A USEFUL ASSESSMENT TOOL: RUBRICS IN SECOND / FOREIGN LANGUAGE CLASSROOMS

Devrim HÖL

If students know what the learning target is, they are better able to hit it (Stiggins, 2001).

1. Introduction

Do these words sound familiar? 'How come I didn't get an A or a 100? Why exactly did I lose points? Is there anything else I can do to enhance my grades?' Current foreign language teaching students appear to be commonly concerned about grading and assessment. Information is obtained, formed, and distributed at such a quick rate in today's society that it has become a core priority and a key attribute of wealthy civilizations (Sert, 2008), and as a result, individuals must be evaluated in all aspects of our lives, including achievement, actions, and relationships. When we focus on the educational side of the assessment, each participant—teacher and learner—requires clear assessments in order to understand what they will do or have already done during the course of their learning and teaching process. Some claim that questioning marks are a natural effect of poor teaching or the use of questionable assessment instruments, while others argue that it is a result of academic standards and a student population used to honoring fewer achievements (O'Donnell et al., 2011). Teachers may be concerned about having their authority called into question, but they may also have some concerns about creating an efficient grading procedure and may rely on 'instinct grading'. As a consequence, when asked to present evidence for academic assessments, they are unable to respond. This assessing with uncertainty is an important aspect that has a vital effect not just in learning, but also in testing in the second language. From this perspective, testing affects learning in EFL classrooms either in a positive or a negative way, so this makes, language testing a critical issue in the realm of education (Ali et al., 2020) because tests and tasks not only assist in identifying problems that students are experiencing, but also in improving methods of teaching and learning. In this critical period, an effective tool that will uncover this obscurity is required, and 'rubrics' can be viewed as one of the most effective tools in education. Rubrics may be defined as an explicit set of guidelines used for grading a specific sort of work or accomplishment (Saritha, 2016), and they can also function as a scoring instrument that gives the standards for a piece of content (Andrade, 2000). It also shows the student's grades, which range from excellent to poor (Ali et al., 2020).

2. Definition of Rubric

A rubric is primarily a grid-based instrument with levels of success and clearly stated criteria for each level. Rubrics differ from basic checklists and grading scales in that each criterion for each level of achievement is described (Taylor & Galaczi, 2011). Educators have widely disparate conceptions of the term rubric. (Reddy & Andramade, 2010). A common definition is a statement that expresses the objectives for a task by specifying the parameters or what contributes and provides degrees of performance ranging from extraordinary to poor (Arter &

Chappuis, 2007; Stiggins, 2001). Rubrics are analytical grading techniques developed by teachers or other assessors to aid in the investigation of the consequences or processes of students' actions (Moskal, 2000). It is the most commonly utilized technique for evaluating performance and flaws. A rubric is a tool used to define a level of performance on a scale and to identify specific criteria for evaluating speaking or writing skills (Crusan, 2015).

A rubric must include three components: a) assessment parameters, b) grade standards, and c) a grading mechanism (Popham, 1997). An assessor analyzes assessment methods while assessing the grade of a student's performance. The criteria, often known as a collection of indicators or a list of recommendations, represent the procedures and content deemed important (Parke, 2001). Performance definitions explain in detail what a learner needs to do in order to show a skill, competency, or criteria in order to obtain a specific degree of achievement, such as poor, fair, good, or excellent. The performance standards address the need to differentiate between good and poor responses for scoring and providing feedback to students. Rubric scoring procedures include the use of a measure to interpret judgments of a product or activity. Teachers frequently use rubrics to judge student work, but many authors contend that they can also serve another, more significant function: Rubrics can both teach and evaluate when utilized by students as a component of a formative evaluation of their tasks in preparation (Arter & McTighe, 2001; Stiggins, 2001). Rubrics, when used as a component of a student-centered method of assessment, can assist students to grasp the goals of their learning and the quality requirements for a specific assignment, as well as make reliable assessments regarding their own performance that can influence modification and progress. Although some studies (Song, 2006) investigate how diagnostic feedback obtained by using a rubric might be utilized to pinpoint areas for development in training, other studies use rubrics merely to assess student performance (Campbell, 2005; Tunon & Brydges, 2006). Rubrics, according to Andrade (2005), can be used for both teaching and evaluative reasons. Several academics describe such applications and present evidence that rubrics aid in learning and teaching (Andrade & Du, 2005; Osana & Seymour, 2004; Powell, 2001; Reitmeier et al., 2004; Schneider, 2006). Rubrics are also utilized in the assessment of programs (Dunbar et al., 2006; Knight, 2006; Oakleaf, 2006; Reddy & Andrade, 2010).

3. Why Rubrics?

There is no disagreement that students' scores and performance must be reviewed. It is also vital to accomplish this with both simple comments and grading in this process. Grading must be done on some criteria, and the teacher cannot do it randomly (Janssen et al., 2015). A rubric is one of these fundamentals (Ghosh et al., 2016). Rubrics may help learners improve and strengthen their ability to think critically (Andrea, 2000; Arter & Chappuis, 2007; Stiggins, 2001; Wong, 2015). It can also provide students with an efficient way to assess their own students' writing skills and discover their writing weaknesses and limitations, resulting in improved writing outputs and higher scores (Silva, 2014). Scoring rubrics are useful because they define the construction to be completed and evaluated. Rubrics "help explain terms and

clarify expectations" (Crusan, 2010, p. 43). It describes various subcategories in education, such as evaluation criteria, learning evaluation, and learning gradients of a set of guidelines (Brown & Kondo-Brown, 2012). Weigle (2002) describes how the grading process utilizing rubrics can be especially "critical because the score is ultimately what will be used in making decisions and inferences about writers" (p. 108). Rubrics can also serve to alleviate the well-known issue of rater variability (Bachman et al., 1995; McNamara, 1996).

4. The Importance of a Reliable Scoring Procedure

As Weigle (2002) explicitly indicates, while conducting timed assessments, it is critical to verify that the judgments we make regarding students' abilities are appropriate. As a result, developing a trustworthy method for deriving meaningful inferences is critical—without such a method, we cannot properly estimate the component we are attempting to measure. Reliability, as conventionally defined, is the assessment of regularity and irregularity in students' performance test scores (Feldt & Brennan, 1989). 'Consistency' implies that students should be capable of performing at a level similar to equivalent tests, and if they are not, the assessment procedure's reliability is called into doubt.

'Educating the examiners' is the primary and most critical technique to assure reliability. Educating assessors in the usage of the measure so that they recognize how other assessors give scores and why they do so contributes to the reliability of this assessment technique (Grabe & Kaplan, 1996). Indeed, Vaughan (1991) argues that "researchers looking at holistic assessment have often assumed that given a scale that describes the characteristics of an essay at each level, trained raters will assess the essays, in the same way, every time" (p. 112). Irrespective of the rating system employed, it is critical to have standards to assess the dependability of scores provided to a piece of task by different assessors. It is because if two assessors cannot be proven to properly evaluate individuals based on observable actions, then further studies of those justices' evaluations will produce incorrect results (Stemler, 2004). Establishing reasonable limits of inter-rater concordance and, as a result, reliability is required. Stemler (2004) proposes three methods for calculating inter-rater agreement:

1. Estimates of consensus—the level to which assessors offer a relatively similar rating.

2. Estimates of consistency—the level to which the distribution of ratings is equivalent among assessors.

3. Estimates of measurement—the level to which grades can be credited to prevalent grading instead of to error elements.

Three additional approaches may be employed to supplement the reliability proof (East, 2009). To begin, intra-rater reliability measurements provide additional evidence of the stability of a one-rater scoring the very same documents twice. When 2 or more samples of paper by the same student are provided, the assessment of one can be matched to the evaluation of the other.

If, for example, the writing tasks are comparable, this evidence is especially essential since it is reasonable to expect that student performance on two activities that are equivalent but not identical will reveal comparable proficiency levels and a reliable scoring system should be capable of detecting this. Finally, manuscripts and grades can be analyzed. If a scoring criterion is regularly used, it is reasonable to believe that the writing of those who receive higher grades will be manifestly greater on a qualitative basis. The extent to which scores raise in path with advancements in scope and complexity of lexis and grammar accuracy, if assessors are affected by variables such as scope and complexity of lexis and grammar, and if a given scale score enables assessors to differentiate productions in these two aspects, indicate reliable scoring (East, 2009).

5. Rubric in Foreign Language Teaching

5.1. The significance of using rubrics in assessment, testing, and evaluation process

A rubric is an essential assessment tool that explains performance objectives for a portion of the output. Rubrics are routinely used to assess students' progress, but they can also fulfill another, possibly more important, aim: they can both teach and assess (Andrade & Du, 2005). Rubrics have the ability to assist students to build a "vision of success" and "making reliable judgments about the level of their own performance" when utilized as a component of a formative, student-centered method of evaluation (Stiggins, 2001, p. 11). Furthermore, rubrics are composed of predefined assessment scales and a list of criteria that aid in objectively evaluating the outcome (Saritha, 2016), and when properly developed, they can facilitate learning in the main objectives and preparation (Yen, 2018). Furthermore, beyond regular assessment, rubrics can help students learn skills and comprehension, as well as establish appropriate assessments of the quality of their own performance (Andrade et al., 2008; Brown & Abeywickrama, 2010). According to Beyreli and Ari's (2009) study, the analytic rubric should be utilized to analyze textual phrases. Thus, it helps teachers to discover deficiencies in students' writing skills prior to the beginning of the school semester, react to these deficiencies, and employ a suitable strategy. Furthermore, analytic scoring assists instructors in their classroom instruction and learners in their writing process, giving them more knowledge about the current situation and assisting in the identification of the strong and weak parts of students' writing owing to more ordered and thorough feedback, according to Crehan (1997). Rubrics are essential for assessing students' performance effectively and objectively. They help to ensure fairness and consistency among students. They are also useful in relieving students' concerns over grading and evaluating. "Rubrics are especially valued in the language classroom because they contribute to student learning and bring transparency to the assessment process" (Jeong, 2015, p.1). They also help teachers save time. Teachers can save time by restricting the field of evaluation to targeted learning outcomes using rubrics. Rubrics can be used by teachers to assess a diverse variety of tasks. According to Saritha (2016), rubrics aid in categorizing essays, research papers, portfolios, and pieces of art, as well as recitals, public speaking, presentations, and team projects. Furthermore, rubrics improve testing reliability and validity (Dawson, 2015).

A few teachers in foreign language education utilize standardized tests to evaluate students, which has no good impact on the students. According to Blaz (2013), when teachers begin using performance evaluation, nearly all of them report that the level of their students' performance increases. This is attributable to a variety of reasons, including:

1. Clarity: When learners understand what is expected of them, they are more prone to perform it.

2. Confidence: Recognizing the assignment requirements aims to provide students with proper instructions and deadlines for their assignment, and they feel certain that they are completing the task correctly.

3. High expectations: A well-written rubric specifies what a great product is and communicates the strong message that by following those criteria, anybody can create products of that level.

4. Student engagement: Students are more likely to complete assignments when the performance evaluation is particularly genuine.

5. Parent understanding: Performance evaluations are invaluable for validating the teacher's assessment during parental meetings or public houses.

Another factor rubrics help with evaluation is 'consistency'. Standard test theory, according to Mislevy and Sheehan (1989), is 'inconsistent' with modern developments in educational and cognitive psychology. According to the study (Blaz, 2013), instead of training that stresses memorizing, students should be taught how to arrange data, make it simple to recall, and use it in real-world circumstances. This is why the new state guidelines emphasize and specify the students' progress rubric, which should be tested most of the time. Finally, rubrics are useful for tracking students' engagement in classroom activities. According to Buttner's (2013) study, teachers should utilize the rubric as a device to address behavioral concerns with kids having familiarity and the tendency for the evaluation process and criteria as well as to encourage other students' outstanding achievements.

Teachers should utilize rubrics because they help students learn. There is little proof that rubrics influence student comprehension and evaluation, in addition to the influence of rubrics on students' progress (Andrade, 2000; Andrade & Du, 2005; Hafner & Hafner, 2003). Another study discovered that rubrics can help teachers, students, and parents track development (Ayhan & Türkyılmaz, 2015). Furthermore, when it involves evaluating, assessing, and testing skills, they are essential in language education and learning. These abilities form rubrics. We can administer many types of assessments to our pupils to learn more about their language skills. Following the test, we can conduct a routine evaluation. Thomas et al. (2004) recognize the significance of evaluation and emphasize its benefits for instructors and students in a variety of ways:

1. It gives important knowledge to help teachers enhance their teaching skills.

2. Teachers can keep track of their students' learning progress and help them improve.

3. It gives the teacher important information regarding effective teaching approaches for each student.

4. Assessment and comments can help students better grasp the learning process.

5. It allows students to build and improve their self-assessment skills, as well as to perceive assessment and evaluation as aspects of the education process.

6. Because of assessment, students can make informed judgments.

7. It aids students in their preparation for foreign tests such as the TOEFL, IELTS, and PTE Academic.

Teachers are more likely to maintain the primary lesson objectives center and front when they select educational methods and build educational settings that allow students to attain these objectives when they have clearly expressed their objectives for student progress in the shape of a rubric (Arter & McTigue, 2001). Rubrics have the ability to assist learners of color, first-generation students, and others from non-traditional backgrounds to learn more effectively. A sometimes overlooked virtue of rubrics is their ability to make instructional goals or preconceptions about the assignments themselves more transparent (Andrade & Ying, 2005). In educational circles, we frequently act on implicit cultural beliefs about student competence and conduct requirements, assuming that all students know the same knowledge and understanding.

Another critical aspect of using rubrics in courses is that teachers must include students in the evaluation process. While participating in the procedure, students assess themselves and receive the teacher's feedback. This saves time and requires greater effort on the part of the learner (Turgut & Kayaoğlu, 2015). After all of these consequences are considered, one teacher needs to establish a routine of utilizing rubrics.

Rubrics provide more than one stone to the building of learners. Raising students' comprehension of rubrics can be a major aid in comprehending the objectives of the organization and teachers, and will eventually significantly enhance their achievements (Maghsoudi & Haririan, 2013). The rubric helps students learn about their strengths and weaknesses (Saritha, 2016). One of the apparent benefits of rubric grading is that it enables learners to assess the present quality of their assignments and identify areas for improvement (Black & Wiliam, 1998; Brough & Pool, 2005; Huba & Freed, 2000; Huber & Hutchings, 2004; Walvoord & Anderson, 1998). A rubric lets students learn where they should concentrate their efforts and time-saving.

6. Rubric Types

Rubrics are scoring scales (rather than checklists) that are used in performance evaluations. They are explicitly labeled as scoring guidelines, which are made up of particular or before performance indicators and are used to evaluate students' progress on performance exams. Rubrics are the most common type of scoring tool used for assessing students' performance or items created as a consequence of an assessment task (Mertler, 2000). There are two different types of rubrics that are commonly used in performance evaluation: holistic and analytic (Mertler, 2000; Moskal, 2000). (see Figure 1).
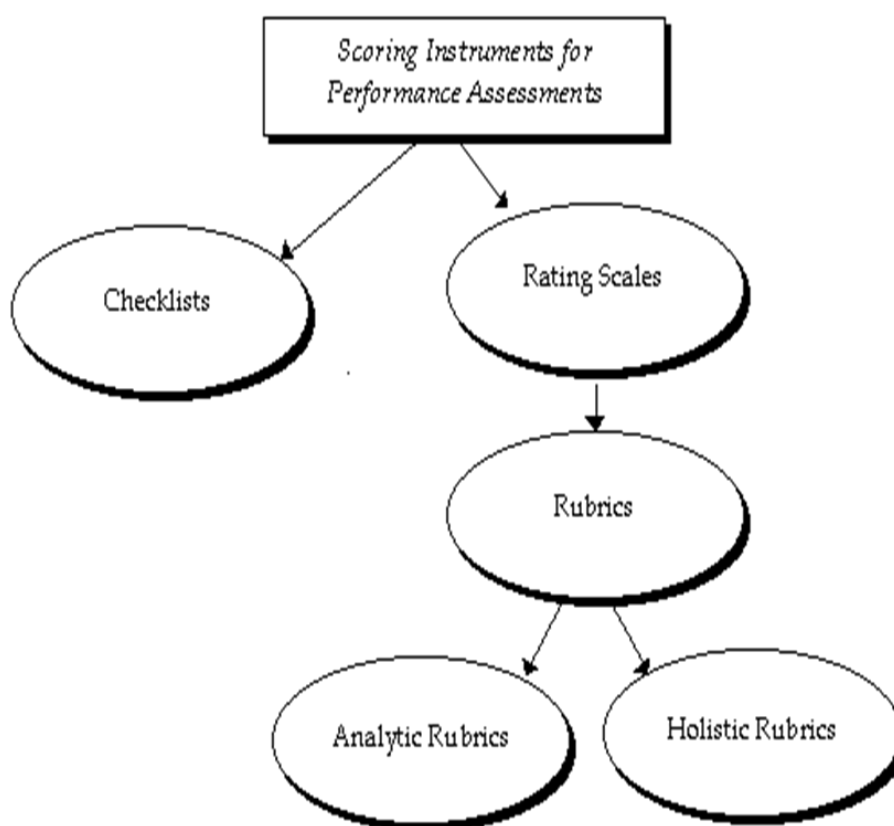

Figure 1. Types of scoring instruments for performance assessments

6.1. Holistic rubric

A holistic rubric examines the whole performance subjectively. One of the standardized tests, the IBT TOEFL, is one of the greatest examples of a rubric that rates speaking performance comprehensively. In a holistic rubric, all evaluation criteria, such as readability, style, and mechanics, are weighted on a particular scale (Brookhart, 1999). A holistic rubric delivers one complete score, typically on a range of 1 to 4 or 1 to 6 marks, based on an overall evaluation of the student's progress. Examples of popular include the words 'great,' 'good,' 'fair,' and 'bad'.

The tester compares a learner's overall achievement to a one-scale descriptor. Moskal (2000), in his debate of the distinctions between holistic and analytic rubrics, relates to circumstances where the holistic rubric is beneficial, such as when the standards set for the assessment of the various variables coincide, and a holistic scoring rubric may be beneficial to an analytic grading rubric. Furthermore, Chase (1999) noted the following on the area of application of this sort of rubric. When faults in some aspects of the procedure can be accepted as the total quality is high, holistic rubrics are commonly used. In addition to the ideas expressed by Moskal (2000) and Chase (1999), Niko (2001) says that holistic rubrics are more suited for use in instances when there is no one, correct solution, such as performance assignments that allow students to develop their own replies. According to Nitko (2001), the use of holistic rubrics also can result in a scoring method that is a little more effective than the utilization of analytic rubrics. As a result, when the goal of the performance assessment is summative, standard holistic rubrics are usually used. This is because it is critical to evaluate the entire work. When performance activities are scored using this method, the student receives a very limited quantity of feedback. Table 1 gives a framework for creating holistic grading rubrics.

Table 1. Template for Holistic Rubric (Basit, 2022)

Template for Holistic Rubric

Score	Description
4	Demonstrates exceptional understanding of the material. All requirements are met and some are exceeded.
3	Demonstrates consistent understanding of the material. All requirements are met.
2	Demonstrates partial understanding of the material. Some requirements are met.
1	Demonstrates minimal understanding of the material. Few requirements are met.
0	No response; Task not attempted.

6.2. Analytic rubric

Using an analytical rubric, many types of abilities are judged based on the student's level of competency. Teachers will create a grading system for students' progress in an analytical rubric. For example, one scale could have the subcategories 'Needs Improvement,' 'Developing,' 'Sufficient,' and/or 'Above Average.' This grading system is formatted as a grid. On top is a grading system, and on the left is qualifying data. The middle portions of the rubric can be left empty or completed with instances of how each performance level fits the requirements. An

analytic rubric, which is functionally similar to a checklist, allows for an independent evaluation of each criterion. Each criterion is assessed using a unique description scale (Brookhart, 1999). Analytic rubrics also categorize achievement and assign points earned to each level of behavior. The points are next added up to create a quantitative performance assessment. For instance, a speaking rubric may include features such as pronunciation, proper tense utilization, transitioning, vocabulary, and fluency. A template for analytic scoring rubrics is shown in Table 2.

Table 2. Template for Analytic Rubric (Basit, 2022)

ANALYTIC RUBRIC

From "Rubrics: The Heart of Assessment"

Pam Stephens

Objective

After defining the concept of art criticism, each student will write a 500-word essay with a beginning, middle, and end that accurately describes, analyzes, interprets, and judges a selected work of art.

	Value 3	Value 2	Value 1	Score
Describe	Provides a complete and accurate description of the key subject matter and elements seen in the artwork.	Provides a partial but mostly accurate description of the subject matter and/or elements seen in the artwork; some key components overlooked.	Provides an incomplete, unclear, or inaccurate description of subject matter and/or elements seen in the artwork; many key components overlooked.	
Analyze	Accurately relates how the structures of art function together to make a complete composition.	Relates with limited proficiency how the structures of art function together to make a complete composition; overlooks some important components.	Has trouble relating how the structures of art function together to make a complete composition.	
Interpret	Suggests a logical and/or symbolic meaning expressed in a work of art; supports idea with multiple points of visual evidence found in the piece.	Suggests a literal meaning expressed in a work of art; supports idea with limited points of visual evidence found in the piece.	Finds it difficult to interpret the meaning of the work; guesses meaning without visual support.	
Evaluate	Uses multiple criteria to judge the quality of a finished work of art; avoids personal opinion.	Uses a limited range of criteria to judge the quality of a work of art; personal opinion shown.	Uses personal opinion to judge the quality of a finished work of art.	
Technical	Finished paper follows rules of grammar and essay writing; is in publishable form.	Finished paper contains minor flaws in grammar and essay writing; needs editing.	Finished paper has numerous flaws in grammar and does not follow conventions of essay writing; needs rewriting.	
Notes to student				Total score

Analytic rubrics then provide thorough scores on each category. As noted earlier, the grading procedure utilizing an analytic rubric can take substantially longer than when employing a holistic rubric. According to Mertler (2001), this is because assessing various separate abilities or attributes separately contains various evaluations of the material by the teacher. Nitko (2001) refers to scenarios involving application areas where the analytic rubric is preferred, such as when a reasonably concentrated kind of answer is needed. Especially for performing assignments when there are just a few right responses and originality is not required. Furthermore, analytical rubrics generate a large number of ratings, accompanied by a summed overall score. Their use suggests a thorough examination (Mertler, 2001).

6.3. Analytic versus holistic rubrics

The information supplied thus far about the holistic rubric and the analytic rubric helps make it plainly evident that each form of rubric includes favorable and detrimental aspects. To begin, unlike analytic rubrics, which are two-dimensional and contain learning outcomes as columns and assessment methods as rows, holistic rubrics are one-dimensional, single-criterion rubrics used to evaluate students' total achievement on a task or product based on a set learning outcomes and allow the teacher to assess the achievements of students using a range of criteria. Furthermore, when employing a holistic rubric, performance statements are conveyed in paragraphs and often in full sentences, but performance statements are given in table style when utilizing analytical rubrics. When comparing the amount of time needed for each type of rubric, holistic rubrics assist assessors to gain time by lowering the number of assessments they must make, whereas analytic rubrics do not. More work is needed to formulate and execute a more holistic rubric. In terms of reliability, a holistic rubric may be routinely implemented by trained assessors, which boosts reliability. Nevertheless, unless that mark for each criterion is well-filled, assessors may not achieve the same score when using an analytic rubric. This is because a holistic rubric enables assessors to analyze the big perspective. When the feedback provided by each form of the rubric is considered, holistic rubrics do not offer particular improvement guidance, however analytic rubrics do supply useful feedback concerning aspects of weakness and strengths. Moreover, criteria in holistic rubrics cannot be evaluated; whereas, criteria in analytic rubrics can be measured to highlight the relative value of each component. Furthermore, with holistic rubrics, choosing the ideal descriptions may be difficult when student performance is at diverse levels encompassing the criteria points.

Lastly, whether one scoring rubric is an acceptable evaluation method is determined by the assessment's goal. Irrespective of which of these two procedures is used, each offers two different advantages during the review process. For starters, they examine the extent to which the set criteria have been met. Second, they provide comments to students on how to enhance their abilities. If these benefits correlate to the assessment's purpose, then a scoring rubric is most likely to be a successful evaluation. Even though the terms are susceptible to change, the goal is to assess a student's level of skill.

7. Work in Progress

7.1. Adopting and / or adapting rubrics

Rubrics assist teachers in assessing a student's total progress, saving time, providing appropriate feedback on the learner's development, clarifying the objectives for both the teachers and the students, and explaining why a student is given a specific grade. By creating a rubric as a guideline for pupils, teachers may give scaffolding. Rubrics also assist teachers in better monitoring students and developing lesson plans based on their requirements. When students understand what is expected of them and how they will be assessed, their learning and the quality of products improve. They will be able to reach their goal more effectively if they are familiar with it. A rubric is also a tool for teachers and students to interact since they align expectations and, as a result, create openness in assessment, which is very crucial for students. Furthermore, Menendez-Varela and Gregori-Giralt (2018, as cited in Kocakülah, 2021) found that involvement in the rubric development and moderating debates aided in the improvement of evaluation abilities, and rubrics may encourage classroom discussion if viewed as educational materials.

Airasian (1991) outlines the following stages for creating a successful rubric:

a. Determine the overall action or assignment to be evaluated, and practice it or picture yourself completing it;

b. Identify the main components of the presentation or output;

c. Attempt to restrict the number of performance indicators, thus they can all be evaluated throughout a performance of the student;

d. If feasible, have groups of educators go through the key traits contained in the assignment; **e.** Define the performance requirements in lines of visible behavioral issues or abilities.

f. Avoid unclear terms that obscure the content of the performance requirements; and

g. Order the performance requirements in the sequence in that they are most probably to be detected.

7.2. Some rubric samples

Educators and organizations create numerous rubrics kinds and variations. These rubrics give insight on the evaluation process, making it apparent and comprehensible, as well as providing clear directions on the assessment criteria. The following are some rubric samples:

Table 3. Independent Speaking Rubric Sample (Retrieved from
https://www.ets.org/content/dam/ets-org/pdfs/toefl/toefl-ibt-speaking-rubrics.pdf)

UNIVERSITY *of* **CAMBRIDGE**
ESOL Examinations

Experts in Language Assessment

Assessing Speaking Performance – Level B2

B2	Grammar and Vocabulary	Discourse Management	Pronunciation	Interactive Communication
5	• Shows a good degree of control of a range of simple and some complex grammatical forms. • Uses a range of appropriate vocabulary to give and exchange views on a wide range of familiar topics.	• Produces extended stretches of language with very little hesitation. • Contributions are relevant and there is a clear organisation of ideas. • Uses a range of cohesive devices and discourse markers.	• Is intelligible. • Intonation is appropriate. • Sentence and word stress is accurately placed. • Individual sounds are articulated clearly.	• Initiates and responds appropriately, linking contributions to those of other speakers. • Maintains and develops the interaction and negotiates towards an outcome.
4	*Performance shares features of Bands 3 and 5.*			
3	• Shows a good degree of control of simple grammatical forms, and attempts some complex grammatical forms. • Uses a range of appropriate vocabulary to give and exchange views on a range of familiar topics.	• Produces extended stretches of language despite some hesitation. • Contributions are relevant and there is very little repetition. • Uses a range of cohesive devices.	• Is intelligible. • Intonation is generally appropriate. • Sentence and word stress is generally accurately placed. • Individual sounds are generally articulated clearly.	• Initiates and responds appropriately. • Maintains and develops the interaction and negotiates towards an outcome with very little support.
2	*Performance shares features of Bands 1 and 3.*			
1	• Shows a good degree of control of simple grammatical forms. • Uses a range of appropriate vocabulary when talking about everyday situations.	• Produces responses which are extended beyond short phrases, despite hesitation. • Contributions are mostly relevant, despite some repetition. • Uses basic cohesive devices.	• Is mostly intelligible, and has some control of phonological features at both utterance and word levels.	• Initiates and responds appropriately. • Keeps the interaction going with very little prompting and support.
0	*Performance below Band 1.*			

Table 4. Speaking Rubric Sample. (Retrieved from https://www.cambridgeenglish.org/images/168619-assessing-speaking-performance-at-level-b2.pdf)

iBT/Next Generation TOEFL Test
Independent Speaking Rubrics (Scoring Standards)

Score	General Description	Delivery	Language Use	Topic Development
4	The response fulfills the demands of the task, with at most minor lapses in completeness. It is highly intelligible and exhibits sustained, coherent discourse. A response at this level is characterized by all of the following:	Generally well-paced flow (fluid expression). Speech is clear. It may include minor lapses, or minor difficulties with pronunciation or intonation patterns, which do not affect overall intelligibility.	The response demonstrates effective use of grammar and vocabulary. It exhibits a fairly high degree of automaticity with good control of basic and complex structures (as appropriate). Some minor (or systematic) errors are noticeable but do not obscure meaning.	Response is sustained and sufficient to the task. It is generally well developed and coherent; relationships between ideas are clear (or clear progression of ideas).
3	The response addresses the task appropriately, but may fall short of being fully developed. It is generally intelligible and coherent, with some fluidity of expression though it exhibits some noticeable lapses in the expression of ideas. A response at this level is characterized by at least two of the following:	Speech is generally clear, with some fluidity of expression, though minor difficulties with pronunciation, intonation, or pacing are noticeable and may require listener effort at times (though overall intelligibility is not significantly affected).	The response demonstrates fairly automatic and effective use of grammar and vocabulary, and fairly coherent expression of relevant ideas. Response may exhibit some imprecise or inaccurate use of vocabulary or grammatical structures or be somewhat limited in the range of structures used. This may affect overall fluency, but it does not seriously interfere with the communication of the message.	Response is mostly coherent and sustained and conveys relevant ideas/information. Overall development is somewhat limited, usually lacks elaboration or specificity. Relationships between ideas may at times not be immediately clear.
2	The response addresses the task, but development of the topic is limited. It contains intelligible speech, although problems with delivery and/or overall coherence occur; meaning may be obscured in places. A response at this level is characterized by at least two of the following:	Speech is basically intelligible, though listener effort is needed because of unclear articulation, awkward intonation, or choppy rhythm/pace; meaning may be obscured in places.	The response demonstrates limited range and control of grammar and vocabulary. These limitations often prevent full expression of ideas. For the most part, only basic sentence structures are used successfully and spoken with fluidity. Structures and vocabulary may express mainly simple (short) and/or general propositions, with simple or unclear connections made among them (serial listing, conjunction, juxtaposition).	The response is connected to the task, though the number of ideas presented or the development of ideas is limited. Mostly basic ideas are expressed with limited elaboration (details and support). At times relevant substance may be vaguely expressed or repetitious. Connections of ideas may be unclear.
1	The response is very limited in content and/or coherence or is only minimally connected to the task, or speech is largely unintelligible. A response at this level is characterized by at least two of the following:	Consistent pronunciation, stress, and intonation difficulties cause considerable listener effort; delivery is choppy, fragmented, or telegraphic; frequent pauses and hesitations.	Range and control of grammar and vocabulary severely limit (or prevent) expression of ideas and connections among ideas. Some low-level responses may rely heavily on practiced or formulaic expressions.	Limited relevant content is expressed. The response generally lacks substance beyond expression of very basic ideas. Speaker may be unable to sustain speech to complete the task and may rely heavily on repetition of the prompt.
0	Speaker makes no attempt to respond OR response is unrelated to the topic.			

8. Conclusion

In conclusion, rubrics have two significant contributions to all stakeholders including teachers, learners, institutions, and regulatory bodies in the learning process. From learners' perspectives, rubrics inform pupils of what is expected from them. As professors explicitly specify their expectations and objectives in the rubrics, students know where they need to work harder and which elements they should prioritize. They explicitly inform them about the assessment criteria so that they can focus on the expected outcomes rather than being lost and feeling

anxious in a gloomy box. This will be a kind of agreement that was decided by the decision makers, mostly teachers or educational institutions in the educational context, on what the assessment requirements are, and this will make the picture (learning-teaching-assessment) as clear as possible. The second contribution of the rubrics to learning is the backwash effect of assessment criteria. With the effective use of rubrics, students may also judge their own work and take more responsibility for their learning and get feedback from the evaluation of their performance with a concrete assessment tool, a rubric. This effect will contribute both to their in-class learning and will pave the way for their extramural learning outside the classroom. Another vital benefit of rubrics is that rubrics enable teachers to quickly monitor a student's learning process and develop the lesson by taking it into account. They assist teachers in explaining to students why they received the grade that they did and provide an unbiased assessment.

9. Suggestions for Future Research

It is unknown whether teachers embrace, adapt, and apply rubrics in their classes, which should be researched. Furthermore, their demands, both in-service and pre-service, should be revealed so that gaps in language and evaluation can be overlapped and bridged. Another idea is for important groups, either governmental or private, to create standardized rubrics to help EFL teachers. Rubrics must be used to assess any skill or assignment; however, they must outline specific learning objectives. When creating a rubric, the emphasis should be on the learning objectives rather than the job itself. The learning objectives must not be ignored by teachers.

REFERENCES

Airasian, P. W. (1991). *Classroom assessment.* Mc Grow-Hill.

Ali, M. M., Yasmin, T., & Afzaal, J. (2020). Rubric effects on assessment: An analysis. *Pakistan Social Sciences Review, 4*(2), 793-804.

Al-Rabai, A. (2014). Rubrics revisited. *International Journal of Education and Research Rubrics*, *2*(5), 473-484.

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational leadership*, *57*(5), 13-19.

Andrade, H. G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College teaching*, *53*(1), 27-31.

Andrade, H. L., Du, Y., & Wang, X. (2008). Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educational Measurement: Issues and Practice*, *27*(2), 3-13.

Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research, and Evaluation*, *10*(1), 3-11.

Arter, J. A., & Chappuis, J. (2007). *Creating & recognizing quality rubrics*. Pearson Merrill Prentice Hall.

Arter, J., & McTighe, J. (2001). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance*. Corwin Press.

Ayhan, Ü., & Türkyılmaz, M. U. (2015). Key of language assessment: Rubrics and rubric design. *International Journal of Language and Linguistics*, *2*(2), 82-92.

Bachman, L. F., Lynch, B. K., & Mason, M. (1995). Investigating variability in tasks and rater judgments in a performance test of foreign language speaking. *Language testing*, *12*(2), 238-257.

Basit, Q. (2022). 40 free rubric templates (assessment examples): Word, PDF, Templates for Microsoft® Word. Available at: https://www.wordtemplatesonline.net/40-free-rubric-templates-assessment-examples-word-pdf/ (Accessed: October 28, 2022).

Beyreli, L., & Ari, G. (2009). The use of analytic rubric in the assessment of writing performance--inter-rater concordance study. *Educational Sciences: Theory and Practice*, *9*(1), 105-125.

Black, P., & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi delta kappan*, *92*(1), 81-90.

Blaz, D. (2013). *Collections of performance tasks & rubrics: Foreign Languages*. Routledge.

Brookhart, S. M. (1999). Teaching about communicating assessment results and grading. *Educational Measurement: Issues and Practice*, *18*(1), 5-13.

Brough, J. A., & Pool, J. E. (2005). Integrating learning and assessment: The development of an assessment culture. *Curriculum integration K-12: Theory and Practice*, 196-204.

Brown, H. D., & Abeywickrama, P. (2004). Language assessment. *Principles and Classroom Practices. White Plains, NY: Pearson Education*.

Brown, J. D., & Kondo-Brown, K. (2012). Rubric-based scoring of Japanese essays: The effects on generalizability of numbers of raters and categories. *Developing, using, and analyzing rubrics in language assessment with case studies in Asian and Pacific languages*, 169-182.

Buttner, A. (2013). *Activities, Games, Assessment Strategies, and Rubrics For The Foreign Language Classroom*. Routledge.

Campbell, A. (2005). Application of ICT and rubrics to the assessment process where professional judgement is involved: the features of an e-marking tool. *Assessment & Evaluation in Higher Education*, *30*(5), 529-537.

Chase, C. I. (1999). *Contemporary Assessment for Educators*. Addison-Wesley Longman.

Crehan, K. D. (1997). A Discussion of Analytic Scoring for Writing Performance Assessments. Educational Research Association.

Crusan, D. (2010). *Assessment in the second language writing classroom*. University of Michigan Press.

Crusan, D. J. (2015). Dance, ten; looks, three: Why rubrics matter. *Assessing Writing*, *26*, 1-4.

De Silva, R. (2014). Rubrics for assessment: Their effects on ESL students' authentic task performance. In *Center for English language communication 4th symposium proceedings*.

Dunbar, N. E., Brooks, C. F., & Kubicka-Miller, T. (2006). Oral communication skills in higher education: Using a performance-based evaluation rubric to assess communication skills. *Innovative Higher Education*, *31*(2), 115-128.

East, M. (2009). Evaluating the reliability of a detailed analytic scoring rubric for foreign language writing. *Assessing writing*, *14*(2), 88-115.

Feldt, L. S., & Brennan, R. L. (1989). Reliability. In: R. L. Linn (Ed.), *Educational measurement* (3rd ed.). (3rd ed., pp. 105–146). Washington, DC: The American Council on Education and the National Council on Measurement in Education.

Ghosh, S., Bowles, M., Ranmuthugala, D., & Brooks, B. (2016). Authentic assessment in seafarer education: using literature review to investigate its validity and reliability through rubrics. *WMU Journal of Maritime Affairs*, *15*(2), 317-336.

Grabe, W., & Kaplan, R. (1996). *Theory and practice of writing: An applied linguistic perspective.* New York: Longman. Grobe, C. (1981). Syntactic maturity, mechanics, and vocabulary as predictors of quality ratings. *Research in the Teaching of English, 15,* 75–85.

Hafner, J., & Hafner, P. (2003). Quantitative analysis of the rubric as an assessment tool: an empirical study of student peer-group rating. *Int. J. Sci. Educ.*, *25*(12), 1509-1528.

Huba, M. E., & Freed, J. E. (2000). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning*. Allyn & Bacon

Huber, M. T., & Hutchings, P. (2004). *Integrative learning: Mapping the terrain. The academy in transition.* Association of American Colleges and Universities.

Janssen, G., Meier, V., & Trace, J. (2015). Building a better rubric: Mixed methods rubric revision. *Assessing writing*, *26*, 51-66.

Jeong, H. (2015). What is your teacher rubric? Extracting teachers' assessment constructs. *Practical Assessment, Research, and Evaluation*, *20*(1), 6.

Knight, L. A. (2006). *Using rubrics to assess information literacy.* Reference services review.

Kocakülah, A. (2021). Development and use of a rubric to assess undergraduates' problem solutions in Physics. *Participatory Educational Research*, *9*(3), 362-382.

Maghsoudi, M., & Haririan, J. (2013). The impact of brainstorming strategies Iranian EFL learners' writing skill regarding their social class status. *International Journal of language and Linguistics*, *1*(1), 60-67.

McNamara, T. F. (1996). *Measuring second language performance*. Longman.

Mertler, C. A. (2000). Designing scoring rubrics for your classroom. *Practical assessment, research, and evaluation*, *7*(1), 25.

Mertler, C. A. (2001). *Using performance assessment in your classroom.* Unpublished manuscript, Bowling Green State University.

Mislevy, R. J., & Sheehan, K. M. (1989). The role of collateral information about examinees in item parameter estimation. *Psychometrika*, *54*(4), 661-679.

Moskal, B. M. (2000). Scoring rubrics: What, when and how?. *Practical Assessment, Research, and Evaluation*, *7*(1), 3.

Nitko, Anthony J. (2001). *Educational assessment of students (4th ed.).* Upper Saddle River.

Oakleaf, M. J. (2006). *Assessing information literacy skills: A rubric approach*. The University of North Carolina at Chapel Hill.

O'Donnell, J. A., Oakley, M., Haney, S., O'Neill, P. N., & Taylor, D. (2011). Rubrics 101: a primer for rubric development in dental education. Journal of Dental Education, 75(9), 1163-1175.

ón, J. T., & Brydges, B. (2006). A study on using rubrics and citation analysis to measure the quality of doctoral dissertation reference lists from traditional and nontraditional institutions. *Journal of Library Administration*, *45*(3-4), 459-481.

Osana, H. P., & Seymour, J. R. (2004). Critical thinking in preservice teachers: A rubric for evaluating argumentation and statistical reasoning. *Educational Research and Evaluation*, *10*(4-6), 473-498.

Panadero, E. (2011). *Instructional help for self-assessment and self-regulation: Evaluation of the efficacy of self-assessment scripts vs. rubrics.* Unpublished doctoral dissertation). Universidad Autónoma de Madrid, Madrid, España.

Parke, C. S. (2001). An approach that examines sources of misfit to improve performance assessment items and rubrics. *Educational Assessment*, *7*(3), 201-225.

Popham, W. J. (1997). What's wrong and what's right with rubrics. *Educational leadership*, *55*(2), 72-75.

Powell, T. A. (2002). *Improving assessment and evaluation methods in film and television production courses*. Capella University.

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & evaluation in higher education*, *35*(4), 435-448.

Reitmeier, C. A., Svendsen, L. K., & Vrchota, D. A. (2004). Improving oral communication skills of students in food science courses. *Journal of Food Science Education*, *3*(2), 15-20.

Resnick, L. B., & Resnick, D. P. (1989). *Tests as Standards of Achievement in Schools.* Center for Technology in Education

Robles, D. C., Angulo, J. S., & De La Serna, M. C. (2014). Federated eRubric service to facilitate self-regulated learning in the European university model. *European Educational Research Journal*, *13*(5), 575-583.

Saritha, K. (2016). Rubric for English language teaching research. *Research Journal of English Language and Literature*, *4*(2), 725-731.

Schneider, F. J. (2006). Rubrics for teacher education in community college. *The Community College Enterprise*, *12*(1), 39.

Sert, N. (2008). Constructivism in the elementary School Curricula. *Journal of theory and Practice in Education, 4*(2), 291-316.

Song, K. H. (2006). A conceptual model of assessing teaching performance and intellectual development of teacher candidates: A pilot study in the US. *Teaching in Higher Education*, *11*(2), 175-190.

Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research, and Evaluation, 9*(1), 4.

Stiggins, R. J. (2001). *Student-involved classroom assessment*. Prentice Hall.

Taylor, L., & Galaczi, E. (2011). Scoring validity. Examining speaking: Research and practice in assessing second language speaking, 30, 171-233.

Thomas, J., Allman, C., & Beech, M. (2004). *Assessment for the diverse classroom: A handbook for teachers.* Florida Department of Education, Bureau of Exceptional Education and Student Services.

Turgut, F., & Kayaoğlu, M. N. (2015). Using rubrics as an instructional tool in EFL writing courses. *Journal of Language and Linguistic Studies*, *11*(1), 47-58.

Vaughan, C. (1991). Holistic assessment: What goes on in the rater's mind. *Assessing second language writing in academic contexts*, 111-125.

Walvoord, B. E., & Anderson, V. J. (2011). *Effective grading: a tool for learning and assessment in college*. John Wiley & Sons.

Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.

Yen, A. C. (2018). Effectiveness of using rubrics for academic writing in an EFL literature classroom. *The Asian Journal of Applied Linguistics*, *5*(1), 70-80.

To Cite this Chapter:

Höl, D. (2022). A useful assessment tool: Rubrics in second / foreign language classrooms. In Önal, A. & Büyükkarcı, K. (Eds.), *Essentials of foreign language teacher edition*, 224-242. ISRES Publishing.

ABOUT THE AUTHOR

Asst. Prof. Dr. Devrim HÖL

ORCID ID: 0000-0001-5151-2581

devrimh@pau.edu.tr

Department of Foreign Language Education at Pamukkale University

Dr. Devrim Höl is currently an assistant professor in the Department of Foreign Language Education at Pamukkale University, Turkey. He holds a master's degree in ELT from Pamukkale University, and a Ph.D. in ELT from Çanakkale Onsekiz Mart University. He worked in various fields. His research interests are Assessment, Evaluation and Testing in English as a Second Language, national and international standardized English language Proficiency tests.