learning. As a result, scholars are now paying specific attention to effective teaching and learning preparation, AfL as fundamental to classroom experience, and what students do after the assessment (Lee, 2007; Liu & Xu, 2017).

Relatedly, some systematic literature reviews have been conducted in the area of language assessment, washback and AfL. This study would fill up what Cheng's et al., (2015) suggestion. Cheng et al. (2015) analyzed the literatures on washback effect of high-stake interest tests during 1993 to 2013, For future research, the synthesis of this analysis of empirical washback studies offers the following guidelines: (1) stakeholder inclusion, (2) interplay mechanism, and (3) perspectives from relevant fields and hypotheses. Nimehchisalem & Bhatti, (2019) constructed literature review on empirical studies in the teachers' assessment literacy. A distinct review somehow combines the importance of technology and assessment in the past 20 years (Chapelle & Voss, 2016), and a systematic review on the method for evaluating rating quality in language assessment (Wind & Peterson, 2018). Schildkamp et al., (2020) reviewed 54 studies on classroom formative assessment. The results show that (1) knowledge and skills (2), psychological factors and (3) social factors influence the use of formative assessment.

In addition to above mentioned studies, Ha (2019) examined the effect of assessment on students' motivation, behavior, and accomplishment in the classroom using a variety of research methodologies; including surveys, interviews, case studies, and experiments from various journals. The review demonstrates that the number of studies is still limited, and the findings are still inconclusive. In contrast, where Cheng (2015) took only the articles from highly reputable journal, Ha (2019) included articles from many sources, and it was not mentioned clearly the number of the articles being reviewed.

The systematic literature review on language assessment effect in language test proven the importance of the field in language assessment especially in the classroom learning process. As impact of classroom-based language assessment (CBLA) is context-dependent, Ha (2019) suggested future study should be undertaken in distinct contexts in order to completely understand how assessment operates and produces beneficial effects on

learning and teaching (AFL). Thus, reviewing the impact of CBLA and looking at how it contributes to the way teaching conducted and how students learn would fill this demand. In addition, Cheng et al. (2015) and Ha (2019) contended that findings on the implications and effect of classroom-based teacher-led formative tests are only beginning to appear. Based on the previous study, implication that existing assessment and review methods are largely large-scale and high-stakes in nature. Thus, future studies are proposed to continue to use testing to drive curriculum, and can thoroughly understand how language assessment works and produces beneficial results on learning.

While the impact of classroom assessment on student performance has been recognized (Perrone & College, 2011; Stoyhoff, 2012; Wach, 2012), few comprehensive literature review studies have specifically explored how classroom evaluations could be used to support the students' learning attitude, comprehension of language points, and teaching practice. Addressing this obvious constraint, the present study investigated the effect of a classroom assessment on students and teachers. This systematic review would give attention to the results of empirical studies in investigating the effect of CBLA as the continuation of the previous review; on student learning attitude and motivation, teachers' assessment practice in higher education context. Finally, it focuses whether most of CBLA considers both AFL and AoL. CBLA is chosen in this review because it is the most used term in the context of language assessment.

SYSTEMATIC REVIEW FRAMEWORK AND METHODOLOGY

This study is largely focused on an analysis of CBLA in the past ten years; start from 2011 to 2020. The review issue is formulated by referring to PRISMA framework to formulate eligibility criteria in systematic reviews model. Considering the model, the formulated systematic analysis question is posed here: "How are the current language assessment practices in terms of CBLA impact on EFL teaching and learning in higher education?".

The elements of the query were generated into key literature search terms and the appropriate inclusion and exclusion criteria for the selection of articles. Therefore, the trends and patterns of research studies on CBLA in EFL higher education can be taken from the findings of this analysis in order to respond to the inquiry.

A. Procedures

The purpose of this review was to determine the existing approaches of language assessment in terms of the effect of CBLA on EFL higher education teaching and learning. The research used the methods described by Petticrew and Roberts to carry out systematic social science literature reviews (2006). Following the formulation of our study questionnaire, the writers defined the key phrases, selected literature sources, and started looking for publications. First, the writer proposed inclusion requirements that formed the basis for the selection of pertinent publications. In order to compare the same information units from each chosen publication, a meticulously constructed data extraction form was used to scan through all of the accessible sources. Finally, to meet the study's requirements, the findings from publications that had been assessed to be of reasonable quality were combined.

B. Databases and Search Terms

Five databases, namely from the Education Resources Information Centre [ERIC], Web of Science, Scopus, PsychINFO and Picarta used the same search words. Initial search words included 'EFL assessment' and 'language classroom assessment impact/effect' as found in a thesaurus and/or terminology used in other similar publications, 'assessment for learning' and identical terms. The search string was added to the term 'higher education' and formative assessment to narrow down the findings of CBLA-related publications. The word "classroom" and related words were added to the search string to limit the search. All the retrieved articles were exported to Excel, where the inclusion criteria were used to make a systematic selection.

C. Inclusion Criteria and Data Extraction

The writers formulated the following inclusion criteria to arrive at a valid set of publications: 1) The articles were published in a peer-reviewed scientific journal, and it did not include books, book chapters, papers, and conference proceedings, as the consistency of these publications is more difficult to determine. 2) The studies published on the findings of the research, not include theoretical articles. 3) The study was carried out in higher education. 4) In classroom practice, the research explored impact of assessment on learners and teachers. 5) The research centered (at least partly) on the effect of assessment to teachers and learners.

Petticrew & Roberts's (2006) data extraction form has been tested and altered several times to guarantee the accuracy and usefulness of data

extraction. The following parts were included in the final data extraction form: a) information on the authors, such as their names and countries of origin; b) research questions and methods; c) sample size; such as number of schools, instructors and students; d) CBLA approaches, such as specifications Aol, Asl, and/or AfL; and e) results, such as evidence of the teacher's role.

D. Search and Selection Result

The investigation is limited to search the research in the past ten years (2011-2020). This period is chosen to see how the studies' concern on CBLA is still getting much respect. The growing interest in the effect of assessment research on applied linguistics and, in particular, on language testing has been fueled this time around. However, various terminologies on CBLA were included for the purpose of searching relevant articles in the online databases before further eliminating articles that do not fit the specified criteria. Two types of results of CBLA studies are discovered according to their publication formats: research paper (including journal articles, book chapters, and research monographs) and review papers on this topic. However, for research, the writer only reviewed all scholarly journals and databases. Studies started looking in two famous journals: Language Testing and Language Assessment Quarterly. These journals were picked because they focus exclusively on high-quality research in the field of language assessment. It was then extended into applied linguistics and language learning papers in order to localize the research. The extended quest included journals including TESOL Quarterly, Applied Linguistics, and System.

In addition, Google scholar was also the used search engine to find relevant articles followed by specifying the articles from highly reputable journals, i.e., Language Assessment Quarterly journal, Modern Language Journal etc. Initially, searching for articles from Web of Science, Language Assessment Quarterly Journal was chosen. Using the keywords classroom-based English language assessment and without limiting the study's year, 416 articles were found in the journal. From those articles, the writer only got two articles that mainly discussed CBLA. Furthermore, the google scholar search engine was used to obtain more comprehensive sources and avoid missing the relevant articles from another high reputable journal. Typing "classroom-based language assessment in EFL higher education," 16.900 articles were found. Accordingly, returning to use time limitation from 2011 to 2020 and additional keywords assessment for learning and learning assessment, the writer, found 276 articles. 161 findings were defined on the basis of data extraction, then this number was narrowed down by identifying

the most important papers applicable to the study issue that were available in full text, screened for validity using the inclusion criterion. 117 articles were either considered to be of poor quality or lacked too much knowledge from the portion of the system to determine the study's quality. Such journals have been omitted from the selection. After making the limitation only reports, not books, from the trusted journal and relevant to the scope, the writer took the remaining 44 articles focusing on CBLA in EFL higher education in this review.

Based on a comprehensive and systematic analysis of the literature on classroom-based language assessment, the proposed framework has been established. The writers studied classroom language assessment practices in EFL contexts published in peer-reviewed journal publications and examined them thematically in order to situate the process in the EFL setting (Garner & Ragland, 2015; Guest et al., 2012). Qualitative meta-analysis was subsequently extended (Schreiber et al., 1997; Timulak, 2009) in our effort to add to practice analysis to the current awareness and ideas of CBLA.

FINDINGS AND DISCUSSION

On the basis of the thematic study and meta-analysis of English language assessment practices in higher education settings in the EFL classroom (CBLA), three themes were found. It was also observed that not all research used a clear context which, in some cases, interchangeably induced the use of the term language evaluation effect with other words such as washback. The discussion on the impact of CBLA on the learners and teachers includes three categories, first, how students' perception on the CBLA. Second, the students' motivation in so called CBLA offline or online and their language achievement. And the third, teachers' assessment literacy and practice. All of the categories are closely related to the term assessment for learning (AfL) and assessment as learning (AaL).

A. CBLA and Students' Perception

Learners are potentially the most impacted by the assessment of all stakeholders. The effect on learners, though, is most likely influenced by the effects of the test, creators of textbooks, and classrooms. The test takers would be affected by planning for and taking the test, as well as the subsequent effects. It may be contrary to the result of experiments and test improvements. An example of an analysis conducted by Gan et al. (2017) found that

interactive-informal performance assessment and student self-assessment were the strongest predictors of the intrinsic motivation and optimistic attitudes of learners towards the EFL course. Although language learners may be the stakeholders most impacted by the language assessment, studies have concentrated more on researching the experiences and behaviors of learners, with a few observational studies detailing the viewpoint of learners' behavior correlated with pre- and post-test effect. Results are contradictory about the impact of exam planning on student test scores. By inspiration, interpretation and comprehension of the test, and previous interactions, the impact on learners tend to be mediated more internally. Regarding this thesis, four articles were to be reviewed to know how students' perception on CBLA. However, much of the research up to now has been descriptive in nature.

Basically, students have positive perceptions on the existence of CBLA and how it effects on their learning. These studies ascertain how learners can be open toward the presence of CBLA (Agcam & Babanoglu, 2016; Galikyan et al., 2019; Wallace, 2018; Suryoputro & Akip, 2016). The importance of self- and peer assessment (Matsuno, 2017), the role of diagnostic (Alderson et al., 2014) or dynamic (e.g., Poehner, 2014) evaluation in facilitating teaching and learning, and technology role in learning and evaluation have been explored in studies focusing more on learners (Mohamadi, 2018; Spivey & McMillan, 2014). Finally, using evidence from two foreign language classes, Hill (2017) proposed a technique of CBLA process study that addressed issues linked to instructors, such as what teachers do, what teachers aspire for, and what principles or standards they employ. Interestingly, teachers looked at how students view language acquisition and evaluation from their perspective.

An interesting comment in Agcam & Babanoglu (2016) was an example how AfL works; "If our growth was not evaluated, I would simply not try to learn English or study it outside the classroom" (p.73). The study discovered that learners believe testing and assessment are appropriate for language instruction and place a premium on listening and speaking over other language skills, indicating that they intend to use the language primarily to communicate with those with whom they do not share a mother tongue. Additionally, this study recommended that EFL educators expand the variety of question forms that are likely to develop students' cognitive abilities and make learning more meaningful and enduring. Suryoputro & Akip (2016) indicate, not differently from Agcam & Babanoglu (2016), that EFL learners viewed portfolio evaluation as a 'novel' approach for evaluating their reading

and a helpful learning tool; thus, they used portfolio evaluation for potential guidance for teaching and learning.

Moreover, Wallace (2018) examines students' perceptions on the fairness and justice in L2 classroom assessment. The findings provide useful insight into how learners may have viewed the fairness of their classroom test. An assessment is deemed fair from the classroom justice literature if learners interpret 1) the methods used in the test to be equitable for all test takers, 2) the score to equally reflect the success of the learners, 3) the contact with the language teacher to be polite, and 4) the test and score details to be timely.

The findings suggest that during a single test administration, language learners in this sample thought that they were treated equally by their language teacher. The fairest element of the aspects of classroom justice that the students focused on was the way their respective teachers dealt with them during their test administration (interactional fairness). In this case, this could be a welcome outcome for language instructors, since expectations of unequal or unfair treatment of students by teachers in other cases have been correlated with perceptions of the instructor's lower reputation. For instance, Chory (2007) revealed how students felt they were handled poorly by their teacher, they viewed him as being of low integrity, character, and empathy. The author indicated that it is important to handle learners in a polite and reasonable manner to prevent the risk of certain derogatory views of the teacher and, by implication, the instructional program. Classroom teachers themselves, who have established faith and consideration during the learning process (including evaluation) as a significant criterion for fair classroom assessment, further support this advice (Tierney, 2014). Finally, the review on this theme has revealed that learners is the best resource to evaluate how teachers provide appropriate CBLA for their learning experience.

B. CBLA, Students' Learning Motivation and Achievement

17 articles found discussing how CBLA give impact on students' motivation and students achievements. Aspects of the classroom assessment environment are closely correlated with the motivation and accomplishment of student learning (Brookhart, 1997). Teachers who make evaluation decisions, including styles or formats of evaluation, determining evaluation purposes, assigning evaluation assignments, assessing success and giving input, and measuring student learning results, largely build the classroom evaluation environment (Hao & Johnson, 2013). Brookhart (1997) suggested that classroom assessments should motivate students to learn more, work harder,

and have a greater sense of self-efficacy. They should also boost students' involvement by lowering anxiety. This implies that students' conceptions of the function of evaluation, their abilities to conduct it, and the motives for performing it may vary. As a result, classroom assessment is regarded to form the foundation for students' expectations of what is important to learn and how learning efforts should be directed (Liu & Xu, 2017). The Brookhart paradigm has gained support from a variety of scientific studies. Research findings of CBLA effects on students' motivation are mixed. Some studies show positive effects of language tests on motivation on most students (Gan et al., 2017, 2019; Khalil, 2016; Moser, 2017).

The first study conducted by Gan et al. (2019) found that, while teacher-controlled performance-oriented assessment was the most frequently used type of assessment in the EFL classroom, interactive-informal teacher-student evaluation and student self-assessment were the best predictors of students' intrinsic motivation and optimistic attitudes toward the EFL course. As a result, not all methods of classroom assessment were equally compelling. Pan & Newfields (2012) discovered that tests had a negligible effect on students' motivation. However, his research demonstrated that motivation was not examined in terms of degree but rather the type, or more precisely, motives for learning English. A theory notes that practices of language classroom assessment are closely related to the learning motivation of students, and various practices of classroom assessment forecast the learning motivation to a different degree.

Furthermore, CBLA brings positive impact on students achievement in some skills, such as English language learners' oral performance (Kermad & Kang, 2019, Muñoz & Álvarez, 2010). CBLA impact on students' writing ability in portfolio-based classroom (Lam, 2016; Lee et al., 2019; Lee & Coniam, 2013). And, how CBLA influences students' second language acquisition (Perrone, 2011). Lee et al. (2019) used data obtained from interviews with two involved teachers and six chosen students, as well as classroom observations. The results indicate that the very first effort by the teachers to bring assessment for learning (AfL) into effect gained advantages for themselves and the students. Interestingly, Perrone (2011) found that the position of the assessment of the classroom had a distinct effect on how the two students viewed the new learning level, affecting the consistency and speed at which the new form of target language was processed. The study findings found that there were distinct differences in student learning at the level of the learners. In addition, Perrone (2011) asserted that, however, few

longitudinal studies have investigated the influence of a role of classroom assessment on the language acquisition of individual EFL learners. The ongoing investigation leads to the field's understanding of how these various tests can be applied in different ways by examining this interaction in order to support particular language learners on their pathways to second language acquisition.

As has been mentioned earlier, the format of test influences how students prepare facing the test (Winke & Lim, 2017). The way tests are conducted also brings impact on the students learning (Benettayeb-Ouahiani, 2016; Han & Finkelstein, 2013; Spivey & McMillan, 2014). Spivey & McMillan (2014) stated that online examination can offer some benefits that traditional paper-and-pencil research does not provide. To begin, online assessment enables the test to be completed during flexible training periods. Second, online software may be programmed to randomly sequence questions and gather replies to multiple-choice and matching questions. Thirdly, it delivers a range of feedback, including a test score, a test score with right answers, and a test score with comprehensive solutions. Additionally, where direction is provided, instructors have authority (e.g. immediately, set to a particular date and time after all questions are completed). Fifth, preferred online review platforms may be established to offer recommendations as to where the text might assist in answering a question. Spivey & McMillan (2014) also argued that the assessment protocol affected neither test efforts nor course results. The writers, however, considered a clear positive association between the students' efforts and their course results. In the same way, a research by Mohamadi (2018) suggests that an important way to make learning successful is to use stimulating technologies and methods together with suitable appraisal strategies.

Accordingly, Assessments can have a beneficial or negative impact (Xu & Liu, 2018). When a test is well-designed and accurately assesses what should be examined, a beneficial outcome occurs. Meanwhile, Ali & Hamid (2020); Kirkpatrick & Gyem (2012) have explored the detrimental effect of EFL assessment. Negative effects arise for various reasons, including the fact that language teaching is influenced not only by assessment elements but also by a complex of social psychological, political, economic, and data-driven accountability aspects that are incorporated in testing and teaching (Ali & Hamid, 2020; Furaidah et al., 2015). These effects may occur regardless the kind of testing used (e.g., paper-based to online testing settings) (Cox et al., 2018). (Cox et al., 2018). Additionally, comparative assessments of online formative and summative assessment deal with test behavior issues such as test

confidence, anxiety level, (Cassady & Gridley, 2005), and students' participation (Han & Finkelstein, 2013).

The findings imply that EFL teachers should be cognizant of the efficacy of appropriate assessment procedures that lead to the finest potential student motivation and learning outcomes. This enables teachers to concentrate on the ideal classroom assessment approaches that enhance students' active involvement, autonomy, and accountability for their learning. Teachers should be encouraged to utilize assessment not only as a measure for determining student performance at the completion of a unit, but also as a tool for stimulating students' motivation to learn by immersing them in learning dialogues with one another and with their teachers. Additionally, this analysis demonstrates that EFL teachers should be better informed about the value of self-evaluation, interactive-informal evaluation, suitable assessment techniques, and in-class diagnostic evaluation in order to assist students in developing self-regulation.

C. CBLA and EFL Teachers

The review of CBLA and EFL teachers' assessment practice leads to three distinction themes; first the teachers' perception and conception of CBLA, second is teachers' assessment literacy, and the third is the teachers' CBLA practice.

1. Teachers' Perception and Conception of CBLA

A large body of literature around EFL and L2 teachers' cognitions and experiences with classroom-based assessment has been gradually building up. It is a requirement for understanding teachers' perceptions and the (in)effectiveness of CBLA in EFL classrooms. The teacher framework of classroom assessment is used as an umbrella word to refer to an organized value structure that accepts everything a teacher feels about the meaning and intent of evaluation, including values, ideas, attitudes and preferences (Thompson, 1992 in Liu & Xu, 2017). Research investigating the teachers' perception on CBLA could be found having similar result (Chen et al., 2013; Gan et al., 2017; Nasr et al., 2018; Oz & Atay, 2017).

The assessment conceptions and perceptions of teachers are increasingly shaped by their previous experiences of both being evaluated as learners and implementing assessment as teachers. Three messages are used in our explanations of the creation and role of conceptions. First, assessment principles for teachers tend to be outside of assessment preparation. There

were no statistically meaningful variations between teachers with varying degrees of assessment preparation experience in the mean scores of teachers' evaluation. Second, 'observational apprenticeship' established in teachers' own schooling experiences would have a major effect on their attitudes towards assessment. Third, teacher assessment perceptions appear to be an inextricable aspect of teacher teaching and learning conceptions, both of which are influenced by one's epistemological values and views of learning (Brown, 2004).

Conducting a qualitative study, Nasr et al. (2018) disclosed that the majority of EFL teachers considered the use of evaluation as helpful and effective. However, in classroom studies, minor differences were noticed in terms of tracking procedures for assessment. In comparison, the perceived supervision and perceived scaffolding of appraisal for instructional activities by EFL teachers was not substantially different in terms of their years of teaching experience, academic degree, and teaching ability levels. Nasr's analysis is in line with Oz & Atay (2017), which also revealed that there is not much correlation between the interpretation of knowledge and assessment. Such studies offer insights into the promotion of assessment for learning (AfL) culture among EFL teachers. It also has major consequences for teacher educators and scholars to investigate innovative methods of incorporating teaching appraisal into teaching as a method of maximizing student performance. Their assessment conceptions, however, need to be considered "as an indispensable starting point for any further professional development on the subject" (Remesal, 2011), p. 474). Gan et al. (2017) point to the value of successful professional growth that is important to the acquisition of the opportunity to give positive guidance and encourage the students' active participation in learning and assessment.

2. Teachers' Assessment Literacy

As the assessment process is embedded in classroom experience, questions and other educational practices, teachers, on the one hand, need to facilitate and foster productive classroom interactions and, on the other hand, establish assessment-embedded classroom activities and assignments. Teachers need to ensure that these practices are accurate in order to use evidence generated by such a framework for their own understanding of student learning and decision-making. Relatedly, 19 articles were found to have discussed the teachers' assessment literacy (TAL), teachers' perception toward CBLA and teachers' practice in CBLA. TAL deals with how EFL teachers' knowledge on the assessment. Language teacher assessment literacy discussion

has continued to concentrate on the more planned and formal forms of assessment (Elshawa et al., 2017; Levi & Inbar-Lourie, 2020; Ölmez-Öztürk, 2018; Thirakunkovit, 2019; Zulaiha et al., 2020).

Concomitantly, CBLA is inseparable from teachers' assessment literacy. If the teachers do not have enough knowledge to carry out a proper assessment, it may give a bad impact on student outcomes. It could be that students' abilities are measured in inappropriate ways. The importance of EFL teachers' assessment literacy has been studied by some scholars, starting from Giraldo (2018); Hill (2017); Levi & Inbar-Lourie (2020); Nimehchisalem & Bhatti (2019); Oz & Atay (2017); Tavassoli & Farhady (2018); Thirakunkovit (2019); and Zulaiha et al. (2020). These studies focused on how teachers combine assessment knowledge and teachers' assessment practice. EFL teachers have tried to properly apply classroom-based language assessment principles. Hill (2017) for instance, revealed that teachers tend not to fulfil the CBLA principles due to time constraints and the large number of burdens in the syllabus that must be fulfilled in one period. Nimehchisalem & Bhatti (2019) added that teachers try their best to apply their knowledge, but there are reasons for it to be done imperfectly. Furthermore, Oz & Atay (2017) divulged that, while most teachers are familiar with simple classroom assessment, there is a disparity between performance comprehension and contemplation in the classroom when it comes to classroom experience. These findings offer a need of teachers' language assessment professional development to support their assessment practice.

Research problems often raised is whether teachers differentiate their assessment between assessment for learning (AfL) and assessment as learning (AaL) from the assessment of learning AoL. Referring to various survey results, and the results are relieving, it turns out that EFL teachers consciously emphasize the occurrence of AfL rather than the AoL. This is proven by Chen (2020) when he conducted the contrastive analysis between the use of AfL and AoL. Using different research method, Ashraf & Zolfaghari (2018) revealed that teachers' reflection on their assessment practice was based on their assessment literacy.

Liu & Xu's (2017) thesis is the first to analyze the standard of assessment courses among the small number of review studies (Ha, 2019; Chen 2013). They measure different course characteristic variables (e.g., instructors, content, and students). It is found that while the language evaluation courses sufficiently addressed important elements of language testing philosophy and application, instructional and psychological

measurement and pupil classroom practice paid far less attention to them. While this research somehow ensured the consistency of language assessment courses from the viewpoint of educators, it is not yet understood if language teachers are adequately literate in assessment to cope with problems emerging from the experience of classroom evaluation. Despite these major contributions, more methodological research are desperately required to explore the aspects of teachers' assessment literacy, including the seven competency criteria prescribed by the Criteria for Teacher Competence in Student Educational Achievement (hereinafter the Standards) (AFT, NCME, & NEA, 1990 in Liu & Xu, 2017) they are:

- 1). In choosing assessment approaches suitable for educational decisions, teachers should be qualified.
- 2). In designing assessment approaches useful for educational decisions, teachers should be qualified.
- 3). Teachers should be experienced in conducting, scoring, and interpreting the effects of both externally produced and appraisal processes produced by teachers.
- 4). When making decisions about specific students, preparing instruction, designing curriculum, and school development, teachers should be trained using test outcomes.
- 5). In designing valid grading processes, teachers should be qualified.
- 6). In transmitting test outcomes, parents, other lay audiences, and other educators, teachers should be professional.
- 7). Teachers should be able to recognize unethical, immoral, and improper methods of assessment and the use of information from assessment.

3. Teachers' CBLA Practice

The authors established in this review that teachers are aware of what constitutes excellent assessment, that they have their own assessment criteria, and that they understand how to implement them. Clearly, they are not passive recipients but but active practitioners (Gonen & Akbarov, 2015). Related to CBLA practice, a large body of empirical studies have been conducted (Barnes, 2017; Benettayeb-Ouahiani, 2016; Chen et al., 2013; Chong, 2018; Dann, 2014; Fives & Barnes, 2020; Gonen & Akbarov, 2015; Noori et al., 2017; Wach, 2012; L. Wang, 2020; X. Wang, 2017). As ARG

(Assessment Reform Group, 2002) mentioned 10 principles in assessment for learning, the study of Gonen & Akbarov (2015) do not confirm this application of the principles. The findings of the questionnaire revealed that the concepts were not applied perfectly. Throughout the report, the teachers stated that due to the institution's central grading method, intensive syllabuses to be completed in a short period, and the educational experiences of students they took from the primary and secondary school education systems, they did not have a chance to bring their values into effect in some cases.

Moreover, Zulaiha et al. (2020) demonstrated that teachers had good understanding of the concepts of assessment and adapted such knowledge to work in the classroom. However, the gap between the expertise of teachers and its application in classroom activities has been established, especially in the phases of implementation and monitoring. Some issues impacted the teachers' assessment practice including local or school policies, the utilization of non-achievement indicators (e.g., attendance and attitudes of students), and the role of parents in their children's education. In Zulaiha et al.'s context, the results lead to a deeper understanding of the teachers' assessment literacy as they make sense and communicate with assessment resources and related stakeholders. In addition, Gan et al. (2017) indicated a strong positive association between the teachers' AfL experience and their propensity to follow an approach of achieving. Results also showed that a surface approach to learning was negatively associated with AfL experience, indicating that the more AfL characteristics integrated into classroom teaching and assessment processes in high education setting, the less likely students are to follow a surface approach to learning. This research further offers proof of the possible variations related to different academic settings in the responses of students to AfL.

Furthermore, Turner & Purpura (2015) have suggested a multidimensional approach to CBLA, entitled Learning-oriented Assessment (LOA), in which the goal is to consider the dynamics of how assessment (e.g., evaluations, findings, class interactions, naturalistic talking-interaction, peer reviews, self-assessment, initiatives, portfolios, performative assessments) can be interpreted. Importantly, this approach considers that assessment is multifaceted in universities, containing not only several different aspects (e.g., context) that apply to the learning experience, but also many agents (e.g., students, teachers, peers, computers). In this respect, Turner and Purpura shift towards an approach that characterizes assessment, learning, and teaching, though separate, as fundamentally interconnected, without dichotomous

representations of CBLA as formative/summative or AoL/AfL. Seven critical interconnecting dimensions of the CBLA are described (Purpura, 2016). This includes: the relational dimension (i.e., the social, cultural or intellectual context of learning), the dimension of elicitation (i.e., the process used to evoke performance), the dimension of ability, the dimension of cognition or understanding (i.e., the socio-cognitive features underlying success and learning), the dimension of intimacy (i.e., the personality and interpersonal arrangements involved).

The writers concur with Turner and Purpura (2015) that a prerequisite for CBLA and AfL to be successfully implemented is the teachers' assessment practice which relies on their assessment literacy. From these empirical studies on CBLA and AfL teacher practice, some inferences may be made. It is, first of all, dynamic and situated. Thus, without a contextual interpretation of such practice, it is not easy to judge the consistency of the CBLA practice of an instructor. Second, even though the teachers already have sufficient knowledge of assessment, in practice, limitations are often found. And third, the teachers need to reflect on the assessments that have been carried out. Awareness of reflection and developing the ability to conduct assessments are an obligation for teachers.

Based on this systematic review, researchers have conducted studies showing how the importance of CBLA has a major influence on the assessment and learning process. In addition, the teacher's philosophical perceptions and concepts also play an important role in the formation of a prolific assessment process. In addition, research on CBLA has also been shown to have a major impact on students' motivation to learn a language; CBLA generally refers to AfL rather than AoL. In short, CBLA generally, both from perspective and practice, has consequences for teachers to continue improving their assessment skills and literacy. This review is also a plea for a broader interpretation of CBLA from a broader viewpoint.

First, though CBLA is a global movement and is invariably supported by many Western countries, more empirical research and discussion are required for its viability in local cultural contexts (e.g., Indonesia). Second, considering the crucial role of teachers in CBLA implementation, future studies need to create ties between the CBLA practice and the teachers' assessment literacy. Third, taking into account the role of social environments in influencing CBLA, it has been influenced by institutional laws, power ties between various players (e.g., managers, teachers, students, etc.), and school culture. Studies

with an ethnographic attitude and listen to various voices would be more helpful to explore the essence of the issue.

CONCLUSION

This study questions the application of CBLA in higher educational settings on the basis of an overview of CBLA studies in English language classrooms. It claims that when implementing and adapting CBLA, the difficulty of the evaluation of contextual problems at the macro-, meso- and micro-levels needs to be taken into account. Previous research shows that a common approach that encourages AfL in universities is to design appraisal activities that contribute to high-quality learning and that students find meaningful because these tasks give them a feeling that they are learning something significant for their long-term personal growth. As evidenced by the literatures provided by scholars, it is found that students' expectation and motivation in CBLA, teachers' perceptions and appraisal principles, teachers' CBLA experience, and teachers' assessment literacy show that successful CBLA implementation could be accomplished in their hearts and minds through teacher improvements to evaluation. In promoting the transition of teachers, support from policy makers, teachers, educators, and institutional administrators is equally important. Some paths are therefore proposed for potential CBLA studies.

REFERENCES

- Abbad, M., Morris, D., & Nahlik, C. de. (2009). Looking under the bonnet: Factors affecting student adoption of e-learning systems in Jordan. *International Review of Research in Open and Distance Learning*, 10(2), <https://doi.org/10.19173/irrodl.v10i2.596>
- Agcam, R., & Babanoglu, M. P. (2016). Students' perceptions of language testing and assessment in higher education. *Unibulletin*, 5(1-2), 66-77. <https://doi.org/10.22521/unibulletin.2016.512.6>
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115-129. <https://doi.org/10.1093/applin/14.2.115>

Fitriyah, I., & Widiati, U., A Systematic Literature Review on the Impact of EFL 93
Classroom-Based Language Assessment

Ali, Md. 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, Z. (2015). A case study of tertiary students' experiences using Edmodo in language learning. *International Journal of Language Education and Applied Linguistics (IJLEAL)* Copyright © Penerbit Universiti Malaysia Pahang Print, 2289-9294 Online, 2(2015), 39–48.

Ashraf, H., & Zolfaghari, S. (2018). EFL teachers' assessment literacy and their reflective teaching. *International Journal of Instruction*, 11, 12.

Assessment Reform Group. (2002). *Assessment for learning: 10 principles*. Cambridge University Press.

Barnes, M. (2017). Washback: Exploring what constitutes “good” teaching practices. *Journal of English for Academic Purposes*, 30, 1–12. <https://doi.org/10.1016/j.jeap.2017.10.003>

Benettayeb-Ouahiani, D. A. (2016). Assessment in the EFL university classroom: between tradition and innovation. *Revue Des Études Humaines et Sociales -B/ Littérature et Philosophie*, 15, 8.

Brown, H.D., (2004). *Language assessment; Principles and classroom practices*. Longman, New York

Brookhart, S. M. (1997). A Theoretical Framework for the Role of Classroom Assessment in Motivating Student Effort and Achievement. *Applied Measurement in Education*, 10, 161-180. https://doi.org/10.1207/s15324818ame1002_4

Cai, H. (2012). E-learning and English teaching. *IERI Procedia*, 2, 841–846. <https://doi.org/10.1016/j.ieri.2012.06.180>

Cassady, J. C., & Gridley, B. E. (2005). The effects of online formative and summative assessment on test anxiety and performance. *The Journal of Technology, Learning and Assessment*, 4(1), 31.

Chapelle, C. A., & Voss, E. (2016). 20 Years of technology and language assessment in language learning & technology. *Language Learning*, 20(2116–128), 13.

- Chen, H. (2020). A contrastive analysis of classroom-based language assessments. *English Language Teaching*, 13(5), 110. <https://doi.org/10.5539/elt.v13n5p110>
- Chen, Q., May, L., Klenowski, V., & Kettle, M. (2013). The enactment of formative assessment in English language classrooms in two Chinese universities: Teacher and student responses. *Assessment in Education: Principles, Policy & Practice*, 21(3), 271–285. <https://doi.org/10.1080/0969594X.2013.790308>
- Cheng, L., Sun, Y., & Ma, J. (2015). Review of washback research literature within Kane’s argument-based validation framework. *Language Teaching*, 48(4), 436–470. <https://doi.org/10.1017/S0261444815000233>
- Chong, S. W. (2018). Three paradigms of classroom assessment: implications for written feedback research. *Language Assessment Quarterly*, 15(4), 330–347. <https://doi.org/10.1080/15434303.2017.1405423>
- Cox, T. L., Malone, M. E., & Winke, P. (2018). Future directions in assessment: Influences of standards and implications for language learning. *Foreign Language Annals*, 51(1), 104–115. <https://doi.org/10.1111/flan.12326>
- Daniels, M., Sarte, E., & Cruz, J. D. (2019). Students’ perception on e-learning: A basis for the development of e-learning framework in higher education institutions. *IOP Conference Series: Materials Science and Engineering*, 482, 012008. <https://doi.org/10.1088/1757-899X/482/1/012008>
- Dann, R. (2014). Assessment as learning: Blurring the boundaries of assessment and learning for theory, policy and practice. *Assessment in Education: Principles, Policy & Practice*, 21(2), 149–166. <https://doi.org/10.1080/0969594X.2014.898128>
- Doculan, J. (2016). E-Learning readiness assessment tool for Philippine higher education institutions. *International Journal on Integrating Technology in Education (IJITE)*, 5(2). <https://doi.org/10.5121/ijite.2016.5203>
- Elshawa, N. R. M., Nadzimah Abdullah, A., & Md Rashid, S. (2017). Malaysian instructors’ assessment beliefs in tertiary ESL classrooms. *International Journal of Education and Literacy Studies*, 5(2), 29. <https://doi.org/10.7575/aiac.ijels.v.5n.2p.29>

- Fitriyah, I., & Widiati, U., A Systematic Literature Review on the Impact of EFL 95 Classroom-Based Language Assessment
- Fives, H., & Barnes, N. (2020). Navigating the complex cognitive task of classroom assessment. *Teaching and Teacher Education*, 92, 103063. <https://doi.org/10.1016/j.tate.2020.103063>
- Furaidah, F., Saukah, A., & Widiati, U. (2015). Washback of English national examination in the Indonesian context. *TEFLIN Journal - A Publication on the Teaching and Learning of English*, 26(1), 36. <https://doi.org/10.15639/teflinjournal.v26i1/36-58>
- Galikyan, I., Madyarov, I., & Gasparyan, R. (2019). Student test takers' and teachers' perceptions of the TOEFL Junior® standard test. *ETS Research Report Series*, 2019(1), 1-15. <https://doi.org/10.1002/ets2.12264>
- Gan, Z., He, J., & Liu, F. (2019). Understanding classroom assessment practices and learning motivation in secondary EFL students. *The Journal of AsiaTEFL*, 16(3), 783-800. <https://doi.org/10.18823/asiatefl.2019.16.3.2.783>
- Gan, Z., Liu, F., & Yang, C. C. R. (2017). Assessment for learning in the Chinese context: Prospective EFL teachers' perceptions and their relations to learning approach. *Journal of Language Teaching and Research*, 8(6), 1126. <https://doi.org/10.17507/jltr.0806.13>
- Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Profile: Issues in Teachers' Professional Development*, 20(1), 179-195. <https://doi.org/10.15446/profile.v20n1.62089>
- Gonen, K., & Akbarov, A. (2015). Instructors' principles and practices in the implementation stage of classroom-based language assessment in higher education in Turkey. *Journal Of European Education*, 5(3). <https://doi.org/10.18656/jee.04624>
- Ha, N. T. T. (2019). A literature review of washback effects of assessment on language learning. *Journal of Science Ho Chi Minh City Open University*, 9(5), 13.
- Han, J. H., & Finkelstein, A. (2013). Understanding the effects of professors' pedagogical development with Clicker assessment and feedback technologies and the impact on students' engagement and learning in higher education. *Computers & Education*, 65, 64-76. <https://doi.org/10.1016/j.compedu.2013.02.002>

Hao, S., & Johnson, R. L. (2013). Teachers' classroom assessment practices and fourth-graders' reading literacy achievements: An international study. *Teaching and Teacher Education*, 29, 53–63. <https://doi.org/10.1016/j.tate.2012.08.010>

Hill, K. (2017). *Understanding classroom-based assessment practices: A precondition for teacher assessment literacy*. 6(1), 17.

Johnson, M., & Shaw, S. (2019). What is computer-based testing washback, how can it be evaluated and how can this support practitioner research? *Journal of Further and Higher Education*, 43(9), 1255–1270. <https://doi.org/10.1080/0309877X.2018.1471127>

Kermad, A., & Kang, O. (2019). Effect of classroom assessment stakes on English language learners' oral performance. *TESOL Journal*, 10(2), e00392. <https://doi.org/10.1002/tesj.392>

Khalil, R. M. R. (2016). Assessment as a learning tool in a flipped English language classroom in higher education. *Arab World English Journal*, 17.

Kirkpatrick, R., & Gyem, K. (2012). Washback effects of the new English assessment system on secondary schools in Bhutan. *Language Testing in Asia*, 2(4), 5. <https://doi.org/10.1186/2229-0443-2-4-5>

Lam, R. (2016). Assessment as learning: Examining a cycle of teaching, learning, and assessment of writing in the portfolio-based classroom. *Studies in Higher Education*, 41(11), 1900–1917. <https://doi.org/10.1080/03075079.2014.999317>

Lee, I. (2007). Assessment for learning: integrating assessment, teaching, and learning in the ESL/EFL writing classroom. *Canadian Modern Language Review*, 64(1), 199–213. <https://doi.org/10.3138/cmlr.64.1.199>

Lee, I., & Coniam, D. (2013). Introducing assessment for learning for EFL writing in an assessment of learning examination-driven system in Hong Kong. *Journal of Second Language Writing*, 22(1), 34–50. <https://doi.org/10.1016/j.jslw.2012.11.003>

Lee, I., Mak, P., & Yuan, R. E. (2019). Assessment as learning in primary writing classrooms: An exploratory study. *Studies in Educational Evaluation*, 62, 72–81. <https://doi.org/10.1016/j.stueduc.2019.04.012>

Fitriyah, I., & Widiati, U., A Systematic Literature Review on the Impact of EFL 97 Classroom-Based Language Assessment

- Levi, T., & Inbar-Lourie, O. (2020). Assessment literacy or language assessment literacy: learning from the teachers. *Language Assessment Quarterly*, 17(2), 168–182. <https://doi.org/10.1080/15434303.2019.1692347>
- Liu, J., & Xu, Y. (2017). Assessment for learning in English language classrooms in china: Contexts, problems, and solutions. In H. Reinders, D. Nunan, & B. Zou (Eds.), *Innovation in Language Learning and Teaching* (pp. 17–37). Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-60092-9_2
- Matsuno, S. (2017). Adoptability of Peer Assessment in ESL Classroom. *Creative Education*, 08(08), 1292–1301. <https://doi.org/10.4236/ce.2017.88091>
- McKinley, J., & Thompson, G. (2018). Washback effect in teaching English as an international language. In J. I. Liontas, T. International Association, & M. DelliCarpini (Eds.), *The TESOL Encyclopedia of English Language Teaching* (pp. 1–12). John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118784235.eelt0656>
- Mohamadi, Z. (2018a). Comparative effect of online summative and formative assessment on EFL student writing ability. *Studies in Educational Evaluation*, 59, 29–40. <https://doi.org/10.1016/j.stueduc.2018.02.003>
- Moser, J. (2017). Creating in-class self-directed learning through can do objectives, portfolio use, and formative assessment. *The Journal of AsiaTEFL*, 14(4), 674–686. <https://doi.org/10.18823/asiatefl.2017.14.4.6.674>
- Muñoz, A. P., & Álvarez, M. E. (2010). Washback of an oral assessment system in the EFL classroom. *Language Testing*, 27(1), 33–49. <https://doi.org/10.1177/0265532209347148>
- Nasr, M., Bagheri, M. S., Sadighi, F., Rassaei, E., & Heidari-Shahreza, M. A. (2018). Iranian EFL teachers' perceptions of assessment for learning regarding monitoring and scaffolding practices as a function of their demographics. *Cogent Education*, 5(1), 1558916. <https://doi.org/10.1080/2331186X.2018.1558916>

- Nimehchisalem, V., & Bhatti, N. (2019a). A Review of Literature on Language Assessment Literacy in last two decades (1999-2018). *International Journal of Innovation*, 8(11), 17.
- Noori, A., Shafie, N. H., Mashwani, H. U., & Tareen, H. (2017). Afghan EFL lecturers' assessment practices in the classroom. *Imperial Journal of Interdisciplinary Research*, 3(10), 14.
- Ölmezer-Öztürk, E. (2018). Toward measuring language teachers' assessment knowledge: Development and validation of Language Assessment Knowledge Scale (LAKS). *Language Testing in Asia*, 8(20), 15.
- Oz, S., & Atay, D. (2017). Turkish EFL instructors' in-class language assessment literacy: Perceptions and practices. *International Association of Research in Foreign Language Education and Applied Linguistics ELT Research Journal*, 6(1), 25-44.
- P. Wallace, M. (2018). Fairness and justice in L2 classroom assessment: perceptions from test takers. *The Journal of AsiaTEFL*, 15(4), 1051-1064. <https://doi.org/10.18823/asiatefl.2018.15.4.11.1051>
- Pan, Y.-C., & Newfields, T. (2012). Tertiary EFL proficiency graduation requirements in Taiwan: A study of washback on learning. *Electronic Journal of Foreign Language Teaching*, 9(1), 108-122.
- Perrone, M., & College, H. (2011). The effect of classroom-based assessment and language processing on the second language acquisition of EFL students. *Journal of Adult Education*, 40(1), 14.
- Petticrew, M., & Roberts, H. (2006). *Systematic Reviews in the Social Sciences*. Oxford.
- Purpura, J. E. (2016). Second and foreign language assessment: second and foreign language assessment. *The Modern Language Journal*, 100(S1), 190-208. <https://doi.org/10.1111/modl.12308>
- Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education*, 27(2), 472-482. <https://doi.org/10.1016/j.tate.2010.09.017>

- Fitriyah, I., & Widiati, U., A Systematic Literature Review on the Impact of EFL 99 Classroom-Based Language Assessment
- Ridhwan, M. (2017). Understanding formative and summative assessment for EFL teachers: theoretical reflections on assessment for learning. *JSHMIC: Journal of English for Academic*, 4(1), 40–50. [https://doi.org/10.25299/jshmic.2017.vol4\(1\).505](https://doi.org/10.25299/jshmic.2017.vol4(1).505)
- Schildkamp, K., Kleij, F. M. V. der, Heitink, M. C., Kippers, W. B., & Veldkamp, B. P. (2020). Formative assessment_ A systematic review of critical teacher prerequisites for classroom practice. *International Journal of Educational Research*, 103((2020) 101602), 16.
- Spivey, M. F., & McMillan, J. J. (2014). Classroom versus online assessment. *Journal of Education for Business*, 89(8), 450–456. <https://doi.org/10.1080/08832323.2014.937676>
- Stoyonoff, S. (2012). Looking backward and forward at classroom-based language assessment. *ELT Journal*, 66(4), 523–532. <https://doi.org/10.1093/elt/ccs041>
- Suryoputro, G., & Akip, R. (2016). Exploring EFL Students' perceptions of washback of portfolios in reading assessment. *International Journal of Language and Linguistics*, 3(2), 6.
- Tavassoli, K., & Farhady, H. (2018). Assessment knowledge needs of EFL teachers. *Teaching English Language*, 12(2), 45–65.
- Thirakunkovit, S. (2019). Exploring writing assessment literacy and classroom practices of Thai university instructors. *The Journal of AsiaTEFL*, 16(4), 1135–1151. <https://doi.org/10.18823/asiatefl.2019.16.4.5.1135>
- Turner, C. E., & Purpura, J. E. (2015). *Learning-oriented assessment in the classroom. Handbook of second language assessment*. DeGruyter.
- Wach, A. (2012a). Classroom-based language efficiency assessment: A challenge for EFL teachers. *Glottodidactica. An International Journal of Applied Linguistics*, 39(1), 81. <https://doi.org/10.14746/gl.2012.39.1.8>
- Wagner, E. (2020). Duolingo English test, Revised Version July 2019. *Language Assessment Quarterly*, 17(3), 300–315. <https://doi.org/10.1080/15434303.2020.1771343>
- Wang, L. (2020). Chinese university EFL teachers' beliefs and practices of classroom writing assessment. *Studies in Educational Evaluation*, 66, 11.

- Wang, X. (2017). A Chinese EFL Teacher's classroom assessment practices. *Language Assessment Quarterly*, 14(4), 312-327. <https://doi.org/10.1080/15434303.2017.1393819>
- Wihastyanang, W. D., & Latief, M. A. (2018). The impact of electronic feedback on students' writing quality. *The Online Journal of New Horizons in Education*, 8(4), 8.
- Wind, S. A., & Peterson, M. E. (2018). A systematic review of methods for evaluating rating quality in language assessment. *Language Testing*, 35(2), 161-192. <https://doi.org/10.1177/0265532216686999>
- Winke, P., & Lim, H. (2017). The effects of test preparation on second-language listening test performance. *Language Assessment Quarterly*, 14(4), 380-397. <https://doi.org/10.1080/15434303.2017.1399396>
- Xu, Q., & Liu, J. (2018). *A study on the washback effects of the Test for English Majors (TEM): Implications for testing and teaching reforms*. Springer Singapore. <https://doi.org/10.1007/978-981-13-1963-1>
- Zulaiha, S., Mulyono, H., & Ambarsari, L. (2020). An investigation into EFL teachers' assessment literacy: Indonesian teachers' perceptions and classroom practice. *European Journal of Contemporary Education*, 9(1), 189-201. <https://doi.org/10.13187/ejced.2020.1.189>