6 ICLLCE 2016-012 Kim Thanh Tuyen

Developing and Validating Scoring Rubrics for Assessing Research Papers Writing Ability of EFL/ESL Undergraduate students

Kim Thanh Tuyen[a]*, Shuki Bin Osman[a]
Nor Shafrin Binti Ahmad[a], Thai Cong Dan[b]
[a] School of Educational Studies, Universiti Sains Malaysia (USM)
[b] School of Social Sciences and Humanities Campus II, Can Tho University, Vietnam
*Corresponding author: Kimthanhtuyen80@yahoo.com.vn

ABSTRACT

To avoid the subjectivity and a risk of reliability, selecting the most suitable method of scoring written texts is important in the evaluation of academic writing ability. Thus, previous authors proposed several methods of marking, namely holistic method, the impression method, error count method, analytic method, etc., for scoring academic writing texts in different genres at various levels. However, very little attention has been paid to identifying effective methods of scoring research papers writing ability. Therefore, this paper aims to propose using analytic method including an attempt to separate various aspects of a composition. Specifically, in this study, analytic scale is proposed to evaluate research papers writing ability of EFL/ ESL undergraduate students. According to this marking scale, writing components, namely content (relevant ideas), organization (structure and coherence), language use (vocabulary and grammar choice), and mechanics use (punctuation and spelling based on APA style) are graded separately. Delphi technique (DT) was used to validate it through the interviews of experts including two boards of ten experienced and qualified lecturers of TESOL and curriculum studies in Can Tho University (CTU), Vietnam and Universiti Sains Malaysia (USM). The results from the interviews revealed that scoring rubrics for the assessment of research papers writing ability were determined. This paper is, therefore believed to contribute to practical applications for research papers reviewers, lecturers, examiners and undergraduate and postgraduate students in EFL/ ESL contexts.

Keywords: Research papers writing ability; scoring rubrics; EFL/ ESL undergraduate students

Introduction

Objectives

Due to the importance of English academic writing in academic life as well as at workplace, it is necessary to identify valid and reliable ways to test undergraduate students' academic writing ability. Indeed, testing plays a vital role in the educational system as well as in the teaching and learning process. It is actually required in all classrooms and becomes a key predictor of future professional or academic success. It is also important to students due to its impacts on their education and their daily life (William 1996; Brown 1996; White 1994; Sahin 2007). According to Hughes (1989), testing can provide both lecturers and learners with beneficial backwashes as well as harmful backwashes. A good classroom assessment can bring benefits to lecturers in some ways; for example, it helps the lecturers see how well learners have learnt. Moreover, a valid assessment can also help the lecturers to evaluate the effectiveness of the syllabus as well as the methods and materials the lecturers are using. Consequently, the lecturer can select their own teaching methods to make sure that it corresponds to the syllabus, the objective of the curriculum,

learners' needs and learners' levels. Additionally, a good classroom assessment provides a good feedback for learners; as a result, once the task has been scored and evaluated, learners can learn something about their areas of strengths and weaknesses. Learners also have a better chance to assess and self-regulate their learning process. In other words, assessment plays a decisive role in deciding the final result of students' learning process.

However, assessment can lead to harmful backwashes if it lacks validity and reliability. It is challenging for teachers of English to evaluate students' assignments in a reliable way, especially students' writing ability (Alderson et al., 1995) because it needs to take into considerations task variables, test taker variables, rater variables, and rating scales (Bachman & Palmer 1996). Therefore, to avoid the subjectivity and a risk of reliability of assessment, selecting the most suitable method of scoring students' assignments including English written assignments is necessary. Methods of scoring have been presented in previous literature (e.g., Hamp-Lyons, 1991; Shohamy, 1995) to improve both accuracy and consistency (Brown, 1996; Wiseman, 2012). Nevertheless, there is a need to take into account the rating scale. According to Park (2004), in order to assess academic writing ability, scoring rubrics should be used. In this sense, three types of scoring rubrics, namely analytic, holistic and primary trait are introduced (Bachman & Palmer 1996; Weigle 2002; Alderson, 1995). In addition to the use of such scoring rubrics as scoring guilds to evaluate writing ability, the impression method, and error count method are used (Heaton, 1974; Cohen, 1994) for scoring academic writing texts in different genres at various levels. Yet, Becker (2010/2011) revealed that U.S. universities usually adopted an existing scale, and very few educational institutions designed their own scoring rubrics. Especially, very little attention has been paid to identifying effective methods of scoring research papers writing ability. For this reason, this paper aims to propose a valid and reliable scoring rubric to assess research papers writing ability of EFL/ESL undergraduate students in Can Tho University who major in English.

Research Question

Based on the above objectives, the research question of this study is stated as follows:

  "Which core components of English academic writing should be proposed in the scoring rubric of assessing research papers writing ability of English majored students of Can Tho University, Vietnam?"

Theory

Which core components of English academic writing should be assessed? Scoring rubric is known as a marking scheme or marking guide. Cohen (1994) suggested grading aspects of academic writing such as organization, ideas, spelling and so on separately. Meanwhile, Park, T. (2004) suggested grading more aspects of writing components such as content, organization, cohesion, register, vocabulary, grammar, or mechanics. In this study, analytic scale adapted from Carroll and West (1989, as cited in Tribble, 1996) is proposed to grade research papers writing ability of EFL/ ESL undergraduate students. According to this marking scale, writing components, namely content (relevant ideas), organization (structure and coherence), language use (vocabulary and grammar choice), and mechanics use (punctuation and spelling based on APA style) are graded separately.

Methodology

Instrument

Using the Delphi method has a wide range of advantages (Brill et al., 2006; Hasson et al., 2000; Lambrecht, 2007) such as expressing opinions without anxiety. Thus, Delphi method, as the most effective method, has been used to gain experts' ideas in a reliable way. Interestingly, this technique does not require us to directly contact the participants. Thus, Delphi method is allowed to gather experts into a panel despite their different backgrounds so that these experts could contribute to the discussions regardless of their geographic location (Linstone & Turoff, 2002). However, it was suggested by Sappe (1984) that institutions should be flexible when using this method. It means that, instead of depending on outside groups' ideas, these educational institutions are possible to change or adjust the content of the program to fit the real contexts. Similarly, self-analysis is also encouraged to be used for the application of this method.

For the mentioned reasons, in this study, the Delphi Technique (DT) was conducted within two rounds to validate a scoring rubric for evaluating research papers writing ability of EFL/ESL undergraduate students. To save time, a questionnaire with structured questions was used through face to face discussions to survey two groups of experts' opinions about the criteria description. The questionnaire was developed by researcher based on literature on scoring rubrics of academic writing ability. Three point scale responses (Yes, No, Undecided) were used for each item. The questionnaire includes necessary criteria of evaluating research paper writing and thus categorized into 4 structures. Structure 1 contains items involving evaluation criteria related to the content of RPW. Structure 2 contains items involving evaluation criteria related to the organization of RPW. Structure 3 contains items involving evaluation criteria related to the language use of RPW. Structure 4 contains items involving evaluation criteria related to the mechanics use. The scoring rubric of evaluating research paper writing ability through the questionnaire was then introduced to the expert jury. These experts were asked to evaluate and validate it based on the proposed questionnaire. They were also encouraged to make any addition, omission, corrections, and change if necessary. After all their comments and suggestions were consolidated, the scoring rubric was adjusted by the researcher and then sent to the expert jury again to confirm the final scoring rubric of evaluating research paper writing.

Participants

In this study, a scoring rubric was designed to evaluate research papers writing ability of EFL/ESL undergraduate students, who were the third year university students in academic year (2016-2017), majoring in English. They are studying in English Department, School of Education, Can Tho University (CTU), a multidisciplinary university, which is the biggest public university in the Southeast of Vietnam. It has a mission to develop scientific research projects and get an access to scientific and technological knowledge for problematic solutions to science, technology, economics, culture and society in the region. These student are taught four skills of English as their major subjects (i.e., speaking, listening, writing, and reading). They are compulsory to study genres of academic writing (i.e., sentences writing, letters, paragraphs writing, essays writing etc.) in the first and second academic year. In the third academic year, these students are encouraged to attend a course of research paper writing to prepare for their thesis writing in the final year of their undergraduate program.

Sutphin (1981) proposed using nine experts, whereas Huber and Delbecq (1972) suggested using at least five of them. According to Dobbins (1999), determining the number of experts should be based on the study, and the cost. An expert as an individual

who has special skills or knowledge (Gove, 1981). Therefore, in this research, ten experts including two boards of experienced and well-qualified lecturers of TESOL and curriculum studies in CTU and Universiti Sains Malaysia (USM) were interviewed to confirm the components of the scoring rubric of assessing research paper writing ability of EFL/ESL undergraduate students. Six lecturers got Ph.D degree, and four lecturers got Master degree. All of them are enthusiastic and experienced in teaching academic writing for ESL/ELS university students.

Literature review

What is an effective scoring rubric of assessing English academic writing ability? Scoring rubrics are descriptive scoring schemes that are developed by teachers or other evaluators to guide the analysis of the products or processes of students' efforts (Brookhart, 1999). Scoring rubrics are typically employed when a judgement of quality is required and may be used to evaluate a broad range of subjects and activities. One common use of scoring rubrics is to guide the evaluation of writing samples. Judgements concerning the quality of a given writing sample may vary depending upon the criteria established by the individual evaluator. One evaluator may heavily weigh the evaluation process upon the linguistic structure, while another evaluator may be more interested in the persuasiveness of the argument. A high quality essay is likely to have a combination of these and other factors. By developing a pre-defined scheme for the evaluation process, the subjectivity involved in evaluating an essay becomes more objective. Schafer (2004) views rubrics as tools that are used to assess the quality of student work in a range of excellent to poor performances. A rubric has a criteria that corresponds to a scale of possible points to score spoken or written performances. The highest point refers to the best performance, whereas the lowest one refers to the worst performance on the scale. Various levels of proficiency are included in the scale. It can be generic enough to be used with various types of writing. Four common scoring types which use rubrics include holistic scoring, analytical scoring, weighted trait scoring, and primary trait scoring (Campbell, Melenyzer, Nettles, & Wyman, 2000). Herman, Aschbacher, and Winters (1992) proposed four characteristic features of a rubric such as criteria, standards, scale, and examples. An effective rubric has a clear criteria for the test-takers to know what is expected and for the raters to assess the answers. It also has a good standard for various levels of performance, and has a valid and reliable scale to meet the standard of writing performance. It is possible to include an example of expected performance at the different levels on the scale.

Types of Assessment Rubrics

Advantages and disadvantages of holistic scoring rubrics. Relevant literature (e.g., Cumming, 1990; Weigle, 2002; East & Young, 2007) refers to the important role of three basic assessment rubrics for evaluation, namely primary trait analytic, and holistic scoring rubrics in the evaluation of written proficiency. There is a difference in impact, discriminatory power, inter-rater reliability, the degree of bias, and the cost-effectiveness in terms of time, effort and money among three basic assessment rubrics (Kuo, 2007). Holistic scoring takes into account the general response of written work and provides an overall score to the performance (White, 1994; Weigle, 2002; Hyland, 2002). According to Park, T. (2004), holistic scoring is economical compared to analytic scoring because a single score is used to grade writing work. A single score gives useful ranking information but no details. Therefore, holistic scoring is widely used to assess writing in large-scale due to its a general guideline that defines good performance at each score point. According to White (1994) and Cohen (1994), holistic scoring has more advantages than

disadvantages. In addition to its advantages, holistic scoring rubric has its drawbacks. For example, diagnostic information about students' writing cannot be provided because specific aspects of writing (e.g. organization, grammar, vocabulary etc.) are not provided. Moreover, it lacks reliability because it scores the text generally (Ferris & Hedgcock, 1998; Park, T., 2004)

Advantages and disadvantages of analytic scoring rubrics**.** Using analytic rubrics to assess students' writing may lead to the improvements in learners' writing skills (Anderson, 1995) and critical thinking subskills of writing (Becker, 2011). In this regard, writing components (i.e., unity, coherence, flow of ideas, formality level and so on) are analyzed. Based on this type of scoring rubric, aspects of writing such as organization, ideas, spelling and so on are graded separately (Cohen, 1994). According to Park, T. (2004), writing components such as content, organization, cohesion, register, vocabulary, grammar, or mechanics are graded. In this sense, more detailed information about a test taker's performance in different aspects of writing is provided. Unlike holistic scoring, analytic rubrics provide more useful diagnostic information about students' writing abilities including information about the strengths and weaknesses of students. Therefore, instructors and curriculum developers could give the instruction to meet the students' needs. It is stated that L2 learners get more benefits from analytic scoring, especially for some learners who may perform content and organization of writing well, but may have more errors on grammars; others may perform sentence structure well, but may not good at organize their writing coherently. Likewise, less inexperienced raters find it easier to grade students' writing based on analytic scale. Finally, a clear analytic scoring with clear criteria helps writers get consistent and direct feedback. However, it has several disadvantages. It is impossible to evaluate a piece of good writing based on 3 or 4 criteria. Each scale may not be used separately. It is difficult to grade writing performance effectively based on the description for each scale (e.g. what does 'adequate organization' mean?'). Indeed, if the 'idea' scale has high scores, other scales are influenced.

Advantages and disadvantages of primary trait scoring**.** Primary trait scoring was developed for scoring the performance of essay writing (Lloyd-Jones, 1977). Primary trait scoring, known as holistic scoring, is less common than other methods (Becker, 2011). It is similar to holistic scoring; however, it is required to focus on individual characteristics of writing task. It deals with the core features of particular genre of writing; for instance, by considering differences between several types of essays (Weigle, 2002). According to Lloyd- Jones (1977), like analytic scales, primary trait scoring yields a quite reliable score thanks to sufficient training which is provided for raters. Primary trait scoring guides focus on the rater's attention to the features of a piece of writing which are relevant to the kind of discourse (i.e., audience, speaker role, purpose, and subject). A unique quality of primary trait scoring is that scoring guides are constructed for a particular writing task set in a full rhetorical context. However, it also has its drawbacks due to this characteristics. For example, it will ignore other mistakes or errors on other aspects of writing. Moreover, it takes a longer time to design because the scoring guilds are long. Perkins. K (1983) stated that it has not been widely used in the classroom due to inadequate feedback about students' writing.

Findings
Core Components in the Scoring Rubric of Assessing Research Papers Writing Ability

From the obtained data, all of ten experts were agreed with core components in the scoring rubric of assessing research papers writing ability of EFL/ESL undergraduate students proposed by the researcher of this study. As mentioned earlier, scoring rubric has

also been known as a marking scheme or marking guide. In this study, analytic method, including an attempt to separate various aspects of a composition was used. Specifically, analytic scale adapted from Carroll and West (1989, as cited in Tribble, 1996) was proposed to grade research papers writing ability of EFL/ESL undergraduate students. According to this marking scale, writing components, namely content (relevant ideas), organization (structure and coherence), language use (vocabulary and grammar choice), and mechanics use (punctuation and spelling based on APA style) are graded separately.

Based on the evaluation system of Can Tho University, the total point of all subjects is 10.0 points. Thus, in this paper, the total point of research writing ability is 10.0 points including 4.0 points for content, 2.0 points for organization, 2.0 points for language use, and 2.0 points for mechanics use. The score is converted into 4 levels. Excellent to good level includes Level A and level B+. Level A is in the score range of 9-10 points. Level B+ is in the score range of 8.0-8.9 points. Fair level includes level B and level C+. Level B is in the score range of 7.0-7.9 points. Level C+ is in the score range of 6.5-6.9 points. Average include level C and level D+. Level C is in the score range of 5.5-6.4 points. Level D+ is in the score range of 5.0-5.4 points. Poor level includes level D and level F. Level D is in the score range of 4.0-4.9 points. Level F is below point 4.0. Detailed scoring rubric of assessing research papers writing ability is presented in Figure below.

Writing components	Score value	Level/ Description of criteria
Content	9-10 (A) 8.0-8.9 (B+) 7.0-7.9 (B)	EXCELLENT TO GOOD: Excellent to very good treatment of the subject• Content relevant to the topic • Balanced presentation of relevant and legitimate information that clearly supports a central purpose or argument and shows a thoughtful, in-depth analysis of a significant topic•Readers gain important insights.
	6.5-6.9 (C+) 5.5-6.4 (C) 5.0-5.4 (D+)	FAIR: Adequate treatment of the topic• Most content relevant to the topic• Information provides reasonable support for a central purpose or argument and displays evidence of a basic analysis of a significant topic•Readers gain some insights.
		AVARAGE:Treatment of the topic is hardly adequate • Some irrelevant content, relevant to the topic• Analysis is basic or general•Reader gains few insights.
	4.0-4.9 (D) <4 (F)	POOR: Inadequate treatment of the topic • almost no useful detail• Reader is confused or may be misinformed•Central purpose or argument is not clearly identified•Analysis is vague or not evident.
Organization	9-10 (A) 8.0-8.9 (B+) 7.0-7.9 (B)	EXCELENT TO GOOD: Fluent expression • ideas clearly stated/supported • Appropriately organized paragraphs or sections• Logical sequenced (coherence)• connectives appropriately used (cohesion).

Writing components	Score value	Level/ Description of criteria
	6.5-6.9 (C+) 5.5-6.4 (C) 5.0-5.4 (D+)	FAIR: Uneven expression • but main ideas stand out• Paraphrasing or section organization evident• Logical sequenced (coherence)• Connectives appropriately used (cohesion).
	4.0-4.9 (D) <4 (F)	AVERAGE: Very uneven expression• Ideas difficult to follow stand out• Paraphrasing or section organization does not help the reader, logical sequence difficult to follow (coherence)•Connectives largely absent (cohesion).
		POOR: Lack fluent expression, ideas very difficult to follow• Little sense of paraphrasing / organization• No sense of logical sequence (coherence) • Connectives not used (cohesion).
Language use (Vocabulary+ grammar)	9-10 (A) 8.0-8.9 (B+)	EXCELLENT TO GOOD: accurate word and usage• Appropriate selection to match register• Confident handing of appropriate structures • Hardly any errors of agreement, tense, number, word order/function, articles, pronouns, prepositions, meaning never obscured.
	7.0-7.9 (B) 6.5-6.9 (C+)	FAIR:• Occasional mistakes in word choice and usage• Register not always appropriate• acceptable grammar- but problems with more complex structures•Mostly appropriate structures• some errors of agreement, tense, number, word order/function, articles, pronouns, prepositions, meaning some times obscured
	5.5-6.4 (C) 5.0-5.4 (D+)	AVERAGE: A noticeable number of mistakes in word choice and usage• Register not always appropriate• insufficient range of structures with control only shown in simple constructions• Mostly appropriate structures•Frequent appropriate structures• Some errors of agreement, tense, number, word order/function, articles, pronouns, prepositions, meaning some times obscured
	4.0-4.9 (D) <4 (F)	POOR:Uncomfortably frequent mistakes in word/ idiom choice and usage• Register , no appropriate sense of register• Major problems with structures-even simple ones• frequent errors of negation agreement, tense, number, word order/function, articles, pronouns, prepositions, meaning often obscured

Writing components	Score value	Level/ Description of criteria
Mechanics use skills	9-10 (A) 8.0-8.9 (B+) 7.0-7.9 (B) 6.5-6.9 (C+) 5.5-6.4 (C) 5.0-5.4 (D+) 4.0-4.9 (D) <4 (F)	EXCELLENT TO GOOD: The writing is free or almost free of errors on punctuation, spelling and APA style. FAIR:There are occasional errors on punctuation, selling and APA style, but they don't represent a major distraction or obscure meaning. AVERAGE: The writing has many errors on punctuation, spelling and APA style, and the reader is distracted. POOR: There are so many errors on punctuation, spelling and APA style that meaning is obscured.

Figure 1. Scoring rubrics of assessing research paper writing
Sources: Adapted from Carroll and West (1989)'s scoring profile (as cited in Tribble, 1996)

Limitations

From the results obtained from Delphi technique, this study is limited to propose the scoring rubric perceived as essential for EFL/ESL lecturers and examiners to evaluate research paper writing ability of EFL/ESL undergraduate students, who major in English and study four skills of English: listening, speaking, reading, writing. Especially, they have already learnt basic academic writing skills and academic writing language in the previous courses. This scoring rubric is, therefore, not generalized to the assessment of research paper writing ability of EFL/ESL undergraduate students in all fields.

Recommendations, and Conclusions

The results from the interviews revealed that scoring rubrics for the assessment of research papers writing ability were determined. Thus, it is believed that this paper will significantly contribute to the practical applications for examiners, lecturers, undergraduate and postgraduate students in EFL/ESL contexts. Noticeably, it is suggested that Higher Education institutions in general and lecturers in specific should adjust or modify the description of each criteria as possible to meet the requirements of their own educational institutions.

Acknowledgements

References

Alderson, J. C., Clapman, C., & Wall, D. (1995). *Language test construction and evaluation.* Cambridge: CUP.
Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice.* Oxford: Oxford University Press.

Becker, A. (2010/2011). Examining rubrics used to measure writing performance in U.S. intensive English programs. *The CATESOL Journal*, *22*(1), 113-130. Retrieved from http://www.catesol.org/Becker%20113-130.pdf

Brill, J.M., Bishop, M.J., & Walker, A.E. (2006). The competencies and characteristics required of an effective project manager: A web-based Delphi study. Educational Technology Research and Development, *54*(2), 115-140. Doi:10.1007/s11423-006-8251-y.

Brookhart, S. M. (1999). *The Art and Science of Classroom Assessment: The Missing Part of Pedagogy.* ASHE-ERIC Higher Education Report (Vol. 27, No.1). Washington, DC: The George Washington University, Graduate School of Education and Human Development.

Brown, J. D. (1996). *Testing in language programs.* Upper Saddle River, NJ: Prentice Hall Regents.

Campbell, D. M., Melenyzer, B. J., Nettles, D. H., & Wyman, R. M. Jr. (2000). *Portfolio and performance assessment in teacher education.* Boston: Allyn and Bacon.

Cohen, A. D. (1994). *Assessing language ability in the classroom.* Boston, MA: Heinle & Heinle.

Cumming, A. (1990). Expertise in evaluating second language compositions. *Language Testing*, *7*(1), 31-51.

Dobbins, T.R. (1999). Clinical experiences for agricultural teacher education programs in North Carolina, South Carolina and Virginia. (Doctoral Dissertation, Virginiar Polytechnic Institute and State University, 1999). Dissertation Abstracts International. 60/11. 3875

East, M., & Young, D. (2007). Scoring L2 writing samples: Exploring the relative effectiveness of two different diagnostic methods. *New Zealand Studies in Applied Linguistics*, *13*, 1-21.

Ferris D & Hedgcock JS (2005) *Teaching ESL Composition: Purpose, Process, and Practice.* Mahwah: Lawrence Erlbaum.

Hamp-Lyons, L. (1991). Scoring procedures for ESL context. In L. Hamp-Lyons (Ed.), *Assessing second language writing* (pp. 241-277). Norwood, NJ: Ablex.

Herman, J. L., Aschbacher, P. R., & Winters, L. (1992). *A practical guide to alternative assessment.* Alexandria, VA: Association for Supervision and Curriculum Development.

Huber, G.P.& Delbecq, A. (June, 1972). Guidelines for combining the judgements of individual members in decision conferences. *The academy of Management journal*, *15*(2), 161-174.

Hyland, K. (2002). *Teaching and research Writing.* Harlow, England: Pearson Education Limited.

Klimova, F. B. (2011). Evaluating writing in English as a second language. *Procedia-Social and Behavioral Sciences,* 28,390-394.

Kuo, S. (2007). Which rubric is more suitable for NSS liberal studies? Analytic or holistic? *Educational Research Journal*, *22*(2), 179-199.

Gove, P.B. (Ed) (1981). Webster's Third New Internatinal Dictionary. Springfield. MA; G & C C Merriam Co.

Lambrecht, J.J. (2007). Business Education Delphi study of future directions for the field. *Delta Pi Epsilon Journal*, *49*(1), 15-25. Retrieved from ERIC database. (EJ779205).

Lloyd-Jones, Richard. (1977). "Primary Trait Scoring." Charles Cooper and Lee Odell (Eds.). Evaluating Writing: Describing, Measuring, and Judging, pp.33-66. Urbana, IL: NCTE.

Linstone, H.A. & Turoff, M. (1975). The Delphi Method: Techniques and applications. Reading, Massachusetts: Addison-Wesley Publishing Co.

Odell, Lee, and Charles IL Cooper. (1980). "Procedures for Evaluating Writing:Assumptions and NeededResearch." College English 42(1), pp. 35-43.

Park, T. (2004). *Scoring procedures for assessing writing*. Retrieved August, 2016, from http://www.tc.columbia.edu/academic/tesol/webjournal/park_Forum.pdf

Perkins K (1983) On the use of composition scoring techniques, objective measures, and objective tests to evaluate ESL writing ability. *TESOL Quarterly* 17: 651-671.

Tribble, C. (1996). Writing. Oxford: Oxford University Press.

Sahin, A. E. (2007). İlköğretim bölümü mezunlarının başarılarının mezun oldukları lise türlerine göre karşılaştırılması.*Egitim Araştirmalari-Eurasian Journal of Educational Research*, 29, 113-128.

Sappe, J.H.(1984). A national Delphi study of selected terminology used by curriculum, University specialists in vocational education curriculum development. (Doctoral Dissertation, University of George, 1984). Disseration Abtracts Internaltional , 45/09, 2852.

Schafer, L. (2004). *Rubric*. Retrieved February 9, 2015, from http://www.etc.edu.cn/eet/articles/rubrics/index.htm

Shohamy, E. (1995). Performance assessment in language testing. *Annual Review of Applied Linguistics*, *15*,188-211.

Spandel, V. and R. Stiggins. (1980). Direct Measures of Writing Skill.Portland, OR: Northwest RegionalEducational Lab.

Sutphin, H.D. (1981). Positions held by teachers, teacher educators, and state supervisors about selected national issuesa in agriculture education. (Doctoral Dissertation, the Ohio State University, 1981). Disseration Abtracts Internaltional, 42/10,4257.

Webber, K. L. (2012). The use of learner-centered assessment in US colleges and universities. *Research in Higher Education*,

53, 201-228.

Weigle, S. C. (2002). *Assessing Writing*. Cambridge: Cambridge University Press.

Schafer, L. (2004). *Rubric*. Retrieved February 9, 2015, from http://www.etc.edu.cn/eet/articles/rubrics/index.htm

William, D. J. (1996). *Assessing writing. Preparing to teaching writing*. Belmond: Lawrence Erlbaum Associates.

White, E. M. (1994). *Teaching and assessing writing*. San Fransisco: Jossey-Bass Publishers.

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.