

EVALUATION CRITERIA FOR STUDENTS' ACADEMIC WRITING: AN ACTION RESEARCH

Hisamitsu Wada

[Key Words] Criteria, Academic writing, Evaluation, Feedback

Abstract

In order to establish criteria for students' academic writing that is appropriate for both evaluation and feedback by teachers, this action research determined 15 criteria for evaluating students' academic writing. These criteria were then applied in the evaluation of writings that a university freshman had written on the 7 question topics that appeared in TWE in TOEFL. Finally, in order to examine both the validity of the criteria employed in this research and the improvement of the writing skills of the student, the student was solicited to revise her own writing on every topic twice; as a result, she wrote 3 essays on each topic, i.e., she wrote 21 essays in all. The results indicate that the criteria work well not only in evaluation but also in feedback by the teachers, and that writing essays several times on the same topic clearly contributes to the improvement of students' writing skills in the perspective of academic writing.

1. Introduction

Nearly a quarter of a century has passed since the “process-or-product” argument first emerged in the field of TESOL (Teaching English to Students of Other Languages) in the early 1980s, and it is true that this argument, in general, has contributed to changing the attitude of many teachers of English language toward their feedback on students' writings from evaluating students' essays as products into focusing on the process that their students write English essays. In fact, most textbooks of English writing that have been published in the US since the 1990s employ titles/subtitles including the word “process,” indicating that these textbooks adopt process approach.

However, in reality, concrete criteria for teachers' feedback on students' writings that can evaluate the “process” of writing have not been established. As a result, it seems that

although many English teachers desire to give their students feedback that is beneficial for improving the process of writing English, their feedback appears different from teacher to teacher. Some teachers might focus on the holistic organization, while others might concentrate on grammatical mistakes.

Therefore, in this action research, in order to establish concrete criteria for the evaluation and feedback on students' English writing in the area of academic writing, 15 evaluation criteria were initially determined based on the evaluation criteria for academic paragraph writing proposed by Wada (2005), adding criteria for organization and coherence of the whole essay.

Next, 21 academic English essays written by the participant were evaluated with the 15 criteria in order to examine to what extent and in what ways the criteria are beneficial for teachers' evaluation and feedback.

2. Rationale

As Wada (2005) points out, the criteria for the evaluation and feedback on students' academic writing by teachers should be derived from the nature of academic writing. Therefore, the 15 temporary criteria that are to be examined in this action research are based on the criteria that Wada (2005) proposed regarding academic paragraph writing. Since this study aims to examine criteria for an English essay, 5 criteria concerning organization and 4 criteria concerning coherence of the whole essay were newly determined.

In grading the essay based on the 15 temporary criteria determined in this study, the question of the grading balance between these criteria arises. Some teachers may claim that emphasis should be laid on grammatical accuracy, while others may prefer to focus on the organization and coherence of the essay. In order to solve this problem, this research employed a quantitative approach based on the hypothesis that the more a student writes an English essay the more the student's writing skills improve. To be concrete, each criterion was regarded as having an equivalent value because the evaluations on the 21 essays by the 15 criteria (315 in all) can be expected to indicate whole validity of these criteria. For example, if the first essay that the participant wrote got 15 points (full mark is 45 since each criterion is graded at 4 levels: 0, 1, 2, and 3), her 10th essay got 28 points, and her 21st essay got 40 points, it could indicate that her writing skills improved on the whole, and thereby it can be said that the criteria have validity to some extent. On the other hand, if the score greatly fluctuates and shows no statistically significant increase, it can be said that the criteria have no validity.

Therefore, the results of this action research are holistically analyzed so that the general validity of the 15 criteria are examined; that is, the validity of each criterion is not analyzed, which would be the theme of further research.

Finally, this action research has two reasons why a freshman who just entered the university with no experience of writing an English essay was chosen as the participant of this study: 1) The results would be more likely to indicate the trait of the participant's improvement in writing skills than the results of a participant who has considerable experience of writing; in other words, the scores would more clearly indicate the evaluation validity of the 15 criteria, and 2) It could be expected that by being given the 15 criteria in revising her own essay, the criteria could work as feedback; that is, the participant can raise her consciousness of such criteria of academic writing.

3. Method

3.1 Participant

The participant is a freshman of a Japanese private university. At the outset of this action research (April 2005), she had no experience of writing an essay in English, nor had she learned basic skills of academic writing at the high school level. Thus, in the initial phase of this action research, the author introduced her basic concepts and rudimentary skills of English academic writing in two 3-hour private lessons such as the concept of topic sentence and unity of a paragraph, the skills of paragraph development, the concept of organization and coherence, and the skills to write the Introduction and Conclusion parts of an English essay. Next, the participant was solicited not only to write English essays on 7 question topics that appeared in TWE (the Test of Writing English) in TOEFL (the Test of English as Foreign Language) but also to revise each of her own essays twice; as a result, she wrote 21 academic English essays during nearly 6 months (from April 2005 to September 2005). The 7 essay topics employed in this action research are shown in Appendix.

3.2 Criteria

The 15 criteria employed in this study are as follows:

<Organization>

1. Does this essay have an Introduction, Body, and Conclusion?
2. Is the Introduction appropriate?
3. Is the Conclusion appropriate?
4. Does the Body have more than one reason and/or example to support the thesis?

5. Is the length of this essay enough (around 300 words)?

<Coherence>

6. Is the overall development of the thesis coherent without contradiction?

7. Does this essay have appropriate reasons and/or examples to support the thesis?

8. Is the usage of transitions in this essay appropriate?

9. Are paragraph development skills used appropriately in accordance with the topic sentence?

<Topic Sentence, Style, and Syntax Variety>

10. Does each paragraph have a proper topic sentence?

11. Is the balance between Simple Sentence, Compound Sentence, and Complex Sentence appropriate? Does this essay use many Complex Sentences?

12. Is the average length of the sentence in this essay appropriate to academic writing (the ideal length is 25 words per sentence on the average)?

13. Does this essay use various syntax?

<Surface-level Errors>

14. Does this essay have errors that hinder the reader from comprehending what the writer really intends to say?

15. Does this essay have errors in grammar, spelling, or punctuation?

3.3 Grading procedure

All the 21 English essays that the participant wrote were graded with the criteria as follows:

1) Each criteria, except for Criterion #15, was scored at 4 levels: 0 means “not written” or “incomprehensible”; 1 means “does not satisfy the criterion”; 2 means “needs improvement although roughly satisfies the criterion”; and 3 means “Satisfies the criterion and no improvement is necessary”.

2) As for Criterion #15 (errors in grammar, spelling, or punctuation), scoring was based on the number of errors per sentence. If the essay has more than 1.0 error per sentence on average, it is equivalent to every sentence having at least one surface-level error. Therefore, the grading scale is determined as follows:

0: Average number of errors per sentence (N) = 0.7 or above

1: $0.7 > N \geq 0.4$

2: $0.4 > N \geq 0.2$

3: $0.2 > N$

4. Results and findings

Firstly, the results indicate overall evaluation validity of the 15 criteria that were employed in this action research. Since the participant was solicited to revise each of her own essays 2 times, let us examine the scores of the final version of the 7 essays in order to examine the overall improvement of her writing skills. **Table 1** indicates total scores of the 7 final versions. Although the first 3 scores indicate no significant improvement, it is evident that on the whole the participant's writing skills improved.

Table 1 Total score of the final versions

1st	2nd	3rd	4th	5th	6th	7th
28	25	28	39	38	37	38

Secondly, in order to inspect the hypothesis that the more an EFL student writes in English the more surface-level errors decrease, let us examine the score of Criterion #15 (errors in grammar, spelling, or punctuation) of the "first" version of the 7 essays because it is natural that the "second" and "final" versions should contain revisions of surface-level errors, and because it can be assumed that the "first" version is more likely to indicate the writer's spontaneous errors. **Table 2** indicates the scores of Criterion #15 of the "first" versions of the 7 essays.

Table 2 Sores of surface-level errors in the first versions

1st	2nd	3rd	4th	5th	6th	7th
1	1	1	1	1	1	1

This result may indicate that as far as surface-level errors are concerned, the participant's writing skills did not improve adequately. However, since this was her first experience writing an English essay, she might have concentrated more on organization or idea generating than on surface-level errors. Or it may be said that the grading standard was too severe for the beginner writer. If the participant keeps on writing English for several years, her surface-level errors might decrease enough to satisfy the criterion in this research.

On the other hand, as **Table 3** indicates, the scores of surface-level errors in the "final" versions of the 7 essays were relatively good. Therefore, it is possible that the participant focused less on surface-level errors than on other criteria such as organization and syntax variety.

Table 3 Sores of surface-level errors in the final versions

1st	2nd	3rd	4th	5th	6th	7th
1	2	2	2	2	2	3

Thirdly, instead of examining the improvement of writing skills of the participant from the viewpoint of surface-level errors, let us examine how the scores regarding organization and coherence of the participant's essays changed along with this action research. **Table 4** and **Table 5** indicate the total scores of the 5 organization criteria and those of the 4 coherence criteria concerning the "final" version of the 7 essays.

Table 4 Total scores of 5 organization criteria
(Full Mark is 15) in the final versions

1st	2nd	3rd	4th	5th	6th	7th
12	10	8	14	12	14	15

Table 5 Total sores of 4 Coherence criteria
(Full Mark is 12) in the final versions

1st	2nd	3rd	4th	5th	6th	7th
9	7	10	12	12	11	10

Interestingly enough, these results clearly indicate that the participant's writing skills concerning organization and coherence improved to a great extent. Moreover, since these scores are high enough (indicating full marks 3 times), it can be said that the participant focused on the organization and the coherence of her essay rather than on the surface-level errors. In other words, the participant appears to realize that what is important in writing an English essay is not grammatical accuracy but how to convey her thoughts and ideas in an organized and coherent way. It goes without saying that grammatical and spelling accuracy is required in academic writing. However, the fact that in this research the participant's total scores improved on the whole (**Table 1**) while the score of surface-level errors did not improve at all (**Table 2**) evidently implies that criteria on students writings for evaluation and feedback should lay emphasis not only surface-level errors but also organization and coherence of the essay.

Finally, let us examine the scores regarding Topic Sentence, Style, and Syntax Variety in order to examine the validity of the criteria for these items. **Table 6** indicates the total

scores of these items of the “final” versions of the 7 essays.

Table 6 Total sores of 4 criteria (Full Mark is 12)
for Topic Sentence, Style, and Syntax
Variety in the final versions

1st	2nd	3rd	4th	5th	6th	7th
6	5	7	9	10	9	10

Like organization and coherence criteria, in these criteria the scores clearly indicate improvement in the writing skills of the participant. This also means that the participant appears to realize the importance of the Topic Sentence and Syntax Variety in writing an essay in English.

5. Discussion and Conclusion

In this action research, 15 criteria on students’ academic writings for evaluation and feedback by teachers were determined. As the results indicate, the grading standard of surface-level criteria might have been too strict for the beginner writers. In fact, taking into consideration the fact that Japanese college students in general are given no experience writing an English essay at the high school level, it would be better for teachers who teach academic writing at the university level to modify the standard so that it works in the practical setting.

Also, if the scoring standard of surface-level errors is too strict, it is feared that students may focus more on such errors than on other criteria. Insofar as one of the main objectives of teaching academic writing is to have the students acquire the skills for expressing their ideas and thoughts in an organized and coherent way with proper style, it is risky to lay too much emphasis on surface-level errors.

The 15 criteria seem to have worked well as a feedback device to the extent that the participant was able to improve her evaluation score by revising her own essay based on the criteria. Likewise, since the total scores of the 7 essays significantly improved, it can be said that on the whole the criteria worked well as an evaluation tool. In this research, however, the validity of each criterion was not statistically analyzed because this study employed a quantitative approach. Therefore, further research is necessary to identify the validity of each criterion with a greater participant pool.

References

- Casanave, C. P. (1993). Student Voices: The Insiders Speak Out on Journal Writing in C. P. Casanave (Ed.), *Journal Writing: Pedagogical Perspectives*. Keio University, SFC
- Cohen, A. D. & Cavalcanti, M. C. (1994). Feedback on Compositions: Teacher and Student Verbal Reports, in *Second Language Writing: Research Insights for the Classroom* (Edited by Kroll, B.). NY: Cambridge University press
- Horowitz, D. (1986). Process not product: Less than meet the eye. *TESOL Quarterly*, 20, 141-144.
- Lim, P.L. & Kurtin, M. (1992). *TOEFL Grammar Workbook*. NY: Arco Publishing
- Markline, J., Hawkins, R., & Isaacson, B. (1995). *Thinking on Paper: A writing process workbook with readings*. Florida: Harcourt Brace
- Oshima, A. & Hogue, A. (1991). *Writing Academic English*. 2nd ed. NY: Addison-Wesley
- Oshima, A. & Hogue, A. (1988). *Introduction to Academic Writing*. NY: Addison-Wesley
- Reid, J. M. & Lindstrom, M. (1985). *The Process of Paragraph Writing*. NJ: Prentice Hall Regents
- Rubin, J. & Thompson, I. (1994). *How to Be a More Successful Language Learner*. Boston: Heinle & Heinle Publishers
- Wada, H. (2005). *An attempt to Establish Feedback Criteria for Paragraph Writing: In the Perspective of Academic Writing*. Tokyo: Journal of Bunkyo Gakuin University Department of Foreign Languages and Bunkyo Gakuin College, Vol. 4
- Wada, H. (1997). *Process Approach in the Perspective of Academic Writing: An Action Research Paper*. NY: Master's thesis accepted by Teachers College, Columbia University
- Williams, M. & Burden, R. M. (1997). *Psychology for Language Teachers: A Social Constructivist Approach*. NY: Cambridge University Press
- Zinsser, W. (1990). *On Writing Well* (4th Ed.). NY: Harper Perennial

Appendix

Appendix: The questions from TWE that are employed in this research

- Q1. People attend college or university for many different reasons (for example, new experiences, career preparation, increased knowledge). Why do you think people attend college or university? Use specific reasons and examples to support your answer.
- Q2. Do you agree or disagree with the following statement? Parents are the best teachers. Use specific reasons and examples to support your answer.
- Q3. It has been said, "Not everything that is learned is contained in books". Compare and contrast knowledge gained from experience with knowledge gained from books. In your opinion, which source is more important? Why?
- Q4. A company has announced that it wishes to build a large factory near your community. Discuss the advantages and disadvantages of this new influence on your community. Do you support or oppose the factory? Explain your position.
- Q5. If you could change one important thing about your hometown, what would you change? Use reasons and specific examples to support your answer.
- Q6. How do movies or television influence people's behavior? Use reasons and specific examples

to support your answer.

Q7. Do you agree or disagree with the following statement? Television has destroyed communication among friends and family. Use specific reasons and examples to support your answer.