Literature Review of Student Self-assessment and Its Effects on the Second Language Writing

Bing Xu
Xiamen University Tan Kah Kee College
Zhangzhou, China

Abstract—**This paper is a detailed review of the theories and previous studies that investigate self-assessment and second language writing. After the introduction of the definition of self-assessment, a review of the validity and reliability of self-assessment will be presented, followed by its potential benefits to students, teachers and the conventions of testing. The previous studies on the effects of self-assessment on second language writing are reviewed at the end of this paper.**

Keywords—self-assessment; validity; reliability; second language writing

I. INTRODUCTION

A common view of the nature of language teaching and learning is that students should have some input in the complete learning cycle. Being part of the complete learning cycle implies being involved in the assessment process, since assessment is now recognized as a component in the learning process. Traditional strategies for assessing students' work exclusively rely on testing. However, a more balanced approach in which classroom tests and examinations are supplemented with alternate forms such as portfolio assessment, classroom observation, and self-assessment, which are called formative assessment, has interested many researchers.

II. DEFINITION OF SELF-ASSESSMENT

Self-assessment is defined as the involvement of learners in making judgments about their own learning, particularly about their achievements and the outcomes of their learning (Boud & Falchikov, 1989). The judgments they make may be about what they have done, what they should be doing or why they should be doing it, and what they are going to do next. In this definition, one factor is crucial, which is the identification of criteria or standards to be applied to one's work. Without appropriate criteria, self-assessment will not be objective or effective.

With regard to the self-assessment of writing, Schendel and O'Neill (1999) consider it as a means by which teachers help students to become more metacognitive about their writing and writing process and an attempt by which the negative impacts of grading or test writing can be avoided.

Despite the fact that assessment and evaluation are different terms and some researchers have tried to distinguish between self-assessment and self-evaluation (MacGregor, 1993), many other researchers do not suggest a systematic difference between self-assessment and self-evaluation. Therefore, in this paper, the two terms self-assessment and self-evaluation are regarded synonymously. Self-assessment refers specifically to student self-assessment rather than teacher self-assessment or self-assessment of institutions.

III. VALIDITY AND RELIABILITY OF SELF-ASSESSMENT

Two qualities that most affect the value of an assessment measure are validity and reliability.

A. Validity of Self-assessment

Validity refers to the extent to which the data collection procedure measures what it supposed to measure (Seliger & Shohamy, 1999; Genesee & Upshur, 2001). A number of different types of validity have been identified, each of which presents a different perspective on collecting and interpreting data. The most widely used types of validity are content validity, criterion-related validity, and construct validity. Among all these types of validity, there is one closely related to the issue of self-assessment, which is named concurrent validity, a subcategory of criterion-related validity. Concurrent validity is established when comparing one instrument with another procedure for the purpose of measuring the same skill. If the two procedures correlate with each other, that will provide evidence for the validity of the instrument (Seliger & Shohamy, 1999). Concurrent validity is usually demonstrated by a statistical method using correlation. The resulting correlation coefficient will provide evidence for the concurrent validity of the assessment.

A lot of research studies have explored the issue on the validity of self-assessment (Bachman & Palmer, 1989; Stefani, 1994; Patri, 2002; AlFally, 2004). Some studies found that self-assessment had a relatively high validity. For example, Oskarsson's (1980) study in Sweden indicated a correlation of 0.60 between student self-assessment and teacher assessment. Bailey (1998) described her study carried out at the University of California which involved forty-five international teaching assistants who were non-native English speakers. The participants were asked to complete a self-assessment form in which they evaluated their own use of English in terms of grammar, pronunciation, fluency, vocabulary, and overall ability. These categorical values were correlated with the trained raters' assessment of

each participant's English. The results indicated a correlation of 0.64 on grammar, 0.61 on pronunciation, 0.62 on fluency, 0.58 on vocabulary, and 0.63 on overall ability. Based on the results, Bailey concluded that there was a moderately strong correlation between the international teaching assistants' self-assessment and the evaluations conducted by the trained raters. LeBlanc and Painchaud (1985), working at the University of Ottawa in Canada, also calculated the correlation between student self-assessment of their language skills through a questionnaire and their scores on a standardized English proficiency test. This study obtained a total correlation of 0.53.

However, the validity of self-assessment can be influenced by many factors, such as the design of the self-assessment questionnaires, the training on self-assessment, student's language proficiency and cultural background, and personal traits of students, gender differences, etc (AlFally, 2004). Based on a review of sixteen self-assessment studies, Blanche and Merino (1989) identified five factors that could threaten the validity of self-assessment: (1) insufficient training of learners in how to perform the task of self-assessment; (2) a lack of criteria or standards suitable for both leaner self-assessment and teacher's interpretation; (3) a conflict between the cultural backgrounds of the learners and the culture on which the self-assessment is based; (4) learners' inabilities to reflect on their performance, and (5) the effects of personal subjective influence of the learners. The first three factors are interrelated because training will be crucial in establishing a shared criterion both to learners and teachers, and effective in dealing with the issues of different cultural backgrounds of the learners (Cohen, 2005).

B. Reliability of Self-assessment

Reliability provides information on the extent to which the data collection procedure is consistent and accurate (Seliger & Shohamy, 1999). It is concerned with freedom from nonsystematic fluctuation (Genesee & Upshur, 2001). Different types of reliability have been identified, depending on where the research suspects that inaccuracies in the data collection procedure could occur (Seliger & Shohamy, 1999). Generally, there are two kinds of reliability in second language research studies: rater reliability and instrument reliability. The defining characteristics of rater reliability are that scores by two or more raters or between one rater at time X and that same rater at time Y are consistent. Based on this defining characteristic, rater reliability can be further grouped into two types: interrater reliability which refers to the consistency in scores between /among two or more raters and intrarater reliability which considers the consistency in ratings produced by one rater at different times. Not only do researchers make sure that raters perform in a consistent way, they also need to ensure that the instrument they are using is reliable. There are three types of instrument reliability: test-retest reliability, parallel-forms reliability, and internal consistency reliability. Test-retest reliability occurs when the same test is given to the same (or similar) samples on two different occasions. Parallel-forms reliability will be used when two versions of the same test are administered to the same individuals to ensure that the two versions are really equal. In internal consistency reliability, researchers administer one single measurement instrument to the same group of individuals on one occasion. Statistical methods like Split-half, Kuder-Richardson 20 and 21, and Cronbach's alpha are commonly used to determine the internal consistency reliability (Mackey & Gass, 2005).

Reliability is indicated by the reliability coefficient, ranging from 0.00 to 1.00. The higher the coefficient, the more reliable the procedure is.

Many research studies have focused on the issue of reliability of self-assessment (Bachman & Palmer, 1989; Stefani, 1994; Ross, 1998; Orsmond, et al., 2000; Patri, 2002; AlFally, 2004). Some studies proved that self-assessment had a high degree of reliability. For instance, Bachman and Palmer reported high reliability coefficient ranging from 0.857 to 0.917 across individual scales with exception of grammatical ability. Therefore they concluded "the obtained reliabilities are much higher than had been expected" (1989: 22). AlFally (2004) had the similar findings in his study. He investigated the role of some psychological and personality traits of EFL students in the accuracy of self-assessment by means of a 14-statement questionnaire developed by Patri (2002). The reliability coefficient he calculated was 0.934.

IV. POTENTIAL BENEFITS OF SELF-ASSESSMENT

In recent years, there has been an increasing interest in the role of self-assessment in language learning and teaching (LeBlanc & Painchaud, 1985; Bachman & Palmer, 1989; Blanche & Merino, 1989; Oscarsson, 1989). Such interest is a logical outcome of the advantages of self-assessment. There is a general consensus in the literature that student self-assessment offers numerous advantages to students, teachers, and the conventions of testing.

A. Benefits of Self-assessment to Students

Self-assessment involves students to a much higher degree than other traditional assessing instruments. Being part of the complete learning cycle, students should be involved more in the assessment process (Leblanc & Painchaud, 1985). Being encouraged to assess their own work, students feel they have more responsibility for their learning and consequently make improvements more or less in their work.

LeBlanc and Painchaud (1985) investigated the use of self-assessment as a second language placement instrument. In this research project, students were selected randomly to fill out a self-assessment questionnaire. The questionnaire contained four parts covering the four basic skills: listening, reading, speaking, and writing. Each part consisted of 10 statements. Students were asked to read each of the 40 statements and to give themselves a score on a scale ranging from 1 to 5. The total scores on the self-assessment questionnaire were compared with those on the proficiency tests, suggesting a correlation of 0.53 between the two total scores. The researchers concluded that students could assess their own knowledge of the second language to some extent and self-assessment, as a placement instrument, could be

considered to be a very valuable tool in which students found themselves with more responsibility for their own learning.

Self-assessment helps students to specify their goals in different phases throughout the whole learning process. It has been demonstrated that being clear about goals makes a positive contribution to performance (Locke, Shaw, Saari & Latham, 1981). When students are asked to do the task of self-assessment, they, first of all, need to be very clear about what they are supposed to perform. Reexamining their performance may help to indicate whether or not their objectives in learning have been achieved. According to Falchikov and Boud (1989), self-assessment can be a valuable learning activity and can provide strong feedback to students about leaning standards.

Ross, Rolheiser, Hogaboam-Gray (1999) explored the effects of self-assessment training on narrative writing. In this quasi-experimental study, 148 students in grade four to six were selected as a control group, while another 148 students in similar grades with self-assessment training over an 8-week period were regarded as an experimental group. All subjects completed a battery of instruments at the beginning of the project in the following sequence: firstly, they completed a questionnaire, and then they wrote a short essay, finally they evaluated their short essays, shared their attitudes towards self-evaluation and completed a goal orientation survey. At the end of the project, they repeated the writing and self-evaluation tasks. The researchers found that students in the experimental group could evaluate their work more accurately because uncertainty about the criteria for a piece of writing was reduced.

Self-assessment can facilitate students' autonomous learning by reflective thinking. It is nowadays a popular practice for enhancing student autonomy in learning. Self-assessment equips the students with the skills that encompass critical awareness and reflectivity. Thinking reflectively is always associated with the outcomes of self-assessment (Freeman & Lewis, 1998; Boud, 2000). Ross et al. (1999) believe self-assessment is unique in asking students to reflect on their own performance. According to Baron (1981), thinking is the most essential expression of intelligence and by thinking well, a person may learn to manage his own capacity limits.

Dearnley and Meddings (2007) conducted a preliminary exploration of the impact of student self-assessment on learning. Six students participated in the study who assessed their work against the given criteria and assigned themselves a mark, and then filled out a self-assessment form. Five members of the teaching team did the same procedure as students did. The obtained data were analyzed in two stages: content analysis and constant comparative analysis. The outcomes of this study proved that students should be taught the rules of self-assessment and given sufficient self-assessment practice. The findings indicated that self-assessment made students think.

B. Benefits of Self-assessment to Teachers

Self-assessment, as a kind of formative assessment, not only benefits learners, but also offers great help to teachers.

Self-assessment can provide teachers with invaluable information about students which may not be obtainable otherwise. According to Ross et al. (1999), self-assessment offers rich details about learners to teachers. Self-evaluation instruments which elicit information about students' effort, goal orientations, strengths and weakness, and beliefs about their competence give teachers a full understanding of the reasons why students performed as they did. With the help of this information, teachers are able to anticipate learners' difficulties and impediments to learning.

Self-assessment saves teachers plenty of time and energy. Responding to students writing is one of the most challenging jobs writing instructors face and it is certainly the most time-consuming. Ferris (2007: 165) commented that the process of giving feedback to students' writing was "frustrating" and "filled me with anxiety". Similarly, when Guénette watched her students threw their corrected essays into the wastepaper basket before leaving the classroom, she asked herself "should teachers spend hours correcting their students' written productions?" (2007: 40) At least self-assessment may shift teachers' effort from marking to planning and moderating assessment activities (Boud & Falchikov, 1989).

Self-assessment increases dialogue between teachers and learners. Traditional means of assessment, such as examination papers are result-oriented, namely, both teachers and students focused on the final results of examinations. There is rarely mutual communication between teachers and students in exams. However, self-assessment provides the process that emphasizes mutual-communication between teachers and students. Denscombe and Robins asserted that "self-assessment facilitates communication between tutor and students, discussion about the course, and a general atmosphere in which there is a genuine exchange of opinion and it develops an atmosphere of openness and frankness" (1980: 68).

Self-assessment also helps teachers become more aware of their strengths and weaknesses, which in turn may assist teachers to improve their performance. By viewing student self-assessment, teachers may reflect on what has or has not worked well for students and try to work out to solve those problems. In this way teachers are helped to build a more positive outlook of their jobs and develop their abilities of "reflecting on, reviewing, evaluating, and revising their teaching practices" (Weiser, 2000: 140).

C. Benefits of Self-assessment to the Conventions of Testing

Self-assessment contributes to the traditional culture of testing in many ways. Compared with other standardized tests, self-assessment may offer several advantages. Self-assessment reduces time for testing. LeBlanc and Painchaud (1985) found in their studies that a 60-item self-assessment questionnaire took students about 20 minutes to complete, compared with the 100 minutes needed to complete the standardized tests.

The use of self-assessment also eliminates the need for proctors against cheating. In fact, students can do this kind of assessment at any place and at any time. It is no longer

necessary to establish strict testing schedule and to find proctors. It is believed that due to this advantage, self-assessment is especially efficient to an institution that registers numbers of language students (LeBlanc & Painchaud, 1985).

V. SELF-ASSESSMENT AND SECOND LANGUAGE WRITING

For most of the cases, students complete a writing assignment without later valuing it. They leave the total work of assessment to teachers, believing that it is the teacher's job to mark essays. Even if they get the feedback from the teacher, they glance at it quickly, paying attention to the grade they got, never taking into account of the reasons why they deserve such a grade. If it goes on like this, students may make a slow and insignificant progress in their writing. However, self-assessment may serve as an aid in dealing with the dilemma students face in that it familiarizes students with the features of a good piece of writing product and makes students reconsider their own work and be self-conscious about the parts that need improvements.

Self-assessment ensures that students are familiar with the basic requirements of writing. The quality of students' writing may be improved when students understand the details of the criteria against which their work will be judged. If students apply these criteria to assess their work, the effects will be stronger. Ross et al. (1998) found that students liked self-evaluation because it increased clarity about expectations and gave students sufficient feedback which they may use to improve the quality of their work. Students reported that with the help of self-assessment, they focused on what they needed to improve instead of the final grade they got because they had already been clear about the criteria for assessing a piece of writing. In the study carried out by the same researchers (Ross et al., 1999), they found that self-evaluation had a much larger impact on the performance of students who wrote poorly at the beginning of the study. The reason may be that self-evolution training give poorer writers explicit feedback on what they need to improve and on what is more meaningful to them than the feedback they usually receive from the teacher. Students pay more attention to self-evaluation because they understand the criteria, they feel ownership of the data, and they feel empowered because the teacher trusts them to rate themselves fairly.

Self-assessment makes students self-aware of their own work, motivating further revision. According to Miller, "taking a stand about the relation of a new event to one's prior authentic values completes that event by associating it with our past experiences" (1982: 182). Ferris (2007) pointed that self-evaluation was not just cop-outs for lazy or exhausted teachers since the mere act of rereading and rewriting one's own essay usually resulted in at least some improvements. Research studies (Fathman & Whalley, 1990; Ferris & Roberts, 2001) yielded the findings that even students who simply rewrote or self-edited with no input from anyone improved their end products at least somewhat. Weaker students will benefit more because any ambiguity about the criteria of writing can be avoided in the self-

assessment process. Moreover, the self-evaluation training provides poor writers with explicit feedback on what they need to improve, which is more meaningful to them than the feedback they usually received from the teacher, since students may make changes according to the criteria (Cohen & Cavaleanti, 1990). With this help of feedback, students add, delete, substitute and rearrange their work. Miller finally concluded that "self-evaluation experiencing—the quality of one's writing in relation to subjective standards—is crucial to the development of an individual's perception of writing as an important and natural way to investigate problems and represent ideas" (1982: 182).

VI. CONCLUSION

Although a considerable number of studies have offered valuable insights into the effects of student self-assessment, most of these studies have been carried out in English context. In such a context, native English students have less difficulty in assessing their own compositions. However, in China, the concept of self-assessment is viewed as inconsistent with the values of oriental culture. Littlewood (1999) suggests that East Asian students generally believed that teachers, as the symbols of authority, should be responsible for students' learning. Therefore, up till now, there are only a few studies concerning the effects of self-assessment on students with respect to the English speaking skills in China. Besides, few studies concerning the effects of self-assessment on students' writing have been carried out in Chinese context. Therefore, future studies intend to explore the possible effects of self-assessment on students' EFL writing and confirm the findings of previous studies conducted in English context.

REFERENCES

[1] AlFally, I., "The role of some selected psychological and personality traits of the rater in the accuracy of self- and peer- assessment," System, 32, 407-425, 2004.

[2] Bachman, L. F., Fundamental considerations in language testing. Oxford: Oxford University Press, 1990.

[3] Bachman, L. and Palmer, A., "The construct validation of self-ratings of communicative language ability," Language Testing, 6, 14-19, 1989.

[4] Bailey, K. M., Learning about language assessment: dilemmas, decisions and directions, Beijing: Foreign Language Teaching and Research Press, 1998

[5] Baron, J., "Reflective thinking as a goal of education," Intelligence, 5, 291-309, 1981

[6] Blanche, P., & Merino, B.J., "Self-assessment of foreign language skills: implications for teachers and researchers," Language Learning, 39 (3), 313-340, (1989).

[7] Boud, D., "Sustainable assessment: rethinking assessment for the learning society," Studies in Continuing Education, 22 (2), 151-167, 2000.

[8] Boud, D., & Falchikov, N., "Quantitative studies of student self-assessment in higher education: a critical analysis of findings," Higher Education, 18, 529-549, 1989.

[9] Cohen, A. D., Assessing language ability in the classroom. Beijing: Foreign Language Teaching and Research Press, 2005.

[10] Cohen, A. D., & Cavaleanti, M. C., Feedback on compositions: Teacher and student verbal reports. In B. Kroll (Ed.), Second language writing: Research insights for the classroom (pp.155-177). Cambridge: Cambridge University Press, 1990.

[11] Dearnley, A., & Meddings, F. S., "Student self-assessment and its impact on learning—a pilot study," Nurse Education Today, 27, 333-340, 2007.

[12] Denscombe, M., & Robins, L., "Self-assessment and essay writing," Teaching Sociology, 8 (1), 63-78, 1980.

[13] Falchikov, N., & Boud, D., "Student self-assessment in higher education: A meta-analysis," Review of Educational Research, 59 (4), 395-430, 1989.

[14] Fathman, A., & Whalley, E., Teacher response to student writing: Focusing on form versus content. In B. Kroll (Ed.), Second language writing: Research insights for the classroom (pp.178-190). Cambridge: Cambridge University Press, 1990.

[15] Ferris, D., "Preparing teachers to respond to student writing," Journal of Second Language Writing, 16 (3), 165-193, 2007.

[16] Ferris, D., & Roberts, B., "Error feedback in L2 writing classes: how explicit does it need to be?" Journal of Second Language Writing, 10, 161-184. (2001).

[17] Freeman, R., & Lewis, R., Planning and implementing assessment. London: Kogan Page, 1998.

[18] Genesee, F., & Upshur, J.A., Classroom-based evaluation in second language education. Beijing: Foreign Language Teaching and Research Press, 2001.

[19] Guénette, D., "Is feedback pedagogically correct? Research design issues in studies of feedback on writing," Journal of Second Language Writing, 16 (1), 40-53, 2007.

[20] LeBlanc, R., & Painchaud, G., "Self-assessment as a second language placement instrument," TESOL Quarterly, 19 (4), 673-687, 1985.

[21] Littlewood, W., "Defining and developing autonomy in East Asian context," Applied Linguistics, 20 (1), 71-94,1999.

[22] Locke, E., Shaw 69-1980. Psychological Bulletin, 90, 125-152. 1981.

[23] MacGregor, J., "Student self-evaluation: fostering reflective learning," New Directions for Teaching and Learning, 56, 35-46, 1993.

[24] Mackey, A., & Gass, S. M., Second language research methodology and design. Mahwah, New Jersey: Lawrence Erlbaum Associates, 2005.

[25] Miller, S., "How writers evaluate their own writing," College Composition and Communication, 33 (2), 176-183, 1982.

[26] Orsmond, P., Merry, S., & Reiling, K., "The use of student derived marking criteria in peer and self-assessment," Assessment and Evaluation in Higher Education, 25, 23-38, 2000.

[27] Oskarsson, M., Approaches to self-assessment in foreign language learning. Pergamon: Oxford, 1980.

[28] Oskarsson, M., "Self-assessment of language proficiency: Rationale and applications," Language Testing, 6, 1-13, 1989.

[29] Patri, M., "The influence of peer feedback on self- and per-assessment of oral skills," Language Testing, 19, 109-131, 2002.

[30] Ross, S., "Self-assessment in second language testing: A meta-analysis of experiential factors," Language Testing, 15, 1-20, 1998.

[31] Ross, J. A., Rolheiser, C., & Hogaboam-Gray, A., "Skills training versus action research in-service: Impact on student attitudes to self-evaluation," Teaching and Teacher Education, 14, 463-477, 1998.

[32] Ross, J. A., Rolheiser, C., & Hogaboam-Gray, A., "Effects of self-evaluation training on narrative writing," Assessing Writing, 6 (1), 107-132, 1999.

[33] Schendel, E., & O'Neill, P., "Exploring the theories and consequences of self-assessment through ethical inquiry," Assessing Writing, 6 (2), 199-227, 1999.

[34] Seliger, H. W., & Shohamy, E., Second language research methods. Shanghai: Shanghai Foreign Language Education Press, 1999.

[35] Stefani, L. A. J., " Peer, self, and tutor assessment: relative reliabilities," Studies in Higher Education, 19, 69-75,1994

[36] Weiser, I., Self-assessment, reflection, and the new teacher of writing. In J. Smith & K. Yancey (Eds.), Self-assessment and development in writing (pp.139-156). Cresskill, NJ: Hampton, 2000.

[37] Weiss, C. H., Evaluation research: Methods for assessing program effectiveness. Englewood Cliffs, NJ: Prentice-Hall, 1972.