Constructive alignment of assessment practices in English language classrooms

Lynrose Jane Dumandan Genon a,1,*, Chezka Bianca P. Torres b,2

a, b Mindanao State University-Iligan Institute of Technology, Andres Bonifacio Ave, Iligan City, 9200 Lanao del Norte, Philippines
1 lynrosejane.genon@g.msuiit.edu.ph, 2 chezkabianca.torres@g.msuiit.edu.ph
* corresponding author

ARTICLE INFO

Article history
Received 04 July 2020
Revised 17 August 2020
Accepted 14 December 2020

Keywords:
language assessments
language classrooms
constructive alignment

ABSTRACT

This qualitative study identified the language assessment practices in terms of purpose, type, and timing in four elementary language classes in the Philippines. It then evaluated the constructive alignment and content validity of the assessment and described how the constructive alignment reflects the quality of teaching and learning in these language classrooms. Findings revealed that the assessment practices are entrenched in the teachers’ pedagogy serving various purposes (like monitoring, facilitation and motivation) other than evaluation. Generally, the assessment content and tasks show alignment to targeted learning goals. Several assessments, however, do not target the intended skill but rather tapped on the sub-skill, and they reveal alignment issues as well as teaching-learning conditions in the classrooms. Findings uncover, albeit a snapshot, of the gaps not only in assessment practices, but also in the articulation of curriculum goals leaving teachers to interpret and operationalize these on their own. These issues affect the students’ achievement since alignment affects achievement and opportunity to learn is at optimum levels when lesson plans are aligned with benchmarks, standards, and assessments, thus increasing academic achievement.

This is an open access article under the CC-BY-SA license.

1. Introduction

While teaching and learning are fundamental areas in language acquisition, the picture is not complete without assessment, which serve as mediating channels between the lesson and the learner. Assessment should be in the limelight because it provides a basis for implicit or explicit judgments. It determines whether or not the goals of teaching are met. It affects decisions about grades, advancement, syllabus, curriculum, and instruction. It evaluates how much of the teaching is taken in by learners and if the students are learning what they are supposed to learn. It has an indispensable influence and role in language classrooms as it affects both teaching and learning. A study on assessment vis-à-vis learning and teaching will allow for a more insightful investigation of significant classroom phenomena (Alderson & Banerjee, 2001). The power of assessment in transforming language learning and teaching may be more potent than expected (Jin, 2017).

The study examines the strength and usefulness of assessment in second language teaching and learning and infers the views and principles practiced by language teachers in local classrooms through their assessment practices. Its main focus is evaluating whether the assessment task targeted the intended skill in the learning objectives through looking at its content validity. The study examined the assessment practices in four elementary English language classrooms to see how these reflect the language teaching and learning. It specifically sought answers to the following questions:
What are the purposes, types, and timing of the language assessment practices investigated? What is the quality of the constructive alignment and the content validity of the language assessment practices investigated? How does the quality of constructive alignment of the assessment practices reflect teaching and learning in these language classrooms?

1.1. Language Learning in the Philippines

The English language education in the Philippines is frequently faced with low academic performance among Filipino students. Despite the extensive effort of the Department of Education in the learning of the English language, pupil performance was generally low in subjects that relied on English (Bautista, Bernardo, Ocampo, 2009). In the Program for International Student Assessment (PISA), the Philippines ranks last among 79 countries in reading comprehension, and also ends up in the low 70s in mathematics and science. Low functional literacy skills is apparent in the assessment results (PISA 2018 results, 2019). This problem can be attributed to many factors, such as poverty, lack of school facilities, learning resources and qualified teachers who can teach English in diverse contexts, languages, levels, and regions (Madrunio, Martin, & Plata, 2016). Bautista, Bernardo, Ocampo (2009) also suggested that this is due to poor teacher quality, and irrelevant learning materials in the classroom. Looking into the said language learning problem and failure in achieving English competence as ideally projected in the curriculum’s aim, it is imperative to investigate the language classroom and examine the major components of language classroom, teaching, learning and assessment, to understand the problem in the grassroots.

1.2. Assessment and the Teaching-Learning Process

In language classroom, teaching and learning and assessment share a cyclic relationship. Assessment is an integral component of teaching and learning. Its influence on the teaching and learning process makes it a crucial component of school improvement (Banda, 2005). Assessment for learning guides and facilitates the teacher in monitoring and empowering students in their learning, in a way assessment supplements teaching and learning. Assessment can also evaluate and at the same time improve teaching and learning. Ideally, knowing how to assess students in order to improve instruction is a core principle of effective teaching (Bautista, Bernardo, Ocampo, 2009). According to the Department of Education the purposes of school-based assessment are: (1) to improve the teaching and learning process, (2) to identify student’s strengths and weaknesses, (3) to determine the student’s subject area performance and/or achievement, and (4) to report student progress to parents.

Learning and teaching can be explored by examining assessment. The manner and the choice of assessment on a given set of learning objectives can yield insights to the quality of learning and teaching that is transpiring. By investigating the method, timing, content and format of assessment and evaluating its constructive alignment, the results may shed light to the realities and challenges both teachers and learners face in language classrooms. Assessment practices, content, and quality of constructive alignment, are key indicators to assess whether the experiences in the classroom have met the target learning objectives. They help in identifying gaps in the teaching and learning process and understanding language classroom. According to (Banda, 2005), teacher’s knowledge and skills, academic qualification, experience, perception and the educational program is an influencing factor in classroom assessment practices together with the class size, teaching and learning resources, student-student interaction and student-teacher interaction.

In the context of the Philippines, teachers have the responsibility and the ability to shape the language classroom, as exemplified by the Department of Education cited in Mariñas and Ditapat (2011) stating that “The curriculum is designed to be interpreted by teachers and implemented with variations. Schools are encouraged to innovate and enrich or adapt, as long as they have met the basic requirements of the curriculum” (p. 114). However, in a study conducted by Deocampo (2013) on the alignment of pedagogical task and test task, results showed that there are glitches in pre-service language teachers’ administration of the classes. Results of the study show that problematic practices in design assessment such as incoherence between the learning objective and the employed task, mismatch between instruction and test, and testing other skills not specified in the construct. Therefore, giving the teachers the responsibility and ability in shaping the language
classrooms entails the necessity to educate and train them about effective and useful instructional methodologies in language pedagogy.

1.3. Assessment in Language Classrooms

Taking the interconnected relationship of teaching, learning and assessment, alignment is then essential in optimizing the learning conditions for quality learning. According to Biggs (as cited in McLaughlin, 2001), if there is a match between the assessment tasks, learning activities and objectives, the student will learn what is intended. The use of an aligned design process ensures that there is consistency between objectives, learning activities and assessment. Assessment is vital to effective language teaching (Brown, 2001). An understanding of this interwoven relationship requires examination of the alignment of test constructs, learning objectives, and of the many other facets and qualities that make a test or assessment useful. This alignment pertains to teaching and testing in the same way and with the same objective Brown (cited in Deocampo, 2013). This suggests the parallelism of the instruction, desired outcomes, and assessment.

According to Carol Chapelle and Geoff Brindley (2002), “assessment refers to the act of collecting information and making ‘judgments’ about a language learner’s knowledge of a language and ability to use it” (p. 267), without the learner feeling a sense that he or she is being constantly judged according to his or her degree of knowledge or competence (Brown, 2001). Assessment can either be conducted through the use of formal measurements, tests, portfolios, journals, and the like, or by an informal method which may be planned or unplanned without a need for explicit or fixed judgments about a student’s competence for the purpose of recording results. Assessment complements teaching and learning. Assessment both formative and summative facilitates the feedback process of improving learning. Through assessment, learners are able to evaluate their responses and make adjustments on what and how they learn. It serves as an avenue in the language classroom from which students can identify gaps in knowledge, important information, and even connect procedural errors or misconceptions. Through assessment, they can chart their own development. On the other hand, this process allows teachers to make adjustments on what and how they teach.

Assessment as opposed to tests cover a much wider domain. Tests are formal tools of assessment that consist of specified tasks through which language abilities are elicited. This subset to assessment, as stated by Brown (2001), is a “method of measuring a person’s ability or knowledge in a given domain”. It is a prepared instrument which is specifically designed to tap into the learner’s storehouse of skills and knowledge. It occurs at identifiable times when learners are aware that their performance is being measured and evaluated. Essential to useful tests is construct and content. These are the two important aspects of test validity. As demonstrated in Palmer and Bachman’s Model, validity pertains to the degree to which the test actually measures what it claims to measure, and it is the extent to which interpretations made on the basis of test scores are appropriate and meaningful. Furthermore, according to Coombe (2010), a test is said to be valid if it tests what was taught and how it was taught. Content validity pertains to how an assessment calls for performance that matches that of the established set of goals or instructional objectives in a particular course or unit. Examples of these objectives are the written descriptions of what students are expected to be able to know and be able to do at a specific stage of their education, and these objectives are typically reflected in the curriculum or lesson plan.

1.4. Constructive Alignment

Biggs (2003) introduced the constructive alignment system which is an approach to curriculum design that puts prime importance to quality learning. According to Biggs (as cited in McLaughlin, 2001), if there is a match between the assessment tasks, learning activities and objectives, the student will learn what is intended. It is also expounded in McLaughlin’s (2001) study entitled ‘Inclusivity and Alignment: Principles of pedagogy, task and assessment design for effective cross-cultural online learning’ that the use of an aligned design process ensures that there is consistency between objectives, learning activities, and assessment.

Basically, constructive alignment has two major aspects according to the Biggs (2003)—the constructive aspect and the alignment aspect. The ideas that students construct through meaningful learning activities is the focus of the constructive aspect, and alignment aspect centers on what the
teacher does. The use of this aligned process ensures that there is consistency between objectives, learning activities, and assessment. And this aligned process has an effect on students’ achievement. The relationship of alignment to student achievement was discussed by Squires (2012) emphasizing that research on curriculum alignment suggests that alignment significantly improves student achievement. This is true in the case of American schools according to Cohen (as cited in Squire, 2012). It was found out that the lack of excellence in American schools is not caused by ineffective teaching but mostly by misaligning what teachers teach and what they assess as having been taught.

This established connection between assessment and achievement is also reflected in Bloom’s model, cited in Squire’s study, which showed that when curriculum, instruction, and lesson planning are aligned with curriculum-embedded tests (and there is a teach, test, re-teach, test model in place), students’ test results can improve dramatically. In his study on Designing a Unit Assessment Using Constructive Alignment, (Lawrence, 2019) concluded that the correlation between teaching, learning outcomes and assessment helps make the overall learning experience more transparent and meaningful for students. By aligning the assessment with the learning outcomes means that students know how their achievement will be measured.

In Australian universities, the roles of achieving alignment of university strategic plans with curricula and graduate attributes are increasingly assigned to faculty and discipline educational developers and designers according to Oliver (as cited in Gesa, Olubukola & Linda, 2019), highlighting the crucial role of teachers in achieving alignment in design of teaching and learning. Research findings on curriculum design and course delivery call for more emphasis and detailed focus on ‘...engaging pedagogies, professional development for instructors, course designers and administrators should give alignment the long over-due attention…so that proper implementation of the constructive alignment can be ensured in practice’ (Wang et al. 2013, p. 488).

2. Research Method

This study used naturalistic enquiry combined with qualitative methods in data gathering to examine the usefulness of assessment in four selected elementary school classrooms. The researchers focused on studying classroom situations as they unfold naturally; it is non manipulative and non-controlling; and the researchers were open to whatever emerges (there is no predetermined constraints or findings). The aims of this study were well-suited for qualitative-descriptive research methodology. The instruments used yielded qualitative data.

2.1. Locale of the Study

The research took place in Iligan City, an urbanized city situated in Northern Mindanao – Region 10. Cebuano is the major language in the city and the majority of the population speaks English as their second language. The selection of the locale of the study is purposive. The public elementary school selected for this research is Tambo Central School, which is formerly known as Iligan City East Central School (ICECS) and is situated in Tambo, Hinaplanon, Iligan City, located in Southern Philippines. This school is one of the top performing elementary public schools in Iligan City, making it fitting to draw inferences from the teaching-learning practices in the classroom.

2.2. Research Participants

The participants in this study were chosen based on a purposive sampling method. Grade levels three to six were chosen since the said grade levels use English as a medium of instruction thus making second language learning relevant. The participants are students, who belong to the achievers’ class and their teachers. The students belong to the section for high achievers in grades three to six. The basis for the students’ sectioning is their academic standing. The grade schoolers are between 9-12 years old, mostly coming from an indigent or average earning family.

The English teachers of these students, on the other hand, are regular English teachers. These teachers have been in the teaching profession for more than 15 years and have graduated with a degree of Bachelor of Science in Elementary Education. One of the bases for assigning the teachers to the achievers’ class is their seniority and the evaluation of the school principal. They are also usually recommended by senior Master Teachers. These teachers are also considered to be one of the best teachers in their school based on their credentials and qualifications. The observed class for
each grade level is the section for the high achievers. Each class consists of an average number of 40 students, wherein two thirds is female and one third is male. The classes are only observed during their English class which is scheduled in the morning. The class covers a one-hour duration per session.

2.3. Data Gathering Procedures

Data gathering was initially done through conducting a pilot testing of the observation form and interview questionnaire. These instruments were then modified based on the results of the pilot testing and of the feedback of the teachers. The participants of the pilot phase were not allowed to participate in the actual data gathering phase. In answering question number one (1), the language assessment practices were gathered through classroom observation in a span of one week during the third quarter in grades three to six. This also includes their periodical exam on the said quarter. During the actual data gathering, the researchers gathered field notes of all forms of assessment practices per grade level in a span of one week, within the third quarter of the academic year. The observation covered five class sessions per grade level, wherein each session lasted for an hour. Furthermore, the data were gathered through conducting a direct observation within the targeted time frame. The researchers were watching rather than taking part of the class. Observations were done objectively excluding personal biases and perceptions of the observers. The assessment practices are then presented with their purpose, type, timing, corresponding description, the intended learning objective, and also the test specifications for formal assessment.

The researchers also procured copies of the third periodical tests and a set of learning objectives per grade level. Periodical tests are summative assessment administered quarterly. It is not a teacher-made test, rather it is a division-made test. It is a centralized test tailored from the budgeted skills or the objectives in the teacher’s lesson plan provided by the Department of Education (DepEd). All the designed third periodical exams for each grade are used in all elementary public schools in Iligan City. To aid in the analysis of assessment practices and the administered tests, the researchers inferred the constructs of these tests. These were later validated by three experts in the field of Language Teaching and Linguistics.

The data gathering for question number two (2), on the other hand, takes its foundation from the input about constructive alignment. This part looks into the constructive alignment of the gathered assessment practices presented in the first question. This aligned design process ensures that there is consistency between objectives, learning activities, and assessment. For the purpose of evaluating the alignment of the objectives and the corresponding tasks, the researchers gathered the assessment tasks administered in the classrooms as well as the learning objectives which said tasks are based on. The researchers inferred the constructs or the abilities that tasks measure and conducted interviews with the teachers and focus group discussion with the students.

The focus group discussion with the students and interviews of the teachers were done to supplement the data gathered in the observation phase. Ten students, five male and five female, per grade level, were chosen randomly for the focus group discussion. The questions asked were about the activities and assessment practices conducted during the lessons covered in the observation and matters about their periodical test—who prepared the test, how reviews were conducted, and their attitude towards the test. Teachers were also interviewed about their lessons and about their periodical tests. This study used a one-on-one non-structured interview to fill in the possible missed-out assessment practices during the direct observation and to elicit information about their periodical test. This is a confirmatory process as to why such activities and assessment tasks were done in the classroom.

The learning objectives were taken from the teacher’s lesson guide that is provided by the Department of Education (DepEd). All teachers in public elementary schools have the same lesson guides; however, teachers still have the autonomy in tailoring the provided lesson guide in the actual classroom. The answer to the last question takes into account all the issues that have been discovered in evaluating the constructive alignment of the assessment tasks administered in the four classrooms giving a picture of the numerous gaps in the curriculum, its operationalization in the classroom, and its assessment. This is used as a basis in making inferences about the teaching and learning in these four English language classrooms.
3. Findings and Discussion

3.1. Problem 1: What are the language assessment practices in terms of purpose, type and timing?

Assessment Practices in Four Grade Levels

In the classroom observation for grades four to six, with a total duration of 20 hours, there are a total of 98 assessment practices gathered, both formal and informal assessment. Every classroom has an average of 25 assessment practices in a week. Given this number of assessment practices, we can infer that every classroom has a good number of assessment practices which reflects that assessment practices are of much importance in the language classroom. It is also observable that in every session, the majority of the time is allotted to various assessment practices—collaborative work, interactive discussion, and oral recitation. Research conducted that the use of tasks in language classrooms that encourage learners to work together and collaborate in the solution of their language-related problems pushes learners to talk about the language they are using and pool their individual resources to solve the linguistic problems they encounter (Dobao, 2014). The data also reveals that most of the assessment practices are informal and are part of the teaching activities.

Moreover, as to the timing and purpose, it is revealed in the data that the assessment practices that are administered before the discussion are used to gauge students’ needs. The assessment practices during the discussion facilitate learning and are part of the teaching activities. These assessment practices monitor students’ progress and encourage engagement, participation, and collaboration. And those which are administered after the discussion are used to evaluate how much is learned from the lesson.

Third Quarter Periodical Test in Four Grade Levels

The periodical tests in the grade levels ranges from 40-50 items which are multiple choice except for the writing part. This test is administered at the end of the third quarter thus covers the learning objectives of the entire quarter which comprises nine weeks and is equivalent to 45 sessions with a duration of 45 hours, since one session is equal to one hour. From the given data it is evident that the items of the test are distributed to four macro skills listening, speaking, reading and writing. This tells us that the test taps the four macro-skills in the language learning, however, the items are not evenly distributed. Most of the items are concentrated on speaking and reading. The less assessed skill in the test is writing and listening which comprises an average of five items in the said tests. Thus, it can be inferred in the data that the macro-skills are not equally assessed. Testing all four macro-skills is crucial in language learning, not only for fairness to the learners but also to their proficiency profiles. Taking the interconnectivity of teaching, learning, and assessment into account, what is tested can affect what is taught as well as what is learned. Selective testing can cause uneven attention paid to language skills resulting to uneven profiles of proficiency skills of learners (Powers, 2010).

3.2. Problem 2: How are these assessment practices constructively aligned and the assessment content valid?

Looking closer at the test for the purpose of the evaluation of the alignment of the objectives and the corresponding tasks, the researchers gathered the learning objectives from which the assessment tasks are based on; the researchers inferred the constructs or the abilities that tasks measure and conducted interviews with the teachers and focus group discussion with the students. The evaluation of alignment takes its foundation from the input about constructive alignment. According to Biggs (as cited by McLoughlin, 2001), the student will learn what is intended if there is a match between assessment tasks, learning activities, and objectives. This is called constructive alignment. In addition, Kurz, Talapatra and Roach (2012), also exemplified that the test must comprise items that sample exclusively across the constructs expressed in the intended curriculum which students (presumably) had the opportunity to learn during classroom discussion.

Drawing upon the content of the assessment tasks, it is observable that there are assessment tasks that are aligned to the targeted objectives. In examples 6.1 to 6.5 shown in Appendix D, the learning objective intends students to demonstrate their ability to distinguish advertisements from propaganda. This skill is subsumed under the speaking macro-skill. In the data, five (5) observable assessment practices were identified. All these assessment practices elicited performance for
learners to demonstrate the desired skill to be obtained by the end of the lesson. In terms of content validity, each task is relevant and representative of the construct expressed in the learning objective. There is a match between the task and the content or subject area being assessed. The performance of the students indicates that they have successfully achieved the intended learning outcome, showing that the assessment practices are constructively aligned.

However, some assessment tasks are not aligned with the intended learning objectives identified in the study. Misaligned assessment tasks include tasks that are inadequate in assessing the four macro-skills of language as shown in the data. It also includes tasks that are insufficient operationalization of the intended learning objectives. There are language classes observed where learning objectives were not reflected in the assessment tasks, misinterpreted and translated to erroneous tasks, there are also misplaced tasks and, in some cases, there are too many tasks targeting the enabling skills failing to target the intended skill. Additionally, some tasks target the intended learning objectives, however, still considered misaligned for performance-based learning objectives are reduced into knowledge-based assessment tasks, and some tasks are inauthentic.

**Inadequacy in Assessing the Four Macro-Skills of Language Learning**

There are assessment tasks that the items are inadequately distributed to assess the intended knowledge and ability of students expressed in the learning objective. In the periodical tests, it is evident that the items of the test are distributed to four macro skills—listening, speaking, reading, and writing. The data shows that the test taps the four macro-skills in the language learning, however, as to the distribution of the items it is evident that items are not evenly distributed. Most of the items are concentrated on speaking and reading. The less assessed skill in the test is writing and listening which comprises an average of five items in the said tests. Thus, it can be inferred in the data that the macro-skills are not equally assessed.

For example, the Grade Three periodical test, almost all test items are in a multiple-choice type of test, except for the writing part, which constitutes five points of the exam. The data as shown in Appendix D shows that the items of the test are not evenly distributed to the four macro-skills. There are only eight items attributed for listening, 16 for speaking, ten for reading, and six for writing. It is evident in the data that the test does not measure the four macro-skills equally. Though enough number of items are attributed to speaking and reading skills, it is still important to point out that an eight-item test is insufficient in measuring listening skill and a six-item task is insufficient to measure writing skill. Considering that summative decisions that include passing or failing students and certifying their level of ability will be based on this test, it is important to take these issues into account. Testing all four macro-skills is crucial in language learning, not only for fairness to the learners but also to their proficiency profiles. Taking the interconnectivity of teaching, learning, and assessment into account, what is tested can affect what is taught as well as what is learned. Selective testing can cause uneven attention paid to language skills resulting in uneven profiles of proficiency skills of learners (Powers, 2010).

**Lack / Insufficient Operationalization of the Intended Learning Objectives**

In the data, these refer to assessment tasks which are insufficient in terms of the extent to which it covers the intended learning objectives. These tasks do not comprise the number of items that sample exclusively across the constructs expressed in the intended curriculum and do not sufficiently gauge the depth and breadth the objective outlined. For example, in assessment 3.9 shown in Appendix D, the objective that it targets is both the ability to read and write “oa” digraph but the assessment tasks focus on the recognition and reading ability only. In this part, the objective of reading and writing words, phrases, and sentences with “oa” diphthong is not reflected in the assessment tasks. No assessment task intends to facilitate and assess the skill expressed in this objective. This shows that there are learning objectives that are not reflected in the assessment of tasks.

It is also evident from the test that there is a faulty interpretation of constructs made by the test-maker. The examples 5.24 and 5.25 in Appendix D, which intends to measure the ability of students to infer the general mood of the selection. The mood as interpreted in the test refers to the literal emotion of the character or the prevailing emotion manifested in the situation which is erroneous since mood as a literary device does not solely pertain to the emotion or feeling. Given this faulty interpretation, it affects how the construct is being measured and also limits the learning and
performance of the students. Some tasks are also misplaced for they are just randomly inserted in the discussion which can potentially create confusion for the pupils. For example, assessment task 3.10 in Appendix D is the transition of the previous and new lesson but was administered as the third assessment tasks in class. This should have been the first task to be administered since it marks the end of the previous lesson. The two preceding tasks are formative assessments for the new lesson.

There are also instances where too many tasks are targeting the enabling skills. This leads to failure to meet the target skill and insufficiency in the operationalization of the intended objective. This is evident in examples 3.1, 3.2, and 3.3 in Appendix D. Although unlocking of difficulty is necessary to prepare and enable the students to comprehend the text that is to be discussed in class, allotting three assessment tasks in achieving this purpose is excessive. Unnecessary tasks consume time that would have been allotted to other more important tasks that will target the intended learning objective.

Performance-based Learning Objectives and Knowledge-based Assessment Tasks

Performance-based learning objectives require assessment tasks where students are engaged in certain activities or create products to demonstrate their academic knowledge and abilities, (Alsardary, Pontiggia, Hamid, & Blumberg, 2011) for the level of engagement affects subsequent language development (Zabih, 2020). In this case, most of the items insufficiently measure the intended learning objective as it only taps the recognition knowledge of the pupils. For example, the learning objectives that require students to produce sentences and use words in sentences are reduced to recognizing and identifying words. In this case, the tasks fail to cover the extent of the intended learning objectives. This is also evident in the misaligned writing tasks. Writing assessment for Grade 5 pupils stops at the format level and lacks a rubric for evaluation and scoring. In the learning objective where students are expected to write a telegram, the assessment task asks learners to write the jumbled parts of the letter properly while observing correct punctuation marks and capitalization as shown in example 5. PT. 46-50 in Appendix D.

As stated in the K-12 curriculum, for Grade five (5) level, the students are already able to plan a 2-3 paragraph composition writing and they already can compose clear and coherent sentences using appropriate grammatical structures but the test only assesses their ability to arrange the given parts of a letter following a prescribed format. In examples 3.15, 3.16, and 3.17 in Appendix D where the learning objectives focused on the writing and speaking skills of the students, the assessment practices in this part are aligned but it is not able to completely cover the intended learning objectives. In these assessment tasks, reading, writing phrases and sentences with “oa” diphthong is assessed by just underlining diphthongs in the given text. In this particular example, we can infer

Low-level comprehension tasks are also evident in the data. Reading comprehension tasks stop at the inferential level and most of the tasks assessing the students’ reading skills were not as intellectually demanding as it should be. Assessment tasks 32, 3.5, and 3.6 are shown in Appendix D facilitate comprehension which is an enabling skill of writing because it is necessary for students to fully comprehend the text before they can create a written output based on the text. However, what makes these problematic is that these comprehension tasks stop at the inferential level which is just level 2 of the dimension of reading comprehension based on Gray, Gates Smith, and Barret’s model. The students are yet to go through levels 3 and 4 which are critical and integrative before they can reproduce the text information in a different form, which is level 5 and is the target of the lesson objective. We can infer then that these tasks fail in preparing the pupils to do the task which requires them to write a short note of advice to the country mouse.

Among the recurring issues observed from all the summative assessment tests being examined, this test shares the common issue of how the construct as expressed in the learning objects are reduced to be measured only in terms of its cognitive aspect. This is well-exemplified examples 6. PT.13-16 in Appendix D. This shows how the content of the test items do not sufficiently cover and represent the intended learning objective. Inadequacy of the test content mirrors the extent to which the learning objectives are measured. It also mirrors how students’ demonstration of their level of ability is constrained or enabled (Lewkowicz, 2000). If such is the case, then we can infer based on the content that there is a weak constructive alignment of the test. The test tasks contained in the tests do not fully match the learning objectives specified.
Inauthentic Assessment Tasks

It is also important to point out that most of the tasks, though aligned, are devoid of context. Bachman and Palmer (as cited in Phakiti, 2008) stated that for a test to be authentic, given language test tasks should correspond to the features of a “target language use” task. Learning becomes meaningful and effective if the students are allowed to link their school activities to real-life experience. Therefore, if the tasks are contextualized, most likely there will be better retention (Lewkowicz, 2000). So, in this case, where assessment tasks are devoid of context, less retention will be expected from the students.

In this study, these are tasks devoid of context to demonstrate and apply the extent to which they were able to acquire the intended skill. In example 5.22 in Appendix D, the assessment task is aligned to the learning objective, however, it is devoid of context. The students were tasked to write five sentences based on a picture that is randomly picked by the Teachers. In this example, though this task will make the student write it is still not considered meaningful learning for the students cannot make a connection between the classroom and real-world tasks.

3.3. Problem 3: How do the quality of constructive alignment in these assessment practices reflect teaching and learning in these language classrooms?

Taking into account all these issues that have been discovered in the evaluation of constructive alignment of the assessment tasks administered in the four classrooms, we can have a picture, albeit a snapshot, of the numerous gaps in the curriculum, its operationalization in the classroom, and its assessment. Thus, we can infer that the constructive alignment of the assessment practices in these classrooms are problematic. As stated by McLoughlin (2001) and Jervis L.M. & Jervis L. (2005), the use of an aligned design process ensures that there is consistency between objectives, learning activities, and assessment. However, there are assessment tasks that are not aligned with the intended learning objectives identified in the study. These include tasks that are inadequate in assessing the four macro-skills of language as shown in the data. It also includes tasks that lack or are insufficient operationalization of the intended learning objectives. There are language classes observed where learning objectives were not reflected in the assessment tasks. Insufficient operationalization appears in the form of misinterpretation of learning objectives misplaced tasks and, in some cases, there are too many tasks targeting the enabling skills failing to target the intended skill. Additionally, some tasks target the intended learning objectives, however still considered misaligned for performance-based learning objectives are reduced into knowledge-based assessment tasks, and some tasks are inauthentic.

The issues that are revealed in the assessment practices reflect the state and condition of the teaching and learning in the classrooms. Though the four macro-skills are interrelated, each contributes uniquely to an individual’s overall communicative ability, thus must be assess equally. It is important to test for each of these four skills individually because each is a critical aspect of communicative competence. Direct evidence of specific skills can provide an indirect evidence of other skills. More comprehensive and integrated assessment of language skills provides several sources of information and yields better decisions individual’s communicative proficiency (Powers & Powers, 2015). Assessment tasks that lack or are insufficient operationalization of the intended learning objectives are crucial in language teaching and learning for what is tested is very likely to affect not only what is taught, but how it is taught. Alderson and Wall (1993) posited that a test affects the rate and sequence, and the degree and depth of how language is taught and learned. Messick (1996) also added that the test influences language teachers and learners to do things they would not otherwise do that consequentially promote or inhibit language learning.

The tasks that are devoid of context suggest that the learning experience of the students in the classroom is not as effective and as meaningful when they could connect their activities in the classroom to the real world. In these classrooms, students spend most of their time in knowledge-based assessment tasks. Moreover, on the issues of weak alignment between the learning objective and the test items, misaligned writing tasks, and misinterpretation of the lesson guide or learning objectives, we can infer that there is a gap between the curriculum developers and the teachers. These issues would reflect that teachers and curriculum developers have a different understanding and perception of the constructs articulated in the learning objectives. These issues also reflect that
students become exhausted in doing several tasks in the classroom without even learning what is intended to be learned.

As for reading comprehension, the assessment practices reveal that the reading comprehension tasks are at literal and inferential levels. These tasks do not assess students’ critical, integrative, and creative comprehension ability. And also, for the writing ability, as shown in the assessment practices, writing tasks are more focused on the format rather than the content, thus we can see how writing is perceived as conforming to a suggested format not an actual production of sentences and an expression of students’ thoughts. The lack of rubric would make the inferences from the result of the assessment tasks unreliable. In a bigger picture, all these issues affect the students’ achievement since as stated by Mohamud and Fleck (2010), alignment affects achievement and ample evidence from research on alignment of standards, instruction, and assessments suggests that students perform better on tests that measure content they have had the opportunity to learn. Opportunity to learn is at optimum levels when lesson plans are aligned with benchmarks, standards, and assessments, thus increasing academic achievement. Thus, the problematic constructive alignment and weak content validity of the assessment practices could be contributory to the language proficiency of the students in the language.

4. Conclusion

According to Schmidt, et. al (as cited in Squires, 2012), there is a significant connection between achievement and alignment of the taught and tested curriculum. The findings of the study point out how the quality of constructive alignment of the identified assessment practices can either enable or constrain the learning process and acquisition of the students. At the system level, the first curricular challenge as stated by Kurz, Talapatra, and Roach (2012) is the alignment between the intended and assessed curriculum. In a well aligned system, the teaching instruction and activities would facilitate well to achieve and assess the desired outcomes of the curriculum, as articulated in the set of learning objectives provided by the system. However, a misalignment of the taught and assessed curriculum could limit the learning opportunities of the students. Therefore, the goal for improved educational opportunities and outcomes requires us to examine the assessment practices as these reveals much about the teaching and operationalization of the intended curriculum. By examining the quality of constructive alignment and validity of the content in assessment practices, this study provides a snapshot of the quality of teaching and learning in the language classrooms. The study looks into the three components of constructive alignment: the teaching activities, assessment practices in classrooms – particularly the informal assessment, and the learning objectives, in a span of one week. These components by itself already reveal a lot about the practices in the language classroom. Aside from that, the study also takes into account the administered periodical tests as it will reveal a broader picture of the teaching and learning.

However, this study is only limited to the honor’s class of grades four to six in Tambo Central School. Since the Mother Tongue Based-Multilingual Education (MTB-MLE) has been implemented, English only begins as a medium of instruction in these grade levels and it is only then that the second language learning becomes relevant. Based on the findings, policy makers can look into assessment practices in connection to the educational system’s pursuit for an optimized learning condition for quality learning. Investment on teacher training addressing the problems of alignment of curriculum, pedagogies, and assessment practices within classrooms and across the system is seen as one of the significant policy implications of this research. Policy-makers need to be made aware of the significance of assessment for learning and the essential roles of teachers to its implementation. Likewise, teachers collectively need to participate in broader debates about system accountability in relation to assessment of learning. Moreover, in improving constructive alignment and content, it is also important to look into instructional strategies and evaluate which seems to be working well and which are not. This will make a clearer diagnosis and would determine the focus of professional efforts.
REFERENCES


Götzinger, J. (2009) E8 Writing Validation Report for E8 Writing


McLoughlin, C. (2001). Inclusivity and alignment: Principles of pedagogy, task and assessment design for effective cross-cultural online learning, Distance Education, (22)1, 7-29, https://doi.org/10.1080/01587910022020102

Mohamud A. & Fleck D. (2010) Alignment of Standards, Assessment and Instruction: Implications for English Language Learners in Ohio, Theory Into Practice, 49(2), 129-136, https://doi.org/10.1080/00405981003626643


Appendix A

Observation Protocol

Date: Name of the School:
Time of Observation: Start: ____ End: ____
Grade Level: Teacher:
Number of Students: Boys: Girls:
Classroom Number:
A. Kind of formative assessment task given and the way it was structured or administered

<table>
<thead>
<tr>
<th>Kind of Assessment</th>
<th>Type and Purpose</th>
<th>Administration</th>
<th>Timing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>I</td>
<td>P</td>
</tr>
<tr>
<td>Oral Recitation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reporting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Answering textbooks</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Answering worksheet questions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quiz</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Role Play</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collaborative Activity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interactive Discussion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others:</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B. Other Comments (Additional information necessary to capture the assessment activities including comments on any feature of the class.)
Appendix B

Focus Group Discussion Questionnaire

Focus: Test Alignment to Teaching Instruction and Learning Objectives

1. Was the test easy or difficult for you? Why?
2. What is the most difficult part of the test? Why? Did you discuss this in class?
3. What is the easiest? Why? Did you discuss this in class?
4. How did the teacher discuss the lesson?
5. What lessons were not included in the test that were discussed in class?
6. Did you study for the test?
7. Did the teacher conduct a review before the test?

Appendix C

Interview Questionnaire

1. Who made the test?
   How did they make it?
2. How did you prepare the students for the test given that you did not make the test?
3. Are the lessons/objectives covered in the test sufficient?
4. Do you think that the lessons covered in the test are the most important?

Objectives of the test should be clarified.
   a.) What are the objectives covered in the test?
   b.) What specific objective/s is/are targeted in this part of the test?
   c.) What are not included?
   d.) How much is covered by the test?
### Appendix D  Constructs and Content of All Forms of Assessment Tasks

<table>
<thead>
<tr>
<th>Thematic Classification</th>
<th>Assessment No.</th>
<th>Learning Objectives Based on Teacher's Lesson Guide</th>
<th>Assessment Task</th>
<th>Inferred Constructs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constructively aligned tasks to Learning Objectives</td>
<td>6.1, 6.2, 6.3, 6.4, 6.5</td>
<td>Distinguish advertisement from propaganda</td>
<td>Interactive Discussion</td>
<td>Ability to orally demonstrate examples of advertisements</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Students were asked to volunteer and reenact an advertisement from the radio or television</td>
<td></td>
</tr>
<tr>
<td>Lack/ insufficient operationalization of the intended learning objectives</td>
<td>3.9</td>
<td>Read and write words with “oa” digraph as in goat</td>
<td>Answering Textbooks</td>
<td>Ability to recognize words and associate images to them.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identify several effects based on a given cause</td>
<td></td>
<td>Ability to associate written word with the image it signifies</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Read and write words, phrases and sentences with “oa” diphthong</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack/ insufficient operationalization of the intended learning objectives</td>
<td>5.24, 5.25</td>
<td>Infer the general mood of expression in a selection</td>
<td>Interactive Discussion</td>
<td>Ability to define general mood</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The teacher asks questions to the students such as: What do you think is a general mood? Why do you think we should adjust to the general mood of our environment?</td>
<td>Ability to communicate ideas orally</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Collaborative Activity</td>
<td>Ability to portray different moods</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The students are tasked to sing “Joy to the World” expressing the following moods: happy, sad, angry</td>
<td>Ability to express ideas a within groups</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Follow up question: How did you feel singing the song?</td>
<td></td>
</tr>
<tr>
<td>Lack/ insufficient operationalization of the intended learning objectives</td>
<td>3.10</td>
<td>Read and write words with “oa” digraph as in goat</td>
<td>Role Play</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The students are tasked to perform their role play about “What the City Mouse should do”</td>
<td></td>
</tr>
</tbody>
</table>

Lynrose Jane Dumandan Genon & Chezka Bianca P. Torres (Constructive alignment of assessment practices)
Identify several effects based on a given cause in class. This task was given as an assignment to them.

Read and write words, phrases and sentences with oa diphthong

*no rubrics

*excellent clap is given after student's performance

<table>
<thead>
<tr>
<th>Lack/ insufficient operationalization of the intended learning objectives</th>
<th>Identify several effects based on a given cause</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Read and write words, phrases and sentences with oa diphthong</td>
</tr>
</tbody>
</table>

| 3.1, 3.2, 3.3 |

Understand that contentment and the ability to adapt are important

Write a note on what to be done by the Country Mouse

**Spelling Quiz**
The students are tasked to spell 10 words. The teacher reads the words aloud and use it in a sentence.

**Words:** Province, city, mouse, dinner, curious, food delicious, terrible, scratching, corner

**Interactive Discussion**
The teacher shows a picture to the class (Picture of a Mouse) and the teacher asks the students, randomly, to tell something about the picture.

**Collaborative Activity**
The teacher calls 10 students to hold the two sets of strips of paper, the first set are the words and the second set are the meaning of the words. The students are tasked to arrange to arrange the strips of words with their corresponding meanings. The words are selected text for the lesson.

**Interactive Discussion**
The teacher asks motivational questions to students related to the story to be discussed such as: *Have you been in the province? What are the experiences you have in there?*

**Writing**
Students are tasked to write the jumbled parts of the letter properly while observing correct punctuation marks and capitalization.

**Interactive Discussion**
The teacher asks motivational questions to students related to the story to be discussed such as: *Have you been in the province? What are the experiences you have in there?*

**Writing**
Students are tasked to write the jumbled parts of the letter properly while observing correct punctuation marks and capitalization.

**Interactive Discussion**
The teacher asks motivational questions to students related to the story to be discussed such as: *Have you been in the province? What are the experiences you have in there?*
| Performance-based learning objectives are reduced into knowledge-based assessment tasks. | 6.PT.13-16 | Use verbs in the active and passive voice | Multiple Choice. Students choose from the given options the passive version of an active voice and vice versa. Students are tasked to identify if the sentence is in a passive or active voice. **Oral Recitation** The teacher asks comprehension questions to the students (literal and inferential). Communicate ideas orally. Ability to recall what the text says. Ability to read between the lines or to infer information from the text. |
| Performance-based learning objectives are reduced into knowledge-based assessment tasks | 3.15, 3.16, 3.17 | Identify and write complex sentences | Read and write words, phrases and sentences with “oa” diphthong. **Oral Reading and Boardwork** The student are tasked to underline the “oa” diphthong in the sentences on the board and they read it after. **Boardwork** Random students are tasked to underline the “oa” diphthongs in the sentences written on the board. Speaking skills; ability to identify and recognize sentences in the active and passive voice. Ability to identify words with “oa” diphthong. Ability to pronounce words with “oa” diphthong. |
| Inauthentic Assessment Tasks | 5.22 | Write a sentence using descriptive words | **Writing Sentences** The students are tasked to write five sentences based on the picture given by the teacher. Ability to identify words with “oa” diphthongs in sentences. Ability to write sentences about a picture. |