DOI: 10.15613/hijrh/2024/v11i2/48141
ISSN (Print): 2349-4778
***HuSS*: International Journal of Research in Humanities and Social Sciences,** *Vol 11(2), 51-57, 2024* *ISSN (Online): 2349-8900*

Assessors' Bias in the Assessment of Writing Skills

R. Kalpana[1*] and Suma Kodandaram[2]

[1]Professor (CAS), Department of English, PSG College of Technology, Coimbatore - 641004, Tamil Nadu, India; rkp.english@psgtech.ac.in

[2]Freelance Language Assessment Consultant and Teacher, Coimbatore - 641004, Tamil Nadu, India

Abstract

This paper is based on a study conducted to understand if there is an assessor's bias while scoring writing skills at the college level. It starts with the hypothesis that the assessor's bias plays a role in assessing writing skills. A group of experienced teachers of English at a reputed higher educational institution in Coimbatore, India, were asked to respond to a questionnaire. The results were converted into a pie chart and analysed by comparing percentages using four different criteria – ideas and topic development, cohesion and paragraphing, grammar and punctuation, spelling and vocabulary, and whole vs the parts. The analysis and discussion of results seem to prove our hypothesis right. There seems to be a positive bias favouring ideas and topic development and grammar and punctuation over cohesion and paragraphing, spelling and vocabulary.

Keywords: Assessor, Bias, Grammar, Ideas, Paragraphing, Vocabulary

1. Introduction

Assessment of writing skills is one of the most common ways of assessing English language skills. From diagnostic tests to evaluation and end-semester exams, tests focus much on learners' writing skills, certainly more than the other language skills. There are also several competitive exams and international language tests that learners are required to take when in pursuit of higher academic and career-oriented aspirations and goals. Generally, writing is "regarded as a 'high stakes' activity performing a gate-keeping role in university education as students need to demonstrate the required standard of academic writing if they are to succeed in their studies[1-4].

The test taker is often at the centre and is advised about preparing for and performing well in writing tests. However, it is imperative to note that there is another side to it – that of the assessor. Not that the role of the assessor is ignored usually, undoubtedly not. The role of the assessor and the criteria for assessment are the most crucial part of any test of writing skills, just as with the assessment of other language skills. The most trusted international language tests lay out an elaborate list of standards, also stipulating different levels of proficiency and specifying benchmarks for each. So, if a score is awarded to a certain written script, the assessor knows fairly well as to which yardstick is applied and so does the test taker. The internet is replete with the standardized assessment criteria for each of the standardized international tests, granting easy access to the assessor, the test taker or anyone else.

Everything seems to be perfect, in theory, – the criteria is laid out, the test taker and the assessor know what exactly is expected at each level and the test gives a standardised result. Now, this is what is trusted and accepted from any assessment of writing. However, the question is whether there could still be a scope for bias when this is applied in practice. Hence, this preliminary study on the role of the assessor's bias in the assessment of writing skills was undertaken. A more detailed study based on the observations made here is intended in future.

2. Hypothesis and Methodology

Generally, it has been observed that "the potential for bias is a concern when assessing student work and that when it does occur scorers are often not aware that the bias is operating"[5]. Going by this, in this study, we were trying to test the hypothesis that the assessor's bias plays a role in the assessment of writing skills.

**Author for correspondence*

A questionnaire consisting of 20 questions was used with teachers at the college level. It was a sample group of sixty-eight, currently employed as English teachers from reputed higher educational institutions in Tamil Nadu and Telangana. The volunteers were chosen from those teachers who were willing to respond to the set of questions and hence, the convenience sampling method was adopted. Some of them had quite a few years of experience in teaching and assessment, ranging between eleven to thirty-five years. This teacher cohort assesses several student essays, and their assessment is likely to impact the understanding of the learners' writing proficiency. Further, their assessment would also give an impression of their writing competencies to the stakeholders – the companies in which the students would be placed after their graduation or higher education institutions where the students are likely to get admitted. Hence, we felt it was necessary to study if the assessment was done objectively or if any bias was involved.

Broadly, the focus of the questionnaire was on four main areas – ideas and organisation, grammar and punctuation, spelling and vocabulary, and cohesion and paragraphing. This categorization was not overtly made known to the participants and the questions referring to different categories were randomly distributed across the questionnaire with no grouping of questions. Teachers' bias was being judged based on whether they might overlook one criterion for the other, and if one might impress them more towards not considering the other. For instance, could there be a chance for a highly impressive vocabulary to colour the assessor's perception of the organisation of ideas?

The responses from the questionnaire were collated together in a graphical form. The resultant pie charts were closed observed and analysed to arrive at a cohesive and comprehensive conclusion regarding the hypothesis we started with.

3. Analysis and Discussion

This section presents the analysis and of the results categorised separately into four criteria for ease of discussion, while the questionnaire or the interview itself did not follow this categorization. The four categories are as follows: ideas and organisation, cohesion and paragraphing, grammar and punctuation and spelling and vocabulary.

Most of the participants seemed to agree with the view that, as with any other skill, writing as a skill is an integration of all these categories and is a whole in itself, and that writing assessment, ideally, should also look at the integration of all of these. The study tries to understand, if, despite this understanding that each assessor has, there could be a scope for bias. Let us now consider our observations for each category in the following sub-sections.

3.1 Ideas and Topic Development

The attempt was to figure out how much it matters for the assessors to consider ideas presented in the written test and the logical organisation of those ideas. On being asked if they would give marks when ideas are good, relevant and well-extended with examples, even if the essays are not so well-organised, about fifty-eight per cent of them agreed that they would while less than fifteen per cent disagreed, as depicted in Figure 1. Twenty-seven point five per cent chose not to respond to this question.

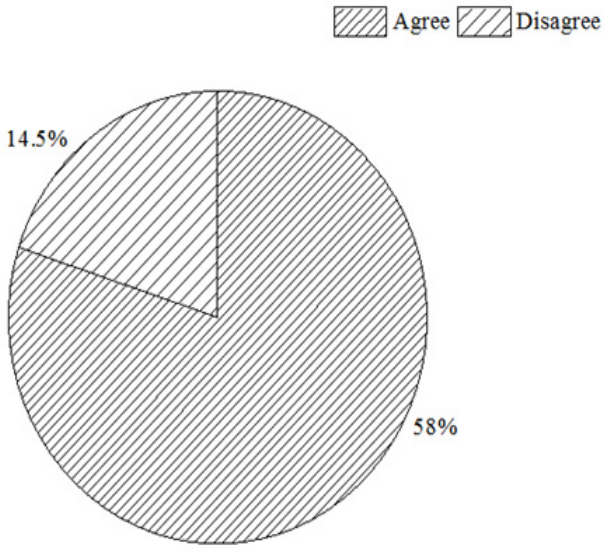


Figure 1. Relevance and organisation of ideas.

Some of the justifications given by the teachers include

"If ideas are not organised properly, how do I know they are relevant?"

"Sometimes students write in bits and pieces. If some ideas are relevant, I do not bother about organisation!

"Organisation should be easy to follow. Otherwise, I can't give marks."

"Organisation is needed to bring the desired effect. It gives meaning."

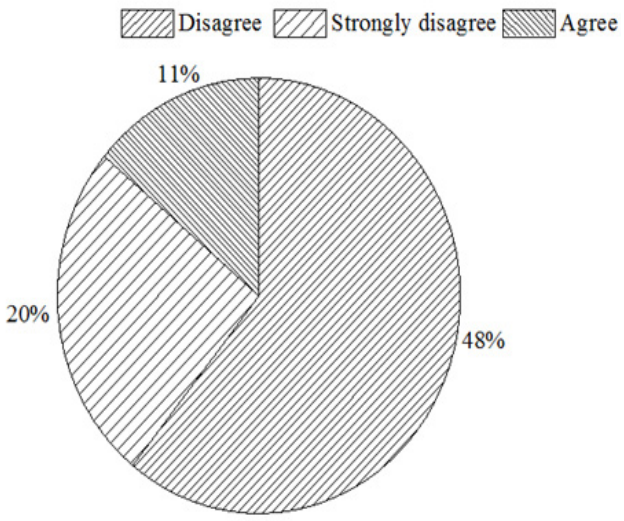


Figure 2. Grammar, vocabulary and development of ideas.

"Clear expression of main ideas is absolutely essential."

"Main ideas have to be presented properly."

The respondents were also asked if they would not reduce marks when grammar and vocabulary are satisfactory, but ideas in the essays are irrelevant or underdeveloped. It is interesting to note that forty eight percent disagreed and twenty percent of them strongly disagreed while only eleven per cent agreed (Figure 2).

Twenty-one per cent of respondents chose not to give their responses. In justification of their choices, some of the respondents came up with the following responses:

"Ideas, content, notation of words all are important."

"I think generating the right ideas is important."

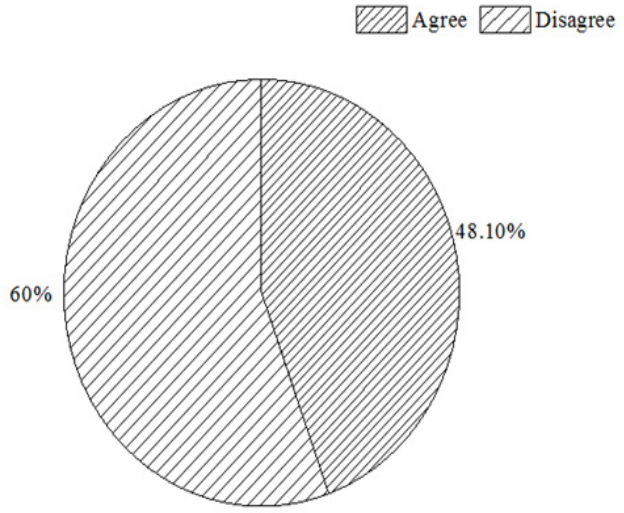


Figure 3. Cohesion - development of ideas.

"Relevance to the key topic is a major concern. It reflects a student's understanding of the topic but the inaccuracy of expression will call for scaling down the grades as accuracy is important for a well-defined writing piece."

"Grammar is indispensable."

"I assess language...not ideas."

"If the content is relevant, we need to give marks."

"I will penalise them in the grammar section for such mistakes. However, to do that the content matter should be clear enough despite issues with sentence structure/ grammar."

Thus, it can be observed that assessors consider ideas and the development of main points to be important criteria while assessing writing skills. There may be a tendency among assessors to be positively biased towards ideas more than other criteria such as grammar or vocabulary. A study conducted by Freedman so found that content and organisation had a greater impact on scores than structure and mechanics of writing[6].

3.2 Cohesion and Paragraphing

Hasan[7] mentions that "cohesion refers to the relations of meaning that exist within the text and is expressed through the structural organization of the text. It occurs where the interpretation of some elements in the text is dependent on that of another". Cohesion and paragraphing could be looked at as a corollary to the organisation of the ideas presented in the written script. One might want to conclude that for ideas and topic development to come through, cohesion and paragraphing are paramount. Therefore, continuing from the previous sub-section, the researchers liked to see if the observation would be further strengthened or weakened by the results found for this sub-section.

As depicted in Figure 3 below, forty-eight point one per cent of the respondents agreed that they would not penalise their students if the cohesion or flow is not so good, provided that they see that ideas are developed and are relevant in an essay/paragraph. We received some duplicate responses for this question.

More than sixty per cent seemed to not penalise their students if the ideas are presented in an organised manner and if the topic sentences are not clear in each paragraph. As Figure 4 displays, fifty two percent of the respondents agreed and twelve percent of them strongly agreed that they would not mark their students less if the ideas are

well-organised and if they can understand the flow, even if the students haven't used cohesive devices/higher level discourse markers.

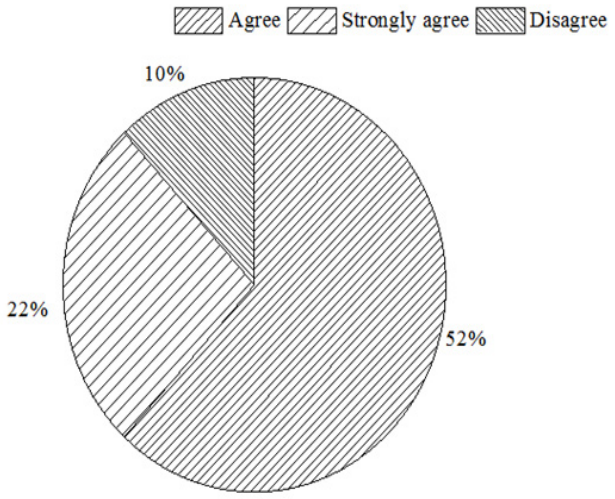


Figure 4. Organisation of ideas.

Sixteen per cent of respondents chose not to respond to this question. Those respondents who considered paragraphing to be important stated the following "logical development is possible only if paragraphing is clear."

"Paragraphing shows the flow of content; there are no paragraphs or marks at the college level. The paragraph reflects content and vocabulary. Lack of para leads to vagueness, and distraction and can disturb coherence."

"Clear para discusses the concept of the essay properly"

It seems that a majority of the assessors take cohesion and paragraphing to be less important than ideas and the development of main points. It seems to strengthen the positive bias that assessors have for ideas that we observed in the previous sub-section.

3.3 Grammar and Punctuation

Grammar and punctuation are important as they help to convey the thoughts accurately. According to Celce-Murcia[8], "plenty of grammar errors in ESL students' academic writing make their performance weaker or sometimes they fail the examinations". We wanted to see where the assessor's bias stands concerning grammar and the other three categories counted in this study.

When asked if the assessors would award marks if the ideas presented by students are relevant to the topic of the essay/paragraph even if their sentences are grammatically incorrect or repetitive, thirty-three per cent of them agreed that they would, and more than forty per cent of them disagreed (Figure 5). Nineteen per cent of participants did not give their responses.

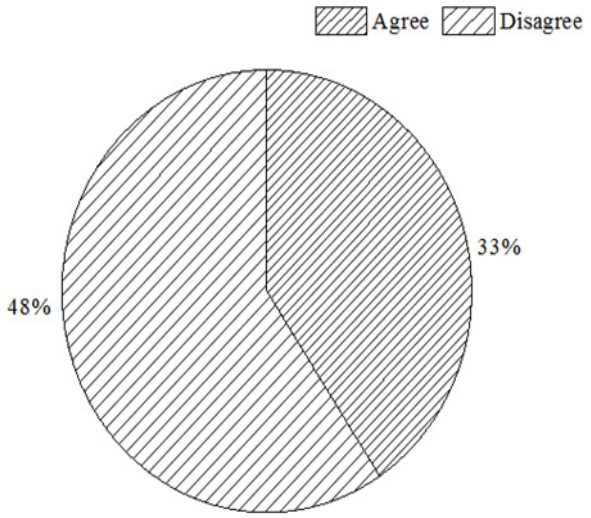


Figure 5. Grammar - relevance to topic.

Fifty-two per cent of the respondents agreed that they tend to penalise (students) more for grammar than for vocabulary (Figure 6). Eight per cent chose not to respond. The justifications of participants are given below:

"To me, grammar communicates meaning. Vocabulary also communicates meaning but not so much as grammar does."

"We can't expect a huge range."

"As a language teacher, I will see both are balanced."

"Students are second language writers, so I compromise."

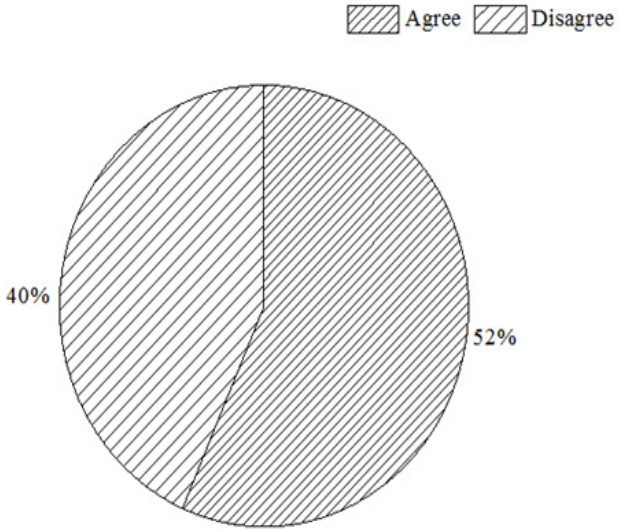


Figure 6. Grammar and punctuation.

"Bad punctuation gives rise to ambiguity."

"If meaning is unaffected, rare errors in punctuation are tolerable."

"Punctuation helps organise thoughts better and helps a reader to interact with the writer's thoughts without misinterpretation. Minor omissions can be overlooked if the essay is originally and brilliantly presented."

The results show that a majority of the respondents have a positive bias towards greater grammatical accuracy while assessing writing skills. The focus seems to be more on the correct use of grammar and punctuation rather than on vocabulary or ideas and paragraphing.

3.4 Spelling and Vocabulary

The general understanding is that assessors look for correct spelling and range of vocabulary used while assessing writing skills. Idiomatic and figurative use of language also counts to reflect the level of proficiency the test taker has attempted to exhibit. As pointed out by Nation[9], "the knowledge and appropriate usage of academic vocabulary in written production is crucial for university students, particularly concerning high-stakes writing and assessment, where academic vocabulary is considered to be a key component of academic success." Here again, we have tried to consider the assessor's bias concerning spelling and vocabulary in the other three categories discussed above.

Do assessors tend to get carried away by high-sounding vocabulary even when they are not relevant to the topic? Because of this, do they end up awarding higher marks? Forty-eight per cent of our respondents strongly disagree with this.

Students sometimes use bombastic/high-sounding, uncommon words. In this situation, do assessors tend to be carried away and mark them up even if those words are not relevant to the topic/theme of the essay? Fifty-two percent of our respondents disagree and forty percent of them strongly disagree (Figure 7). That means a significant ninety per cent of them say that they do not get carried away by vocabulary, but instead focus on the topic and ideas. Eight per cent of participants chose not to respond.

Some responses in their defence are as follows:

"Relevance more important than high-sounding words."

"Focus more important, verbose language is injurious."

"While I am generally impressed with the lexical repertoire of students, I normally don't get carried away to the extent of ignoring their irrelevance."

"I prefer content and appropriate use of words."

"Highfalutin words do not matter in writing. The idea is not to impress."

The study also seeks to find out respondents' perceptions of spelling and vocabulary about grammar. One of the questions seeks to find out if they don't find a range of grammatical structures to be as important as a range of vocabulary. Forty-four point four per cent of them disagree and eleven per cent of them strongly disagree with this point as given in Figure 8. That is, more than half of the respondents feel that grammar is as important as vocabulary.

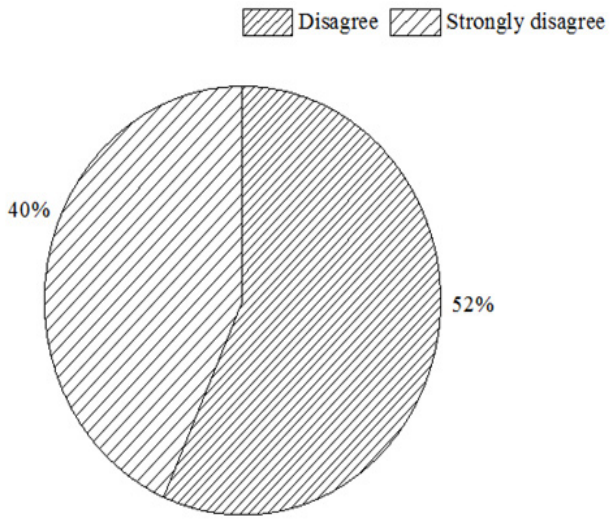


Figure 7. Relevance of vocabulary to topic.

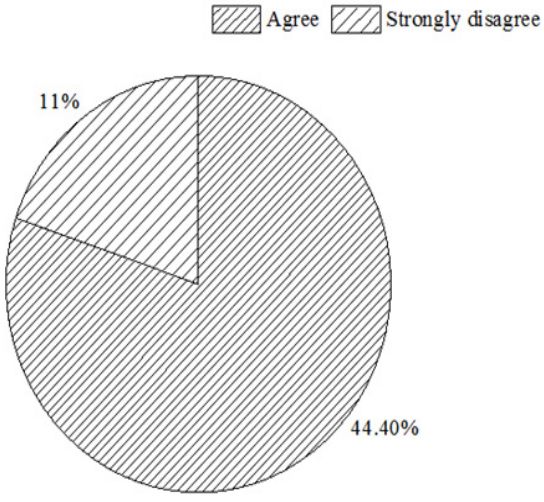


Figure 8. Relevance of vocabulary, grammar and topic development.

Forty-four point six per cent of respondents had no opinions on this aspect. Therefore, it can be understood that assessors do not seem to be coloured by the range of vocabulary used by the candidate and tend to focus equally, if not more, on grammar and topic development.

3.5 The Whole, Not Parts!

It was also important to see if there might be assessors that consider only the overall impression they get from the script they are marking rather than looking for different criteria discussed above.

When asked if they think that instead of looking for individual features such as grammar, vocabulary etc, looking at the essay or paragraph holistically helps (them) mark better, about forty-eight per cent of the respondents affirmed this, as presented in Figure 9. Two per cent of respondents chose not to respond to this question.

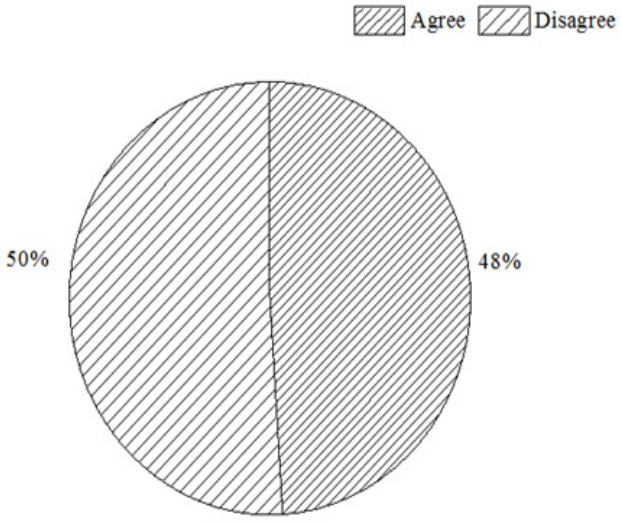


Figure 9. Whole - parts.

Some of the responses that justify their choice are as follows:

"I look at the whole appeal rather than its components."

"The overall thing matters…not components."

Fifty-six per cent of the respondents disagreed with the following point in the questionnaire – "I find the whole job of looking at individual features tiring, time-consuming and not worth it! So, I tend to give them marks based on the overall impression I get after reading an essay/paragraph"- while only twelve per cent said they prefer this. Thirty-two per cent of respondents chose not to respond to this question, as displayed in Figure 10.

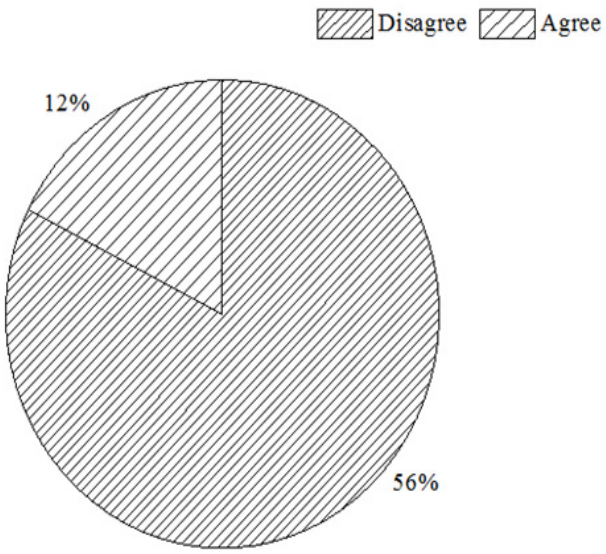


Figure 10. Holistic - analytic.

So, it can be seen that more than fifty per cent of the assessors confirm that they would look for different criteria while scoring the written scripts, thereby retaining the objectivity of the process, rather than going by the overall impression which could make the process of assessing writing tests more subjective.

4. Conclusion

The study started with the hypothesis that 'the assessor's bias plays a role in the assessment of writing skills. The survey was conducted using a questionnaire and gathered the first-hand opinions of experienced assessors of writing skills. The analysis and discussion of the result of this study appear to prove the hypothesis right. There is a positive bias, especially concerning ideas and topic development, and also grammar and punctuation. At the same time, it can also be seen that assessors do not seem to pay as much attention to cohesion and paragraphing; and spelling and vocabulary, thus validating our hypothesis. It can also be understood that more than fifty per cent of the assessors prefer scoring the written tests based on standard criteria for assessment rather than basing their judgement on the overall impression they get by reading the written script.

5. Scope for Further Study

This is a preliminary study as mentioned earlier. There is a need to conduct more advanced studies with

special attention to each criterion to assess the extent of subjectivity and bias in the assessment of writing skills. This can be done by getting a group of experienced assessors to award marks to a set of standardized written scripts to measure the actual variance in each criterion. There can be a follow-up study on how to educate assessors about such bias and measures that could be taken to avoid assessors' bias in the assessment of writing tests. Further, the sample was also small to validate the results for the entire group of teachers. The sample size can also be increased. Moreover, this study was largely quantitative and triangulating the results using qualitative methods of research would yield more valid results.

6. References

1. Flowerdew J. English for Specific Academic Purposes (ESAP) Writing: Making the case. Writing and Pedagogy. 2016; 8(1):5–32. https://doi.org/10.1558/wap.v8i1.30077
2. Lillis T, Scott M. Defining academic literacies research: Issues of epistemology, ideology and strategy. Journal of Applied Linguistics and Professional Practice. 2007; 4(1):5–32. https://doi.org/10.1558/japl.v4i1.5
3. Wingate U. Achieving transformation through collaboration: The role of academic literacies. Journal of Learning Development in Higher Education. 2019; 15:1–9. https://doi.org/10.47408/jldhe.v0i15.566
4. Wingate U, Tribble C. The best of both worlds? Towards an English for Academic Purposes/Academic Literacies writing pedagogy. Studies in Higher Education. 2012; 37(4):481–95. https://doi.org/10.1080/03075079.2010.525630
5. Steinke P, Fitch P. Minimizing Bias When Assessing Student Work. Research and Practice in Assessment. 2017; 12(Winter):88. https://files.eric.ed.gov/fulltext/EJ1168692.pdf
6. Freedman S. How characteristics of student essays influence teachers' evaluations. Journal of Educational Psychology. 1979; 71(3):328–38. https://doi.org/10.1037/0022-0663.71.3.328
7. Halliday MAK, Hasan R. Cohesion in English. London: Longman. 1976.
8. Celce-Murcia M. Importance of learning grammar for academic writing: Difficulties faced by TESL students in academic writing. Discourse analysis and grammar instruction. Annual Review of Applied Linguistics. 1991; 11:135-51. https://doi.org/10.1017/s0267190500002002.
9. Nation ISP. Learning Vocabulary in Another Language. Cambridge University Press. 2001. https://doi.org/10.1017/CBO9781139524759 PMCid:PMC90373