

On English Writing Assessment Under the Guidance of National Criteria

Li Yan

School of Foreign Languages, Yan'an University
Yan'an, Shaanxi, China

Abstract—After years of feasible analysis and amendments, the National Criteria of Teaching Quality for Undergraduate English Majors (henceforth NC) was released last March. As a guideline for English teaching, NC will play a significant role in China's Higher Education and greatly influence the reorientation of course design in English teaching. This paper argues that English writing course should accord with the NC and make modifications in all aspects. It summarizes the status quo of English writing assessment in China and analyzes the problems in present assessment practice, aiming to present some modifications for assessment in English writing course.

Keywords—English writing; Assessment; National Criteria

I. INTRODUCTION

With the launch of the National Criteria of Teaching Quality for Undergraduate English Majors (henceforth NC), universities and colleges in China are endeavoring to combine the latest requirements for English majors with their teaching objectives and planning. NC provides guidelines for fostering English majors in such a new era and aims to cultivate qualified professionals in every respect. Assessment plays a significant part in NC, which includes preface, objectives, proficiency, curriculum, teaching and assessment, staff and facilities. Peng (2016) further identifies assessment as one of the three functions (qualification, construction and assessment) of NC. Therefore, assessment is inevitable in consideration of NC and English teaching.

As one of the five basic skills (listening, speaking, reading, writing and translation) for English majors, writing showcases the output capability in language learning. It is never pointless to stress the importance of writing skill for an English major, for writing reflects what is your mind capable of, such as structuring, reasoning, comparison & contrast, narrating, describing, expositing, and arguing, etc. Hence, how to assess one's writing ability is a key issue for representing one's mind work. As for a teacher who teaches English writing, assessments are to some extent a type of feedback, which helps with improving, revising and modifying his/her teaching practice.

This paper is organized as follows: section 2 presents the status quo of English writing assessment in China via reviewing related literature and particularly evaluating the weakness in assessments. In section 3, we explore the ties between NC and assessment types and purposes. In section 4, we apply the designed writing assessment principles for

English majors in a provincial university, and we summarize the modifications considering the NC. Section 5 is the conclusion. More specifically, we claim that multiple assessments should be organically combined with NC in order to enhance English writing teaching in every respect.

II. STATUS QUO OF ENGLISH WRITING ASSESSMENT IN CHINA

A. Literature review

Traditional English writing assessments overemphasize summative evaluation. The performance of students attending an English writing course is largely determined by how much they score in their finals which range from 0 to 100. The operation of traditional assessment works largely with the following presumptions: the quality of writing can be defined and ascertained; different assessors can make consistent judgement over students' writing; students' writing ability can be decoded and represented by score; reliability among different assessors can be maintained in stability and standardization with spatial and contextual variation; the merits of an assessment are merely confined to clear objection and evaluation tools. This oversimplification of assessment ignores the characteristics of writing as a process rather than a stable state. Writing a complex work in which brainstorming, organization, diction, attitudes, cognition cwork at the same time, a single test could hardly reflect the real writing ability of a writer with certain spatial, time, topic and genre constrains.

English writing assessment has become a hot topic in scientific research since 1990s. Writing in United States linked to large-scale assessment ever since college entrance examinations evolving from oral tests of literacy abilities to written ones (Brereton 1995; Trachsel 1992), and it is still a component of entrance evaluations at most institutions of higher education. In China, English composition is frequently a required section in English written tests of different levels, through rising-junior exams, graduation tests, and other competency certifications. Self-assessment, peer assessment, teacher assessment, computer-assisted assessment, portfolio assessment, corpus assessment, as well as dynamic assessment are greatly discussed and argued over the past 30 years. Huot (2002) made a contrast between traditional writing assessments and new ones in terms of procedure, purpose and presumption independently. Butler & Lee (2006) made a

comparative study to show the power of two types of self-assessment. Davidson (2009) focused on some prominent issues in teacher assessment. Poehner (2009) studied the application of group dynamic assessment in L2 classroom. However, just as Delaware (2010:24) puts it “Perfect assessments and perfect assessors do not exist.” Each text has its own range of assessability, each assessment tool has its limitations and biases, and each assessor in the procedure is limited to his or her own understanding of the text and grading principles. Great achievements have been made throughout all these discussions and insights. However, there is neither perfect assessment theory nor perfect method. What’s more, no perfect assessment and assessor do not mean there is nothing we could do to optimize the teaching/learning performance.

Huot (2002) proposes an instructive assessment, in which the principles of the evaluation include: get students involved in the process of assessment, enhance their motivations for revising, make the students real writers, to improve one’s writing proficiency by way of assessment. He held that a common way for instructive assessment for students is portfolio assessment. Brown (2002) recommended distinguishing between norm-referenced tests and criterion-referenced tests so as to ensure that the best type of test can be used to make the best possible decision. Norm-referenced evaluations, which aim to disperse the performance of learners in a normal distribution, are most often adopted to diversify the relative performance of a broad spectrum of examinees. (Brown 2002:2) The normal distribution of scores is also known as the ‘bell-curve’. The criterion-referenced evaluations, on the other hand, do not pay much attention on the comparative performance of examinees as the norm-referenced evaluations do. To the opposite, the criterion-referenced tests aim to check whether the examinee is qualified to the pre-existing standards or not and clarify the gap between the performance of examinees and the pre-fixed criteria. Thus, norm-referenced tests show a relative rating results while criterion-referenced tests present a fixed one in terms of criteria. Finch (2004) argued that adopting peer-assessment can be an integral part in the student-centered learning process. Formative feedback among students was exchanged on each other’s sites via notice boards and the final assessment was submitted to their teacher. Finch’s study showed that most of the students have positive views towards the use of explicit peer assessment through cooperative interactions including organized instruction in the writing process.

B. Dilemmas in English writing assessment

English writing assessment studies have gained more and more attention from researchers across China, and assessment models and systems are developed so as to enhance the English writing proficiency as precise as possible. However, some existing problems concerning assessments inevitably cause low efficiency in practice. First, writing assessment itself includes a large extent of subjectiveness. A scientific design of writing evaluation is critical in practice, but the distribution between subjectiveness and objectiveness in assessment is not always easy to balance. As we all know,

writing assessment goes far more beyond a required assignment with time and word limits. The clear-cut writing and grading correlation or matching cannot reveal the fidelity of one’s real proficiency. Writing includes layout, capitalization, word division, diction, sentence pattern, structure, method of organization, cohesion and coherence, ideas, etc. Among all of the above factors, some can be tested by objective questions while others can only be evaluated in a subjective way. Multiple choice, blank filling, judgement of right or wrong can be adopted to objective question design, while topic sentence, cohesion and coherence, ideas can be tested by way of subjective question design. Therefore, the design of assessment should take the factors involved into consideration and well balance the percentage of subjective and objective questions. Second, different assessment measures are not integrated efficiently. A variety of measures in assessment are mushrooming in the past decades. Corrective assessment is employed frequently at the early stage of writing assessment, and it focuses on finding out the errors and then correct them one by one. With the development of information processing, computers software or app is adopted in assessment by automatic grading. Recent years witness the diversified and novel measurements such as self-assessment, peer assessment, group assessment, teacher assessment, as well as dynamic assessment. A design feature of these new treatments is that they implement the traditional teacher’s assessment and shift to multi-perspective assessments. However, distribution dilemma also arises among the multiple perspectives, because whether all measurements should be taken equally or some of them should dominant others actually turn into an argument. Third, principles and criteria for assessment are not clearly prefaced in present assessment. Some evaluations are given and graded completely intuitively by the tester, and some are designed according to the textbooks or lectures, and some others are conceived out of certain kind of test syllabus. Therefore, almost every measure is applied behind some guidelines, but the correlations between are far beyond recognition. Teaching orientations, objectives, contents, methods as well as assessments should work and develop within the framework of guidance.

Although different measures and perspectives are adopted in present assessment practice, the search for a relatively optimal way of writing test is always on the move. Due to the diversified purpose of various tests and the subjectiveness of writing test itself, proper criteria should be customized to individual writing test. Therefore, the aims of criteria-referenced tests are much more precise and approximate to the intention of giving a writing test. In some sense, classroom teaching operates with specific objectives and principles, hence criteria-referenced rating and grading could be a better replacement in assessment practice.

III. NATIONAL CRITERIA AND ENGLISH WRITING ASSESSMENT

Bachman (1990:279) holds that language assessment does not take place in “a values-free psychometric test tube”. So writing assessment design needs comprehensive considerations rather than sheer test papers.

A. A brief introduction to NC

Ministry of Education of China focused on drafting national criteria for undergraduate majors since 2013. As an important part of this nation-level move, national criteria for undergraduate English majors will inevitably influence the construction and development of English major in Chinese universities.

NC depicts guidelines for objectives, proficiency, curriculum, teaching and assessment, staff and facilities. The assessment requirements for undergraduate English majors are stated as follows:

The assessment aims to enhance learning. To provide prompt feedbacks to teaching, the contents and criteria of assessments should base on cultivation plans, scientific models and methods should be innovated, and reasonable application of assessment results should be employed. Formative and final assessment should be organically integrated.

The details of NC outlines 4 basics for assessment requirements. First, it clarifies the initiation and goal of assessments for English majors, that is learning; second, it presents multiple motivations for assessment; third, it emphasizes innovation in models and methods of assessment; fourth, it reminds of integration between two dominant assessments.

B. Purposes of writing assessment

A test paper cannot define one’s English writing proficiency. One of the most important things to consider before designing any assessment tool is its purpose. Bachman and Palmer (1996) hold that the overriding purposes of any language tests are to first make inferences about language ability, and secondly to make decisions based on those inferences. Generally speaking, there are three common types of inferences that a tester might choose from. First, achievement test, which tests the extent to which learning has taken place at the end of a unit or course of study. Second, diagnostic test which analyzes the particular strengths and weaknesses of a group of students at the beginning of a course so that the teacher will be able to adapt instruction to best meet learners’ needs. Third, proficiency test, which evaluates the general language abilities of students. The last type is often used for the purpose of deciding admission to a program or institution.

In order to ensure that the best type of test is used to make the best possible decision, Brown (2002) proposes a distinction between tests that are norm-referenced and those that are criterion-referenced. As for norm-referenced tests, they are used to compare the relative abilities of test takers so as to disperse the performances of testees in a normal

distribution. A wide representation of scores in normal distribution is known as the bell-curve. A norm-referenced test aims to include items that are capable of distinguishing among the relative abilities of a broad spectrum of examinees. The primary purpose of criterion-referenced tests on the other hand, is not to learn how an examinee’s performance compares with that of other test-takers, but to describe the amount that they know of a specific domain of knowledge or set of objectives. (Brown, 2002). As the name suggests therefore, in a criterion-referenced test, an examinee’s performance is referenced by a set of pre-existing external criteria rather than the performances of other test-takers.

C. Purposes of NC and English writing assessment

Taking NC into account, the purpose of English writing assessment lies in the context of English majors in universities or colleges. Undergraduates have already obtained qualifications for higher education, it is not necessary to disperse the performance of students in a bell curve. What is more important in university education is students’ proficiency in their own majors. Thus, a criteria-referenced test seems to be more appropriate for English majors. For this reason, NC directly qualify the base of assessment design.

Hamp-Lyons (1991b) concluded that the depth and breadth of our knowledge are not sufficient for us to be able to judge the appropriateness of any given writing assessment across the board (p. 326). Human languages diverse along different dimensions and an absolute judgement is not easy to give, however, relatively appropriate judgements can be concluded according to certain criteria. Since the purpose of this article is to discuss writing assessment in a classroom setting, the remainder of this article will focus only on the use of criterion-referenced achievement for the purpose of making inferences about students’ ability to perform specific, clearly defined writing tasks.

Influential writing tests and assessments in developed countries like ETS (Educational Testing Service) and CEEB (College Entrance Examination Board) are based on imperialist philosophy, which holds that writing is a fixed ability independent from environment, so psychometrical and statistical principles can also be applied to writing assessment. Within such a framework of philosophy, objectiveness is a key criterion for assessment, and fidelity is the eternal topic of assessment researches. However, complete objectiveness does not really exist in assessment, and the exploration for a feasible assessment theory is quite necessary. A theory aiming at enhancing teaching and learning should follow some principles. First, assessment should be designed according to situational requirements. Second, local institutions should be responsible for revising, renewing and fidelity verdicts. Third, the authenticity of writing assessment should be ensured by teaching objectives as well as social-cultural factors. Fourth, the importance of idea presentation should be emphasized in assignments and grading criteria. Five, students should be well aware of the entire process of assessment.

English writing assessments can be generally categorized into overall assessment and analytical assessment in China. In most occasion, overall assessment gives a general evaluation

by some specific expert and the results are graded as A, B, C or D. This evaluation well presents the relative writing proficiency and orientation within a certain range, but it fails to notify what is really wrong with the writing, whether it is a grammatical or cohesive or structural insufficiency. For the analytical way, it is so time-consuming and detailed that a balance among different parts with diversified criteria is hardly well struck, though the strengths lie in an excellent appraisal of performance in different aspect of writing requirements. Factors of assessment objectives, contents, techniques, criterion are taken into considerations, but the role of each factor and the distribution are not easy to balance. Multiple measurements are adopted in present-day English writing assessment, while most evaluation bias on a certain treatment, which leads to unreliability. However, with the guiding framework of NC, writing proficiency assessments will focus on the enhancement of students' learning, including triggering of students' learning motivations, diverse and scientific stimuli of language input, skills in enhancing language processing, and finally perfecting production in writing form.

IV. MODIFICATIONS OF ENGLISH WRITING ASSESSMENT

Modifications are carried out during our teaching practice taking into consideration to amplify students' writing ability in classroom settings. First, in order to enhance students' awareness in writing as way of thinking, creative thinking is taken into account for grading. Writing basics take up relatively small proportion in proficiency tests, for they are "musts" for an English writing course. Second, English teachers are required and trained to develop specific assessment methods for writing, various on-line evaluation, peer-evaluation or group-evaluation methods and assessment criteria for grading are introduced to allow teachers as well as learners on the one way to supervise others' work and on the other way to observe their own writing. Exchanging peer-evaluation helps students improve their writing skills, peers may negotiate their scores. Third, multimedia is a good way for conducting assessment considering the lack of time in the L2 classroom. Computer-mediated communication apps increase the need and effectiveness of peer feedback in the L2 writing environment. Hence when students exchange their work through Wechat or Tencent QQ or the bulletin board on the electric teaching platform, they can save time and stimulate their motivation for perfection. And what's more it is crucial for learners to receive efficient assessment using proper criteria on L2 writing in the classroom, which helps students develop their English writing proficiency .

V. SUMMARY

This paper attempts to showcase the influences of NC bringing about and how assessments should be conducted in the context of Chinese higher education. Some modifications has been carried on in terms of the inner core of NC, and detailed elaborations for assessment will be introduced in our forthcoming research papers.

ACKNOWLEDGEMENT

This article is sponsored by Teaching Reform Program of Yan'an University (2017-20) entitled Assessment Modification for English Writing Course under the Guidance of National Criteria; and by Innovation Program entitled Translation of History of Oxford English and Its Wechat Operation (S201910719023); and by Counseling and Cultivation Program of Yan'an University (YDZZYB17-063).

REFERENCES

- [1] Bachman, L.F. 1990. *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- [2] Bachman, L.F., & Palmer, A. 1996. *Language testing in practice*. Oxford: Oxford University Press.
- [3] Bachman, L.F. & Palmer, A. 2010. *Language assessment in practice*. Oxford: Oxford University Press.
- [4] Brown, J.D. 1995. *The elements of language curriculum*. Boston: Heinle & Heinle.
- [5] Brown, J.D., & Hudson, T. 2002. *Criterion-referenced Language Testing*. Cambridge: Cambridge University Press.
- [6] Butler, Y. G. 2009. How do teachers observe and evaluate elementary school students' foreign language performance? A case study from South Korea. *TESOL Quarterly*, 43(3), 417-444.
- [7] Hamp-Lyons, L. 1991. *Issues and directions in assessing second language writing in academic contexts, Assessing second language writing in academic contexts*, NJ: Ablex Publishing Corporation.
- [8] Huot Brian. 2002. *Rearticulating Writing Assessment for Teaching and Learning*, Utah: Utah University Press.
- [9] Peng, P. 2016. On relationships between National Criteria of Teaching Quality for Undergraduate English Majors and university criteria. *Foreign Language Teaching and Research*. Vol. 48(1).
- [10] Sword, H. 2016. *The Writer's Diet*. The University of Chicago Press.
- [11] Thompson, G., & Hunston, S. 2000. *Evaluation: An introduction*. In S. Hunston & G. Thompson (Eds), *Evaluation in text: Authorial stance and the construction of discourse*. Oxford: Oxford University Press.
- [12] Waring, H. Z. 2008. Using explicit positive assessment in the language classroom: IRF, feedback, and learning opportunities. *The Modern Language Journal*, 92(4), 577-594.