

# Designing a Rubric for the Summative Evaluation of Students' Translation at University Level

Shadiya S. Mohammed<sup>1</sup> and Alya' M. H. A. Al-Rubai'I<sup>1</sup>

<sup>1</sup> Department of Translation, University of Duhok, Duhok, Kurdistan Region, Iraq.

**ABSTRACT:** This quantitative and qualitative study aims to determine the extent to which translation teachers accept the proposed criteria for evaluating their students' translation according to academic rank and the differences between translation teachers' viewpoints on the proposed criteria according to years of experience for the purpose of designing a rubric for the summative evaluation of students' translation at the university level. The study addresses the following questions: To what extent the university translation teachers accept the proposed criteria for the summative evaluation rubric according to academic rank? Do the translation teachers view the proposed criteria for the summative evaluation rubric distinctly according to teaching experience? Eight translation teachers completed the Pilot Study questionnaire, and 45 translation teachers received the Main Study questionnaire. The data were statistically analysed to test the study's validity and reliability. The study found out that the reliability and validity were generally acceptable; translation teachers' acceptance of the suggested criteria varied. However, there were no appreciable discrepancies in their opinions about the criteria based on the latter variables. The most important conclusions were as follows:

- Translation teachers' approval of the suggested criteria for evaluating their students' translation varies according to academic rank and years of experience in teaching.
- No significant discrepancies in the translation teachers' perspectives on their application of the proposed criteria according to academic rank and years of experience in teaching.
- Summative evaluation of translation can be effectively accomplished using a rubric including criteria for Equivalence, Translation Method, Mechanics of Writing, Naturalness, and Readability.

**Keywords:** Rubric, Evaluation, Assessment, Summative, Formative, Translation Quality, Students' Translations.

## 1. Introduction

Rubrics are important for evaluating students' performance and products including translations. They are helpful tools for expounding study aims, augmenting curriculums and class assignments, making teachers' expectations more comprehensible, encouraging students to perform better and evaluate their own work, helping teachers improve their feedback to students regarding their strengths and flaws, and providing consistent, objective, accurate, and faster marking. Elmgrab (2014) noted that diverse perspectives on translation have developed a variety of frameworks for evaluating translation quality. Teaching and learning might be negatively impacted, students can get confused, and course objectives can be obscured due to the significant discrepancy in teachers' evaluations of the criteria and instruments. The purpose of this study is to elicit responses from university teachers on how they view or accept the criteria that the researcher developed after consulting the relevant people. Therefore, the following study questions were set: To what extent the university translation teachers accept the proposed criteria for the summative evaluation rubric according to academic rank? Do the translation teachers view the proposed criteria for the summative evaluation rubric distinctly according to teaching experience? The study aims at:

- identifying the extent of translation teachers' acceptance of the proposed criteria for evaluating their students' translation according to academic rank for the sake of reliability and objectivity of the designed rubric; and
- identifying the differences between translation teachers' viewpoints on adopting the proposed criteria according to years of experience in teaching for the sake of reliability and objectivity of the designed rubric.

The current study is limited to teachers who specialize in translation and who were lecturing in translation departments at the Iraqi universities of Al-Iraqia, Mustansiriyah, Mosul, Baghdad, Tikrit, Duhok, Al-Esraa, Northern Technical, Basrah, Kerbala, Sulaimani, and Nawroz in the academic year 2021-2022. This study can also be helpful to translation students because it includes the standards they should be familiar with before being judged on their translation abilities.

## 2. THEORETICAL BACKGROUND AND LITERATURE REVIEW

According to Medadian and Mahabadi (2015), without evaluation and assessment, student trainees will not know how to enhance the calibre of their translations in a methodical manner, and translation teachers will not have a concrete road map to provide their students with objective marks.

Haryadi (2019) concurred that translation evaluation is one of the most important subfields of applied translation studies, which focuses on evaluating translations, developing evaluation standards, and assessing translator performance. Different frameworks and rubrics have been developed by academics to assess the quality of translation. In order to teach translation and improve students' performance and competency in this area, a scoring rubric is essential. It provides lecturers and students with feedback, acts as a scoring tool, and provides faculty with suggestions for curricular improvement.

### STEP 1. Rubrics

Popham (1997) stated that a rubric is an evaluation tool that includes criteria and explains degrees of quality. Richards and Schmidt (2011) defined it as a set of criteria or rules that are used in grading a test taker's output or performance.

A rubric clearly identifies the success criteria for each section of the student's written, spoken, or visual work, as a teaching tool and a method of evaluation. It can be used to grade homework, record class participation, or calculate final marks.

#### 2.1.1 Types of Rubrics

Rensselaer Polytechnic Institute (n.d) and Rockin Resources (n.d.) classified rubrics into several categories, as follows:

- **Analytical (Skill-based):** It is used to evaluate the specific talents that students possess and to concentrate on particular writing aspects including word choice, sentence structure, syntax, and mechanics. It is suitable for pinpointing weaknesses and highlight strengths. It needs more time to score but offers judgement that is more thorough.
- **Checklist:** It is used when an assignment completeness is evaluated rather than quality and competence.
- **Holistic:** It is employed as all of the evaluation's factors are taken into account simultaneously (e.g., clarity, organization, and mechanics). Based on an overall assessment of the student work, the rater assigns a single point scale (often 1 to 4 or 1 to 6), such as Excellent, Good, Average, Fair, or Poor.
- **Reflective:** It is used to guide students in analysing their writing in light of particular standards or content and identifying their accomplishments and shortcomings.
- **Standard-based:** It is applied to guarantee that student work is in accordance with standards. This helpful tool enables the teacher to identify which students still need to practice to become proficient in various areas.
- **Generic (General):** It is according to Arter and McTighe (2001) applied to all tasks including similar performances. It uses distinct evaluations for each criterion.
- **Task-specific:** It is adopted to evaluate a particular task and distinctive criteria independently. However, taking into account every characteristic associated with a specific assignment might not be possible (Arter & McTighe, 2001).

### 2.1.2 Characteristics of Rubrics

Different studies described diverse features of rubrics. A well-done rubric is both an educational and evaluative mechanism. TeachersFirst (n.d.) website listed characteristics for creating a purposeful rubric as shown below:

- Clarity: A rubric should have plain language to be construed reliably.
- Consistency: A consistent rubric produces trustworthy outcomes that are influenced by the degree of detail offered to students, clarity of qualities and descriptions, and training of its users.
- Continuity: The quality difference between a score point of five and a score point of four should be equivalent to that between three and two. All descriptors should reflect the continuous degrees of continuity.
- Criteria: A set of criteria should be included in an effective rubric so that students understand exactly what the teacher is looking for.
- Descriptions: The wording used in effective rubrics is quite descriptive. The rubric outlines precisely what constitutes high-quality work. The descriptors' specificity enables students to validate and recognize their scores.
- Fairness: A rubric must make it possible to draw conclusions that are pertinent, fair, and applicable to all students.
- Generalizability: A rubric needs to make it easier to generalize about students' accomplishments. It can improve the generalizability of the results by increasing the volume and diversity of students' evaluated work.
- Gradations: Depending on how well a standard has been met, several quality echelons are needed (a scale). The performance level should be described in each gradation. A rubric typically has 4-6 gradation levels.
- Matching Learning Objectives: A rubric is employed for efficient addressing of significant learning outcomes.
- Reliability: When all teachers arrive at comparable results, the rubric is considered reliable (for a given task). Time is another aspect of reliability when the rubric allows evaluating a product with the very criteria for judging the first one).
- Usability: A rubric should produce results that can inform important decisions to ameliorate instruction and learning.
- Validity: Rather than rewarding what is obvious and easy for the instructor to mark, a valid rubric rewards what is fundamental to the performance and task.

### 2.1.3 Reliability and Validity of Rubrics

Riazi (2003), Stevens and Levi (2004), and Khanmohammad and Osanloo (2009) regarded rubrics as catalysts for grading students' performance in relation to the learning and development objective. They also provided detailed judgement on the nature and depth of the students' current learning. Similarly, they thoroughly explain acceptable and unacceptable performance levels. To help students understand how the assignment links to the course material and to promote their authority in the classroom through transparency, teachers can share rubrics with them when the tests are administered.

Williams (2009) perceived reliability as the degree to which an evaluation yields the same results when administered repeatedly to the same population under the same circumstances. An evaluation is reliable if the criteria are stable and the evaluators' decisions are consistent. According to him, validity is the extent to which an evaluation measures what it is designed to measure, such as translation skills. Content validity is the extent to which an evaluation covers the skills necessary for performance.

### 2.1.4 Principles of Rubrics

University of Sussex (n.d.) presented principles of rubrics from three perspectives as shown below:

From the student's perspective

- If students are aware of the rubric beforehand, they will be able to manage their learning, as they will know how their work will be evaluated.
- Students will better comprehend their marks if the rubric is clear.

- Detailed marking criteria can assist students in understanding any further written or narrative judgement.
- They serve as instruments for students to measure their own work as well as that of others.

From the teacher's perspective

- They expressly link evaluation tasks and scoring standards to module learning objectives.
- They involve the students in evaluation, marking or module learning outcomes.

From the marker's perspective

- They contribute to accurate and consistent marking.
- They shorten the time spent for comments and marking.
- They supplement personalized feedback, not replace it.
- They facilitate marking by using thoughtful criteria.

## STEP 2. Translation Evaluation versus Translation Assessment

Many scholars have drawn attention to the distinctions between the evaluation and assessment of translation even if they sometimes overlap slightly.

Evaluation, according to Angelo and Cross (1993: 427), is based on grades, classroom activities, course material, and student competence levels, as well as discussions, collaboration, attendance, and speaking abilities.

On the other hand, assessment provides data to improve teaching and learning with a focus on results. Assessment is "an interactive process between students and professors that informs a teacher about the effectiveness of their students' learning what they are teaching," they argued. In order to help students improve their learning and study techniques, evaluation information is frequently made available to them. This information is "learner-centred, course-based, frequently anonymous, and not scored," Angelo and Cross (1993: 427) added. According to Straight (2002), evaluation is fixed because it rewards success and punishes failure; summative because it is done at the end of an academic phase to determine quality; product-oriented because it reveals what has been learned; prescriptive because it imposes external standards; judgmental because it gives an overall mark or score; and competitive because students and/or teachers compete. Summative evaluation covers the overall scope of a course or academic program. It is used to evaluate the effectiveness and results of the final academic phases; it displays the taught material, establishes outside standards, makes judgments, provides overall grades, honours achievement, and separates "passing" students from "failed" students.

Straight (2002) designated assessment as:

- Process-oriented because it exposes how learning is progressing based on measurement focus.
- Formative since it is intended to improve learning in terms of timing.
- Reflective in that it establishes internal standards or objectives based on the relationship between the administrator and the recipient.
- Diagnostic in that it points up opportunities for development in terms of findings and applications.
- Flexible because it changes as problems are explained, taking into account the continual modifiability of criteria.
- Absolute because it pursues the ideal results in terms of measuring standards.
- Collaborative because students and/or teachers can learn from one another regarding the relationships between items.

Taras (2010:125) asserted that evaluation is merely "the gathering and combining of performance data with a weighted set of goal scales to yield either comparative or numerical ratings, and in the justification of (a) the data-gathering instruments, (b) the weightings, and (c) the selection of goals". Colina (2011) argued there is a major difference between evaluating translations and assessing them. She stated that evaluation refers to the process where we judge and classify according to some set criteria therefore subjectivity may be involved.

In contrast to formative assessment, summative evaluation occurs at the halfway point and/or end of the academic year to define the degree to which curriculum objectives have been met and acts as a benchmark for successful students through scores or marks. Evaluation is a thorough process that looks at someone or something's performance to tell the difference between intended and actual results. To evaluate is to cast judgment on an

evaluated person in order to determine their aptitudes, qualities, or proficiencies. A person's ability to apply their knowledge is evaluated during an exam, and decisions are made in light of the outcomes. The evaluation process is used to compare the skills and traits of two or more persons. The evaluation's subject ought to adhere to the evaluator's prerequisites.

In the light of the scant research on summative evaluation of student translations in comparison to the evaluation of professional translators' and published translations, evaluation is opted here. The conceptual and functional qualities of evaluation are in accord with the purpose of this study, which is another justification.

### **STEP 3. Previous Studies**

Although this study has opted for the evaluation of students' translations, overlapped evaluation-assessment terminology will be used as they are stated in some of the following studies:

Martínez Melis (1997) outlined various aspects of student translation assessment as follows:

- Aims (academic, educational, speculative).
- Types (product/process/quality assessment);
- Functions (diagnostic, formative, summative);
- Objects (student translator competency, study plans, programs); and
- Means (translations, evaluating criteria, grading scales, tests, exercises, questionnaires).

Al-qinai (2000) relied on Textual Typology, Formal Correspondence, Thematic Coherence, Reference Cohesion, Lexical-Syntactical Properties, and Grammatical/Syntactic Equivalence as assessment factors. His eclectic practical model targets textual/functional or pragmatic compatibility (i.e. quality of linguistic conversion) rather than the logistics of management and presentation (i.e. quality of service). He (2000: 499) said, "The assessment of a translated text seeks to measure the degree of efficiency of the text with regard to the syntactic, semantic and pragmatic function of ST within the cultural frame and expressive potentials of both source language and target language".

Al-qinai (2000) used a variety of criteria for evaluation, including Grammatical/Syntactic Equivalence, Thematic Coherence, Reference Cohesion, Lexical-Syntactical Properties, Formal Correspondence, and Textual Typology. His heterogeneous hands-on model focused more on functional/ textual or pragmatic congruence (i.e., linguistic conversion quality) than presentation and management quality. He (2000: 499) said, "The assessment of a translated text seeks to measure the degree of efficiency of the text with regard to the syntactic, semantic, and pragmatic function of ST within the cultural frame and expressive potentials of both source language and target language".

Hady (2015) criticized Al-qinai (2000) for the following reasons:

- Failing to justify his choice of the seven criteria or indicate their significance or weight variation.
- Failing to provide any way for students who do not speak Arabic to weigh the evaluation procedures.
- Failing to offer an overall statement of adequacy.
- Failing to explain how to judge a translation if it only acts well on a few criteria.
- Relying excessively on market research to determine the quality of the translation as well as time-consumption.

Waddington (2001) divided the translation evaluation techniques employed by professors at European and Canadian institutions into two categories: error-based and holistic, or a combination of the two. According to him, the majority of Translation Quality Assessment (TQA) research has been descriptive and theoretic, and it has mainly concentrated on developing standards for good translation, identifying translation errors, correlating the severity of errors with psycholinguistic theory of Scenes and Frames and textual levels, developing scales to characterize different levels of translation proficiency, and analysing text linguistically. He stated that subjective error analysis approaches partially relied on the corrector's personal judgment and measured negative rather than positive aspects of an error. He concluded that combining holistic evaluation and mistake analysis is more advantageous. Surveyed teachers, according to Waddington (2001), evaluated student translations using error analysis, holistic method, and a combination of both. He claimed that they subjectively appraised because they

believed that objective standards would not be possible to implement in academic environments. He said that almost all contributions in TQA have centred on:

- Creating standards for a good translation assessment rubric.
- Type of translation errors.
- Using text linguistic analysis to support quality evaluation.
- Establishing hierarchically different textual levels and connecting the significance of errors to these levels.
- Assessment based on Scenes and Frames psycholinguistic theory.

Karimnia and Shahrabi (2011) criticized this model for ignoring translation shifts typically performed by a translator due to differences in the rank structures of the languages. Furthermore, because evaluation varies from person to person, the holistic method C's description revealed a substantial degree of generality and subjectivity, undermining the researcher's endeavour to objectify the work. Waddington's model had some drawbacks, according to Asgari (2021), one of which was that the unit of translation was not identified.

Medadian and Mahabadi (2015) argued that any model for evaluating student translation should be:

- Practical and Manageable: A multi-parameter model is far too unmanageable and intricate for a real-life summative evaluation of student translations.
- Objective: Postmodern and response-based views are very subjective and impressionistic in educational settings.
- Tailor-made: It is used for evaluating student translators' proficiency summatively as opposed to published translation and professional translators' proficiency.
- Non-Reductionist: It uses a single criterion, such as Skopos or analogous response, or a binary criterion, such as free/literal or overt/covert, to evaluate a comparatively complicated phenomena. Reductionism undermines the model's viability.
- Bidirectional: It takes both the translation skills and source/target languages into account.
- Objective and Holistic: It warrants accuracy.

To quantify the student translations, Medadian and Mahabadi (2015) suggested a model with five equivalence types at different language levels as a guide for a corrective scale and five matching mistake gravities in the marking scale (70 percent by error analysis and 30 marks by holistic evaluation of the translation quality). Concerning their model, certain observations can be made. Its viability and suitability were not reinforced by the insights and feedback of translation teachers they were not statistically tested. The study did not mention the academic environment and level of the targeted students (i.e. undergraduates, MA, or PhD), as well. What distinguishes the current study is that it aims at designing a workable rubric for Iraqi undergraduates' translation summative evaluation in consultation with university translation teachers. Backed by the insights and feedback of translation teachers, the study was reliable, valid, and statistically tested. Further, this study is distinct from its antecedents in terms of aims, place of study, targeted population, research instrument.

### **3. METHODOLOGY**

This part of the research describes the methodology for conducting the experimental part of the quantitative and qualitative study including Research Design, Sample, Research Tool, Reliability and Validity, Pre-Pilot, Pilot and Main Study, and Procedure and Data Collection.

#### **Step 1. Research Design**

To fulfil the aims of this study, the researcher designed quantitative and qualitative questionnaires to determine the extent of the translation teachers' acceptance of the criteria proposed for translation according to academic rank and the differences between translation teachers' viewpoints about the criteria based on years of experience in teaching (See Appendices A and C).

**Step 2. Sample**

Teachers from translation departments at thirteen Iraqi universities who took part in the two rounds of this study were given a poll question to answer. Thirty-one complete responses-out of thirty five-were received; three were duplicates, and one was blank. Eight translation teachers replied to the Pilot Study questionnaire, while forty five completed the Main Study questionnaire.

**Step 3. Research Tool**

A poll question was administered to teachers from translation departments at thirteen Iraqi universities who took part in the two rounds of this study. Thirty-one complete responses – out of thirty-five – were received; three were duplicates, and one was blank. Two questionnaires were then used. The first questionnaire was designed based on criteria for translation evaluation. The second one was constructed based on the respondents' answers to the first questionnaire. (See Appendices A, C and D)

**Step 4. Validity and Reliability**

Test-Retest Reliability was employed to evaluate the consistency and accuracy of the questionnaire instrument by carrying out the same procedure on the same respondents twice. The statistical technique was applied to the Pilot Study sample of eight teaching staff members, which was then reapplied after 14 days to calculate the sum of each test. The results of the computation of the Pearson Correlation Coefficient are shown in Table 1.

**Step 5. Pre-Pilot**

Thirteen Iraqi universities' translation faculty members who took part in the two phases of this study were given the following poll question: **What standards should university students' translation be judged against, in your opinion?** Thirty-one complete responses – out of thirty-five – were received; three were duplicates, and one was blank.

Five translation teachers were then selected to serve as jury members in order to give the study questionnaire more face and content validity.

Table 1: Pre-Pilot Questionnaire Design Criteria

	Criteria	Fit	Unfit
1	.....		

**Step 6. Pilot and Main Study**

As pointed out above, teachers from translation departments at thirteen Iraqi universities who participated in the two rounds of the study were asked the following poll question: **What standards should university students' translation be judged against, in your opinion?** Five teachers were then chosen to serve on the jury.

The same eight teachers, who make up 10% of the Main Study sample, were sent the Pilot Study questionnaire twice, with a two-week interval in between.

Table 2: Pilot Study Questionnaire

	Criteria	In my opinion, this statement is of		
		Most Importance	Moderate Importance	Least Importance
1	.....			
2	.....			
3	.....			

Although the teachers from the Pilot Study were not part of the final sample, the questionnaire's reliability was measured using data collected from the same sample twice. According to statistics, a questionnaire is considered reliable when its value is 0.8 or higher.

Forty-five translation teachers from various universities in Iraq were then given a questionnaire for the Main Study. Based on their values, the responses were analysed to show the relevance of the criteria in order from most to least important. Based on the objective criteria of Equivalence, Translation Method, Mechanics of Writing, Naturalness, and Readability, a rubric for summative evaluation of translation was proposed. (See Appendices A and B).

**Step 7. Procedure and Data Collection**

- Posing a poll question to teachers in the Translation Departments of thirteen Iraqi universities over the course of the two phases of the study, as indicated above.
- Receiving thirty-one complete responses, three duplicate responses, and one blank response to the poll question.
- Creating Pre-Pilot questionnaire design criteria and sharing them with the five-translation teacher jury, as illustrated above.
- Creating a questionnaire-style list of statements based on the teachers' responses and criteria for evaluating translation.
- Sending the Pilot Study questionnaire to the same eight teachers twice, separated by a two-week interval, to measure the reliability of the questionnaire.
- Suggesting a rubric for summative evaluation of translation based on the objective criteria of Equivalence, Translation Method, Mechanics of Writing, Naturalness, and Readability.
- Distributing the Main Study questionnaire to forty five translation teachers at various Iraqi universities.

Table 3: Test-Retest Reliability according to Pearson Correlation Coefficient

Test	Mean	Sum	Pearson Correlation Coefficient	Sig. (2-tailed)
Test 1	20.50	164	0.81	0.013
Test 2	19.75	158		
N	8	8		

Table 3 shows that the two tests had a Pearson Correlation Coefficient of 0.81, which is statistically significant and indicates that the instrument is reasonably stable. The most common formula was applied:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \tag{1}$$

The Main Study sample by University is displayed in the table below:

Table 4: Main Study sample according to university

University	Frequency	Percentage
Al-Iraqia	2	4.4
Mustansiriyah	2	4.4
Mosul	15	33.3
Baghdad	1	2.2
Tikrit	5	11.1
Duhok	11	24.4



Al-Esraa	1	2.2
Northern Technical	1	2.2
Basrah	1	2.2
Kerbala	2	4.4
Sulaimani	1	2.2
Nawroz	3	6.7
<b>Total</b>	<b>45</b>	<b>100.0</b>

#### 4. DATA ANALYSIS AND DISCUSSION

This section is devoted to the analysis of the data gathered by the two questionnaires and the suggested rubric, in addition to the discussion of the study findings in light of the pertinent research questions and aims.

**Questions 1:** *To what extent the university translation teachers accept the proposed criteria for the summative evaluation rubric according to academic rank?*

**Aim 1:** *Identifying the extent of translation teachers' acceptance of the proposed criteria for evaluating their students' translation according to academic rank, for the sake of reliability and objectivity of the designed rubric.*

Table 5: Translation teachers' acceptance of the proposed criteria according to academic rank

Academic Rank		Appropriate Equivalence at the Micro-and Macro levels of the Text	Appropriate Translation	Mechanics of Writing	Naturalness	Readability	Mean
Asst. Lecturer	Mean	3.91	4.05	4.05	4.05	4.27	4.064
Lecturer	Mean	3.75	3.75	3.87	4.13	4.13	3.925
Asst. Professor	Mean	3.69	3.85	3.69	3.92	3.85	3.800
Professor	Mean	3.50	3.50	3.00	2.50	2.50	3.000

Table 5 shows that, with the exception of Naturalness, Assistant Lecturers had the greatest Means for accepting the suggested criteria for rating their students' translation in accordance with academic rank, with a subtotal Mean of 4.064. This can be ascribed to the typical propensity of this set of teachers to increase their experience, assert their academic status, and advance their brief careers by evincing interest in learning, carrying out academic duties, and applying evaluation criteria. As a result, Lecturers, Assistant Professors, and Professors came after them, with Means of 3.925, 3.800, and 3.000 that decreased consecutively, for their acceptance of the suggested criteria. The

translation teachers differed in the acceptance of the proposed criteria for evaluating their students' translations according to academic rank.

**Questions 2:** Do the translation teachers view the proposed criteria for the summative evaluation rubric distinctly according to teaching experience?

**Aim 2:** Identifying the differences between translation teachers' viewpoints in terms of their adoption of the proposed criteria according to years of experience in teaching for the sake of reliability and objectivity of the designed rubric.

Table 6: Translation teachers' viewpoints in terms of their adoption of the proposed criteria according to years of experience in teaching

Criteria	Years of Experience in Teaching	N	Mean Rank	Kruskal Wallis Test		Sig. (2-Tailed)	
				Calculated	Tabular		
Appropriate Equivalence at the Micro-and Macro levels of the Text	1-5	9	25.94	1.418	7.81	0.701	
	6-10	9	25.11				
	11-15	11	22.18				
	16 and Over	16	20.72				
Appropriate Translation	1-5	9	24.56	0.460			0.928
	6-10	9	20.89				
	11-15	11	22.95				
	16 and Over	16	23.34				
Mechanics of Writing	1-5	9	25.28	1.164			0.762
	6-10	9	25.06				
	11-15	11	20.41				
	16 and Over	16	22.34				
Naturalness	1-5	9	23.22	0.390			0.942
	6-10	9	22.61				
	11-15	11	24.82				
	16 and Over	16	21.84				
Readability	1-5	9	23.28	6.767		0.080	
	6-10	9	31.00				
	11-15	11	17.00				
	16 and Over	16	22.47				
Total	1-5	9	24.89	0.738		0.864	
	6-10	9	24.72				
	11-15	11	23.00				

	16 and Over	16	20.97		
	<b>Total</b>	45			

The Kruskal-Wallis test was performed to examine whether there are statistically significant differences between the translation teachers' viewpoints on their adoption of the suggested criteria based on years of experience teaching variables, as revealed in the below formula. Table 6 demonstrates that there are no differences for either variable and that for every criteria, the Calculated value is less than the Tabulated value.

$$K = \left[ \frac{12}{(n+1)} \sum \frac{R_i^2}{n_i} \right] - 3(n+1) \quad (2)$$

## 5. FINDINGS

The study has made the following findings:

1. Translation teachers differed in the acceptance of the proposed criteria for evaluating their students' translation according to the academic rank variable.
2. Translation teachers differed in the acceptance of the proposed criteria for evaluating their students' translation according to the years of experience in teaching variable.
3. There were no significant differences between the translation teachers' viewpoints on their adoption of the proposed criteria according to the academic rank variable.
4. There were no significant differences between translation teachers' viewpoints on their adoption of the proposed criteria according to the years of experience in teaching variable.

## 6. CONCLUSIONS

Based on the research's findings, it was feasible to come to the following conclusions:

- Translation teachers' approval of the suggested criteria for evaluating their students' translation varies according to academic rank and years of experience in teaching.
- There are no appreciable discrepancies in the translation teachers' perspectives on their application of the proposed criteria according to academic rank and years of experience in teaching.
- Summative evaluation of translation can be effectively accomplished using a rubric including criteria for Equivalence, Translation Method, Mechanics of Writing, Naturalness, and Readability.
- The reliability and validity of the proposed rubric are generally acceptable.
- This study can be helpful for students majoring in translation because it will provide them the opportunity to improve their translation abilities after making them aware of rubrics and evaluation standards.
- The rubric can help translation teachers evaluate students' translations more effectively and objectively.
- This study can be useful to translation academics, practitioners, commissioners, rubric designers, and researchers.
- Regardless of their academic rank or years of teaching experience, teachers of translation can use this designed rubric as a well-organized and thorough instrument to evaluate the translation work of their students.
- This study is distinct from its antecedents in terms of aims, place of study, targeted population, research instrument, rubric design, summative evaluation, university translation teachers' contribution, reliability, validity, and statistical test. This study can bridge the gap in the systematic evaluation of translation in the Iraqi academic context.

## REFERENCES

1. Al-qinai, J. (2000). Translation quality assessment: strategies, parameters and procedures. *Meta*, 45(3), 497-519. <https://doi.org/10.7202/001878ar>
2. Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers*. San Francisco: Jossey-Bass. <http://www.sinclair.edu/about/learning/assessment/resources/atm/cat/>
3. Arter, J. & McTighe, J. (2001). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance*. Thousand Oaks, CA: Corwin Press. [https://scholar.google.com/scholar\\_lookup?author=J.+A.+Arter&author=J.+McTighe+&publication\\_year=2001&title=Scoring+Rubrics+in+the+Classroom%3A+Using+Performance+Criteria+for+Assessing+and+Improving+Student+Performance](https://scholar.google.com/scholar_lookup?author=J.+A.+Arter&author=J.+McTighe+&publication_year=2001&title=Scoring+Rubrics+in+the+Classroom%3A+Using+Performance+Criteria+for+Assessing+and+Improving+Student+Performance)
4. Asgari, M. (2021). Relationship between two translation quality assessments: Holistic rating and Waddington's model of assessment. *International Journal of Language and Translation Research*, 1(2), 27-46. [https://DOI.org/10.12906/978389966720\\_002](https://DOI.org/10.12906/978389966720_002)
5. Baker, M. (2018). *In other words: A coursebook on translation* (3rd ed.). London and New York: Routledge. <https://doi.org/10.4324/9781315619187>
6. Cruse, D. (1986). *Lexical semantics*. Cambridge: Cambridge University Press. [https://www.academia.edu/47995936/D\\_A\\_Cruse\\_Lexical\\_Semantics\\_Cambridge\\_Cambridge\\_University\\_Press\\_1986\\_310\\_s](https://www.academia.edu/47995936/D_A_Cruse_Lexical_Semantics_Cambridge_Cambridge_University_Press_1986_310_s)
7. Elmgrab, R. (2014). Possible criteria for evaluating students' translation errors. *International Journal of Humanities and Cultural Studies*, 1(3), 131-145. [https://www.academia.edu/13961599/Possible\\_Criteria\\_for\\_Evaluating\\_Students\\_Translation\\_Errors](https://www.academia.edu/13961599/Possible_Criteria_for_Evaluating_Students_Translation_Errors)
8. Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. L. Morgan (Eds.), *Syntax and semantics, speech acts*, (3, pp. 41-58). New York: Academic Press. <https://www.ucl.ac.uk/lis/studypacks/Grice-Logic.pdf>
9. Hady, A. (2015). *Assessing translation quality of polemical texts* (Master's Thesis). Faculty of the American University of Sharjah, College of Arts and Sciences. <https://dspace.aus.edu/xmlui/bitstream/handle/11073/8788/29.232-2015.18%20Ahmad%20S.%20Abdel%20Hady.pdf?sequence=1&isAllowed=y>
10. Haryadi, A. (2019). *An analysis of scoring rubric use to assess the student translation at D3 English Department, Politeknik Negeri Malang [Malang State Polytechnic]* (Master's Thesis). University of Muhammadiyah Malang. <https://eprints.umm.ac.id/57479/1/NASKAH.pdf>
11. Karimnia, A. & Shahraki, A. (2011). Waddington's model of translation quality assessment: A critical inquiry. *Elixir Ling. & Trans*, 40, 5219-5224. [https://www.researchgate.net/publication/266084381\\_Waddington's\\_model\\_of\\_translation\\_quality\\_assessment\\_a\\_critical\\_inquiry](https://www.researchgate.net/publication/266084381_Waddington's_model_of_translation_quality_assessment_a_critical_inquiry)
12. Khanmohammad, H. & Osanloo, M. (2009). Moving toward objective scoring: A Rubric for translation assessment. *Journal of English language Studies*, 1(1), 131-153. <https://www.sid.ir/en/journal/ViewPaper.aspx?id=221139>
13. Koller, W. (1979). *Einführung in die Übersetzungswissenschaft* [Introduction to the study of translation]. Heidelberg: Quelle und Meyer. [https://kupdf.net/download/koller-werner-einfuehrung-in-die-ubersetzungswissenschaft-pdf\\_58c58d8edc0d605608339033\\_pdf](https://kupdf.net/download/koller-werner-einfuehrung-in-die-ubersetzungswissenschaft-pdf_58c58d8edc0d605608339033_pdf)
14. Medadian, G. & Mahabadi, D. (2015). A summative translation quality assessment model for undergraduate student translations: Objectivity versus manageability. *Studies about Languages*, 40-54. <https://doi.org/10.5755/j01.sal.0.26.12421>
15. Martínez Melis, N. (1997). *Évaluation et Traduction: Cadre de Recherche sur L'évaluation dans la Didactique de la Traduction* [Evaluation and translation: A research framework on evaluation in translation didactics (Doctoral Dissertation)]. Universitat Autònoma de Barcelona [Autonomous University of Barcelona]. <https://www.thesisred.net/handle/10803/5251#page=2>
16. Popham, W. J. (1997). What's wrong - and what's right - with rubrics. *Educational leadership*, 55, 72-75. <https://www.ascd.org/el/articles/whats-wrong-and-whats-right-with-rubrics>
17. Rensselaer Polytechnic Institute (n.d.). What are rubrics? <https://provost.rpi.edu/learning-assessment/assessment/what-are-rubrics>
18. Riazi, A. M. (2003). The invisible in translation: The role of text structure in translation. *Journal of Translation*, 7, 1-8. <https://translationjournal.net/journal/24structure.htm>
19. Richards, J. C., & Schmidt, R. W. (2011). *Longman dictionary of language teaching and applied linguistics*. Routledge. <https://doi.org/10.4324/9781315833835>
20. Rockin Resources. (n.d.). Five different types of rubrics. <https://rockinresources.com/2022/10/5-different-types-of-rubrics.html>
21. Stevens, D. D., & Levi, A. (2004). Introduction to rubrics: An assessment tool to save grading time, convey effective feedback, and promote student learning. *Stylus Pub LIC*. <http://catalog.ebay.ca/introduction-to-rubrics>
22. Straight, H. S. (2002). *The difference between assessment and evaluation: Teaching assistant orientation*, (PowerPoint slides). <http://www.binghamton.edu/academics/provost/documents/assessment-evaluation-straight.ppt>
23. Taras, M. (2005). Assessment, summative and formative: Some theoretical reflections. *British Journal of Educational Studies*, 53(3), 466-478. <https://doi.org/10.1111/j.1467-8527.2005.00307.x>
24. TeachersFirst. (n.d.). Important characteristics of rubrics. <https://www.teachersfirst.com/lessons/rubrics/characteristics.cfm>
25. University of Sussex (n.d.). Principles of rubrics and grading forms. <https://staff.sussex.ac.uk/teaching/enhancement/support/assessment-design/rubrics-and-grading-forms>
26. Waddington, C. (2001). Should translations be assessed holistically or through error analysis? *Hermes, Journal of Linguistics*, 26, 15-38. <https://doi.org/10.7146/hjclb.v14i26.25637>
27. Williams, M. (2009). Translation quality assessment. *Mutatis Mutandis*, 2(1), 3-23. <http://aprendeenlinea.udea.edu.co/revistas/index.php/mutatismutandis/article/view/1825/160>

## APPENDICES

### Appendix A

#### Questionnaire 1

##### **Designing a Rubric for the Summative Evaluation of Students' Translation at the University Level**

Dear Teacher of Translation

I trust this finds you well,

I am an MA researcher at the Translation Department, College of Languages, University of Duhok. I am writing a thesis entitled: **Designing a Rubric for the Summative Evaluation of Students' Translation at the University Level** supervised by Prof Dr Alya' M. H. A. Al- Rubai'i.

You are kindly invited to respond to the following poll question, based on my supervisor's previous knowledge: **What in your opinion are the criteria for evaluating students' translation at university level?**

Also, please provide your:

1. Name:
2. Email:
3. Academic Title:
4. Type of Translation Taught:
5. Years of Experience in Teaching Translation:
6. University, College, Department:

Thank you for your valuable contribution.

Best regards,

Shadiya Sabri

### Appendix B

#### Questionnaire 2

##### **Designing a Rubric for the Summative Evaluation of Students' Translation at the University Level**

Dear Jury Member,

I trust this finds you well,

I am an MA researcher at the Translation Department, College of Languages, University of Duhok. I am writing a thesis entitled: **Designing a Rubric for the Summative Evaluation of Students' Translation at the University Level** supervised by Prof Dr Alya' M. H. A. Al- Rubai'i.

You are kindly invited to act as a jury member to provide your judging inputs on the criteria of the rubric-related questionnaire below, based on the Iraqi translation teachers' responses already received and my supervisor's previous knowledge. You may add, delete, or modify any item in the space underneath each. Also, please list your:

1. Full Name
2. Specialty
3. Academic Rank
4. Place of Work

Thank you for your valuable contribution.

Best regards,

Shadiya Sabri

Criteria	Unfit	Fit
Equivalence vs Non-equivalence		
Grammatical accuracy		
Mechanics of writing		
Naturalness		
Readability		

This is the way they will appear in the final rubrics: the analytic and the holistic.

**Analytical Rubric**

Criteria	Very poor 1-5	Fair to poor 6-15	Good to average 16-25	Excellent to very good 26-35
Equivalence vs Non-equivalence points (35)	Requirements of achieving equivalence and compensating for non-equivalence by using the appropriate translation strategies and translation method are			
	Noticeably not met	Frequently not met	Somewhat met	Met

Grammatical accuracy (35) points	Errors are very severe	Frequent errors	Some errors	Few, if any, errors
	<b>Very poor</b> 1-2	<b>Fair to poor</b> 3-4	<b>Good to average</b> 5-7	<b>Excellent to very good</b> 8-10
Mechanics of writing (10) Points (Capitalization, if any, because not all languages have capital letters, punctuation, spelling, and paragraphing)	Errors are very noticeable	Frequent errors	Some errors	Few, if any, errors

**A note on Equivalence:**

It is understood that translators in their endeavour to establish equivalence between the ST and the TT take the socio-cultural settings of both languages into consideration. Hence, there is no need to add another type of equivalence, which is cultural equivalence because culture runs through all types of equivalence. The six types are equivalence at word level and above word level (collocations and idioms), grammatical equivalence, textual equivalence (cohesion and information structure), pragmatic equivalence (implicature and coherence), semiotic equivalence (verbal and non-verbal signs), and text-normative equivalence (text type requirements).

**Holistic Rubric**

Criteria	Successful 8-10	Almost successful 5-7	Adequate 3-4	Inadequate 1-2
Naturalness (10) points	All translation reads like a piece originally written in TL	Large sections read like a piece originally written in TL, but others read like a translation	Certain parts read like a piece originally written in TL, but others read like a translation	Almost the entire text reads like a translation
Readability (10) points	All translation is easy to read and understand	Large sections are easy to read and understand, but others are not	Certain parts are easy to read and understand, but others are not	Almost the entire text is not easy to read and understand

**Appendix C**

**Questionnaire 3**

**Designing a Rubric for the Summative Evaluation of Students' Translation at the University Level**

Dear Translation Teacher,

I am a researcher at the Translation Department, College of Languages, University of Duhok, writing an MA dissertation entitled above. Please rate your level of agreement with the translation evaluation criteria per the below rubric scale.

Academic Degree: Master's Degree                       Doctorate

Academic Rank: Asst. Lecturer       Lecturer       Asst. Professor       Professor

Gender: Female                       Male

Years of Service: 1-5     6-10     11-15     16-More

Age: (                      )



University: ( )

**Note: Please keep your responses anonymous.**

Highly appreciate your valuable contribution and time.

### **Explanation of some of the problematic terms used in the rubric**

This explanation is for the teacher to consult in order to see how some problematic terms are defined in the rubric.

#### **1. Equivalence:**

The map of equivalence involves six types. The first five types are taken from Baker (2018) with some minor modifications. The last one is taken from Koller (1979 as cited in Medadian & Mahabadi, 2015, p. 49). It is understood that translators in their endeavour to establish equivalence between the ST and the TT take the socio-cultural setting of both languages into consideration. Hence, there is no need to add another type of equivalence, which is cultural equivalence because culture runs through all types of equivalence. The six types are equivalence at word level and above word level, grammatical equivalence, textual equivalence, pragmatic equivalence, semiotic equivalence, and text-normative equivalence.

##### **a. Equivalence at word level and above word level:**

Equivalence at word level includes **lexical words, function words, and inserts**. However, words also tend to occur in sequences and behave as a combination. Of particular interest to translation are **collocations** and **idioms** because they are language specific.

It should also be noted that words do not occur in a vacuum but in context, which helps determine the meaning intended by the addresser. **Two types of meaning** are particularly important here: **propositional and expressive**. Cruse (1986, p. 277) maintains that they are the ones which the addresser "principally utilizes and directly manipulates in order to convey his intended message." Propositional meaning is the basic meaning, which can be judged as true or false as it refers to or describes someone or something in a real or fictional world. However, expressive meaning cannot be judged in such terms because it is related to the addresser's thoughts or feelings (Cruse, 1968, p. 271).

##### **b. Grammatical equivalence**

Every language has its own rules for the analysis of the structure of words (morphology) and for the use of words in order to construct phrases and sentences (syntax). Equivalence at this level can be seen as choosing a grammatical category in the TL that maps an equivalent meaning in the SL.

##### **c. Textual equivalence**

It is equivalence achieved between the ST and TT in terms of **cohesion** and **information structure**. Cohesion is the grammatical and lexical relationships between the different elements of a text (Richards & Schmidt, 2011). Information structure, on the other hand, is the use of word order in such a way as to highlight certain parts of the message.

##### **d. Pragmatic equivalence**

It is equivalence achieved between the ST and TT in terms of **implicature** and **coherence**. Implicature is what the speaker suggests, implies, or means with an utterance rather than what s/he literally says (Grice, 1975). Coherence is "the network of conceptual relations which underlie the surface text" and helps "organize and make a text" (p. 235).

##### **e. Semiotic equivalence**

It is equivalence achieved between the ST and TT in terms of signs, both verbal and non-verbal (Baker, 2018, p. 279).

##### **f. Text-normative equivalence**

It is equivalence achieved between the ST and TT in terms of text types (Koller, 1979 as cited in Medadian & Mahabadi, 2015, p. 49). No matter what text typology the teacher adopts, text type requirements should be met in translation.

2. **Readability (clarity):** Whether a translation work is easy to understand, including sentence length, redundancy, ambiguity, and lack of structure/clarity.
3. **Naturalness (acceptability):** Adherence to norms of the TL, e.g., the use of its grammatical system or rules, and culture to produce TT, which are not very difficult for the TL readers.

**The Suggested Rubric**

Criteria	Levels of Performance					Marks
	Excellent (9-10)	Very Good (7-8)	Good (5-6)	Average (3-4)	Poor (1-2)	
Appropriate Equivalence at the Micro- and Macro-Levels of the Text						60 (Each of the six types of equivalence 10 marks)
Appropriate Translation Method						10
Mechanics of Writing						10
Naturalness						10

<b>Readability</b>						10
--------------------	--	--	--	--	--	----

**Appendix D**  
**Questionnaire 4**

Please tick in the square of the extent of your agreement on the items of the suggested rubric

	<b>Very Strongly Agree</b>	<b>Strongly Agree</b>	<b>Moderately Agree</b>	<b>Slightly Agree</b>	<b>Hardly Agree</b>	<b>Notes</b>
<b>Appropriate Equivalence at the Micro- and Macro-Levels of the Text</b>						
<b>Appropriate Translation Method</b>						
<b>Mechanics of Writing</b>						

<b>Naturalness</b>						
<b>Readability</b>						