Increasing transparency in assessment to improve students' learning at Language Development Centre of UIN Suska Riau

Dodi Settiawan[a], and Ridho Hilmawan[b]

[a]Faculty of Education and Teacher Training, English Education Department
UIN Sultan Syarif Kasim Riau
[b]Language Development Center of UIN Sultan Syarif Kasim Riau

Abstract: In the foreign/second language classroom, assessment tasks are not just created to measure and document students' performance at a given point in time. Effective assessment should be done through activities and tasks that teachers use to gather evidence and provide feedback about how well the students are performing. Therefore, the results of assessment can inspire the students to improve their learning. This paper is a work-in-progress report that describes several phases in the assessment process that were implemented at Language Development Center of UIN Suska Riau. This paper explores the key concepts for planning an assessment that is useful, valid, reliable, practical, and authentic. The application of transparency in assessment is also presented to provide practical guidance in sharing information about assessments before they occur using the Language Development Center's website. Ultimately, it is hoped that increasing transparency provides students of Language Development Center with chance to perform better or improve their learning because they receive a clear explanation of how the skills, vocabulary, and grammar that they are going to learn will be assessed.

Keywords: language assessment, key assessment concepts, authentic assessment, transparency in assessment.

1. INTRODUCTION

As part of evaluation, assessment in the foreign/second language classroom can be formative or summative. The terms formative and summative are differentiated based on when they are given. Formative assessment is given during the instruction or course while summative assessment is given typically after a semester or at the end of instruction or course. Moreover, formative assessment is not just a way of collecting and documenting information in order to judge students' learning or achievement. In a formative assessment, teachers gather evidence about how well their students are performing. Then, they always provide their students with feedback that contains information for them to focus on and improve. Both types of assessments are often based on different types of instructional tasks.

Assessment tasks and instructional tasks need to be aligned in order to improve students' learning. The instructional tasks are kinds of activities in or out of class that prepare students for assessment tasks. Besides, the assessment tasks are kinds of activities that reveal how well students have learned during or at the end of the instructional tasks. When they are misaligned, students will not learn or practice the skills that

are going to be assessed or the tasks will not measure what they have learned. As a result, the misaligned cannot inspire the students to improve their learning. In other words, they are interchangeable in order to improve students' learning. Students' learning is improved when assessment task is implemented effectively (Campbell and Collins, 2007). In order for the assessment task to be effective, Language Development Center of UIN Sultan Syarif Kasim (Suska) Riau had implemented several phases in its assessment process. This paper explores the institution's key concepts for planning an assessment that is useful, valid, reliable, practical, and authentic. In addition, the application of transparency in assessment to improve students' learning is also presented to provide practical guidance in sharing information about assessments before they occur using the Language Development Center's website.
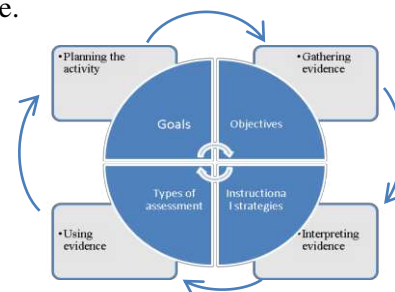
Figure 1. The Assessment Process (NCTM, 1995, p.4)

2. THE ASSESSMENT PROCESS

Several phases in a process of designing effective assessment are needed to help students improve their learning. The Language Development Center adapted the phases of assessment proses as showed in the following figure.

Figure 1 illustrates that assessment has cyclical process. The cycle consists of four phases, namely planning the assessment task, gathering evidence, interpreting evidence, and using the evidence. Teachers must align the phases with course goal or purpose, learning objectives, instructional strategies, and types of assessment. The process is implemented as a guide for creating and developing assessment tools that can improve both classroom instruction and students' learning at the Language Development Center. However, there some key concepts are needed for implementing the process appropriately. The concepts are discussed separately below. However, each concept has connection to support one another. Therefore, they need to be applied together for planning an effective assessment.

Key Concepts for Planning an Assessment

To plan an assessment, there are some concepts that help to ensure the assessment is effective. An assessment is effective when it can consistently measure what teachers want it to measure in an efficient manner. The concepts also help the teacher and students to use the assessment as a valuable source of information regarding learning. The concepts that have been applied to form the basis for planning assessments at the Language Development Center are usefulness, validity, reliability, practicality, and authenticity.

Usefulness

When choosing or designing a test, as an assessment tool, usefulness must be considered (Bachman and Palmer, 1996). The way to ensure the assessment tool is useful is by considering the purpose of the tool before administering it. Moreover, Rogier (2014) explains that to design a test that is useful, teachers must consider the purpose of the test, the group of test takers it is

designed for, and the specific language use the teachers want to evaluate. Therefore, teachers at the Language Development Center of UIN Suska Riau have to match their assessment tool to the specific purpose, the specific language use they want to assess, and their students.

Test		Purpose of the test
Final Test	**Written**	The test measures students' ability to: utilise language use that is appropriate for standard written English and understand academic texts related with the topics they have learned during the course.
Final Test	**Spoken**	The test measures students' ability to use spoken English with confidence in an academic setting and related with the real-life language use they have learned during the course.

Figure 2. Examples of summative assessment tools and their purposes

Figure 2 shows two examples of summative assessment tool used at the Language Development Center that are purposed to measure students' achievement at the end of an English course. By considering the purpose of the test, teachers can choose or design appropriate materials that match what they expect their students to encounter. For example, if a teacher wanted to measure students' ability to understand academic texts, he could not merely use any reading test from the Internet. He would need to be aware in selecting materials that are related with the topics they learned during the course in order to meet the vocabulary that the students have been immersed in. Moreover, the types of questions must match to the reading skills that students had practiced in class.

Validity

Validity refers to "the extent to which inferences made from assessment results is appropriate, meaningful, and useful in terms of purpose of the assessment" (Gronlund, 1998, p. 226). This concept maintains the accuracy of an assessment in which a test has validity when it measures what it is designed to measure. The most important aspect of validity is the appropriateness for the context and meaningful and useful for the test takers and developers. Therefore, in designing a test, a teacher as a test developer should think about what is to be gained by administering the test and how the information will be used. For example, a teacher's goal of assessment is to measure students' reading comprehension skills. Then,

the students are given a test in which they answer questions in written format about a passage they read.

In that case, the vocabulary, sentence structure, and grammar usage in the passage and written questions must be surely not beyond the level of the students. If not, they will be tested on more than just their reading comprehension skills. As a result, the validity of the test as a measure for listening ability is decreasing. According to Rogier (2014), a number of factors can have an adverse effect on validity, including the following:

1. unclear directions
2. test items that ask students to perform at a skill level that is not part of the course objectives
3. test items that are poorly written
4. test length that doesn't allow for adequate sampling or coverage of content
5. complexity and subjectivity of scoring that may inaccurately rank some students

The best way to ensure validity and reliability is to create test specifications and exam blueprints. These will help ensure that tests created and used match what is intended for the course and the students. Figure 3 shows an example of an exam blueprint of a final exam for English-language program course at the Language Development Center.

Reliability

One thing to keep in mind is that a valid test needs to be reliable. Reliability in an assessment refers to "the consistency of the assessment in producing the same score on different testing occasions or with different raters" (O'Malley & Pierce, 1996, p.19). For that reason, a test that can measure what a teacher would like it to measure must produce similar scores consistently. The consistency of test scores is when a student is tested more than once using the same test and the results of the tests are the same. it means the test scores are consistent.

According to Rogier (2014), reliability can be threatened by fluctuations in the learner, in scoring, or in test administration. Fluctuations in the learner are out of the testing administrator's control. For example, teachers cannot control whether a student is sick, tired, or under emotional stress at the time of a test. However, teachers can limit the fluctuations in scoring and test administration. The guidelines for how a test is administered, the length of time allotted to complete the test, and the conditions for testing should be established in advance and written in a test-specifications document. As much as possible, there should be consistency in testing conditions and in how a test is administered each time it is given. Teachers can minimize fluctuations in score by preparing answer keys and scoring rubrics, and by holding norming sessions with those who will be scoring the test.

To improve the reliability of a test, the test should be long enough to sample the content that students are being tested on. The test should be practical in which it provides enough time for most of the students to finish taking the test. Moreover, the items should not be too easy or too difficult, the questions should not be tricky or ambiguous, the directions should be clear, and the score range should be wide. Before administering the test, someone else is needed to take it to see whether he or she encounters problems with directions or content. Then, the person's feedback can be used to see whether the test needs to be improved or not.

Practicality

Practicality refers to how a given test is inexpensive, easy to administer, and efficient in the case of time and scoring procedure (Brown, 2004). Practicality is one of the concepts that help to ensure the assessment is effective. According to Rogier (2014) the issues of practicality include the cost of developing and maintaining the test, time needed to administer and mark the test, ease of marking, availability of suitably trained markers, and administration logistics. There will obviously be a practicality issue with the delivery of the test if the test that is going to be administered to students requires internet connection but the electricity is not available.

The time required for scoring a test always becomes an important practicality issue for

teachers. This issue is important because it is related to what they are testing and resources. For example, in administering a writing test for a group of students, a teacher will consider how to score the test and how much time he needs to score each writing piece analytically. If the test is only a formative assessment, the teacher might apply peer-assessment activities (e.g. asking his students exchange papers and mark them instead of marking each one himself). As a result, consideration of practicality for marking writing helps teachers to think that it is more practical to have students review each other's work and peer edit the first draft than to have the teacher make comments on each initial draft. In conclusion, the practicality in assessment helps teachers to design assessment that can improve students' learning because, as in the example, the consideration to apply the peer-assessment activities allows the students to review the materials.

Authenticity

An assessment is authentic if the tasks of assessment are based on activities that represent real-life settings and classroom (O'Malley & Pierce, 1996). Authentic assessment is designed to challenge the students to use their language to communicate their understandings and applications of knowledge in real-world situation. Therefore, designing assessment tasks that are relevant to real-life contexts in which the language will be used must be considered by teachers or test developers.

Authentic assessment can be designed as performance assessment, portfolios, and students- self assessment. The tasks in the assessment can provide motivation and attitude that improve learning achievement because they reflect real-world situations and contexts in which the language will be used. However, the real-world situations must depend on students' age. For example, if a course is designed for university students who will be a presenter in a seminar, an oral presentation task that use presentation media such as Prezi® or PowerPoint® would be more authentic than a task in which students are asked to watch a video of a seminar and respond to questions

related to the video, or one where the students write the correct forms of verbs in sentence blanks. In conclusion, authentic assessments should be aligned with the purpose of the course or course objectives and instructional tasks that are based on real-life settings.

Transparency in Assessment

Usefulness, validity, reliability, practicality, and authenticity need to be supported with transparency. Transparency refers to the teachers' efforts to provide information to students at least related to what and how they will be assessed. Thus, students will be aware of the skills, vocabulary, and grammar that they will be expected to learn and they receive a clear explanation of how these will be assessed (Rogier, 2014). In other words, transparency in assessment puts the students as a part of testing process. Students are helped to understand what learning outcomes they need to be able to do and what will be tested. They are also helped to understand the test format and how to grade their answers or performance.

Moreover, through transparency, students have the chance to prepare for question types that will be used in a test that they are going to take. Transparency is related with reliability of a test. When a teacher uses a new test format in which his students are not familiar with the format, the test's reliability will be reduced. Increasing transparency will also reduce students' test anxiety and allow them the chance to perform better. To increase transparency, many schools and educational institutions publish their test specifications. For example, the Language Development Center publishes test specifications and blue print of a final exam for English-language program course on its website. Moreover, the scoring rubric for the spoken final exam (see Figure 4) is also published to make the students aware of the skills they will be assessed at the end of a course.

PROSIDING ICTTE FKIP UNS 2015
Vol 1, Nomor 1, Januari 2016
Halaman:

ISSN: 2502-4124

ICTTE
FKIP UNS 2015

CATEGORY	4	3	2	1	0	POINT
Comprehension	Student is able to accurately answer almost all questions posed by examiner about the topic.	Student is able to accurately answer most questions posed by examiner about the topic.	Student is able to accurately answer a few questions posed by examiner about the topic.	Student is unable to accurately answer questions posed by examiner about the topic.	Unable to answer.	0
Vocabulary	Uses vocabulary appropriate for the audience. Extends audience vocabulary by defining words that might be new to most of the audience.	Uses vocabulary appropriate for the audience. Includes 1-2 words that might be new to most of the audience, but does not define them.	Uses vocabulary appropriate for the audience. Does not include any vocabulary that might be new to the audience.	Uses several (5 or more) words or phrases that are not understood by the audience.	Unable to answer.	0
Content	Shows a full understanding of the topic.	Shows a good understanding of the topic.	Shows a good understanding of parts of the topic.	Does not seem to understand the topic very well.	Unable to answer.	0
Uses Complete Sentences	Always (99-100% of time) speaks in complete sentences.	Mostly (80-98%) speaks in complete sentences.	Sometimes (70-80%) speaks in complete sentences.	Rarely speaks in complete sentences.	Unable to answer.	0
Speaks Clearly	Speaks clearly and distinctly all (100-95%) the time, and mispronounces no words.	Speaks clearly and distinctly all (100-95%) the time, but mispronounces one word.	Speaks clearly and distinctly most (94-85%) of the time. Mispronounces no more than one word.	Often mumbles or can not be understood OR mispronounces more than one word.	Unable to answer.	0
					Final Point	0

Figure 4. Oral examination rubric

To date, the transparency has positive impact on students' learning at the Language Development Center. They are provided with chance to practice the test before they take the final exam. As a result, most of them can perform better. Through transparency, they are empowered as learners instead of instilled fear of tests (Rogier, 2014). In addition, integration of media technology in assessment is needed to increase transparency at the Language Development Center. For example, the institution's website provides students with clear explanation of how the skills, vocabulary, grammar that they have learned will be assessed.

3. CONCLUSION

The way to develop assessment skills is not different with the way of developing other skills in our real life. The development just needs extensive and continuous practices. By implementing the assessment process appropriately, analysing whether the tests are useful, valid, reliable, practical, and authentic and increasing transparency we can all become better teachers and test developers. For us, as teachers, skills in creating assessment tools can be developed through real-life teaching in the classroom. The classroom provides us with perfect opportunity to apply assessment fundamentals and to develop assessments that will motivate our students to learn and that will

also motivate us to develop lessons that will facilitate learning.

To sum up, assessment plays an important role in the teaching and learning processes. By reflecting on the tests used in our own classroom, how they are developed, and the results obtained from them, we will become more assessment literate and a better teacher. Testing will not seem so complex or beyond comprehension for students if teachers provide transparency in their assessment. It is because transparency provides students with a better experience, allowing them to improve their learning without making them fear to take the tests.

4. REFERENCES

Bachman, L. F. & Palmer, A. S. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.

Boyles, P. (2006). Assessment literacy. In *New visions in action: National assessment summit papers*, ed. M. Rosenbusch, 18–23. Ames: Iowa State University.

Brookhart, S. M. (2001). The standards and classroom assessment research. Paper presented at the Annual Meeting of the American Association of Colleges for Teacher Education.

Brown, H. D. (2004). *Language Assessment: Principles and Classroom Practices*. New York: Pearson Education.

Campbell, C., & Collins, V. L. (2007). Identifying essential topics in general and special education introductory assessment textbooks. *Educational Measurement: Issues and Practice* 26 (1): 9–18.

Gronlund, N. E. (1998). *Assessment of Students Achievement*. 6th edition. Boston: Alyn and Bacon.

Mertler, C. A. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education* 33 (1): 49–64.

National Council of Teachers of Mathematics. (1995). *Assessment Standards for School Mathematics*. Reston, VA: Author.

O'Leary, M. (2008). Towards an agenda for professional development in assessment. *Journal of In-service Education* 34 (1): 109–114.

O'Malley, J. M. & Pierce, L. V. (1996). *Authentic Assessment for English Language Learners.* New York: Addison-Wesley Publishing Company.

Rogier, D. (2014). Assessment literacy: Building a base for better teaching and learning. *English Teaching Forum* 3: 2