



## Practices of Examiners in Marking English Essay at HSSC Level

Muhammad Ahmad Hashmi<sup>1</sup>, Muhammad Asim Mahmood<sup>2</sup>, Aamir Shehzad<sup>3</sup>, Tanzeel-Ur-Rehman<sup>4</sup>

<sup>1</sup> Department of English, University of Okara, Punjab, Pakistan. Email: muhammadahmadhashmi@gmail.com

<sup>2</sup> Dean, faculty of Arts and Social Sciences, Government College University, Faisalabad, Punjab, Pakistan. Email: masimrai@gmail.com

<sup>3</sup> Assistant Professor, Department of English, University of Okara, Punjab, Pakistan. E-mail: aamirshzd@gmail.com

<sup>4</sup> Department of English, University of Okara, Punjab, Pakistan. Punjab, E-mail: tanzeelnisar89@hotmail.com

### ARTICLE INFO

#### Article History:

Received: October 13, 2021

Revised: November 14, 2021

Accepted: November 14, 2021

Available Online: November 15, 2021

#### Keywords:

Marking Criteria

HSSC Exam

Rubrics

English Assessment

Skill-based Learning

### ABSTRACT

The present study was carried out to spot the raters' practices in marking English essays at the intermediate level in contrast to the marking system of O-level English language assessment in Punjab, Pakistan. The preceding researchers opined that a significant number of students fail in English as a subject. To unravel, the facts a sample of 350 HSSC (Higher Secondary School Certificate) paper raters with mixed experience, academic and professional qualifications have been taken randomly from the different districts of Punjab. Many of them had worked as examiners or sub-examiner for more than 10 years. The researchers formulated, piloted, and self-administered a questionnaire by visiting their workplaces. SPSS has been used to assay the data, and results have been generated. The results deduced that HSSC paper raters never exercised any rubric. They deduct and award scores on their individual judgments about the kind/number of mistakes, length of an essay, and handwriting. There has been found a scarcity of inter-rater reliability. Further, they are self-trained and are ignorant of the objectives; set in the national curriculum for HSSC. The current study is significant as it has implications onboard officials, policymakers, and examiners, which will finally promote skill-based learning.

© 2021 The Authors, Published by iRASD. This is an Open Access Article under the Creative Common Attribution Non-Commercial 4.0

Corresponding Author's Email: [muhammadahmadhashmi@gmail.com](mailto:muhammadahmadhashmi@gmail.com)

## 1. Introduction

Assessment serves as a key to check out the implementation of objectives; its results show the directions and recommendations to mend the policies of the National Curriculum. The present study reports the current English paper marking strategies and reasons behind the failure of students in passing English paper at the HSSC level. Further, the assessment system in Pakistan does not assay the genuine competence of students (Siddique, 2014), but the whole education system pivots on the exam (Khan, 2006). It is the denouement of traditional practices of paper setters in the making of English papers. Unfortunately, assessment in Pakistan just assesses students' knowledge rather than competence (Warsi, 2004). Examiners do not follow any set criteria to mark English papers based on the directions of the national curriculum. Mirza, Nosheen and Mahmood in 1999, reported that examiners should follow a set of standard rubrics to mark English papers at the HSSC level.

Examiners play a crucial role in defining the results of the assessment. A part of the HSSC English paper consists of subjective questions: including, writing an essay, which is supposed to assay the students' ability to recall, integrate and organize the ideas. Therefore, the present study shares its part to explore; the effectiveness of essay type questions in line with the objectives of the national curriculum. In addition, Shirazi (2004) has reported that the

examination system in Pakistan only assesses the students' performance, whereas it should measure the teachers' performance. It is because Board of Intermediate and Secondary Education (BISE) in Pakistan does not conduct regular and scheduled training for its examiners. The outcome of these practices shows very drastic effects that paper checkers only focus on the information and ignore the creative part of writing (Khan, 2012), which can only be checked with a standard set of rubrics.

A rubric holds its significant part to analyze any written discourse systematically (Andrade, 2000). e.g., it provides criteria to score the paper objectively and supports the process of evaluation while assessing a written task (Hafner & Hafner, 2003; Quinlan, 2006; Glickman-Bond & Rose, 2006). According to Gulzar, Buriro and Charan (2017), the situation in Pakistan is quite disappointing, where there is a scarcity of rubrics from HSSC to the master level examination. The present study is significant in exploring the overall scenario of paper checkers in Punjab.

In contrast to HSSC, Cambridge O level exams of English conducted by Cambridge Assessment International Education show an opposite picture in Pakistan. A detailed marking scheme with a given specimen is provided to the paper checker to award suitable marks to the given material. These specimens are accessible online <https://www.cambridgeinternational.org/programmes-and-qualifications/cambridge-o-level-english-language-1123/past-papers/>. The portion of creative writing is marked with the fulfilment of assessment objectives and language band descriptors. These bands include; the understanding of purpose, awareness of the specified situation and audience, appropriateness of text, development of details, well-organized information to justify a personal opinion and appropriateness of tone. On the biases of these descriptors, the writing task of a student is evaluated. This kind of marking scheme is quite different from the marking system of HSSC.

MoE (2006) has emphasized that examiners should explore the learner's strong and weak points while assessing their progress. Secondly, they should focus on the purpose and objective of the test by following specific criteria. To sum up the discussion, it seems that the current marking system of HSSC only checks the students' knowledge about language rather than examining the skill in a specific area. It has promoted; the trend of memorization and reproduction of material among the students to gain higher scores awarded by examiners (Rehmani, 2003). In this way, the present study explores the beliefs of paper checkers while scoring the English essay at the HSSC level. Research Question are;

- What are the current practices of examiners/paper checkers in marking the paper of English at the HSSC level?
- Which language skills are assessed while marking an English essay of HSSC?

The expectations of paper raters always affect the teaching-learning environment, especially at the HSSC level, which defines the students' career and academic future. Therefore, the current research delves the hopes and expectations of paper checkers with the paper of English. The results will also spot the deficiencies and flaws in the English paper marking procedure. Furthermore, it will provide the solutions to ameliorate the current situation.

## **2. Methodology**

### **2.1 Research Tool**

The present research required a survey followed by a questionnaire with 15 most relevant questions about the paper marking of an English paper at the HSSC level. The questionnaire has been piloted to confirm the validity and reliability of every question. Then, the researcher visited every workplace (college) of the respondents carrying a questionnaire and a rubric as a sample to help where it was necessary. The aim was to facilitate respondents to collect the real data.

### **2.2 Data Collection**

The sample of 350 paper checkers including 124 females and 226 males has been selected randomly from different government (94.3%) and private (5.7%) colleges of Punjab province. From the total, 281 paper raters were having B.Ed., M.Ed., and TEFL (PGD) as their

professional degrees while 69 of them have showed the absence of any professional qualification. The sample represented a varied range of academic qualification e.g., 244 (69.7%) had M.A, 102(29.1%) had MPhil and 4(1.1%) were having PhD as their last academic degree.

### 2.3 Experience

The selected paper markers showed a varied range of experience. From the 350 paper markers, 116 were examiners and 234 of them have worked as a sub-examiner. Furthermore, 118 of them were having 0-5 years of experience; 90 (25.7%) were having 6-10 years of experience and 142(40.6%) were showed 11-15 years of experience in marking English papers. This information confirms the validity of sample and reliability of results, which have been deduced from this research.

### 2.4 Data Analysis

The data collected from this sample has been submitted to SPSS and results have been generated. The conclusions have been drawn by discussing each item of questionnaire in detail.

## 3. Results and Discussions

Academic English language exams systematically spot the extent of command on English in real-life language, tasks and topics and the examiner no doubt needs comprehensive training. The respondents recorded an equal response on attending any training as a sub/head examiner, including those 69 respondents who remained involved in the paper development process for more than ten years. It means; that there is quite a render need to reestablish the paper marking system keeping in mind the standard criteria.

**Table 1: Have you attended any training for the sub-examiner or head examiners?**

	Frequency	Percent	Valid Percent	Cumulative Percent
yes	168	48.0	48.0	48.0
No	182	52.0	52.0	100.0
Total	350	100.0	100.0	

**Table 2: If yes then what was the duration of your training?**

	Frequency	Percent	Valid Percent	Cumulative Percent
.0	184	52.6	52.6	52.6
Days	107	30.6	30.6	83.1
weeks	50	14.3	14.3	97.4
Months	9	2.6	2.6	100.0
Total	350	100.0	100.0	

The BISE alleges its duty to sustain reliability across the raters (BISE Training Manual for Head/ Sub-Examiners, 2016, p. 5). Of the 182 respondents, 107 (58.79%) recorded that training was of days but less than a week. On the other hand, 57 respondents responded that it was one-week training. The changing opinion raises questions about the policies and coordination among BISEs in Punjab. The essay highlighted the scarcity of training sessions for paper markers finally to push them to rely on their subjective opinion. It needs constant training sessions; so to achieve a high level of inter-rater reliability (Weir, 2005).

**Table 3: Do you follow any specific rubrics?**

	Frequency	Percent	Valid Percent	Cumulative Percent
Yes	227	64.9	64.9	64.9
No	123	35.1	35.1	100.0
Total	350	100.0	100.0	

The results of the question about the use of rubrics were surprising. According to 123 (35.1%) respondents, they do not follow any rubric. Rubrics are those guidelines, which are developed following the objectives of the curriculum. A rubric allows an examiner to check a set of specific skills otherwise, the examiner could add his judgment. From the 350 respondents, 227 (64.9%) response was positive, which again is a question mark on the efficiency of BISEs in Punjab. How is it made possible to provide a rubric to a specific number

of paper checkers? In contrast, a detailed marking sheet is provided to the paper checkers of O level with an intelligent distribution of marks, which checks the multiple elements of language. Fernandez and Siddiqui (2017) reported that raters are not provided with evaluation criteria in the form of a rubric in scoring an English paper for high stake exam conducted by BISE in Punjab. It also means bypassing the objectives behind designing the paper. Finally, learners' progress cannot be measured on those scales, which have been mentioned in the national curriculum.

**Table 4: Do you credit or discredit on the basis of handwriting?**

	Frequency	Percent	Valid Percent	Cumulative Percent
Yes	227	64.9	64.9	64.9
No	123	35.1	35.1	100.0
Total	350	100.0	100.0	

Responding to the question about the impact of handwriting; 227 (64.9%) opinions endorsed to credit handwriting. It is because the students focus more on the use of pen and markers than the content. The students with average or unclear handwriting get low scores even after writing quality content on paper, e.g., the different scorers have different expectations from a good piece of writing to their subjective opinion sometimes works as a central factor affecting the reliability of marking (Fernandez and Siddiqui, 2017). On contrary, O level marking criteria focuses on achievable objectives, i.e., awareness of the candidate about a topic, required points for development and variety of justifications with appropriate tone and register rather than considering handwriting as a basis for awarding scores.

**Table 5: Do you deduct scores for grammar and spelling mistakes?**

	Frequency	Percent	Valid Percent	Cumulative Percent
Yes	317	90.6	90.6	90.6
No	33	9.4	9.4	100.0
Total	350	100.0	100.0	

Grammar and spellings are part of writing tasks. These are taken into consideration while checking the writing skills of the students. In response to the question about the deduction of scores for grammatical and spelling mistakes, 317(90.6%) paper checkers have said that they deduct scores for these mistakes while 9.4% of the sample responded negatively. This difference of opinion leads to the non-serious behaviour of BISEs in Punjab. This is an important factor to decide set criteria because the raters come from different linguistic backgrounds (Barkaoui, 2010) and have a different attitude towards errors (Huang, 2009). In contrast, the marking criteria of O level English exam consider accurate spelling, varied sentence structures, correct and appropriate tenses, wide and precise vocabulary. There should be conduction of cumulative meeting sessions for raters, where criteria should be set to the mark papers.

**Table 6: How do you deduct scores for mistakes?**

	Frequency	Percent	Valid Percent	Cumulative Percent
By counting	150	42.9	42.9	42.9
Self judgement	200	57.1	57.1	100.0
Total	350	100.0	100.0	

There is again a conflict of opinion on reflecting the question about the criteria for the deduction of scores. From 350 paper raters 150 (42.9%) respondents recorded that they deduct marks by counting the mistakes while 200 (57.1%) responses showed that they do not count the mistakes but it is their self-judgement about the number of mistakes. On the other hand, language testers emphasize the importance of the reliability of scorers (Hughes 1989; Lumley 2002). According to Cho (1999, p.3) "rating discrepancy between raters may cause a very serious impediment to assure the test validation, thereby incurring the mistrust of the language assessment process itself." In this way, there must be a valid and reliable scoring sheet with proper allocation of marks to judge every kind of essay with a variety of elements as in O level marking sheet where language 0-8 bands are given with their comprehensive description.

**Table 7: How do you deduct scores for mistakes?**

	Frequency	Percent	Valid Percent	Cumulative Percent
1 score for 1 mistake	101	28.9	28.9	28.9
1 score for 3 mistakes	205	58.6	58.6	87.4
any other	44	12.6	12.6	100.0
Total	350	100.0	100.0	

This question is an extension of the previous question about the deduction of scores for mistakes. From 350 paper checkers, 101 (28.9%) have agreed to deduct one score for one mistake, and 205 (58.6%) paper markers responded to deducting one score for three mistakes. It shows that students are not equally evaluated. There is a variety of criteria to judge the same paper, and it is luck, not the skill of students that how much scores he/she will get.

**Table 8: Do you distinguish between pen mistakes and serious mistakes? For every option please mention a reason?**

	Frequency	Percent	Valid Percent	Cumulative Percent
Yes	278	79.4	79.4	79.4
No	72	20.6	20.6	100.0
Total	350	100.0	100.0	

Pen mistakes happen in the flow of writing unintentionally. On the other hand, serious mistakes come under the category of error. In response to the question about distinguishing between pen and serious mistakes, 278(79.4%) paper checkers endorsed the statement, while 72 (20.6%) paper raters opposed it. It again highlights the inefficiency of BISE, which did not develop any solid criteria yet to improve the situation. In contrast, the marking scheme of O level counters both of these kinds the slips and errors, which are used to happen during writing, i.e., highly accurate writing; apart from occasional slips and highly appropriate to choose the task type is given band 8. While accurate writing with occasional errors either slips or caused by ambition is given band 7.

**Table 9: Do you receive any instructions regarding scoring of essays, prior to the scoring or during the scoring? If yes than please specify the types of instructions?**

	Frequency	Percent	Valid Percent	Cumulative Percent
Yes	210	60.0	60.0	60.0
No	140	40.0	40.0	100.0
Total	350	100.0	100.0	

There is again a conflict of opinion in response of question about receiving any instructions regarding the scoring of essays. From the 350-paper checkers, 210 (60%) have recorded the response, that they get an instruction sheet for scoring the essay questions. On the other hand, 140 (40%) raters have responded not to get any scoring sheet. The marking criteria of O level English paper include several language qualities including varied sentence structures for particular effects, use of correct and appropriate tenses, use of wide and precise vocabulary, correct use of punctuation, accurate spelling, appropriateness of evidence and unity of paragraphs. To maintain an acceptable level of reliability, Weigle (2002) has recommended the need for a detailed scoring sheet to be followed to mark the composition type of question. Moreover, he recommends the use of a rubric or scoring guide, which serves as a yardstick against the raters' subjective opinion.

**Table 10: Do you make quick judgments by the overall impression the essay has on you?**

	Frequency	Percent	Valid Percent	Cumulative Percent
Yes	217	62.0	62.0	62.0
No	133	38.0	38.0	100.0
Total	350	100.0	100.0	

According to the response of 217(62%) paper checkers, they quickly make the judgement about an essay, while the remaining 38% have claimed to read the entire essay to

make a solid judgment. The O level marking scheme promotes the detailed analysis of writing using multiple variables to construct or develop a final judgement. Unfortunately, the pertaining situation leads to the subjective impression of the raters. Haq and Ghani (2009) have reported that in Pakistan, there are no explicit criteria to score the essay at HSSC, although it is considered as a key component in the high-stake examination.

**Table 11: How did you train yourself as a checker?**

	Frequency	Percent	Valid Percent	Cumulative Percent
Properly trained from Board or University	138	39.4	39.4	39.4
Read a book on testing	61	17.4	17.4	56.9
Self-trained	151	43.1	43.1	100.0
Total	350	100.0	100.0	

The question about training as a checker, 138 (39.4%) paper checkers has recorded to attend proper training from a board or university. On the other hand, 151 (43.1%) have responded that they are self-trained as they have been associated with paper marking for a long time. These are the reasons that students avoid creativity and preferred to cram and reproduce materials in the paper. The result is that the education system in Pakistan only checks the student's knowledge about language, not the performance. This situation can be improved with frequent purposeful training sessions for paper raters.

**Table 12: Which of the following element you deal seriously? Please write a reason**

	Frequency	Percent	Valid Percent	Cumulative Percent
Error	145	41.4	41.4	41.4
Mistake	186	53.1	53.1	94.6
Pen mistake	19	5.4	5.4	100.0
Total	350	100.0	100.0	

Errors are those mistakes about which learners are very unaware. In contrast, students have the realization of mistakes, but they still happen due to some other factors, i.e., shortage of time in writing paper, stressful conditions and absent-mindedness...etc. Here the response shows more percentage (53.1%) in handling the mistakes seriously, while 41.4% response is in favour of dealing with errors seriously. There must be single criteria, which should promote the student's creative ability following the objectives of the national curriculum.

**Table13: How do you guess the length of an essay?**

	Frequency	Percent	Valid Percent	Cumulative Percent
By counting words	62	17.7	17.7	17.7
From paragraphing	143	40.9	40.9	58.6
personal judgement	145	41.4	41.4	100.0
Total	350	100.0	100.0	

In response to the question about the judgement of the length of an essay, two options were generally chosen 143(40.9%) paper checkers have selected the option 'from paragraphing', and 145 (41.4%) paper checkers were in view to apply their judgement. Fernandez and Siddiqui (2017) have reported that the rater's experience and the elements of the essay enable them to guess about the length of an essay. The situation again appeals to have a proper valid and reliable draft to score the English paper in promoting skills-based learning. Generally, Students have been found in cramming the three different essays on the same topic, which includes simple (for low learners), average (for average learners) and complex (for intelligent learners).

A majority of paper raters (50.9%) have perceived themselves as average, while 25.7% and 23.4% of paper checkers perceive themselves as lenient and strict examiners, respectively. Marking of an essay is still taken as a human job, which brings subjective and idiosyncratic evaluation and sometimes leads to discrepancy Fernandez and Siddiqui (2017), e.g., it may raise serious questions on fairness and reliability of marking criteria.

**Table14: How do you perceive yourself as an examiner?**

	Frequency	Percent	Valid Percent	Cumulative Percent
Lenient	90	25.7	25.7	25.7
Strict	82	23.4	23.4	49.1
Average	178	50.9	50.9	100.0
Total	350	100.0	100.0	

**Table 15: How much percentage do you award to fully satisfied essay?**

	Frequency	Percent	Valid Percent	Cumulative Percent
60%	16	4.6	4.6	4.6
70%	84	24.0	24.0	28.6
80%	108	30.9	30.9	59.4
90%	109	31.1	31.1	90.6
100%	32	9.1	9.1	99.7
any other	1	.3	.3	100.0
Total	350	100.0	100.0	

The response to the question about marking the fully satisfied essay shows surprising results, i.e., 109 (31.1%) raters have recorded that they give 90% marks to a fully satisfying essay. In contrast, 108 (30.9%) raters responded to award 80% marks, and 84(24%) raters said to award 70% marks. Fernandez and Siddiqui (2017) conducted the interviews, analyzed the commentaries of raters, and concluded the same results that there is a great deal of difference in marking criteria of English papers at the HSSC level. On the other hand, the bands have been devised in the O level marking system with their proper description in evidence to the specific qualities of language, which measures the writing ability of the student rather than the ability to cram or reproduce the text.

#### 4. Conclusion

The present study was aimed to unfold the practices of paper raters in scoring English paper of HSSC level in Punjab Pakistan; in comparison with the O level marking system. In addition, untrained raters of HSSC, are marking English papers for more than ten years. They are not given any instructions or rubrics to check the creative writing skill of learners. Moreover, there is also a scarcity of inter-rater reliability where HSSC raters mark papers on their likes and dislikes. The results also conclude that criteria to deduct the scores on mistakes change from rater to rater. There has also found a significant involvement of subjectivity where the same essay has been scored differently by different raters. These findings lead that English assessment at the HSSC level does not fulfil its purpose to achieve the objectives set in the national curriculum. In this way, there is a dire need to follow the latest assessment methods to emphasize the proficiency of the learner rather than the language knowledge. There should also be regular training sessions for paper raters. Further, BISE should develop comprehensive rubrics to fulfil the objectives of the national curriculum. The current study addresses the policymakers, board officials and examiners to spot what exactly they should look for while marking the English paper at the intermediate level. References

#### References

- Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational leadership*, 57(5), 13-19.
- Barkaoui, K. (2010). Do ESL essay raters' evaluation criteria change with experience? A mixed-methods, cross-sectional study. *TESOL Quarterly*, 44(1), 31-57.
- Cho, D. (1999). A study on ESL writing assessment: Intra-rater reliability of ESL compositions. *Melbourne Papers in Language Testing*, 8(1), 1-24.
- Fernandez, M., & Siddiqui, A. M. (2017). Markers' criteria in assessing English essays: an exploratory study of the higher secondary school certificate (HSCC) in the Punjab province of Pakistan. *Language Testing in Asia*, 7(1), 6.
- Glickman-Bond, J., & Rose, K. (2006). *Creating and using rubrics in today's classrooms: A practical guide*. Christopher-Gordon Publishers.
- Government of Pakistan (2006). *National Curriculum for English i-xii*. Islamabad. [www.moe.gov.pk/curriculum](http://www.moe.gov.pk/curriculum).

- Gulzar, M. A., Buriro, G. A., & Charan, A. A. (2017). Investigating the Effects of Rubrics on Assessment of Writing Tasks. *International Research Journal of Arts & Humanities (IRJAH)*, 45(45), 191-206.
- 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.
- Haq, N., & Ghani, M. (2009). Bias in grading: A truth that everybody knows but nobody talks about. In *English Language and Literary Forum. Shah Abdul Latif University of Sindh, Pakistan* (Vol. 11, pp. 51-89).
- Huang, J. (2009). Factors Affecting the Assessment of ESL Students' Writing. *International Journal of Applied Educational Studies*, 5(1), 1-17.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press
- Khan, S. (2006). An evaluation of the exercises provided in the English compulsory textbook for class X. Unpublished MA dissertation] Faculty of English Linguistics, University of Karachi.
- Khattak, S. G. (2012). Assessment in schools in Pakistan. *SA-eDUC*, 9(2), 1-13.
- Lumley, T. (2002). Assessment criteria in a large-scale writing test: What do they really mean to the raters?. *Language Testing*, 19(3), 246-276.
- Mirza, M., Nosheen, M., & Mahmood, N. (1999). Impact of Examination System on Teaching Styles of Teachers at Secondary and Higher Secondary Classes. Institute of Education and Research Quaid-EAzam Campus, University of the Punjab, Lahore.
- MoEGP (Ministry of Education Government of Pakistan) (2009) *National Education Policy 2009*. Islamabad: Government of Pakistan.
- Quinlan, A. M. (2006) *A complete Guide to Rubrics', Assessments Made Easy for Teachers*, K-College. Lanham, MD: Rowman and Littlefield Education.
- Rehmani, A. (2003). Impact of public examination system on teaching and learning in Pakistan. *International Biannual Newsletter ANTRIEP*, 8(2), 3-7.
- Shirazi, M. J. H. (2004). Analysis of examination system at university level in Pakistan (Doctoral dissertation, PMAS-Arid Agriculture University, Rawalpindi).
- Siddique, N. (2014). Evaluation of the Assessment Criteria of English Language at Higher Secondary Level in Pakistan. *Evaluation*, 5(4), 12-24.
- Warsi, J. (2004). Conditions under which English is taught in Pakistan: An applied linguistic perspective. *Sarid Journal*, 1(1), 1-9.
- Weigle, S. C. (2002). *Assessing writing*. Ernst Klett Sprachen.
- Weir, C. J. (2005). *Language testing and validation*. Hampshire: Palgrave MacMillan.