



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE HUMANIDADES
DEPARTAMENTO DE ESTUDOS DA LÍNGUA INGLESA, SUAS LITERATURAS E
TRADUÇÃO.
CURSO DE LETRAS INGLÊS

DAYANA KETLEY ARNAUD GOMES RIBEIRO

AN ANALYSIS OF ASSESSMENT MEASURES IN ENGLISH LANGUAGE
LEARNING IN AN ELEMENTARY SCHOOL.

FORTALEZA

2018

DAYANA KETLEY ARNAUD GOMES RIBEIRO

**AN ANALYSIS OF ASSESSMENT MEASURES IN ENGLISH LANGUAGE
LEARNING IN AN ELEMENTARY SCHOOL.**

Trabalho de Conclusão de Curso apresentado ao Curso de Letras-Inglês do Departamento de Estudos da Língua Inglesa, suas Literaturas e Tradução da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Licenciado em Letras - Inglês.

Orientador: Prof. Dr. Lídia Amélia de Barros Cardoso

FORTALEZA

2018

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

R368a Ribeiro, Dayana.

An analysis of assessment measures in english language learning in an elementary school. / Dayana Ribeiro. – 2018.

37 f. : il.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Humanidades, Curso de Letras (Inglês), Fortaleza, 2018.

Orientação: Prof. Dr. Lídia Amélia Cardoso de Barros.

1. Assessment. 2. Measures. 3. Criterias. I. Título.

CDD 420

DAYANA KETLEY ARNAUD GOMES RIBEIRO

**AN ANALYSIS OF ASSESSMENT MEASURES IN ENGLISH LANGUAGE
LEARNING IN AN ELEMENTARY SCHOOL.**

Trabalho de Conclusão de Curso apresentado ao Curso de Letras-Inglês do Departamento de Estudos da Língua Inglesa, suas Literaturas e Tradução da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Licenciado em Letras - Inglês.

Aprovada em: __/__/____.

BANCA EXAMINADORA

Prof. Dr. Lídia Amélia de Barros Cardoso(Orientador)
Universidade Federal do Ceará (UFC)

Prof.. Dr. Pâmela Freitas Pereira Toassi
Universidade Federal do Ceará (UFC)

Prof.. Dr. Samuel de Carvalho Lima
Instituto Federal do Rio Grande do Norte (IFRN)

AGRADECIMENTO

Agradeço primeiramente a Deus, por conduzir todo o trajeto até aqui, permitindo que esse momento chegasse, mesmo com todas as adversidades.

Aos meus pais por todo o suporte não só financeiro, como também emocional pra que eu sempre permanecesse firme nos meus ideais, que permitiram que eu seguisse nessa longa jornada e me motivam a continuar nas que virão.

À Prof. Dr. Lídia Amélia de Barros Cardoso pela excelente orientação e pela paciência em lidar com os prazos e os horários de cada encontro.

Aos professores participantes da banca examinadora Pâmela Freitas Pereira Toassi e Samuel de Carvalho Lima pelo tempo, pelas valiosas colaborações e sugestões.

Aos professores entrevistados, pelo tempo concedido nas entrevistas.

Aos colegas de turma, em especial Jaime Vasconcelos por todo o suporte dado não só na fase de pesquisa, mas em todo o trajeto da graduação até aqui.

RESUMO

Avaliação é um tópico relacionado ao ambiente de ensino- aprendizagem, que levanta muitas discussões entre os professores. Desenvolver um instrumento bom de avaliação que de fato avalie a habilidade dos alunos, ainda é um desafio e preocupação para muitos professores. Essa pesquisa apresenta o conceito de avaliação e as características que a mesma deve apresentar para ser considerada um bom método avaliativo. O estudo também apresenta através da coleta de dados qualitativos, a avaliação em sala de aula na perspectiva dos alunos e os impactos causados neles, através da percepção individual dos testes. Além disso, também mostra as intenções do professor ao desenvolver os testes aplicados em sala. Por fim, o estudo apresenta uma análise dos dados coletados na perspectiva dos alunos, associando os resultados nos testes e a resposta na entrevista, com os aspectos de motivação, atitude e complexidade, e se volta também para a análise das respostas, intenções e atitudes do professor, com os critérios de validade e confiabilidade. Critério estes que devem ser levados em consideração, na análise de um bom instrumento de avaliação.

Palavras-chave: Avaliação. Métodos. Critérios.

ABSTRACT

Assessment is a topic quite related to the learning environment, which raises a great amount of discussions between teachers. Design a good instrument of assessment that in fact assess students' abilities is challenging and still concern a lot of teachers. This research presents assessment concepts and the characteristics that should be taken into account while designing and analyzing and assessment instrument. The study also presents results acquired by data collected through qualitative research in interview format of students' impacts while being tested. Besides that, also shows teachers beliefs while designing a test to be applied at classroom. At the end, the research shows an analysis of data collected from students' perspective, outcomes and answer in the interview with the aspects of motivation, attitude and complexity. Also, relates the teacher's answers, beliefs and attitudes with validity and reliability. These criteria should be taken into account while designing and analyzing a good assessment instrument.

Keywords: Assessment, Measures, Criteria

LIST OF TABLES

Table 1 –	Reasons for Testing	17
Table 2 –	Students' Answers	20
Table 3 –	Level of difficulty	21
Table 4 –	Students' Answers	21
Table 5 –	Questions approved and disapproved	22
Table 6 –	Motivation Aspect.....	25
Table 7 –	Familiarity Aspect.....	26
Table 8 –	Complexity Aspect.....	28

SUMMARY

1	INTRODUCTION.....	13
2	REVIEW OF LITERATURE.....	14
3	METHODOLOGY.....	19
4	FINDINGS.....	19
4.1	Findings of data collected from students.....	19
4.2	Findings from data collected from the teacher.....	23
5	DISCUSSION.....	25
5.1	Students impacts.....	25
5.2	Teacher’s methodology.....	29
6	FINAL REMARKS.....	30
6.1	The role of students.....	30
6.2	The Role of Teachers.....	31
	REFERENCES.....	32
	APPENDIX.....	33
	APPENDIX A – STUDENTS’ PERSONAL INFORMATION	
	QUESTIONNAIRE.....	33
	APPENDIX B – TEACHER’S METHODOLOGIES QUESTIONNAIRE.....	34
	ANNEX.....	35
	ANNEX A – SUMMATIVE TEST (PART 1)	35
	ANNEX B – SUMMATIVE TEST (PART 2)	36
	ANNEX C – STUDENTS’ ROAD MAP.....	37

1. INTRODUCTION

When dealing with learning, one of the topics that usually raises a great amount of discussions and, one that teachers usually concern about, is assessment. Ur (2009) defines assessment as “an activity whose main purpose is to convey (usually to the tester) how well the testee knows or can do something.” Thus, it is important to discuss about the different measures of assessment.

Teaching a second language is directly linked to assessment, once in traditional schools is common to assess students in different way to evaluate their development. The assessment of student’s language abilities is a topic that has been discussed by the teacher, who tends to spend a considerable amount of time thinking and creating activities to evaluate student’s knowledge in a fair way.

Based on this issue and concerning about students’ development during an evaluation, a research at Freitas Mota Primary School from the 5th grade has been done, in order to look for assessment measure applied there, as well as the impacts on students’ learning and abilities.

Besides that, this study aims to gather student’s perceptions to find a relation between tests designed and the rational theories of language teaching through the verification of DATA. At the end, this research presents the teachers’ role while designing tests considering assessment criteria.

2. LITERATURE REVIEW

Concerning about assessment, many teachers spend a fair amount of time discussing about different measures and the reasons to use one method instead of another. First of all, is important to have in mind a definition given by Cambridge Dictionary that assessment is “the act of judging or deciding the amount, value, quality, or importance of something, or the judgment or decision that is made” which means that teachers are always evaluating students in some way, even without a practical test.

Based on that, is necessary to differentiate assessment and evaluation, to understand that both walk together and have a lot of similarities. Hall and Sheehy (2010) have a definition quite similar to the dictionary and concludes that assessment is “a general term which includes all methods used to gather information about student’s knowledge, ability, understanding, attitudes and motivation. While evaluation is defined as a process close to assessment, and according to Ionnanu- Georgiou (2010), this process consists in “gathering information in order to determine the extent to which a language programme meet its goals. Relevant information can be teachers’ and parents’ opinions, textbooks, quality, exam results, and children’s attitudes. Some of the tools of the evaluation process are tests and questionnaires, textbooks analysis, and observation.”

Dealing with second language acquisition, teachers consider the importance of assess students’ language abilities, and as Schoonen (2010) expressed, language ability is “the ability to perform language tasks in real life and real time, that is, the ability to convey or understand a content message through the medium of spoken or written language”.

Besides that, there are some aspects regarding of authenticity of assessment, that is the language used in a real context. Bachman (1990) states that authenticity is seen to comprise: (1) The relationship between test method characteristics and features of a specific language use and situation and (2) The degree to which the test method invokes the test taker’s ability. This is reinforced by Canale and Swain who highlight the importance of a test to be authentic and to consider the sociolinguistic competence, discourse competence, grammatical competence, strategic competence, respectively these competences follow sociolinguistic ability, grammatical ability and strategic ability.

Other aspects that are taken into account when analyzing and designing an assessment instrument, are Reliability and Validity. The first one, reliability, consists in how

the test is measuring the ability that supposed to be tested. There are three factors that contribute to reliability aspects: test factors, situational factors, and individual factors.

In sum, test factors are related to technical aspects, that are the ambiguity and objectivity of the items in a test, and also concern about the freedom of the response, and how the test was designed in terms of instruction, the quality of layout, the familiarity that the respondent will have with the test. Situational factors consist in physical characteristics of the place where the test will be applied, like the lights, the comfort and outside noises. The last aspect, that are Individual factors are divided into two main aspects: transient and stable factors. Transient factors are very individual ones and include the physical and phonological health of the respondent. Stable factors include the ability to use the language, the skills and I.Q.

The second one, validity concerns about what is going to be assessed, whether assessment instrument actually measures what intends to measure. A test can be classified as valid or invalid, if it is presented the purpose intended when designing the test. The assessment of test validity can be conducted in different ways, that are: content validity, criterion related validity, construct validity, systemic validity, and face validity.

Content validity is determined by the adequacy between the test samples, content and objectives of the course area that is being assessed. Valette, (1977: 46 Apud Cohen 1994) states that for the language teacher, the degree of the test validity is not derived from a statistical analysis of test performance, but from a meticulous analysis of the content of each item and of the test as a whole.

Criterion- related validity compares respondent's answers. According to Cohen (1994) it calls for determining how closely respondent's performance on specific sets of objectives on a given assessment instrument parallels their performance on another instrument, or criterion, which is thought to measure the same or similar activities.

Construct validity examines whether the instrument is a true reflection of what is being to be measured. Cohen (1994) complements that in conducting construct validation, we are empirically testing hypothesized relationships between test scores and abilities, an activity that requires both logical analysis and empirical investigation.

Systemic validity concerns about validity construction taking into account the effects of instructional changes. According to Fredriksen & and Collins (1989 Apud Cohen 1984) it consists in:

A systematically valid test is one that induces in the education system curricular and instructional changes that foster the development of the cognitive skills that the test is designed to measure. Evidence for systemic validity would be an improvement in those skills after the test has been in place within the educational system for a period of time.

This specific validity concerns a lot on how assessment instruments affect educational practices and beliefs. Based on this effects Alderson and Wall (1992 Apud Cohen, 1994) consider that a test on teaching or learning calls for looking at

1. The effect of the test on the instructional content
2. The process whereby the content is taught
3. The rate which the material is covered
4. The nature of the sequencing of the material
5. The quantity of material that the students actually learn and the depth at which they learn the given items or objectives.

The last one, the concept of face validity that do not have the same features as the other forms and concerns about the measures of evaluate something, whether a specific test, accomplish the goal of evaluate an ability according to the way the tests is designed and the items organized.

Besides of these main characteristics presented previously, is also important to highlight the teacher's role while evaluating an assessment instrument, since the way that an ability is going to be assessed and the feedback that comes after an evaluation, will cause many effects on students' development. Based on that, Cohen (1994) concerns about to point a set of categories that construct the teacher's role on assessment:

1. The clarity of the instructions to the students- specifically, feedback from students regarding the clarity of the operations they were being asked to perform.
2. The appropriateness of time actually allotted- that is, if the students were pressed for time, was this consistent with the aims of the test?
3. The degree and type of student interest in the task- what parts of the quiz/ test did the students enjoy doing, if any? Students could even be asked to evaluate the instrument anonymously.
4. The level of performance a class basic and by individual students.
5. The meaningfulness of the data retrieved.
6. The appropriateness of the scoring procedures and weighting.

7. The ease of interpretation and evaluation of the score
8. The extent to which the instrument assesses what the teacher set out to asses.

According to Penny Ur (2009), there are some reasons for testing (see table 1).

Table 1: reasons for testing

Table.1 ¹
<ul style="list-style-type: none"> • Give the teacher information about where the students are at the moment, to help decide what to teach next; • Give the students information about what they know, so that they also have an awareness of what they need to learn or review; • Assess for some purpose external to current teaching (a final grade for the course, selection); • Motivate students to learn or review specific material; • Get a noisy class to keep quiet and concentrate; • Provide a clear indication that the class has reached a ‘station’ in learning, such as the end of a unit, thus contributing to a sense of structure in the course as a whole; • Get students to make an effort (in doing the test itself), which is likely to lead to better results and a feeling of satisfaction; • Give students tasks which themselves may actually provide useful review or practice, as well as testing; • Provide students with a sense of achievement and progress in their learning.

Once recognizing the importance of assessment and its division into the two macrocategories, is important to investigate the impacts that this kind of tests would cause in the students. Which one of them would have more positive impacts and could evaluate students’ skills in a more specific and meaningful way.

Assessment as a method of evaluation, is used in many places by many teachers and it can happen in different contexts. Is important to emphasize that, as there are different skills, also depending on the measures taken, different results can be achieved by students.

¹ Source: Ur, P. A course in English language teaching 17th ed, Cambridge 2009.

By doing so, many teachers concern about how to assess language ability, on the one hand, there are teachers who defend that tests should focus on one isolated language or structure. On the other hand, there are teachers who defend that the test should have an integrative feature, with integrative items and procedures to be assessed together.

Teachers have their reason for testing students, and their reasons for choosing a specific choice. It can vary according to the skill that he/ she intends to evaluate, as well as the methodology adopted by the school. And according to Harmer, (2013) These reasons can be divided into four main categories, that are:

Placement test: this one is used to place students in a correct class according to their ages and level. It can be done by testing skills, knowledge and grammar vocabulary.

Diagnostic test: this one, different from the previous one is used to identify any problem and difficulties that students have with their learning and to evaluate students' productive and receptive skills.

Progress or achievement tests: these tests are used to evaluate students' progress in relation to the syllabus created by the teacher. According to Harmer (2013):

“Achievement tests at the end of the term (like progress tests at the end of a unit, a fortnight, etc.) should reflect progress not failure. They should reinforce the learning that has taken place, not go out of their way to expose weakness. They can also help us to decide on changes to future teaching programmes where students do significantly worse in (parts of) the test than we might have expected.”

Proficiency tests: this test is commonly used to get a job or some kind of certification and it can also be used in different stages. Most of them are external exams and students works a lot to pass them.

Portfolio assessment: even with all this styles assessment presented before, that are student that have problems with tests and even with good knowledge, they may find problems with assessments. Teacher concerned about this unfair situation, try to adopt another strategy to evaluate. For that reason, they try allow students to produce a portfolio of their work, over a period of time. As Nunes (2004, Apud, Harmer 2013) states:

“Portfolio assessment of this kind has clear benefits. It provides evidence of student effort. It helps students become more autonomous, and it can ‘foster student reflection and help them to self- monitor their own learning. It has clear validity since, especially with written work, students will have had a chance to edit before

submitting their work, and this approach to assessment has an extremely positive washback effect.”

Establishing that there are many characteristics that should be taken into account, and establishing that there are different ways to evaluate something, is important to point out that while talking about education environment and the period that the evaluation is done, there two main types of tests, that Harmer (1993) defines at Summative and Formative tests.

The first one, Summative test, is very common specially in many schools and are tests applied at the end of a period (month, semester or year) and is also the test that students do in order to pass in a public exam. ENEM is a very good example of summative test.

The second one, Formative test, is related to what teachers do in day- to- day lessons, and the feedback they give in the classroom. Formative assessment suggests that teacher as the students, are developing and consequently changing the focus on the curriculum, which is not well accepted in many schools.

3. METHODOLOGY

In this study 15 students from 5th grade were presented to summative tests applied at the end of each semester and their results were collected. Students were interviewed with a qualitative questionnaire which they should talk about their perceptions about their English tests. After that, the teacher was interviewed in other to fill some gaps of information about what happens in the classroom environment and to discover if there is another evaluation measures besides the written ones in summative tests.

- 1.1 Collect the test used for assessment. See annex 1.
- 1.2 Interview the students before and after the tests (qualitative research) in order to gather perceptions. See appendix 1.
- 1.3 Interview the teacher as an attempt to find some gaps between the test and the content taught in the classroom.
- 1.4 Analyze the results.

4. FINDINGS

4.1 Findings of data collected from students

1. How was your last test? (See table 2).

Students A, B, C and D: “Foi ótima, eu gosto muito de inglês”. “Acho que as provas de são sempre as mais legais e mais fáceis, fora a de artes, é claro.”

Students E, F, G, H and I: “Eu acho que foi boa, mas eu não tenho certeza.” “Eu sabia o conteúdo, mas não fico muito segura quando tenho que fazer provas.”

Students J and K: “Ruim. Acho que eu só acertei a de pintar e completar. “Foi horrível! Toda vez essa é a prova mais chata.” “Eu não sei pra quê a gente tem que estudar inglês, se aqui no Brasil todo mundo fala português.” See table 2.

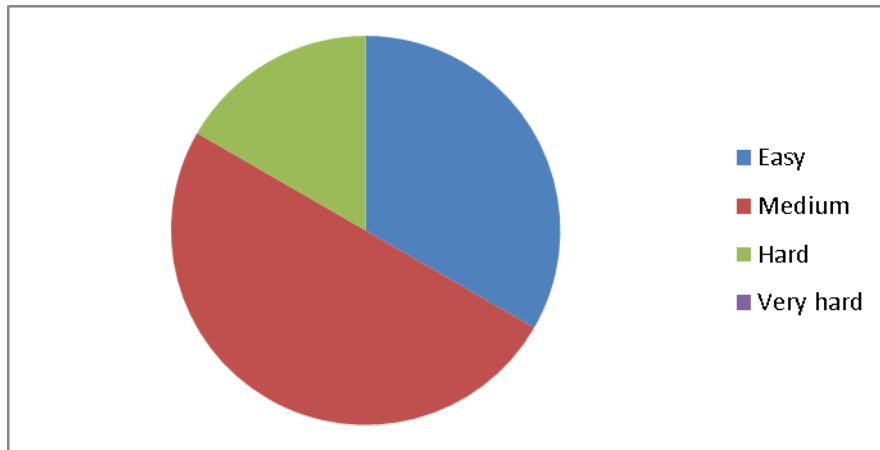
Table 2: students' answer

	Answer 1 (students A, B C and D)	Answer 2 (students E, F, G, H)	Answer 3 (students J and K)
How was your last test?	“Foi ótima, eu gosto muito de inglês”. “Acho que as provas de são sempre as mais legais e mais fáceis, fora a de artes, é claro.”	“Eu acho que foi boa, mas eu não tenho certeza.” “Eu sabia o conteúdo, mas não fico muito segura quando tenho que fazer provas.”	“Ruim. Acho que eu só acertei a de pintar e completar. “Foi horrível! Toda vez essa é a prova mais chata.” “Eu não sei pra quê a gente tem que estudar inglês, se aqui no Brasil todo mundo fala português.”

Source: Research data

2. According to your last test, classify it in: Very easy; easy; medium; hard; very hard

Students had to classify their tests into four categories: easy, medium, hard and very hard. The picture below, shows student’s opinions about the level of difficulty of the assessment tests. See table 3.

Table 3: Level of difficulty

Source: Research data

3. Do you see a connection between the content taught during the classes and the questions in the test? See table 4.

Students A, B, C, E, F, G, H, I, J and K:

“A tia entregou o roteiro pra gente antes da prova, eu acho que tinha tudo que ela falou que ia cair. E ela já tinha dado aula sobre isso antes, só que com questões diferentes das da prova. As questões do livro, não apareceram na prova.”

Student D: “ A prova tinha tudo sim, tudo que a tia falou na sala. Só tiveram umas questões que ela não tinha falado na sala, mas eu lembro que a professora do ano passado tinha ensinado. Então eu acho que ela fez isso, pra ver se a gente se lembrava do que a gente já estudou.

Table 4: student's answer

	Answer	Answer
	(students A, B, C, D, E, F, G, H, I J and K)	(student D)
Do you see a connection between the content taught during the classes and the	“A tia entregou o roteiro pra gente antes da prova, eu acho que tinha tudo que ela falou que ia cair.	“A prova tinha tudo sim, tudo que a tia falou na sala. Só tiveram umas questões que ela não

questions in the test?	E ela já tinha dado aula sobre isso antes, só que com questões diferentes das da prova. As questões do livro, não apareceram na prova.”	tinha falado na sala, mas eu lembro que a professora do ano passado tinha ensinado. Então eu acho que ela fez isso, pra ver se a gente se lembrava do que a gente já estudou.”
------------------------	---	--

Source: Research data

Since most of students classified the tests as easy and medium, was important to look at the kind of questions that they like and do not like, as an attempt to establish a connection between these questions and any aspects of complexity, familiarity or motivation and student's grades. See table 5.

4. What kinds of questions do you like?
5. What kinds of questions you do not like?

Table 5: Questions approved and disapproved

Questions approved	Questions disapproved
Match questions	Translate questions
Draw questions	Change sentence structures questions
Paint questions	
Fill in the gaps questions	

Source: Research data

4.2 Findings of data collected from the teacher

In order to find out these pieces of information, we were asked the following questions:

1. How do you classify your tests?

Teacher: [...] eu acredito que as provas tenham um nível que alcance todos os alunos, desde os alunos com melhor desempenho, até aqueles que são mais fracos. A turma não tem um grande índice de nota baixa, então eu classificaria como provas fáceis [...].

2. Do you think that the tests designed really evaluate students' abilities (reading, writing, speaking and listening)?

Teacher: Eu acho que a intenção principal é essa, de avaliar as habilidades[...] Só que tem que ver que não tem tempo pra avaliar tudo isso num só dia de prova, e que os meninos não compreendem também, o que é reproduzido nos áudios [...] é difícil também querer que os meninos saiam já falando, pra gente avaliar. A gente tem que se preocupar se o conteúdo gramatical está sendo trabalhado, se os alunos estão pronunciando as palavras certas e entendendo o que deve ser feito em cada questão. Consequentemente fazendo isso, as notas são positivas. Mas as provas têm os itens com respostas escritas, trabalhando assim a escrita deles.

3. Is there any other evaluation you do in your classroom besides of summative tests?

Teacher: existe. Além do teste final, são realizadas as provas parciais que são parecidas com as bimestrais. São basicamente questões de gramática e vocabulário [...] eu também pontuo por alguma resposta certa na aula e tradução de alguns dos textos do livro que eu peço pras eles fazerem em grupo e usando o dicionário.

4. Do you design the comprehension activities (learning, reading and speaking) suggested in the book?

Teacher: As atividades de escrita, são as que eu passo com mais frequência, já que estão em maior quantidade no livro. A fala deles é mais praticada com a repetição de palavras, quando eu corrijo a pronúncia deles. Eu não falo as palavras além do livro, porque eles vão ter dificuldade pra entender. O livro é bem tranquilo de entender, usa palavras que eles já conhecem, falam de coisas que eles já usam, como os dias da semana, cores, profissões e etc... Quanto às atividades de compreensão auditiva, geralmente sou eu quem lê os diálogos com a resposta, pra eles identificarem. Não é todo dia que o som tá disponível, então eu tento me virar pra que eles façam todas as atividades. Eu deixo também que um ajude o outro pra que não tenham problemas com mal-entendidos.

5. How do you design your tests?

Teacher: Não tem muito segredo. As provas são elaboradas de acordo com o que é trabalhado em sala de aula. Eu sempre elaboro a partir do roteiro bimestral. Uma vez que eu já digitei e entreguei pra que seja levado por eles, vou pegando as questões de cada capítulo e colocando na prova. É bom porque eles já ficam familiarizados e tem menos chances de errar.

6. Do you intend to evaluate students' integrative skills on your tests?

Teacher: Assim...a minha intenção é de avaliar isso também, já que a gente sabe que é importante e que as habilidades têm que andar juntas no ensino da língua. Mas eu acho que ainda precisa ter um incremento nos métodos, já que eu foco muito na gramática e vocabulário escritos (que é exigido pela escola) e acabo deixando principalmente as habilidades de compreensão auditiva e de fala, de lado.

5. DISCUSSION

5.1 Student's impacts

The aim of this qualitative questionnaire, was through DATA analysis , establish a relation between students' development and grades in English tests, and the following concepts: Attitude/ motivation, familiarity and complexity. For that reason, 3 questions were elicited.

Aspect 1: Attitude and motivation

Question: Fill in the gaps. Score 0,7. (See table 6)

Table 6: Attitude and motivation aspect

Students	Positive outcomes	Negative outcomes
Student A	X	
Student B	X	
Student C	X	
Student D	X	
Student E	X	
Student F	X	
Student G	X	
Student H		X
Student I	X	
Student J		X
Student K		X
Student L		X

Source: Research data

According to the DATA presented in table 6, most of students had a positive outcome. Students reported that questions which they had to complete with the correct answers, were the easiest one. According to them, they felt more comfortable, because they could see the answers written, and had just to fit them in the items. Thus, when they received the test, they started to answer, by this kind of question.

One aspect that should be taken into account, while analyzing these results, is motivation. Motivation is a process, not a aim. Is needed motivation to achieve a goal. Thus, it is considered an important field of learning because is through that people can do something. Brown (1994) believes that: there is no efficient learning without a motivation. In addition, Slavin (1997, 345) describes this process in a general sense: “...motivation is what gets you going, keeps you going and determines where you are trying to go”

As students felt more confident, they felt more confident to solve the problem, it means that they devoted some effort to accomplish this kind of question, this action can be namely as motivation. This set of efforts, can be namely as attitude.

Students who act in a positive way, motivated by something in order to have some learning, have a positive attitude. Smith’s (1971) has a perspective that support this idea: An attitude is a relatively enduring organization of beliefs around an object or a situation, predisposing one to respond in some preferential manner.

Aspect 2: Familiarity

Question: Draw and translate. Score: 0,7. See table 7.

Table 7: Familiarity aspect

Students	Positive outcomes	Negative outcomes
Student A	X	
Student B	X	
Student C	X	
Student D	X	

Student E	X	
Student F	X	
Student G	X	
Student H		X
Student I		X
Student J		X
Student K		X
Student L		X

Source: Research data

Since students started to learn English, they used to deal with translate questions. For that reason, they tend to have a kind of preference while choosing the kind of question as questions they like. So is possible to create a relationship between assessment preferences and familiarity.

Even with some exceptions, students seem to share the same preference for translate questions. First because were adapted to do it, since they started to do it, and also because they could notice a similarity between the vocabulary's contents taught in the classroom, and this translate question.

According to Cambridge Dictionary, familiarity is defined as: "a good knowledge of something, or the fact that you know it so well". As students already got in touch with these questions and this contents in the classroom, they felt able solve them. Thus, they chose translate questions to be one of questions they liked.

Aspect 3: Complexity

Question: Change the sentence structure. Score: 0,8. See table 8.

Table 8: Complexity aspect.

Students	Positive outcomes	Negative outcomes
Student A	X	
Student B		X
Student C	X	
Student D	X	
Student E	X	
Student F		X
Student G	X	
Student H		X
Student I		X
Student J		X
Student K		X
Student L		X

Source: Research data

Different from the aspects discussed previously, this one had different results, and one of the worst outcomes in the whole test. While motivation and familiarity were related to student's confidence and self-esteem, complexity is usually an aspect that makes students feel not able to do it.

Baccarini (1996) says that there are two main types of complexity, there are: organizational complexity which reflects the view that a project is a task containing many

interdependent elements. And technical complexity which deals with complexity related to the transformation processes, which convert inputs into outputs.

Student's reports, suggests that they had an issue with technical complexity, since they had difficulty to change the sentences structures to interrogative and negative form. They were able to notice when the sentences were changed, but they were not comfortable to change it by themselves.

5.2 Teacher's methodology

An interview was important to fill some gaps that only the tests were not able to do, since sometimes may have some spreads between what is taught in the classroom by the teacher and what is learned by students. Without it, would not be possible to know, for instance, if the teacher used any other instrument to evaluate students' comprehension abilities, and without this information is not possible to analyze validity, reliability or authenticity aspects in the tests, as is going to be discussed as following:

Validity aspect

As the teacher reported, the information with the grammar and vocabulary content is given to students before each test, thus they already know what is going to be evaluated in the test. The teacher already explained that the items in the test are similar to what students usually solve in the classroom as an attempt to work with their familiarity.

Based on this information, is possible to see that while talking about validity to students, the test was not valid because they already knew what was going to be assessed. However, in terms of face validity to the teacher, the test did not fit in this aspect, since the teacher had plans to design a test to evaluate students' abilities and according to Cohen (1994) face validity refers to "whether the test looks as if it is measuring what it is supposed to measure." For example, the test presented all the item in Portuguese, so students did not had the opportunity to interpret what they had to do, and a test that aims to measure student's learning ability, should provide questions that test students reading comprehension in a target language. Bachman (1990) clarifies that in terms of face validity, a test should raise an awareness on respondents:

1. Their perceptions of any bias in test content (i.e., whether they perceive the content to a favor a respondent with a certain background knowledge or expertise.)

2. Their understanding of the nature of the task that they are being requested to complete.
3. Their awareness of the nature of their actual performance on the test as a whole and on any particular subtests (for example, the test taking strategies that they employed)

Reliability aspect

Is important to highlight that a test needs to be valid to be reliable. The teacher's major intention is to work with student's written, through grammar and vocabulary, thus is important to be aware that in this aspect, the analysis is just about the written aspect. The other 3 abilities (listening, reading and writing) are not going to be taken into account.

As the reliability is an assess instrument that asks whether an assessment instrument is measuring respondents' abilities are at the same pace that expected results.

6. FINAL REMARKS

6.1 The role of students.

Since the qualitative research was done, was possible to notice that for most of students the tests were not challenging and they did not have problems to solve it. Of course that aspects of complexity, familiarity and motivation were important to preview students' outcomes.

Students had a good attitude toward questions that they had a familiarity. Once they already solved this question in classroom. This information was confirmed by the students and also by the teacher. Students also felt motivated to answer items with lowest degree of complexity, and as the number of mistakes was biggest than those with complex items

All the aspects presented above were integrated in a way that contributed to affect the results in the test. Motivation interfered in two other aspects: the familiarity with questions made students feel motivated and have positive attitude towards the items which they already answered. The complexity presented in items made students feel a little less motivated and insecure to answer.

6.2 The role of the teacher

The interview with the teacher was important to fill some gaps on the pieces of information given by students. In fact, teacher had a plan to evaluate students' abilities, but she did not know how to do it in a effective way. She also demonstrated that the school did not let her provide different instruments to assessment. Many schools especially the private ones, concern a lot about the grades which show students learning grammar content. That is why many teachers have to adjust their class plan and test measures, in order to achieve the learning goals.

Nevertheless, even with school limitations, the teacher had other problems to achieve the validity and reliability aspect. Concerning about the first one, since she had in their plans, to assess students' speaking ability, she should have provided, at least one activity in which students could practice the language. If she had plans to evaluate the listening comprehension, she should have students getting in touch to any activity with audios in the target language. The items presented in the test were all in Portuguese, so students did not have the opportunity to interpret the questions, and also there was no item with texts in English. Students had just to apply grammar and vocabulary content.

In regards of second aspect, If we look at teacher's intention, the test was not valid, since she said that had plans to assess all the four abilities. However, is important to highlight that the test presented all the content provided previously in the curriculum.

The concepts of validity and reliability work together. A test needs to be valid to reliable, in a way that we need to know what the test is going to test, the purpose of the evaluation (validity) in order to be able to measure if the test accomplished the goal (reliability).

REFERENCES

- Brown, H. **Principles of language learning and teaching**. White Plains, NY:Pearson Longman, 2008.
- Harmer, J. **The practice of English language teaching**. Harlow: Pearson Education,2013.
- Shrum, J.; Glisan, E. **Teachers handbook**. Boston: Heinle, 2010.
- Ur, P. **A course in English language teaching**. 17th ed. Cambridge,2009.
- Blaz, Deborah. **A collection of performance tasks and rubrics**- Foreign Languages, 2001.
- Cohen, Andrew D. **Assessing language ability in the classroom**- Heinle & Heinle Publishers A Division of Wadsworth, Inc. Boston, Massachussets, 1994.

APPENDIX A – STUDENTS QUESTIONNAIRE**Student's personal information**

Name: _____

Age: _____ Grade: _____

1. How was your last test?
2. According to your last test, classify it in: Very easy; easy; medium; hard; very hard
3. Do you see a connection between the content taught during the classes and the questions in the test?
4. What kinds of questions do you like?
5. What kinds of questions you do not like?

APPENDIX B- TEACHERS' INTERVIEW**Teacher's methodology**

Name: _____

Grade: _____

1. How do you classify your tests?
2. Do you think that the tests designed really evaluate students' abilities (reading, writing, speaking and listening)?
3. Is there any other evaluation you do in you classroom besides of summative tests?
4. Do you design the comprehension activities (learning, reading and speaking) suggested in the book?
5. How do you design your tests?
6. Do you intend to evaluate students' integrative skills on your tests?

ANNEX A - TEST (PART 1)

CEFM Centro Educacional Freitas Mota
 "Assegure hoje um amanhã brilhante"

Aluno(a):				Nº:
Prof.(a)				Data: / /
Série:	5º ano	Turno:		2º Bimestre
Total de scores:	Esc. Obtidos			Nota:

AVALIAÇÃO BIMESTRAL DE INGLÊS

1) Passe as frases para a forma interrogativa no **inglês**. (07 esc.)

a) She is a dressmaker.

b) He is a doctor.

c) He is a mechanic.

d) They are journalists.

e) She is a dentist.

f) He is a police officer.

g) He is a student.



2) Passe as frases da questão anterior para a forma negativa em inglês. (07 esc)

a) _____

b) _____

c) _____

d) _____

e) _____

f) _____

g) _____

3) Passe para o **português**. (06 esc.)

model _____ waiter _____
 clown _____ nanny _____
 journalist _____ waitress _____

4) Ligue corretamente. (07 esc.)

in	entre
on	sobre
under	dentro
behind	em frente de
in front of	embaixo
beside	ao lado de
between	atrás

5) Desenhe as profissões pedidas, depois escreva a **tradução** delas. (08 esc.)

poet	housewife	soldier	Bus driver

Have a nice test!!!



Inglês

Unidades 4, 5 e 6

Assuntos:

To be; professions; In/On/ Under;
behind/ In front of/ beside/ between;
Who?; More professions; + caderno

