



The Most Suitable Scoring Method to Assess Essay Writing in ESL Classrooms

Arsaythamby Veloo¹, Noor Hashima Abd Aziz^{2*}, Aizan Yaacob¹

¹*School of Education and Modern Languages, Universiti Utara Malaysia, Malaysia*

²*School of Languages, Civilisation and Philosophy, Universiti Utara Malaysia, Malaysia*

Corresponding Author: Noor Hashima Abd Aziz, E-mail: noor934@uum.edu.my

ARTICLE INFO

Article history

Received: March 16, 2018

Accepted: 16, 2018

Published: August 31, 2018

Volume: 9 Issue: 4

Advance access: July 2018

Conflicts of interest: None

Funding: None

Key words:

Classroom-based Assessment,
Holistic Scoring Method,
Analytic Scoring Method,
Primary Trait Scoring Method,
Essay Writing,
ESL Teachers

ABSTRACT

The Malaysian Education Act in 1996 states that the national language which is the Malay language becomes the main medium of instruction in educational institutions and English is accorded the status of a second language. In Malaysian schools, teachers are given the freedom to use their own teaching and assessment methods in assessing their students' writing. However, majority of the English as a Second Language (ESL) teachers use a scoring method adapted from the Malaysian Examination to assess their students' writing. For the Malaysian Certificate of Education (SPM) English Language 1119 subject, students are required to write an essay of more than 350 words and they must be eloquent in their writing style with flawless grammatical errors to obtain the highest score for the subject. The purpose of this study was to explore the ESL teachers' preference towards the most suitable scoring method to assess essay writing in the classrooms. The study used qualitative approach which involved 25 grade 10 ESL teachers at 12 secondary schools in one of the states in Malaysia. The teachers in this study were trained on how to use the three types of scoring method to examine the students' continuous essay writing based on the three types of scoring methods namely Holistic, Analytical and Primary Traits. After the training, a few teachers were selected to be interviewed to get their perspectives of the three scoring methods used for classroom-based assessment. The research involved multiple data collection methods: verbal protocol, documents in the form of students' essays and interviews with the teachers. The findings indicated that the majority of the ESL teachers in this study preferred the holistic scoring method over the other two scoring methods due to its time saving characteristic when it involved a large scale marking and the results were needed to be completed within a short period of time. In addition, the ESL teachers in this study liked the idea of having to give overall evaluation of the essays whereby they could identify not only the overall band, but also the strengths and weaknesses of their students' writing. The teachers discovered that the holistic scoring method helped them to improve their students' learning as they could evaluate their students' overall performance.

INTRODUCTION

There are three phases of changes that have taken place in the Malaysian education system. These changes do not only involve the education curriculum but also the assessment process. Phase One took place between 1957 until 1970. According to Nalliah and Thiyagarajah, "The English language syllabus for Primary 1-6 is taught in three stages. Students in national schools complete Stages 1 and 2 while their peers in national-type schools undergo Stages 1, 2 and 3. The syllabi for secondary schools comprise of Syllabus for Secondary Schools (Malay medium), English and Syllabus for Secondary Schools (English medium): English for two different examinations namely Sijil Pelajaran Malaysia (SPM) or the Malaysian Certificate of Education and Overseas Examination Certificate" (as cited in Selvaraj, 2010, p. 55). In Phase One, the methods of teaching employed were the grammar

translation method, direct method and situational approach (Asmah Haji Omar, 2003). Meanwhile Phase Two of the education reform happened between 1970 and 1990. This is parallel to the implementation of The New Primary School Curriculum (1982) and the Integrated Secondary School Curriculum (KBSM) in 1988. According to Lee (2002), the reform was done "... because of dissatisfaction with the old curriculum which was thought to be too subject content-biased, too much emphasis on rote-learning, too exam oriented, and excessive dependence on textbooks" (p.12). Then Phase Three of the education reform took place from the year 1990 to 2007 where it was done to cater for globalization change such as the development of Information Technology in daily life. According to Carnoy (1999), "The need for skills like languages, mathematical reasoning, scientific logic and programming has emerged as the front runners of a

nation's development and English is the language of globalization, internet, trade and science" (p.57). Since then, more emphasis is given to strengthen the English Language and the use of ICT in the education system.

Assessment of English Examination Papers in Malaysia

Assessment of students has also changed because of the change in its English language curriculum. More challenging English examination papers were introduced to students sitting for the Malaysian Certificate of Education (SPM). SPM is a national examination taken by fifth-year secondary school students in Malaysia, which is equivalent to eleventh grade in American K–12 education and O-Level. It is set and examined by the Malaysian Examinations Syndicate (LPM, 2009). In 1995, the English examination paper known as English 122/322 was upgraded to English 1119. The English 1119 syllabus is based on the O-Level syllabus set by the Cambridge Examination Syndicate in England.

Directed writing, continuous writing and summary writing are among the changes made to the English language syllabus in terms of writing (Ministry of Education Malaysia, 2000). Besides writing, in 2001 the literature component in English subject was introduced with the intention to expose and inculcate reading habits as well as develop creative and critical thinking skills among the Malaysian students. For SPM English Language 1119 subject, students must answer 2 papers: Paper 1 tests on writing skills while Paper 2 evaluates reading skills. In Paper 1, students must answer questions on two sections: Section A question is on guided writing task where students need to understand the scenario given (for example, a teacher's retirement ceremony or a letter to convince a friend to take the national service program) and elaborate on all the points given in the question. On the other hand, Section B is on continuous writing task where students get to choose one title to write on out of the 5 titles given.

Malaysian secondary schools require students to sit for the national examination, SPM which is an exit examination for Fifth form. Some ESL teachers would be selected to assess the English SPM papers. The Malaysian Examination Syndicate is responsible in training these teachers to assess the English language papers using a specific scoring method. However, in schools no scoring method or strategy has been prescribed for the teachers, especially for classroom-based assessment of direct writing in Malaysian secondary schools. Currently, the ESL teachers are using a scoring method adapted from the Malaysian Examination Syndicate. As such, there is a need to do a research to find out the most suitable scoring method to assess essay writing in ESL classrooms, especially when there has been a move to decrease emphasis given to formal examinations. This research is concerned with Form Four and Form Five ESL students' direct writing performance in Malaysian secondary schools. At these two levels, the students are preparing to leave school to pursue their studies at higher institutions of learning or to get a job. Since students need to excel in their writing tasks, teachers need to know how to assess their students' performance efficiently and accurately. This paper

discusses which of the three types of scoring method (holistic, analytical or primary trait) is the most suitable for classroom-based assessment.

LITERATURE REVIEW

Salmani (2014) found that the English as a Second Language examiners tend to be influenced by the essay length and the need to be free from any simple mechanical errors. While there are many scoring methods available in the literature, Malaysian ESL teachers tend to use a scoring method at their convenience (Normah, 2014). According to Hamp-Lyons (1991), the holistic scoring method is problematic for second language examiners. Some examiners may focus on examining the students' ability to elaborate on the content points and organization, but may give limited attention to grammatical accuracy; others may pay extra attention in the number of grammatical mistakes made and vocabulary rather than observing the content of the essays. On the other hand, the weakness of the analytic scoring method is on its time-consuming factor and not suitable for large-scale testing. It requires more time and concentration to determine the right score for each aspect of writing before totaling it up and finding the mean score for each essay. Analytic scoring is time consuming (Weagle, 2002) because it generally requires more time in scoring as the examiners have to make multiple decisions for each essay as well as reading the essay multiple times in order to rate it based on several determined aspects of writing. Not only that, the teacher usually gives fewer essays to the students to practice writing because he or she needs time to mark the essays before returning them back to the students with written feedback (UK Essays, November 2013). With regard to the primary trait scoring method, one of the shortcomings of it is that the primary raters quite often find it hard to focus only on the specified trait as they may unconsciously include other traits in their scoring as well (Salmani, 2014). Similarly, Normah (2014) argues that, even in a study on a group of TESL students at Universiti Pendidikan Sultan Idris (UPSI) found that six out of the seven raters who used the primary trait scoring method to assess directed writing samples admitted that they diverted their attention from the scoring method while assessing.

Types of Scoring Methods

There are three types of scoring methods that can be used to assess students' writing which are Holistic, Analytic and Primary-Traits. The following section details out each scoring method.

Holistic scoring method

The holistic scoring method was introduced in 1960s and it was recognized as the general impression scoring method which became popular in mid 70s (Hunter, Jones and Randhawa, 1996). According to Ghalib and Al-Hattami (2015), the holistic scoring method of assessing written performance is applied in the computer-based Test of English as a Foreign Language (TOEFL), Graduate Record Examination (GRE),

and Graduate Management Admission Test (GMAT). Weagle (2002) explains that the holistic approach is based on a general impression of writing. It considers the overall quality of the product and a single and integrated score or grade will be awarded by the examiner. This approach of scoring advocates the idea that “writing is a single entity which is best captured by a single scale that integrates the inherent quality of the writing” (Hyland, 2003, p. 227).

White (as cited in Salmani, 2014) explains that the holistic method focuses on what the writers “can do well” rather than finding their incompetency and deficiencies in writings. It is like checking whether the writers are really answering the questions and fulfilling the requirements of the task or vice versa rather than scrutinizing on other mechanics of writing. Hyland (2003) asserts that the reliability of the scores through the holistic approach will be more effective when trained examiners mark each essay, because without training and guidance, the examiners might encounter difficulties and uncertainties in assessing as well as determining the right score for the essays according to the specific features laid in the scoring rubrics. Wiseman (2012) states that one of the advantages of holistic scoring rubrics and method in scoring secondary school students’ essays is its time and cost effective. It is the most economical, flexible, practical and applicable assessment. The time required to train the examiners to use the holistic rubrics and to grade the essays using the holistic rubrics is lesser and shorter compared to the analytic scoring. For these reasons, the holistic scoring is the preferred method of scoring in large-scale testing contexts that involve many test takers taking the test at the same time.

Analytic scoring method

Another scoring method is Analytic scoring. Analytic scoring rubric was first introduced in the ESL Composition Profile to measure the performance of the ESL students at North American Universities. It consisted of five different ratings of writing: content (30 points), organization (20 points), vocabulary (20 points), language use (25 points) and mechanics (5 points). Some well-known examples of analytic scales are Test in English for Educational Purposes (TEEP) and the Michigan Writing Assessment Scoring Guide (Ghalib and Al Hattami, 2015). According to Normah (2014), the analytic scoring involves the use of separate scales in assessing different aspects of writing, such as content, organization, vocabulary, grammar and mechanics. This is further supported by Wiseman (2012, p.60) who states that analytic scoring calls for “the separation of the various features of composition into components for scoring purposes”. This type of scoring offers more detailed explanation on the writer’s performance than one single score done in a holistic scoring. Analytic scoring method is preferred compared to the holistic scoring method when comprehensive feedback is needed, especially with small-scale assessment like the classroom assessment. Its detailed feedback assists the examiners (the teachers) in discovering in which aspect their students are good and poor at in essay writing. This would aid the teachers in doing follow-up activities such as consultation, personal coaching and choosing appropriate

exercise for the students in order to help their students to improve their writing. In other words, analytic scoring helps the teachers to discriminate the students’ weak and strong aspects of their writing performance from one test to another. Hence, many researchers have identified analytic scoring as a scoring rubric that has higher discriminating power and it is often used for placement and diagnostic purposes (Ghalib & A-Hattami, 2015).

Primary trait scoring

In addition, Primary Trait Scoring is another type of scoring method which is commonly used. Lloyd-Jones (1977) describes Primary Trait scoring as a type of scoring which involves deciding which one of the aspects of writing acts as the key to success on the task and developing a comprehensive descriptor for performance on that aspect (trait). According to Salmani (2014), this type of scoring is task-specific. It only focuses on a single aspect of writing and goes into detail in that particular aspect. Consequently, the examiner will have a focus in assessing the students’ essay and it is less time consuming compared to the holistic and analytic scoring. Only a score is assigned to the intended criteria for scoring. An example of trait is the development of ideas in an essay, for example, if the guided writing is about “How to Make Fried Noodles”, the trait would be on the clear process of making the fried noodles.

Based on Lloyd-Jones’ (1977) writing, the main advantage of primary trait scoring is that it provides a clear, comprehensive description of a student’s writing ability for a specific rhetorical task. Anderson (1981, as cited in Perkins, 1983) mentions that this type of scoring can be adapted to meet the specific needs of any group of students. Similarly, Normah (2006, p.212) explains that the primary trait scoring method is “very suitable for classroom use because the teachers were given the chance to construct the rubrics depending on what trait they wanted to test on the students”.

METHODOLOGY

This qualitative research involved 25 ESL teachers at 12 secondary schools in one of the states in Malaysia. They were selected based on a few criteria: a) teaching English in secondary school with more than three years of experience, b) teaching English language to Form Four students. The teachers were trained during workshops on how to use the three types of scoring method to examine the students’ continuous essay writing based on the three rubrics scoring methods (Holistic, Analytical and Primary Traits). Additionally, inter-rater reliability was done to look through all the documents with the scores given on essay writing. Initial and on-going rater training is an important way to improve the quality of rater-mediated assessment scheme (McNamara, 2000). After giving the scores, a few teachers were interviewed to get their perspectives of the three scoring methods based on three types of essays: guided writing, summary writing & continuous writing) used for classroom-based assessment.

This research used several data collection techniques: verbal protocol, documents in the form of students’ essays

and interviews with the teachers and students. Best and Kahn (1998) suggest three steps to be taken to analyse qualitative data. In the present study, the three steps such as organizing the data, describing the data and interpreting the data were used. There are two ways to organize data collected: one way is according to individual respondent, and the other way is to group the answers together across respondents. As this paper is part of a larger study, it reports the findings derived from the teachers' interviews, which focused on the teachers' preference of the most suitable scoring method to be used as classroom-based assessment for essay writing.

FINDINGS

Profile of Teachers

Table 1 shows the profile of the participants in the interview. As indicated in Table 1, all of these ESL teachers have a degree in English language teaching, either in TESL, TESOL or English for Communication. The teaching experience ranged from 1 to 20 years, while some of them had some

experience assessing SPM examination papers. Two of them are District language officers.

Teachers' Preference on the Scoring Method

The findings indicated that 85.7 % of the ESL teachers in this study preferred Holistic scoring method, while only 14.3% preferred Primary Trait scoring method. These teachers claimed that they used these scoring methods when assessing their students' guided writing, summary writing and continuous writing. When asked during the interview as to why they preferred holistic scoring method, the teachers mentioned that it is faster, saves time, valid and reliable. Moreover, it gives overall judgement, and it is fair. Table 2 below shows some justifications for the ESL teachers' preference on the most suitable scoring method.

Faster and Saves Time

Most of the ESL teachers in this study preferred the holistic scoring method because it is faster and it can save time. R2

Table 1. Profile of the participants

ESL teachers	Teaching experience	Qualification	Experience assessing SPM papers
R1	12	B.Ed. TESL	5 years
R2	9	B.Ed. TESOL	7 years
R4	19	B.Edu. TESL	none
R9	1	Bachelor of human sciences (English Language & Literature)	none
R13	None	Bachelor of Applied Language (English for Professional Communication)	none
R15	2	Bachelor in English for Professional Comm.	none
R16	20	Bachelor in Education TESL	none
Rater	6	Bachelor in Education TESL	none

Table 2. Teachers' preference and Justifications on the scoring method

Scoring Methods	Themes	Justifications of preference
Holistic	Faster/saves time	Faster, valid and reliable (R1), time saving (R2), a quick method to evaluate based on the reader's general impression (R4), rubric is already distributed by the Malaysian Examination Syndicate (R9), refer to rubric as overall (R3)
	Valid, reliable and consistent	Each band describes performance on several criteria, apply the scoring method consistently, a reliable measurement (R1), the highest construct validity, a list of criteria which include content, style and form.
	Gives overall judgement	Overall judgement of students' performance, marking is done holistically (R1), evaluate quickly without having to look at the marking rubrics many times (R2), the content is there, and they only need to elaborate (R15), no bias since it takes all aspects into account (R3)
	Fairness	Practical, fair judgement, consistency with scoring guide, rubrics aligned with students' need and ability (R4)
	Useful when more essays to mark	Lots of essays from a few classes and get all done before the due date (R2), a big number of papers (R4), practical in assessing hundreds of script (R9)
Primary trait	Focus	Has a focus, less burden, not on grammar or other mechanics of writing (R2), highlight on the content, easier (R13)
Analytic	Time consuming	Time-consuming, focus on separate important traits of writing, refer to the rubrics from time to time, appropriate for a single class (R2)

mentioned that she liked it because it is time saving, while R4 claimed that she could evaluate her students' writing quickly by giving an overall impression and provide the marks in 2 or 3 minutes.

I prefer holistic scoring method for it is time saving. (R2)
As this is a quick method to evaluate a composition based on the reader's general impression of the overall quality of the writing – I can generally read the students' composition and assign a score to it in two or three minutes. (R4)

R3 mentioned that it was time saving because she only had to refer to the overall rubric, while R9 mentioned that the rubric was easy to use because it was distributed by the Malaysian Examination Syndicate to all teachers for assessment purposes. Similarly, R1 also preferred the holistic scoring method because of its time saving and it is reliable.

I refer to holistic scoring method because it saves more time because I only need to refer to rubric as overall. (R3)

Holistic type of scoring method because the rubric is already distributed by Malaysian Examination Syndicate to all teachers for writing assessment. I think it is good to use this type of scoring method compared to others because it can save time and fairer. (R9)

When assessing the students' guided writing, summary writing and continuous writing, the scoring method that has been referred is holistic method. The benefits of holistic method are it is faster, more valid and reliable. (R1)

Useful when Marking Many Essays

The teachers disclosed that they preferred the holistic scoring method when they had to mark many essays. R2 stated that she would use this method when marking essays from a few classes which had to be submitted within a specified deadline.

I prefer this when I have to mark lots of essay from a few classes (involve large scales of students) and have to get all done before the due date. (R2)

Similarly, R4 and R9 stated that holistic scoring saves time when it involves a large number of essays.

Holistic scoring saves a lot of time in assessing a big number of papers. (R4)

This type of scoring method is also more practical if we want to assess hundred scripts. (R9)

Gives Overall Judgement

The participants mentioned that the holistic scoring gives overall judgement of their students' performance. R1 stated that the marking could be done holistically, "not only focusing on the certain particular criteria only but would be based on an overall assessment of performance". Similarly, R2 added that she would rate the overall performance of her students' writing by reading so that she could evaluate the essays "quickly without having to look at the marking rubrics many times".

The method assists the teacher to give overall judgement of students' performance especially in the writing task. Teacher too can concentrate on the strength done by the

students in the writing. Number of decisions that teacher must make to place the band can be minimized as it saves time a lot. Teacher would not mark the writing according to the reference but the marking is done holistically. For example, if the task requires the students to write an open essay about the neighbour, they are free to generate the ideas on their own. Marking is not only focusing on the certain particular criteria only but would be based on an overall assessment of performance. (R1)

Through holistic style of marking, I judge the students' overall performance by reading their writings. (R2)

In addition, R15 mentioned that she preferred the holistic scoring method because it does not focus on a specific part. Besides, the content has been provided and the students are only required to elaborate it.

I prefer to use holistic scoring method for guided writing because this scoring method does not focus on specific part where it is marking on general spec. Students were already been given point to be written. The content is there, and they only need to elaborate on those points given. (R15)

Fairness

Both R4 and R3 claimed that this method is fair and practical because it covers all aspects of assessment. R4 stated that "Holistic scoring provides consistency with scoring guide, or rubrics that are developed aligned with students' need and ability".

This method is very practical in my language class. The students will get a fair judgement from the method as all aspects were taken into account in given marks. Holistic scoring provides consistency with scoring guide, or rubrics that were developed aligned with students' need and ability. (R4)

Other than that, holistic scoring method is more practical and fair. Holistic marking gives us more opportunity to mark the scripts overall and more smoothly. There will be no biased since it takes all aspects into account. (R3)

Valid, Reliable and Consistent

The participants also reported that the holistic method is valid, reliable and consistent. R1 explained that even though teachers had to make overall judgement of the essay, they could match it with several criteria such as organization, vocabulary and language. As such, they could rate the essay consistently and reliably.

In holistic evaluation, teachers make judgments by forming an overall impression of a performance and matching it to the best fit from among the descriptions on the scale. Each band on the scale describes performance on several criteria (e.g. organisation of paragraph + vocabulary + language control). By using this method, experienced teachers tend to apply the scoring method consistently, resulting in more reliable measurement. (R1)

R1 added that when using the holistic scoring, the teacher normally gives his or her students a list of necessary criteria to follow such as the content features and the style, before

giving them the overall impression of the essays. By doing this, they would be able to improve their essays.

Holistic scoring has the highest construct validity when overall attained writing proficiency is the construct to be assessed. When using holistic scoring, a teacher gives students a list of criteria indicating what is necessary to receive a particular score and then assigns a single score for a general impression of the paper as a whole. The list includes content features as well as features of style and form. Students will usually need feedback in addition to the single score in order to understand how they can improve their writing. (R1)

DISCUSSION

In this paper, three types of scoring method are discussed, namely the holistic scoring, analytic scoring and primary trait scoring. Each of them has its own strengths and weaknesses. Holistic scoring method is used to rate a piece of writing generally and is suitable for large-scale assessment as it is time and cost effective. Meanwhile, analytical scoring method is a comprehensive scoring as every language aspect is taken into consideration and rated. It is time-consuming but provides detailed information in discriminating which aspects of writing the students (writers) are good or weak at. Analytical scoring method is suitable for a small-scale assessment for smaller number of students. Last but not least, the primary-trait scoring guides the raters to focus their attention on just a discourse feature in a piece of writing. The scoring for the targeted trait is sharpened and narrowed to the writing task in the question. It is suitable to be used in class-assessment and may differ from a teacher to another.

This paper is written with the aim to explore the ESL teachers' preference towards the most suitable scoring method to be used in the classrooms. The findings indicated that the majority of the ESL teachers in this study preferred the holistic scoring method for its time saving characteristic when a large scale marking needs to be completed within a short period of time. These findings are supported by earlier studies (Becker, 2011; Wiseman, 2012; Ghalib & Al-Hattami, 2015;) which highlighted the benefits of Holistic scoring method. Wiseman (2012) claims that due to its time and cost effective, as well as its economical, flexible, practical and applicable assessment, the holistic scoring is usually the preferred method among teachers, particularly in large scale scoring. In addition, the ESL teachers in this study liked the idea of having to give overall evaluation of the essays whereby they could identify not only the overall band, but also the strengths and weaknesses of their students' writing.

All the participants claimed that the holistic scoring method helped them to improve their students' writing whereby they could evaluate their students' overall performance; hence this helped them to plan their lessons more effectively. As such, initial and on-going training with the teachers in using the scoring rubric is essential in improving the quality of rater-mediated assessment scheme as suggested by McNamara (2000). In this study, the teachers had some experience rating their students' writing based on workshops and seminar they had attended. These trainings eventually

increased their knowledge and skills in assessing students' writing. This is supported by Hyland (2003) who highlights that the reliability of the scores through the holistic approach will be more effective when trained examiners mark each essay. We strongly recommend that training for teachers in scoring their students writing is provided regularly and on long term basis. Without training and guidance, the raters cum teachers might encounter difficulties and uncertainties in assessing as well as determining the right score for the essays according to the specific features prescribed in the scoring rubrics.

CONCLUSION

This study found that the holistic scoring method is the most suitable method to be used for school-based assessment because of its time saving characteristic. Due to the large number of students in each English language class, i.e. 35-40, the ESL teachers need to finish marking the essays within a short period of time to improve their students' performance in writing.

ACKNOWLEDGEMENT

The authors wish to thank the Ministry of Higher Education Malaysia in funding this study under the Fundamental Research Grant Scheme (FRGS) grant, S/O code 13039, and Research and Innovation Management Centre, Universiti Utara Malaysia, Kedah for the administration of this study.

REFERENCES

- Asmah Hj. O. (2003). *Language and Language Situation in South East Asia with a Focus on Malaysia*. Kuala Lumpur: Akademik Pengajian Malaysia, Universiti Malaya.
- Becker, A. (2011). Examining rubrics to measure writing performance in US intensive English programs. *The CATESOL Journal*, 22(11), 113-130.
- Best, J. W., & Kahn, J. V. (1998). *Research in education* (8th ed.). Needham Heights, MA: Allyn and Bacon.
- Carnoy, M. (1999). *Globalization and educational reform: what planners need to know*. International Institute for Educational Planning. UNESCO.
- Ghalib, T. K., & Al-Hattami, A. A. (2015). Holistic versus Analytic Evaluation of EFL Writing: A Case Study. *English Language Teaching*, 8(7), 225.
- Hamp-Lyons, L. (1991). Scoring procedures. In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts* (pp. 241-276). Norwood, NJ: Ablex.
- Hunter, D. M., Jones, R. M., & Randhawa, B. S. (1996). The use of holistic versus analytic scoring for large-scale assessment of writing. *The Canadian Journal of Program Evaluation*, 11(2), 61.
- Hyland, K. (2003). *Second language writing*. NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511667251>
- Lee, Molly N.N. (2002). *Educational Change in Malaysia*. Monograph series No:3/2002. School of Educational Studies. Universiti Sains Malaysia, Malaysia.

- Lloyd-Jones, R. (1977). Primary Trait Scoring. In C.R. Cooper & L. Odell (Eds.), *Evaluating writing: Describing, measuring, judging* (pp. 33-68). Urbana, IL: National Council of Teachers of English.
- Malaysian Examinations Syndicate (LPM, 2009). *Pengumuman Analisis Keputusan SPM 2009*. Putrajaya.
- McNamara, T. (2000). *Language Testing*. Oxford: Oxford University Press.
- Ministry of Education Malaysia (2000). *Sukatan Pelajaran Bahasa Inggeris (English Language Syllabus)*. Kuala Lumpur.
- Normah Othman. (2006). Assessment of Directed Writing by a Group of TESL Student at UPSI. Paper presented at *Prosiding Seminar Penyelidikan*. Universiti Pendidikan Sultan Idris.
- Normah Othman. (2014). The primary trait scoring method for classroom-based assessment of students' direct writing. *International Journal of Learning & Development*, 4(3), 51-61. doi:10.5296/ijld.v4i3.6063
- Perkins, K. (1983). On the use of composition scoring techniques, objective measures, and objective tests to evaluate ESL writing ability. *TESOL Quarterly*, 17(4), 651-671.
- Salmani Nodoushan, M.A. (2014). Assessing writing: A review of the main trends. *Studies in English Language Education*, 1(2), 119-129.
- Selvaraj, B. (2010). English Language Teaching (ELT) Curriculum Reforms in Malaysia, *Voice of Academia*, 5(1), 51-60.
- Weagle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.
- Wiseman, S. (2012). A Comparison of the Performance of Analytic vs. Holistic Scoring Rubrics to Assess L2 Writing. *Iranian Journal of Language Testing*, 2(1), 59-92.