



Using Rubrics for Consistency and Subjectivity Minimization) in The Sudan School Certificate English Examinations (SSC)

Elawad Yagoub Ahemad, Nagla Mohamed Alhasan Mokhtar

¹Designation of 1st Author, ²Designation of 2nd Author, ³Designation of 3rd Author
English language Institute – Jazan University,
Jazan University – Jazan Saudi Arabia

Abstract

This study investigates Sudanese secondary schools teachers' perceptions towards the use of rubrics for assessing writing essay and the procedures they use to rate composition in (SSC) exam as well as in their daily classwork. The participants of this study were 65 English language teachers, who assigned to mark composition question in Sudan Secondary Schools Certificate (SSC) this year June 2022. Descriptive analysis (mean, standard deviation and mean) was used to analyze the questionnaire. The findings show lack of consensus among raters towards the way that they use to assess composition. Moreover, it appears that there is neither clear standard rubric for assessing composition, on exam paper nor the model answer. The study also reveals that English language teachers need more workshops in language assessment and evaluation. The study suggests that using an analytical rubric leads to fair and reliable results. It also enhance learning among learners. It also suggests that the course designers have to pay attention the sequence of writing skill exercises from sentence to paragraph and including new needs for writing such as emails, blogs and short messages

IndexTerms - raters, composition, writing, rubric, Sudan, assessing, assessment

I. INTRODUCTION

Writing assessment in general is considered complicated by many researchers and teachers especially in terms of the scoring side. Throughout the researcher's work as a classroom teacher, he noticed students' dissatisfaction of their given marks in subjective questions. Moreover, on the side of the teachers, He also noticed that some of English teachers do not follow clear, analytical procedure for marking students' writing even though a devoted rubric is designed for the purpose of scoring and marking from fairness perspective. The researcher intends to investigate this issue taking the Sudan School Certificate English Examinations (SSC) as model for many reasons. Firstly, the Sudanese Secondary exam is a main gate for university admission in Sudan. Secondly, few researches have been conducted about (SSC) English exam as. Siddiek (2010), who investigated them from the perspective of content validity and comprehensiveness.

Despite the lack of studies that dealt with assessment in Sudan, the researcher has found out that there is lack in using rubric in investigating the assessment of essay writing among Sudanese raters although it is an integral and essential part of the assessment process. Hence, the researcher endeavors to investigate English language exam raters' prospective and rubric implementation to construct scoring fairness and objectivity. Fulcher (2007) argued "There is a need for a special kind of scoring in ESL writing assessment contexts is to help ensure that scores reflect the salient facets of writing in a balanced way" (p.250).

II. LITERATURE REVIEW

Using rubrics for assessing writing

A scoring rubric is known as a tool for assessing student's performance. It lists certain criteria that describe levels of quality expected from learners. Jonsson & Svingby (2007) define a rubric as "a scoring tool for qualitative rating of authentic or complex student work. It includes criteria for rating important dimensions of performance, as well as standards of attainment for those criteria" (p.131). Although rubrics are mainly used for assessing students work, they can serve important role as well (Heidi, Du & Wang 2008). They can teach and evaluate at the same time by helping students to "make dependable judgments about the quality of their own work" (Reddy & Andrade, 2010, p.437). Heidi et al (2008) claim that, rubrics can increase the students' knowledge about the criteria needed for assessing writing. Moreover, Reddy & Andrade (2010) ask for involving students in co-creating rubrics not only informing them about it. (p444).

Robert, Miller, & Gronlund (2009, p.243) argue that the extended-response question requires complex behaviors that cannot be measured by more objective means. They added "When the evaluation of answers is not guided by clearly defined outcomes and scoring rubrics, it tends to be based on less stable, intuitive judgments"

Types of rubrics

Two types of rubrics are considered for scoring writing, holistic and analytical. Holistic grading scales award a single grade for the entire essay and place more emphasis on the final product than the writing process (Finson, as cited in Taylor, 2013, p.11). On the other hand, according to Puspasari (2011), "An analytical rubric articulates levels for each aspect of assessment so the teachers can assess students' performances on each criterion" (p.7). Comparing the two kinds of rubrics, East (2009) argues that "The strength of the holistic rubric lies in its practicality" (p.90) he adds "Holistic rubric can be problematic because raters will not be in complete agreement on scores. Moreover, if the written work checked by the same rater, it might receive different score. In this case reliability is questionable as Fulcher (2007, p.251) states "The nature of holistic judgments imposes a practical limitation".

Bearing in mind the limitation of holistic score, many researchers prefer analytic scoring despite the long time it takes to apply. According to East (2009, p.91). Weigle (2002) adds "Analytic scoring provides detailed information about the test taker's performance in different aspects of writing" (p.136). Chowdhury (2019) Analytic scoring rubrics can be useful in showing learners' weaknesses and strengths, (p.63). Furthermore, they are highly reliable and provides some degree of diagnostic information, Fulcher (2007, p.254).

Teachers and students' perceptions

A number of studies investigated teachers and students perceptions about using rubrics for assessing writing showing remarkable findings. Sohrabi, Zghanbari and Nabbasi conducted study in (2022) Investigating Iranian EFL teachers' perceptions of writing assessment literacy and came out with these findings, teachers have positive attitude and awareness toward assessing writing but their level of assessment literacy was not reflected in their practice, particularly when it came to the use of assessment rubrics. Moreover, Li (2018) in his research entitled "Constructing and Applying Rubrics in College-Level EFL Writing Assessment in China" examined that rubrics are "not widely used by the college-level English instructors." (p.38). Results also show the effectiveness of using rubrics for assessing writing despite being helpful for students' writing performance and increasing their confidence. Mahmoudi, & Bugra, (2020). In addition Miller (2012) and Li, (2018) claim that teachers considered good writing, the one that is free of vocabulary and grammatical errors.

On students' side, Miller (2012) explains their annoyance about how their writing evaluated before their teacher familiarizes them with the rubric he uses for marking written work. Also, Li, (2018) comes to a similar result that "students feel better about a writing assignment knowing exactly what they will be graded on prior to the assignment" (p.9). Waragh (2019) shows that students' did not know about the criteria used for assessing their writing, which affect their ability to gain high scores.

Teacher training and experience

There is almost unanimity among researchers on the importance of teacher training on how to assess writing and using rubrics for that purpose. Li, (2018) claims "English teachers did not receive adequate training concerning the use of rubrics for writing assessment" (p.66). While Sohrabi, et al (2022) add "Teachers relied on their individual expertise developed throughout their years of practice (p.1). Davis (2016) noticed significant effect of raters' performance before and after training sessions and experience.

The Purpose of the study

The Sudan School Certificate Examination (SSC) is considered the most important exam for students schooling life. Therefore, their admission at university level and future are shaped according to this matter. For these reasons, fair and clear criteria for assessment should be set and shared among stakeholders, teachers and even students. SC English exam lacks clear rubric for assessing students' performance in writing skill.

For the reason stated above, the researcher intends to examine whether rating scale or rubric is used to assess writing. Finding out how teachers rate composition in their ordinary classes, and to investigate teachers' awareness about the use of rubrics and writing assessment by answering the following research questions.

1. Is there a standard scale for rating composition in English language exam?
2. How do teachers normally rate their students' writing?
3. What are teachers' perceptions towards the use of rubrics for marking writing?

III. METHODOLOGY

Participants

The sample of the study was chosen from male and female English language teachers working for Sudan Ministry of education, who participated in rating Sudan School Certificate English Examinations (SSC) this year August 2022. The total number of English language raters were 623 teachers from different states of the country. 76 questionnaires were distributed to around 90 teachers who were assigned to scoring composition, depending on their years of teaching expertise (table 1). 49 out of 65 had experience more than 10 years. Only 65 questionnaires were analyzed in the study excluding 11 questionnaires due to incomplete response. 39 of the participants were graduated from faculty of education and nearly half of them had training course in assessment

Table 1

Demographic information of the participants

	Categories	No	Categories	No
Qualifications	Bachelor	52	1 to 5 yrs.	2
	MA	11	6 to 10 yrs.	4
	PhD	2	Experience	10
Specialization	education	39	11 to 15 yrs.	17
	Arts	21	16 to 20 yrs.	32
	other	5	more than 20	33
Gender	male	37	Training	33
	female	28	course in assessment	32

Instrument

A paper-based questionnaire composed of 24 questions (items) and 4 sections was designed for collecting the data. The first section was for demographic information. 3 point Likert scale was used in section two to collect data about the raters' plan for rating composition in SSC exam. Teachers' daily habit for rating class work was investigated in section three using 5 point Likert scale (1 strongly agree to 5 strongly disagree). The procedure was used in section four which dealt with teachers' perceptions using rubrics for assessing writing.

Procedure

Analytic descriptive approach was used to analyze the collected data. SPSS version 28 was used for the analysis. The questionnaire was distributed to the participants during the break-hour at the English exam rating centre. Cronbach's alpha was used to measure the reliability of the questionnaire's items and it is .802, which indicates high internal consistency.

IV. RESULTS

Table 2

Plan for checking composition

No	Items	Mean	SD	Rank	Degree
1	There is a rubric for rating composition on answer key paper.	1.52	.73	5	Disagree
2	The syllabus contains exercises about how to write a composition.	1.72	.89	2	Disagree
3	The topics are related to the students' needs to write outside.	1.71	.76	3	Disagree
4	Students know how their writing will be assessed.	2.11	.81	1	Disagree
5	Composition topics are suitable for the students' level in regard to the syllabus.	1.66	.79	4	Disagree
	Weighed mean	1.74			Disagree
	Std. deviation		.45		
			2		

To find out an answer for the first question, (Is there a standard scale for rating composition in English language exam?) Descriptive analysis was done as shown in table 2. From which the highest average was given to item 4: (Students know how their writing will be assessed). with mean 2.11 and Std. Deviation .812 While the lowest average was awarded to the first item. (There is a rubric for rating composition on answer key paper). With mean 1.52 and Std. deviation .731. Items 2 and 3 show nearly similar means 1.72 and 1.71. Item 5 shows clearly the variations of the teachers' responses, nearly half of respondents (53.8%) claimed that composition topics are suitable for the students' level in regard to the syllabus while (26.2%) with mean 1.66 and .796 Std. deviations.

From the table above, it shows variations in the participants' responses which lowers the means of each item below the level of acceptance as yes responses. (3 to 2.34). As seen in the table, the weighed mean for the five items is only 1.74.

TABLE 3

FEATURES CONSIDERED IMPORTANT BY EXAM CHECKERS

	FEATURES	YES		NO		TOTAL
1.	VOCABULARY & STRUCTURE	29	44.6%	36	55.4%	65
2.	SPELLING	38	58.5%	27	41.5%	65
3.	NEAT PRESENTATION	42	64.6%	23	35%	65
4.	PUNCTUATION & CAPITALIZATION	26	40%	39	60%	65
5.	COHESION	30	53.8%	35	46.2%	65
6.	ACCURATE LANGUAGE	42	64.6%	23	35.4%	65
7.	ORGANIZATION (SUPPORTING SENTENCES & CONCLUSION ARE LOGICALLY ORDERED)	26	40%	39	60%	65

Table 3 shows that, the raters considered writing good if it has neat presentation and accurate language (64.6%). This result goes with Li (2018) result in the area of grammar but not vocabulary which is not in the top of the participants' priority in this study. It is also noticed that, organization (Supporting sentences &

conclusion are logically ordered) and (Punctuation& capitalization) have got the lowest attention by the teachers (40%).

Table 4

Types of writing tasks appear in exam paper in terms of genres

Writing tasks	Frequency	Percent
Letters	5	%7.7
Texts	45	69.2%
academic essays	15	23.1%
Total	65	100%

From table 4, it appears that the writing section in the exam paper concentrates on one kind of writing, writing an essay,(69.2%) ,(23.1) for academic writing and (7%) of the participants claimed that there were questions about letter writings. It is noticed that, writing emails, lists, taking notes and messages are not included in the exam paper and syllabus as well.

Table 5

Teacher's ordinary habits for rating his class writing

No	Items	Mean	SD	Rank	Degree
1	I always use holistic grading (look for overall grading quality of writing) to assess students' written exams.	2.00	1.061	4	No
2	I give students idea about the rubric before the exam	2.09	1.234	3	No
3	I only focus on spelling and grammatical error when I check a written task.	2.77	1.549	1	neutral
4	I create my own rubric to mark students' assignments.	2.38	1.307	2	No
	Weighed mean	2.3115			
	Std. deviation		.8421		

In table 5, it is noticed that in all items the mean is between (1.80 to 2.59), which is considered (disagree) except item 3 (I only focus on spelling and grammatical error when I check a written task.) with mean (2.77) which lies in the range of neutral (2.60 to 3.39).the weighed mean for this domain tends to (disagree).

Table 6

Teachers' perceptions about using rubrics

No	Items	Mean	SD	Rank	Degree
1	Using a rubric keeps bias to the minimum.	2.02	1.205	7	Disagree
2	Using a rubric promotes students awareness about the criteria needed for a good writing.	2.15	1.202	3	Disagree
3	Using a rubric makes assessment consistent and objective.	2.15	1.093	3	Disagree
4	Using a rubric shows strengths and weaknesses of the students.	2.02	1.038	7	Disagree
5	Using a rubric guides both teachers and	2.28	1.244	2	Disagree

	learners in learning and assessment process.				
6	A rubric elements should vary according to the course objectives.	2.09	1.195	6	Disagree
7	Teachers` experience affects students` score.	2.11	1.252	5	Disagree
8	A rubric can be shared with students for learning as well as for assessing.	2.35	1.397	1	Disagree
	Weighed mean	2.1462			Disagree
	Std. deviation		.8160		

Table 6 shows the weighted average of this section was (2.1462) with std. deviation (.8160), which indicate that the trends of (Teachers` perceptions about using rubrics) is (disagree) according to 5-point Likert scale since (2.1462) is between (1.80 to 2.59) in the interval. The lowest mean (2, 02) was given to item 1 and item 4(Using a rubric keeps bias to the minimum, Using a rubric shows strengths and weaknesses of the students. While the highest mean (2.35) goes to item 8 (A rubric can be shared with students for learning as well as for assessing.).

V Discussion

The first question in this study examined if there is a standard scale for rating composition in English language exam. The results of the study show a clear lack of consensus among the raters in their understanding and the way they assess compositions. This finding goes in align with Li (2018) and Sohrabi et al (2022) which indicates that their assessment literacy was not reflected in their practice .This refers to many reasons: there is neither clear standard rubric for assessing composition on the exam paper nor model answer, the students` books lack systematic tasks with clear criteria to show the learners what is needed and what should be considered in their writing. Moreover, the teachers who rate the composition question are assigned to this task according to their long individual experience in teaching (table 1), this finding also corroborated by Sohrabi et al (2022).it is also noticed that nearly half of the participants received training in language assessment but it seems that it wasn't enough and adequate, as Li (2018) came to the same result.

Table 3 show contradicted findings concerning features that considered important by exam checkers, and these variations in responses emphasize the importance of using analytical rubric to establish consistency and fairness in the exam results.

The second question of this study reveals that teachers do not follow certain rubric, created by them or someone else, to assess the written composition, moreover they do not give students an idea about how their work will be assessed because there is no rubric reveal on the question paper, and it is worth mention that will have positive wash-back effects on the students' performance if they know how their writing will be assessed.

The third question in this study set out to investigate teachers' perceptions towards the use of rubrics for marking writing .It appears from response, teachers are not fully aware of the importance of using rub Based on the above findings, urgent need for improving the Sudan school certificate English examinations (SSC) consistency and fairness. In this case some implications could be suggested. First, exam writers should include a rubric for marking essay question. No doubt it has a positive wash-back on teaching and

for learners to know how their work will be assessed. Secondly, in service training courses in assessment for teachers will develop them professionally and raise their awareness in language testing and assessment. Thirdly, teachers should try to involve students in peer assessment. Thus a solid research base is needed to put (SSC) on the right track as an achievement exam that determine the generation future.

REFERENCES

- [1] 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.
- [2] Chowdhury, F. (2019). Application of rubrics in the classroom: A vital tool for improvement in assessment, feedback and learning. *International Education Studies*, 12(1), 61-68.
- [3] Davis, L. (2016). The influence of training and experience on rater performance in scoring spoken language. *Language Testing*, 33(1), 117–135. <https://doi.org/10.1177/026553215582282>
- [4] East, M. (2009). Evaluating the reliability of a detailed analytic scoring rubric for foreign language writing. *Assessing writing*, 14(2), 88-115.
- [5] Fulcher, G., & Davidson, F. (2007). *Language testing and assessment*. London and New York: Routledge.
- [6] Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2, 130-144. <https://doi.org/10.1016/j.edurev.2007.05.002>
- [7] Li, C. (2018). *Constructing and applying rubrics in college-level EFL writing assessment in China* (Doctoral dissertation, Syracuse University).
- [8] Mahmoudi, F., & Bugra, C. (2020). The Effects of using rubrics and face to face feedback in teaching writing skill in higher education. *International Online Journal of Education and Teaching*, 7(1), 150-158.
- [9] Miller, J. M. (2012). *Fixing mechanics: a study in cross-curricular use of rubrics to grade writing*. https://digitalcommons.lsu.edu/gradschool_theses
- [10] Puspasari, I. D. (2011). *Designing a rubric to assess vocational high school students' writing*. https://eprints.uny.id.ac.id/40286/1/Inggrita%20Dewi%20Pusuasari_062024104

- [11] Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & evaluation in higher education*, 35(4), 435-448. <https://doi.org/10.1080/02602930902862859>
- [12] Robert, L., Miller, M. D., & Gronlund, N. E. (2009). *Measurement and assessment in teaching*. Merrill Prentice Hall.
- [13] Siddiek, A. (2010). Evaluation of the Sudan School Certificate English Examinations. *English Language Teaching*, 3(2), 1-37.
- [14] Sohrabi, Z., Ghanbari, N., & Abbasi, A. (2022). Iranian EFL teachers' perceptions of writing assessment literacy: a countrywide study. *Language Testing in Asia*, 12(1), 1-19.
- [15] Taylor, D. M. (2013). *Writing rubrics as formative assessments in an elementary classroom* (Doctoral dissertation). http://digitalcommons.brockport.edu/ehd_theses/258
- [16] Waragh, Imad. (2019). *Assessment Criteria in EFL Writing Skills*.
<https://www.researchgate.net/publication/332465100>
- [17] Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.