

Understanding Teachers' Knowledge, Skills, and Principles on Language Assessment: A Survey on Teachers' Language Assessment Literacy

Indah Puspawati

*English Language Education Department
Universitas Muhammadiyah Yogyakarta*

Yogyakarta, Indonesia

indahpuspawati@umy.ac.id

Abstract—the Language Assessment Literacy (LAL) refers to language teachers' knowledge, skills and principles on language assessment and evaluation. This paper has investigated the LAL of teachers in a higher education context in Indonesia. The three components of language assessments were the focus of the inquiry in this study. To conduct the research and collect the required data, a self-rated survey LAL and semi-structured interviews were employed. The survey questionnaire was distributed to 30 (N=30) language teachers, while the interview was carried out to five participants. The findings showed that teachers' LAL level was in fair category, meaning that they had sound understanding on knowledge, skills, and principles on language assessments. The study also revealed that the teachers' understanding on and awareness of language assessment principles were the highest compared to their understanding on knowledge and skills on language assessments.

Keywords—*language assessment literacy, assessment knowledge, skills on language assessment, principles of language assessment*

I. INTRODUCTION

As an integral part of teaching, assessment can be regarded as an important aspect in teaching and learning process since the information from assessment is essential for many decisions in teaching and learning. Similarly, stake holders usually use the information from assessment to decide on many educational policies as well. Consequently, designing, administering, scoring, and interpreting the result of assessment needs to be done carefully. To do this, teachers need to have sound knowledge about assessment principles and techniques. Therefore, it is essential that teachers hold sufficient knowledge on assessment literacy.

The term Assessment Literacy (AL) was first proposed by R. J. Stiggins [1] who refers AL to teachers' understanding about principles of educational assessment and their ability to use the understanding to measure various students' achievements. He also emphasizes that all educators must have a good assessment literacy level so that they understand and can use sound assessment practices to assess students' achievement [2]. As this term is highly accepted in the field of educational assessment, many experts have conducted study on the AL to offer its more up to date definition [3],[4],[5]. T. Newfields [3] added the ability to score/grade students fairly and precisely, and this also includes not only an ability to interpret the score/grade effectively, but also an ability to understand the practices of ethical issues of assessment into AL components. In addition, S. M. Brookhart

[4] added knowledge and skills about formative assessment and standard-based reform into the components of assessment literacy. Finally, Y. Xu and G. T. Brown [5] asserted a self-directedness in AL. They argued that assessment literate teachers are those who are aware of and reflect on their assessment practices and are continuously making efforts to improve their understanding of assessment.

In the field of language teaching and learning, the term Language Assessment Literacy (LAL) has already been coined too, as a need to differentiate between LAL and general AL, because general AL does not include the uniqueness of language assessment into the concept [6]. Many definitions have been offered for the term LAL which basically refers to knowledge, skills, and principles in language assessment [7],[8]. F. Giraldo [7] offered a more detailed description as what dimensions and descriptors are there in language testing knowledge, skills, and principles. He explained that language assessment knowledge includes awareness of applied linguistics, awareness of theory and concepts of language testing and assessment, and awareness of own language assessment context. Skills include instructional skills, designing language assessment skills, educational measurement skills, and technological skills. The principles of LAL contain awareness of and action towards critical issues in language assessment.

Studies about teachers' AL/LAL have pointed out that it is important for teachers to have adequate AL. According to T. Newfields, [9] assessment becomes a very crucial component in education because it is something that teachers do all the time and to help them interpret the result of the assessment, AL gives a clear corridor on what step they must take. Furthermore, Y. Xu and G. T. Brown [10] argued that teachers with sufficient AL will be able to make correct interpretation about students' learning based on their assessment results. They will also be able to modify learning instruction appropriately [10]. On the other hand, they argued that teachers with insufficient AL will likely to create invalid and unreliable assessment that the result could lead to false information about students' learning which eventually could lead to 'ill-informed educational decision' [10].

Seeing the importance of teachers' LAL, this study will focus on investigating the LAL of language teachers in a higher education context in Indonesia. Study on this theme is sought to be significant because, in Indonesian context, the teachers generally have the highest independency to decide any form of assessment to measure their students'

achievement compared to those in other education contexts, as stated in Indonesian government regulation no. 74 of 2008. Thus, teachers' LAL in this context is crucial for ensuring reliable and valid measurements. Besides, the newest reform of Indonesia's higher education curriculum in 2016 that highlights the importance of using multiple assessment methods and procedures to assess students' cognitive, psychomotor, and affective competencies requires the teachers to have higher level of LAL. It therefore becomes important to understand teachers' LAL.

While study about teachers' LAL has become more increasing, there are only a few studies that have been carried out in Indonesian context one of which is a study by A. H. Prasetyo [11] that investigated about pre-service teacher LAL development. To determine more proofs concerning this discussed topic, this current study dealt with 1) the university teachers' levels of LAL according to their perspectives and 2) the way they perceive their knowledge, skills, and principles in language assessment.

II. LITERATURE REVIEW

The growing interest about teachers' LAL has resulted in many debates and study to shed lights on understanding it. Many experts also attempted to define LAL, and many instruments are also being developed to measure teacher LAL level. The LAL and study on it are explained below.

A. Language Assessment Literacy

As mentioned earlier, the term LAL basically derived from the general AL concept. The first elaborate description about teachers' AL was offered by American Federation of Teachers (AFT), National Council on Measurement in Education (NCME), and National Education Association (NEA) (1990) AFT in 1990 which proposed 7 standards of teachers' competency in educational assessment of students. The standards stated that teachers should have the ability to 1) choose appropriate assessment method; 2) develop appropriate assessment; 3) administer, score and interpret the results of both externally-produced and teacher-produced assessment methods; 4) use assessment result when making decisions; 5) develop valid student grading procedures; 6) communicate assessment results to stake holders; and 7) recognize unethical, illegal, and inappropriate assessment methods and uses of assessment information.

With the development and changing practices of assessment, M. Brookhart [4] stated that considering the shifting practices on assessment from test-based assessment to multiple methods of assessment, the standards should also consider teachers' knowledge and skills on formative assessment and teachers' knowledge and skills in their own assessment context. Furthermore, Y. Xu and G. T. Brown [5] also added the elements of teachers' self-directed awareness of assessment process that will allow teachers to interpret assessment policies and principles and adapt those into their own classroom context. In addition, A. Scarino [12] also argued that when discussing about AL, teachers should also be aware of their relevant content knowledge when designing assessments. The proposition by A. Scarino basically called out on a more specific, content knowledge-based LA, for example, LAL.

Even though the concept of LAL is quite similar with AL, LAL has different and unique characteristics specific for

language teaching. What differentiate LAL from the general term of AL is that language becomes the construct for assessment [7]. LAL is a combination of AL skill with language specific competency which, then, creates a distinction between general AL and LAL [13]. A. Davies [14] identifies three main components of LAL, i.e., skills, knowledge, and principles. Skills refer to a set of ability to conduct test/assessment, to analyze the assessment, and report the result of the test, and these skills may include writing test items, statistics, test analysis, and developing assessment rubric [14]. Knowledge components refer to relevant background knowledge of language measurements and language measurements methodologies and knowledge of the assessment context [14]. He continues to describe that principles refer to teachers' awareness of issues of ethics, fairness, and consequences of assessments.

Furthermore, a more detail description on each component was offered. To describe more about skill components, F. Giraldo [7] proposes that there are four sets of skills that are important for teachers' LAL, i.e., instructional skills, design skills for language assessment, skills in educational measurements, and technological skills needed for language measurements. For knowledge components, knowledge of applied linguistics, knowledge of theories and concepts of language assessments, and knowledge of teachers' own language assessment context are three kinds of knowledge that teachers should have [7]. Finally, principles basically measure the teachers' awareness and action toward critical issues in language assessment, such as using the results of assessment to decide about the learners, critically analyzing the impact of standardized tests, and also being awareness of applying fair and transparent language assessments. In addition, F. Giraldo [7] has developed sixty-six descriptors for every component to assess teachers' LAL. Because of the detailed description, this study used F. Giraldo descriptors to survey teachers' LAL.

B. Study on Teachers' LAL

Study about teachers' LAL revealed the understanding on what the teachers LA/LAL level is with somewhat similar results. For example, a study by N. Sultana [15] reported that teachers participated in this study had inadequate LAL, and that raised issues about the quality of their knowledge, skills, and principles on language assessment, and this brought out questions on the quality of the language assessments and/or tests that they produced. Similarly, a study by C. Lan and S. Fan [16] revealed that the teachers participated in their study were at "functional level" of their Classroom-Based Language Assessment Level (CBLAL). They described the functional level as the level in which teachers had good understanding of basic terms and concepts about CBLAL, but they were not able to use the knowledge in their CBLAL practice. This level, they claimed, was not adequate for teachers because they should have at least procedural and conceptual levels of LAL not only the understanding on the central concepts of LAL and their use in practice [16]. In addition, a study by R. Lam [17] also revealed that the teachers participated in his study reported to have basic level on LAL despite their extensive work experience. All teachers' LAL levels in those aforementioned studies were based on the self-rating questionnaire that the LAL levels were only perceived by the teachers themselves and this proves that teachers still needed to improve their LAL.

The need of teachers to improve LAL in order to improve teachers' practice in assessment was validated by study that sought out about teachers' LAL and their assessment practices. A study by S. A. Kalajahi and A. N. Abdullah [18], for example, reported that there was a significant correlation between teachers' assessment beliefs and their assessment practices. This finding indicated that the higher level of teacher LAL, the better their assessment practices. A study conducted by M. Mellati and M. Khademi [19] also revealed that there was a significant difference of classroom practice between teachers who had good assessment literacy level (assessment literate teachers) and those who had low assessment literacy level (assessment illiterate teachers). The assessment literate teachers were able to set teaching goals based on students' interests and needs, use dynamic assessments through assignments, and provide feedback to students to monitor their progress in learning. On the other hand, assessment illiterate teachers in their study only focused on conventional paper and pencil test at the end of the course which enabled them to monitor the students' progress during the course. Subsequently, these teachers' teaching practice was less dynamic because instructional activities were merely to prepare students for end of term test.

In addition, study on teachers' LAL revealed a need for training for teachers in developing their LAL. For example, N. Sultana's [15] study indicated that the teachers' expertise as a language testing practitioner was developed based on suggestion from colleagues and experiment. This study also suggested that teachers should attend more training about language assessment to equip them with the dynamic challenging of language assessment and with up-to-date assessment practices in order to improve the teachers' LAL. A study carried out by F. Giraldo [20] also recommended that the teachers in this study could take an advantage in participating in a professional development course that focus on not only theoretical but also practical aspects of language assessment. More importantly, program for developing teachers' LAL should also consider contextual factors as elements that can foster teachers' LAL.

The proposal to integrate contextual factors in developing teachers' LAL also suggested by X. Yan, C., Zhang, and J. J. Fan [21] whose study suggested that LAL training for teachers should pay attention to the context of assessment-teaching. They argued that assessment trainers should examine teachers' current assessment practice and match the training to the need. Relatedly, T. Firoozi, K. Razavipour, and A. Ahmadi's [22] study revealed that the reformed of assessment policy in their study context have created a gap between the teachers' assessment practice and the new assessment policy. Hence, it recalled for taking a new policy into consideration in designing training for developing teachers' LAL. In order to do that, the teachers' LAL needs to be identified.

III. METHODOLOGY

A mixed method design was used in this study to understand teachers' LAL in a higher education context in Indonesia. The use of mixed method in this study aimed at gaining a more comprehensive understanding about teachers' LAL, because, according to J. W. Creswell and J. D. Creswell [23], by comparing different perspectives drawn from

quantitative and qualitative data, a study can gather a more complete understanding of study problems and questions.

Both quantitative and qualitative data were obtained for this study in two phases. The first phase was obtaining quantitative data through a survey using a self-rated questionnaire about teachers' LAL. After that, the quantitative data were analyzed, and the result were used to plan the second phase, i.e., the qualitative data phase. The qualitative data were collected through interviews to explain them in more details about the quantitative data result. According to J. W. Creswell and J. D. Creswell [23], this type of mixed method study which involves two-phase design is called "exploratory sequential mixed method design". The aim of using both quantitative and qualitative data was to justify the reliability of the study by obtaining data for multiple sources, so that in-depth perspectives toward LAL could be obtained.

A. Participants

The participants in this study were teachers in a language faculty in a private university in Indonesia. Out of 39 teachers in this faculty, 30 completed the questionnaire of LAL. Five teachers volunteered to participate in the interview stage. Two of the five teachers had experience in teaching a language assessment course, while the other three did not. With this different experience, more comprehensive data could be obtained to understand teachers' LAL. The participants' study background was either masters' degree or doctorate degree in language education, general education, or linguistics. They had at least a seven year experience in teaching, and were aged ranging from 28 – 55 years old.

Instruments. The instrument used for gathering the quantitative data was a five-point Likert-scale questionnaire developed by F. Giraldo [7]. This questionnaire consisted of sixty-six items, and it is a self-rating questionnaire where participants rated their own LAL level. The questionnaire measures teachers' LAL components such as knowledge, skills, and principles on language assessment. Each component had different dimensions and descriptors; therefore, it has 8 dimensions and 66 descriptors. The overall description of each component, dimensions and descriptors is shown in table 1 below.

TABLE 1 THE COMPOSITION OF THE QUESTIONNAIRE

LAL Components	LAL Dimensions	Number of Descriptors
Knowledge	Awareness of applied linguistics	4
	Awareness of theory and concepts	15
	Awareness of own language assessment context	5
Skills	Instructional skills	11
	Design skills for language assessments	12
	Skills in educational measurement	6
	Technological skills	3
Principles	Awareness of and actions towards critical issues in language assessment	10

The questionnaire was developed by F. Giraldo [7] based on related literature and previous studies on assessment literacy from both general education experts and language education scholars. A core list of assessment knowledge, skills, and principles was developed to make the questionnaire

visible for the construct and content validity. To ensure the content validity of the questionnaire, two experts experienced in the field of language assessment and language education were requested to comment on it. The comments were used to revise the questionnaire accordingly. As for the reliability of the questionnaire, the Cronbach's alpha was used to measure the reliability with a score of 0.980 which means that this questionnaire had high level of internal consistency.

After the questionnaire, it was followed by interviews to five respondents face to face and lasted around 30 – 45 minutes. It was a semi structured interview that allowed the participants to express their opinion openly and freely. The interview guide included 10 questions about their knowledge, skills, and principles on language assessments, and additional follow up questions were added for clarifications. The interviews were voice recorded for data analysis purposes.

B. Data Analysis

For analyzing the quantitative data, descriptive analysis of the data was employed. Means and standard deviations were used to calculate each descriptor. Subsequently, the means were ranked from highest to lowest for each dimension so the level of LAL that the teacher perceived could be seen. The level of LAL used in this study was adapted from C. Lan and S. Fan's [16] study about the literacy level, i.e., 1) very poor: teachers are unfamiliar with language assessment concepts and methods; 2) poor: teachers understand specific terms and concepts related to language assessment, but may indicate misconceptions; 3) Fair: teachers have a wide range of basic concepts and terms related to language assessment; 4) good: teachers understand central concepts of language assessment, and are able to use the knowledge in practice and 5) excellent: teachers understand central concepts of language assessment, and are able to use the knowledge in practice.

For analyzing the interview data, there were some steps employed in this study. The inductive thematic analysis was employed, so that themes could be identified from the data. After the interviews were transcribed, each interview transcript was analyzed to find initial coding. After that, the codes and themes were reviewed to select the ones that best described and answered the study question. Finally, after the transcript was analyzed thoroughly, the final themes that corresponded to the study question were chosen and presented as the findings answering the study questions.

IV. FINDINGS

The level of teachers' LAL was determined by the quantitative data already obtained using the 66-item questionnaire about the teachers' LAL. From the quantitative data analysis, 30 participants in this study rated their overall LAL level at the average of 3.381 which fell into fair category (Table II).

TABLE II GENERAL LAL LEVEL

LAL Level	Percentage
Excellent	0
Good	26.7
Fair	60.0
Poor	13.3
Very poor	0

There were no teachers who rated themselves to have both excellent and very poor LAL level. A total of 26.7% of the respondents rated themselves to have good level of LAL, while 13.3% claimed to have poor level of LAL. The majority of the respondents were categorized into a fair level showing their perception of having a sound basic knowledge, skills, and principles on language assessments, but not in the assessment practices in their teaching.

The questionnaire divided the LAL into three components, i.e., knowledge, skills, and principles. Among those three components, 'principles' (Mean 3.964) and skills (mean 3.248) were the highest and the lowest, respectively.

Each component of LAL has its own dimensions, and overall, there are eight dimensions used in the questionnaire. Among the eight components, the teachers showed the highest confidence level in awareness of and actions toward critical issues in language assessment (Mean 3.968; SD .570). On the contrary, based on their perception, their skills related to measurement were the lowest (Mean 2.607; SD 1.166).

Using the information from the quantitative data, the interview protocol was designed to get more detail explanation about the teachers' LAL. Subsequently, the data from the interviews discussed important findings from the questionnaire related to how teachers perceived their knowledge, skills, and principles on language assessments explored further in the next section.

A. Teachers' Knowledge

The data from the interview revealed that the teachers had a sufficient knowledge on applied linguistics, for example, their awareness of the aspects in second language acquisition and their impacts on language assessment. One participant stated that the most important aspect in designing language assessment was the washback to students' motivation in learning. Another participant also agreed that students' motivation and learning style were two important factors that should be considered in designing language assessment. The remaining participants also revealed that they were also aware of different approaches in language teaching and their impact on the assessment; that is why all participants expressed that they used multiple assessment methods such as test-based assessments and performance-based assessments in their teaching to capture a more comprehensive understanding to student performance.

Besides, participants also expressed that they always made informed decision on the methods of assessment they used in their class paying attention to which assessment that could fulfill the criteria of language assessments such as validity, reliability, practicality, authenticity, and washback. In addition, they were able to evaluate tests and assessment policies by making adjustment in assessment models to suit their needs.

Both quantitative and qualitative data presented participants' poor ability to do statistical analysis for assessment. A participant claimed that she never did statistical analysis on the validity and reliability of the test she created. Another participant supported the importance of statistical analysis for assessment despite her poor ability in statistics.

B. Teachers' Skills of Language Assessment

Among the dimensions of the skills in language assessment, based on the results of the interview, the teachers believed to be good at collecting data for students' language development as reported by a participant that she used both test and other alternative assessments such as observation, interview, etc. to see the students' language development. Another participant argued that by employing different methods of assessment, he wanted to give opportunity to students to succeed in the assessment. However, in descriptor no. 27 that participants are able to give feedback on student assessment, the participants claimed that they had difficulties in giving feedback to students in timely manner because of the workload they had.

However, for the component of skills on language assessment, the participants perceived low ability, especially skills to use statistical analysis. A participant argued that her departments' policy on assessment was to use performance-based assessment for assessing students, "that is why I do not feel the need to have those skills related to test-based assessment". The descriptors in this questionnaire for the dimension of skills of educational measurements indeed discussed about interpreting descriptive statistic data, tests' psychometric analysis, item difficulties, item discrimination, and calculating tests' validity and reliability statistically. Although the participants mentioned that they paid attention to the validity and reliability of assessment, they claimed they never statistically calculated them.

C. Teachers' Principles of Language Assessment

The principles of language assessment discussed about aspects of critical language assessment where teachers should be able to see issues of fairness, power relation, transparency, and effects of language assessments. For this dimension, the participants reflected that they always paid attention to students' feedback on the assessment, and they were also willing to adjust the language assessment based on the feedback from students. One participant stated that it was her way to show respect to their students in terms of language assessment. Besides, the participants also claimed that they provided assessment that was fair for all students and avoided discriminating students, because they were aware that this was an important aspect of language assessment.

When being asked about their ability to criticize external tests both in terms of quality and in term of impact of the assessment toward the students, the participants stated that they were well aware about the quality and the impact of standardized test to students. One participant expressed that a standardized test like National Examination or TOEFL test had negative impact for students because of how high the stake it is for the test takers. Other participants also agreed that assessment/test should not give negative impact for the students since it would also give bad impact to learning. Another participant also claimed that he would adjust the assessment method if it gave negative impacts to students such as making students nervous, giving too much workload to students, etc.

V. DISCUSSION

This study attempted to investigate the level of teachers' LAL and how teachers perceived about their LAL. For the level of teachers LAL, this study revealed that university

teachers participated in this study perceived their LAL as in fair level where they had adequate knowledge, skills, and principles to do assessment-related tasks. The result of this study concerning teachers' LAL is similar to the finding of the study carried out by Y. Xu and G. T. Brown [10]. However, they argued that fair level of teachers' assessment literacy was not sufficient for university teachers since their responsibility on assessment was enormous. A study conducted by R. Lam [17] also revealed the same result that the participants in his study reported to have moderate LAL level. A study by N. Sultana [15] even revealed that his participants' LAL were very low.

Although this current finding of the study cannot be generalized because of the small sample size, this finding adds evidence that teachers' LAL is still become a global issue [17]. Besides, the interview data also supported that, based on the teachers' perception, they still needed to have training to improve their LAL. The needs for additional training in LAL have widely been voiced in different educational contexts [24],[21],[17],[25],[22],[15].

Concerning how teachers' perception on their knowledge, skills, and principles on language assessment, the study indicated that, in terms knowledge component, the teachers reported to have fair knowledge on all three dimensions such as knowledge on applied linguistics, theories and concepts, and assessment context. Both the quantitative and qualitative data showed the same result. Among the three dimensions, the teachers rated highest on their knowledge on applied linguistics, and knowledge on theories and concepts of language assessment as the lowest. This may indicate that the teachers were more confident in their knowledge about general applied linguistics such as language teaching methodology, trends in second language acquisition, etc. than theories and concepts of language assessment itself. This finding is in line with X. Yan, C., Zhang, & J. J. Fan's [21] study that their participants showed little interest in assessment concept and theories.

The findings related to skills component were somewhat similar to the knowledge component that the teachers reported to have better instructional skills such as using multiple method in language assessment, using assessment result for improving teaching activities, etc. than skills in educational measurement. In the skills component, however, there was a difference between the quantitative and qualitative data in the descriptor about giving students' feedback. Based on the survey, it was the highest in this dimension, but the interviews showed that giving feedback was one of the challenges they faced. When asked to clarify this mismatched between the survey and the interview data, one participant said, "I understand that students need the feedback from assessment for their learning and I know how to give feedback, but I often miss the deadline because it take me forever to read the students' works". This mismatch may indicate the mismatch between teachers' belief and their practice in their teaching.

In addition, for the component of skills on language assessment, the lowest mean score was shown by the dimension of skills in educational measurements. When asked about why it was poor, most of them said that their past training did not include subject on statistics, and their current assessment context did not demand such practice. The poor ability in education measurement skills is similar to the study

by F. Giraldo [20] that the participants in his study were also having difficulty related to educational measurement such as item construction and statistical calculation on validity and reliability. Similarly, based on a study by N. Sultana, [15] non-exact science teachers lacked the ability to analyze test items statistically.

Finally, among the components of LAL, the principle of language assessment component was the highest. The teachers reported that they are aware of issues in critical language assessment and took action in order toward unfairness and unethical issues in language assessment, and they rated the highest their ability to treat all students with respects. Both the quantitative and qualitative data confirmed the finding. This finding is similar to the findings of a study by V. Berry, S. Shehaan, and S. Munro [25] that the participants in their study consistently analyzed the impact of the test-based assessment toward their students. However, this finding is conflicting with N. Sultana's [15] study that the teachers in his study were not aware of the quality of a standardized test and its impact to the students. The difference finding regarding the principle component can be caused by the teaching context where the teachers in the current study and V. Berry, S. Shehaan, S. Munro's [25] study taught in university level where teachers had more independent in deciding their assessment. For N. Sultana's [15] participants, they taught in secondary levels where in many contexts, the assessment policy was often regulated by the government so that teachers had only limited independence to set assessment.

VI. CONCLUSION

Although the discussion about teachers' LAL has been around for quite some time, the concerns about teachers' LAL are still valid. Teachers in this study reported to only have fair level of LAL (Mean: 3.381) which was not adequate to help them do assessment-related tasks. The teachers also reported to have more knowledge and skills in teaching, but less knowledge and fewer skills in educational assessment. This may be caused by the previous training in the teacher education where courses were more emphasized on teaching and learning than the assessment. Besides, teachers also claimed that they developed their LAL through learning from colleagues, on the job experience, and predetermined assessment policy from the institutions. Thus, in order to develop teachers' LAL, more training focuses on developing teachers' knowledge and skills in language assessment is needed. In addition, the teachers in this study were also aware that they needed more training for language assessment, so that they could make informed decision about their assessment practices. Therefore, the needs for training programs that chatter teachers' need in their teaching context is emergent.

REFERENCES

- [1] R. J. Stiggins, Facing the challenges of a new era of educational assessment. *Journal of Educational Measurement*, 1991, 4: 263-273.
- [2] R. Stiggins, New assessment beliefs for a new schools mission. *Phi Delta Kappan*, 2004: 22-27.
- [3] T. Newfields, Teacher development and assessment literacy// T. Newfields, M. Glendall, M. Kawate-Mierzejewska, M. Ishida, M. Chapman, & P. Ross (Eds.) *Authentic Communication: Proceedings of the 5th Annual JALT Pan-sig Conference Shizuoka*. JP: Tokai University College of Marine Science, 2006: 48-73.
- [4] S. M. Brookhart, Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 2011, 30: 3-12.
- [5] Y. Xu, & G. T. Brown, Teacher Assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 2016, 58: 149-162.
- [6] M. L. Stable-Havener, Defining, conceptualizing, problematizing, and assessing language teacher assessment literacy. *Language Teacher Assessment Literacy*, 2018, 18: 1-22.
- [7] F. Giraldo, Language assessment literacy: Implications for language teachers. Profile: Issues in Teachers' Professional Development, 2018, 20: 179-195.
- [8] A. E. Yastibaş & M. Takkaç, Understanding language assessment literacy: Developing language assessments. *Journal of Language and Linguistic Studies*, 2018, 14: 178-193.
- [9] T. Newfields, 2018 [NO REFERENCE]
- [10] Y. Xu, & G. T. Brown, University english teacher assessment literacy: A survey-test report from china. *Papers in Language Testing and Assessment*, 2017, 6: 133-158.
- [11] A. H. Prasetyo, Language assessment literacy development: A student-teacher's experiences in teaching practice program. *LLT Journal: A Journal on Language and Language Teaching*, 2018, 21: 199-206.
- [12] A. Scarino, Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, 2013, 30: 309-327.
- [13] O. Inbar-Lourie, Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 2008, 25: 385-402.
- [14] A. Davies, Textbook trends in teaching language testing. *Language Testing*, 2008, 25: 327-347.
- [15] N. Sultana, Language assessment literacy: An uncharted area for the english language teachers in bangladesh. *Language Testing in Asia*, 2019, 9: 1-14.
- [16] C. Lan, & S. Fan, Developing classroom-based language assessment literacy for in-service EFL teachers: The gaps. *Studies in Education Evaluation*, 2019: 112-122.
- [17] R. Lam, *Teacher assessment literacy: Surveying knowledge, conceptions*. System, 2019.
- [18] S. A. Kalajahi, & A. N. Abdullah, Assessing assessment literacy and practices among lecturers. *Pedagogika*, 2016, 232-248.
- [19] M. Mellati, & M. Khademi, Exploring teachers' assessment literacy: Impact on learners' writing achievements and implications for teacher development. *Australian Journal of Teacher Education*, 2018, 43: 1-18.
- [20] F. Giraldo, Language assessment practices and beliefs: Implications for language assessment literacy. *HOW*, 2019, 26: 35-61.
- [21] X. Yan, C. Zhang, & J. J. Fan, "Assessment knowledge is important, but...": How contextual and experiential factors mediate assessment practice and training needs of language teachers. *Elsevier Journal*, 2018: 158-168.
- [22] T. Firoozi, K. Razavipour, & A. Ahmadi, The language assessment literacy needs of iranian efl teachers with a focus on reformed assessment policies. *Language Testing in Asia*, 2019, 9: 1-14.
- [23] J. W. Creswell, & J. D. Creswell, *Study Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Los Angeles: SAGE Publication, Inc, 2018..
- [24] R. Lam, Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 2015, 32: 169 - 197.
- [25] V. Berry, S. Shehaan, & S. Munro, What does language assessment mean to teachers. *ELT Journal*, 2019.