

Self-Assessment of Business English Writing Ability of Thai University Students: Criterion-Related Validity and Gender Differences

Sutthirak Sapsirin

ABSTRACT

The use of self-assessment of language ability has received much attention in the field of second/foreign language learning and assessment. However, studies on its validity have yielded contradictory results. In addition, little research has been conducted into the differences between males and females in regards to self-assessment. Therefore, focusing on business English, this study aims to (1) investigate criterion-related validity of self-assessment by examining the relationship between self-assessment ratings of writing ability and writing achievement test performance, and (2) determine whether there are any differences in self-assessment ratings between males and females. The results of the research show that there is a weak, positive correlation between self-assessment ratings and achievement test scores. In addition, it reveals that both male and female students assessed themselves similarly although females performed significantly better than males on the achievement test.

Keywords: self-assessment, writing ability, criterion-related validity, gender differences

บทคัดย่อ

การประเมินตนเองด้านความสามารถทางภาษาได้รับความสนใจเป็นอย่างดีในการเรียนและการประเมินภาษาเป็นภาษาที่สองหรือภาษาต่างประเทศ อย่างไรก็ตาม ผลการศึกษาด้านความตรงของการประเมินตนเองมีผลที่ขัดแย้งกัน นอกจากนี้การศึกษาด้านความแตกต่างระหว่างเพศในการประเมินตนเองยังมีไม่มากนัก ดังนั้นงานวิจัยครั้งนี้ซึ่งมุ่งเน้นที่ภาษาอังกฤษธุรกิจมีวัตถุประสงค์เพื่อ (1) ศึกษาความตรงตามเกณฑ์สัมพัทธ์ของการประเมินตนเองโดยศึกษาความสัมพันธ์ระหว่างการประเมินตนเองด้านความสามารถในการเขียนภาษาอังกฤษธุรกิจกับความสามารถที่วัดโดยแบบทดสอบวัดผลสัมฤทธิ์ด้านการเขียน และ (2) ศึกษาความแตกต่างในการประเมินตนเองระหว่างนักศึกษาชายและหญิง ผลการศึกษพบว่า การประเมินตนเองของนักศึกษาสัมพันธ์กับคะแนนจากแบบทดสอบวัดผลสัมฤทธิ์ที่ระดับต่ำทางบวก และนักศึกษาชายและหญิงประเมินตนเองในระดับใกล้เคียงกัน แม้ว่านักศึกษาหญิงได้คะแนนจากแบบทดสอบวัดผลสัมฤทธิ์สูงกว่านักศึกษาชายอย่างมีนัยสำคัญ

คำสำคัญ: การประเมินตนเอง ความสามารถในการเขียน ความตรงตามเกณฑ์สัมพัทธ์ ความแตกต่างระหว่างเพศ

INTRODUCTION

Self-assessment can be defined as “procedures by which the learners themselves evaluate their language skills and knowledge” (Bailey, 1998, p. 227), and it has been a topic of interest in second/foreign language (SL/FL) learning for quite some time. According to Blanche and Merino (1989), the first studies of self-assessment in language learning were published in 1976, and since then, self-assessment has continued to develop as a distinct field within SL education and testing.

The use of self-assessment (or self-rating, self-appraisal) has continually gained support in regards to language learning. This is at least in part because the application of self-assessment can give learners opportunity to training in evaluating themselves, which can promote learning (Oscarson, 1989). Such involvement in the learning process can help students become more goal-oriented and more active in their language learning, increase their awareness of course content and individual progress, and foster autonomous learning (Brindley, 1989; Green, 2014; Harris, 1997; Oscarson, 1989; Rolfe, 1990).

In addition, belief in the use of self-assessment is deeply rooted in the principle that not only the teacher but also the learner should be involved in evaluation (LeBlanc & Painchaud, 1985; Nunan, 1988). That is, students should be part of the learning cycle which does not only include teaching and learning but also assessment (LeBlanc & Painchaud, 1985). To be more involved in language assessment, students need to be able to self-assess their performance, bringing meaningful input into the assessment process.

As a language testing tool, self-assessment can be beneficial in several respects. For example, using self-assessment instruments can expand the range of assessment in addition to allowing the students to share the assessment burden with the teacher (Oscarson, 1989). Self-assessment can also be useful in examining learner outcomes. In a study by Brantmeier, Vanderplank, and Strube (2012), a self-assessment questionnaire of listening, speaking, reading and writing was found to be a practical measure of Spanish language ability in terms of test administration and grading. Also, it can provide feedback and reduce student and instructor anxiety and affective barriers. In addition to being used in the classroom, self-assessment can also be used in placement testing of SL students. According to LeBlanc and Painchaud (1985), a self-assessment questionnaire of listening and reading abilities which is based on descriptors with six proficiency levels was able to place students just as well as the standardized proficiency tests that had been used. This allows student placement to be made

much simpler and more secure while promoting more student involvement than using a standardized test.

Several studies have investigated criterion-related validity of self-assessment of language ability by comparing them against teacher ratings and proficiency tests. These studies have found self-assessment of language ability to be both valid and accurate. For example, LeBlanc and Painchaud (1985) reported significant correlations between a self-assessment questionnaire and English proficiency tests. The correlations between each part of the self-assessment (Listening, Reading, Speaking and Writing) and the total test scores ranged from .39 to .53, and the correlation between the total scores of the self-assessment and the tests was .53. These results show that students possess the ability to accurately assess their SL ability to a significant degree. The literature review on self-evaluation of foreign language skills by Blanche and Merino (1989) showed that students' self-estimates can be classified as either generally good or very good. Several studies that compared self-assessment with more objective measures of proficiency found Pearson correlation coefficients from .50 to .60. This indicates that "a set of self-assessments (such as answers to a questionnaire) tends to carry about the same weight as any of the various parts (subtests) of a standardized testing instrument" (p. 324). Bachman and Palmer (1989) using factor analyses found that self-rating instruments of communicative language abilities can be both reliable and valid. Ross (1998) conducted a self-assessment meta-analysis of SL proficiency, and for 60 correlations obtained from 10 studies, he found an average correlation of .63 ($p < .0001$), with considerable variation among studies. He concluded that these results, which are in agreement with Blanche and Merino (1989), suggest that "self-assessment typically provides robust concurrent validity with criterion variables" (p. 16).

Other studies of correlation have yielded similar results. Milleret et al. (1991) found that the total self-assessment scores of grammar, communication and cultural knowledge correlated highly with results of oral interviews to assess Portuguese proficiency (.78) and the results of the standardized Portuguese Speaking Test (.75). AlFallay (2004) reported high correlations between self-ratings and teacher ratings of oral presentations of university students in Saudi Arabia. In addition, Brantmeier et al. (2012) reported that correlations between self-assessment ratings and standardized Spanish language ability tests of advanced learners were significant although, as noted by the researchers, the level of the correlation was not substantial (correlations between self-assessment and test of vocabulary, grammar, listening, reading and total scores were .332, .305, .274, .239, and .341, respectively). Ashton (2014) reported positive statistically significant correlations between the learner self-

assessments and tests of reading proficiency in non-European languages (ranging from .59 to .68), as well as between learner self-assessments and teacher assessments (ranging from .59 to .70).

Similar to studies of self-assessment of general language ability, studies of self-assessment of writing ability have shown positive and significant correlations between self-assessment instruments and writing tests or teacher marks. As mentioned previously, LeBlanc and Painchaud (1985) asked students studying French or English as a SL to respond to a “can-do” self-assessment questionnaire before taking a proficiency test. The results regarding writing ability showed that self-assessment of writing correlated significantly with the overall test score of general English proficiency tests ($r = .50$), indicating that students are able to fairly accurately judge their writing ability. Similarly, Ross (1998) conducted a meta-analysis that showed the average correlation between self-assessment of writing skill and the test criterion to be at the moderate level ($r = .53$).

There are, however, several studies that have raised doubts about the validity of self-assessments of writing ability. For example, Janssen-van Dieten (1989) compared two tests of Dutch as a SL: one was developed in a test format and the other was a self-assessment. The researcher found most of the correlation coefficients between self-assessment of writing and test criterion to be low (e.g. $r = .32, .41, .43$), bringing into question the value of self-assessments. Peirce, Swain and Hart (1993) compared self-assessment of writing of students in French immersion programs against a writing test in which they were asked to express their opinion based on a reading passage. Weak correlations between the two measures were found (correlations ranged from $-.01$ to $.25$). In another study, Saito and Fujita (2004) compared self-ratings with ratings of English as a foreign language (EFL) essay writing at one Japanese college. The analysis of self-ratings and teacher ratings showed that the self-ratings had almost no correlation with teacher ratings ($r = .07$). Matsuno (2009) also examined assessments of Japanese university EFL essay writing ability and found that students in this study tended to rate themselves overly harshly. As a result, the researcher concluded that self-assessment was somewhat idiosyncratic and may not be useful in formal assessment.

As we can see, one of the most frequently used types of self-assessment instrument in SL/FL research has been the questionnaire (e.g., AlFalley, 2004; Bachman & Palmer, 1989; de Saint Leger, 2009). In designing a self-assessment questionnaire, Bachman and Palmer (1989) suggested the use of questions that ask respondents to rate the difficulty of using a trait (e.g., “How hard is it for you to organize a speech or piece of writing in English with

several different ideas in it?”). They found this type of question to be more effective than others including the “can-do” type (e.g., “Can you organize a speech or piece of writing in English with several different ideas in it?”). This indicates that students are more aware of the specific types of difficulty they have than of their general abilities. In contrast, Heilenman (1990) cautioned against the use of difficulty items as there is evidence that wording effects can influence student responses to a self-assessment questionnaire. Some students may not pay close enough attention to the wording of the questions which can lead to errors in response. Thus, “can-do” items may be more appropriate in a self-report.

In developing a task in a self-assessment instrument, certain studies have determined that tasks that are more specific and focused can predict placement scores better than more general ones (Strong-Krause, 2000) and can have a stronger relationship to more traditional measurements (Peirce et al., 1993). In addition, it has been recommended that to gain predictive accuracy of criterion skills, self-assessment of language learning achievement should be designed based on specific curricular content (Ross, 1998).

Apart from studies on validity, question type and task, there is a body of literature relating to the effects of individual characteristics on self-assessment, for example, level of student proficiency (e.g., Alderson, 2006; Ashton, 2014; Brantmeier et al., 2012), native language background (e.g., Strong-Krause, 2000), and language anxiety (e.g., MacIntyre, Noels & Clément, 1997). Among the range of variables, however, one that warrants further exploration is gender. As Falchikov (2003) pointed out, “gender effects occur in a wide variety of social and academic situations and there is no reason to exclude them as a possibility in the context of self... assessment” (p. 107). Also, a review of the literature reveals that previous studies which included gender as a factor in relation to SL/FL self-assessment did not explore the issue in any comprehensive fashion. One study which did address this area is Strong-Krause (2000) who investigated the effectiveness of self-assessment in ESL placement. The participants, 81 students, took a placement exam before or after completing self-assessment questionnaires regarding listening, speaking, reading, and writing ability. Multivariate analyses of variance show that there was no statistically significant difference between how males and females rated their ESL ability and how they performed on their placement exam. The study concluded that gender did not have effect on self-assessment ratings.

As we can see, research on the validity of self-assessment of SL/FL writing ability has yielded contradictory results and studies on gender difference have not been extensively conducted. In addition, to date, most studies have focused on general or academic writing,

and few have investigated writing ability in regards to business communication—a particularly important skill for those seeking success in today’s increasingly globalized business world. It would, therefore, be of benefit to further explore criterion-related validity of self-assessment by investigating how self-assessment ratings of business English writing ability correlate with achievement test performance and whether gender difference exists in self-assessment ratings of such ability. The present study therefore addresses the following questions:

1. What is the relationship between self-assessment ratings of business English writing ability and writing achievement test performance?
2. Is there any difference in self-assessment ratings between male and female students?

METHODOLOGY

Participants

Participants were 416 students in the second semester of their second year of studies in a Thai program at the Faculty of Commerce and Accountancy in one of the largest universities in Thailand in academic year 2013. Of these students, 317 (76.2%) were females and 99 (23.8%) were males. Although the number of students in the two gender groups was not equal, it represents the proportion of students in the faculty. The range of age was from 18-22, with the mean being 19.77 ($SD = .616$).

Instruments

There were two instruments used in this study: the final exam of the Business English Correspondence course and a self-assessment questionnaire of business English correspondence ability. The Business English Correspondence course is a required course for second-year business students, and the final exam consists of four main parts based on the content of the course book¹ covered in the second half of the semester. The test provided prompts which cover various business situations where students need to write an email or informal report (see table 1). There were seven raters who had been teaching the course and marking its achievement tests for at least five years. The raters used a holistic rating scale for each part, except the cloze test for which they agreed on acceptable answers for each blank.

¹ *Writing for the Real World 2: An Introduction to Business Writing* (Bernard, R., & Meehan, A., 2005. Oxford: Oxford University Press)

For the parts which involved two raters, the scores were averaged and added to the scores for the other parts to arrive at the total score (40 points).

Table 1: Final exam format, number of raters and inter-rater reliability

Part	Num ber of raters	Rater	Inter-rater reliability
1. A reminder email (8 points)	1	Rater 1	NA
2. Complaint	1		
2A: A complaint email in a cloze format (3 points)		Rater 2	NA
2B: An email response to the complaint (7 points)		Rater 2	NA
3. Memo & Proposal	2		
3A: A memo in a cloze format (3 points)		Raters 3, 4	NA
3B: A proposal email related to the memo (7 points)		Raters 3, 4	.913*
4. An informal report (12 points)	3		
		Raters 5, 6	.926*
		Raters 5, 7	.942*
		Raters 6, 7	.909*

* significant at $p < .05$

The self-assessment questionnaire consisted of questions about student background e.g., age, major, and gender, as well as 50 items asking students to judge their ability in email and informal business report writing. The items were based on the language functions taught in the lessons. The items appeared in the following format (examples from the section on writing a response to a complaint email):

I can thank the customer who wrote a complaint.

I can refer to the problem and apologize.

I can give an explanation for the problem.

Students were asked to respond to the following items using responses based on a Likert-type five-point scale:

1 (“I cannot do this at all”);

- 2 (“I can do this a little.”);
- 3 (“I can do this fairly well.”);
- 4 (“I can do this well.”);
- 5 (“I can do this very well.”).

The score of each item was summed, resulting in total scores between 50 and 250.

After the questionnaire was developed in English and Thai by the researcher, it was given to two course instructors who have had extensive experience in applied linguistics research to validate. The researcher discussed the questionnaire with the instructors after receiving their feedback and revised the questionnaire accordingly.

Procedure

Data used in the analyses were collected from students’ self-assessments and final exam scores. The self-assessment questionnaires were distributed to 16 classes. Students completed the consent form and Thai version of questionnaire in the last session of class before the final exam. They were told that their responses to the questionnaire would be kept confidential and would not affect their grades. Students took approximately 15 minutes to complete the questionnaire. Of 452 questionnaires distributed, there were 416 with complete data which were used for the analyses (Cronbach’s alpha = .976).

Analysis

An exploratory data analysis was conducted to determine if the self-assessment score distribution and final exam score distribution were normally distributed. Results for the Kolmogorov-Smirnov test for normality (Lowie & Seton, 2013) indicated that both distributions deviated significantly from a normal distribution ($D = .076, p < .001$ and $D = .068, p < .001$, respectively). Therefore, non-parametric statistics were used; that is, Spearman’s rank correlation was performed to address the first research question, and the Mann-Whitney U test for the second.

FINDINGS

RQ1 What is the relationship between self-assessment ratings of business English writing ability and writing achievement test performance?

The student achievement scores in the final examination ranged from 16.51 to 38.13 ($Mdn = 32.13$). As for self-assessment ratings, the total scores ranged from 64 to 240 ($Mdn =$

165). The Spearman's rho revealed a weak, positive correlation between the self-assessment ratings and final exam scores ($r_s = .29$, $n = 416$, $p < 0.001$; see figure 1). Squaring the correlation coefficients indicated that 8.41% of the variance in the self-assessment ratings was explained by the final exam scores.

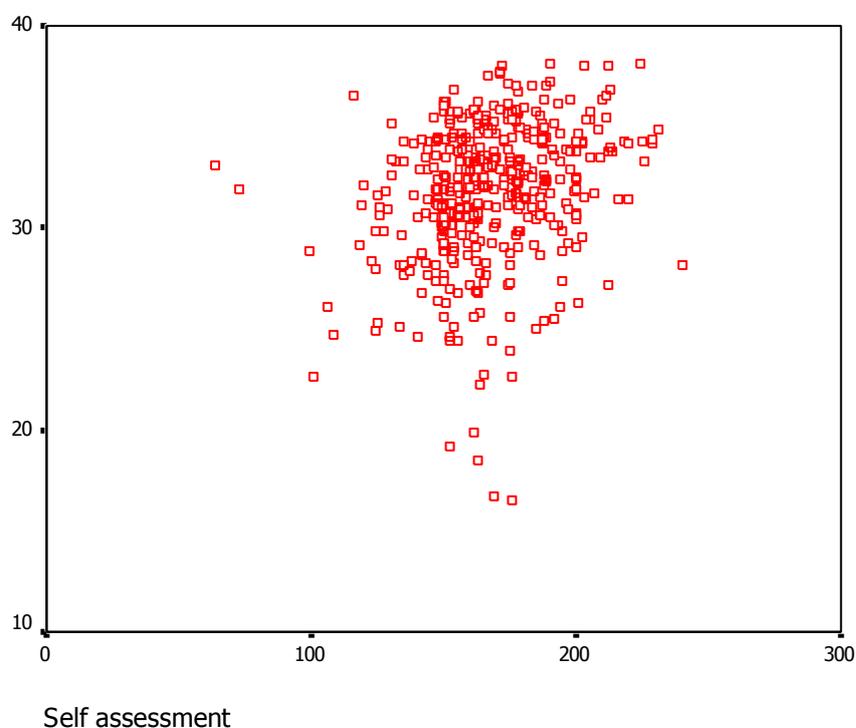


Figure 1. Relationship between final exam scores and self-assessment.

RQ 2 Is there any difference in self-assessment ratings between male and female students?

Before doing an analysis to see whether gender might affect self-assessment ratings, it was necessary to determine whether male and female students performed differently on the final exam. Descriptive statistics showed that female students ($Mdn = 32.75$; mean rank = 228.65) scored higher on the final exam than male students ($Mdn = 30.13$; mean rank = 143.99). Significant differences between groups were found ($U = 9305$, $p < .001$, $r = .30$) and the difference was at the medium level (Cohen, 1988).

Then, the Mann-Whitney test was performed. Table 2 shows that no significant gender differences in self-assessment ratings appeared ($U = 15204$, $p = .641$).

Table 2: Mean ranking in self-assessment ratings of female and male students

	<i>M</i> rank	<i>U</i>	<i>p</i> -value
Group 1 (Female)	210.04	15204	.641
Group 2 (Male)	203.58		

The results from the two analyses of gender difference in the final exam and self-assessments indicated that although females outperformed males on the final exam, there was no significant difference in the way males and females self-assessed their writing ability.

DISCUSSION

This study investigated the criterion-related validity of self-assessment by examining the relationship between self-assessment ratings of business English writing ability and writing achievement test performance as well as the influence of gender on self-assessment ratings. As for the first research question, the findings show that there was a weak, positive correlation between the self-assessment ratings and final exam scores. This indicates that self-assessment is a weak indicator of business English writing ability as tested in the final exam. The correlation being lower than .50 may be interpreted as indicating that the students are not able to accurately assess their own language ability (Oskarsson, 1978 as cited in LeBlanc & Painchaud, 1985).

The findings of this study are consistent with those of Janssen-van Dieten (1989), Peirce et al. (1993) and Saito and Fujita (2004). However, there are several factors which should be taken into consideration in regards to these findings. First, it is possible that students in this study lacked experience in judging their business English writing ability. According to Falchikov and Boud (1989), students in introductory programs may not be able to provide as accurate a self-assessment as those who are in the later years of the programs since the former are less experienced. The business students in the present study were in their second year of the program; therefore, the English courses they had taken previously in their program were two general English courses in the first year and only one course in basic business English writing as required in the previous semester. Also, most students had not had any business-related work experience in which they were required to write different types of emails and reports, which may help explain why they did not have an accurate picture of their business English writing ability.

Another factor influencing results could be the way students interpreted the self-assessment questionnaire. Although both instruments were based on specific curricular content as taught in the course (Ross, 1998), in order to perform well on the final exam, students needed to have a sufficient understanding of the situations described in the prompt. It is possible that when assessing their own ability, students may have relied on their abilities from doing exercises in the past but had not fully considered how to apply their knowledge in

a different situation. Thus, it is possible that if the questionnaire had emphasized that students should assess their writing abilities in light of the different context, the results may have been different.

The findings on the criterion-related validity of the present study's self-assessment questionnaire are contradictory to the findings of a number of previous studies (e.g., LeBlanc & Painchaud, 1985; Ross, 1998). This suggests that more research is needed to explore the usefulness of self-assessment of business English writing ability in the context of language assessment. As the present study employed only quantitative data, it is also recommended that future studies incorporate both qualitative and quantitative data such as interviews to obtain information on the processes involved in students assessing their ability. This may help provide an explanation for the low correlation obtained in the study.

In regards to the second research question, the findings reveal that both male and female students assessed themselves similarly although females performed significantly better than males on the final exam. The results seem to contradict those of Strong-Krause (2000), who found that male and female students were not significantly different in either their test performance or self-assessment ratings, indicating no gender effect.

It should be noted that the results regarding gender difference may be partially explained by the low ability of students in judging their own language ability as reflected in the weak correlation presented above. Also, the results indicate that students' self-assessment ratings did not necessarily reflect actual differences in writing ability. Furthermore, it was not possible for the present study to pinpoint whether male students were overestimating or females were underestimating their ability. Future research is required to investigate this issue. It has been noted that by asking both students and examiners to use similar rating scales to judge the quality of students' writing and providing some training to do so, it is possible to find out to what extent gender bias plays a role in self-assessment (cf. Blackwood, 2013; Langan et al., 2008; Matsuno, 2009; Pallier, 2003).

Another area that is worth investigating is the influence of gender in relation to other variables. As pointed out by AlFallay (2004), self-assessment is influenced by a number of psychological and personality traits. These include proficiency levels (e.g., Ashton, 2014; Brantmeier et al., 2012), anxiety (e.g., AlFallay, 2004; MacIntyre et al., 1997), attitude/motivation (AlFallay, 2004) and so on. It remains unclear whether or to what extent these factors interact with gender and what effect the interplay would have on self-assessment ratings of one's language ability.

In regards to the generalizability of the results reported, *caution must be* exercised due to the following aspects. First, the participants in the present study were Thai undergraduate business students in an EFL context. The data were collected by a self-report questionnaire and an achievement test. In addition, both of the instruments aimed to measure business English correspondence writing ability. Therefore, generalizability of the results should be limited to comparable participants and materials.

CONCLUSION

The present study shows that the use of self-assessment in regards to business English writing ability may not be as valid as indicated in other similar studies. Students seem to have difficulty in accurately assessing themselves and factors in addition to the possibility that gender may play a role in the self-assessment process. However, as argued by a number of researchers, the ability to self-assess can play an important role in language learning (Brindley, 1989; Green, 2014; Harris, 1997; Oscarson, 1989; Rolfe, 1990). The fact that self-assessment may involve subjective estimates “does not necessarily invalidate the practice of self-assessment techniques in language testing and evaluation and... self-assessment may be motivated by reasons that go beyond mere evaluation” (Oscarson, 1989, p. 2). In other words, the main benefit of self-assessment may lie in improving the learning process (Janssen-van Dieten, 1989). As stated earlier, further work on the development and factors involving self-assessment is needed. It is clear that such research may well lead to a better understanding of how to maximize the potential of self-assessment for the benefit of language learners.

REFERENCES

- Alderson, J. C. (2006). *Diagnosing foreign language proficiency: The interface between learning and assessment*. London: Continuum.
- AlFallay, I. (2004). The role of some selected psychological and personality traits of the rater in the accuracy of self-and peer-assessment. *System*, 32(3), 407-425.
- Ashton, K. (2014). Using self-assessment to compare learners' reading proficiency in a multilingual assessment framework. *System*, 42, 105–119.
- Bachman, L. F., & Palmer, A. S. (1989). The construct validation of self-ratings of communicative language ability. *Language Testing*, 6(1), 14-29.
- Bailey K. M. (1998). *Learning about language assessment*. Cambridge, MA: Heinle & Heinle.
- Bernard, R., & Meehan, A. (2005). *Writing for the real world 2: An introduction to business writing*. Oxford: Oxford University Press.
- Blackwood, T. (2013). Business undergraduates' knowledge monitoring accuracy: How much do they know about how much they know? *Teaching in Higher Education*, 18(1), 65-77.
- Blanche, P., & Merino, B.J. (1989). Self-assessment of foreign-language skills: Implications for teachers and researchers. *Language Learning*, 39(3), 313-340.
- Brantmeier, C., Vanderplank, R., & Strube, M. (2012). What about me? Individual self-assessment by skill and level of language instruction. *System*, 40(1), 144-160.
- Brindley, G. P. (1989). *Assessing achievement in the learner-centred curriculum*. Sydney: National Centre for English Language Teaching and Research, Macquarie University.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum.
- De Saint Leger, D. (2009) Self-assessment of speaking skills and participation in a foreign language class. *Foreign Language Annals*, 42(1), 158-78.
- Falchikov, N. (2003). Involving students in assessment. *Psychology Learning and Teaching*, 3(2), 102-108.
- Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, 59(4), 395-430.
- Green, A. (2014). *Exploring language assessment and testing: Language in action*. New York: Routledge.

- Harris, M. (1997). Self-assessment of language learning in formal settings. *ELT Journal*, 51(1), 12-20.
- Heilenman, L.K. (1990). Self-assessment of second language ability: The role of response effects. *Language Testing*, 7(2), 174-201.
- Janssen-van Dieten, A. (1989). The development of a test of Dutch as a second language: The validity of self-assessment by inexperienced subjects. *Language Testing*, 6(1), 30-46.
- Langan, A. M., Shuker, D. M., Cullen, W. R., Penney, D., Preziosi, R. F., & Wheeler, C. P. (2008). Relationships between student characteristics and self, peer and tutor evaluations of oral presentations. *Assessment & Evaluation in Higher Education*, 33(2), 179-190.
- LeBlanc, R., & Painchaud, G. (1985). Self-assessment as a second language placement instrument. *TESOL Quarterly*, 19(4), 673-687.
- Lowie W., & Seton B. (2013). *Essential statistics for applied linguistics*. Basingstoke: Palgrave Macmillan.
- MacIntyre, P. D., Noels, K. A., & Clément, R. (1997). Biases in self - ratings of second language proficiency: The role of language anxiety. *Language Learning*, 47(2), 265-287.
- Matsuno, S. (2009). Self-, peer-, and teacher-assessments in Japanese university EFL writing classrooms. *Language Testing*, 26(1), 75-100.
- Milleret M., Stansfield, C., & Mann-Kenyon, D. (1991). The validity of the Portuguese speaking test for use in a summer study abroad program. *Hispania*, 74(3), 778-87.
- Nunan, D. (1988). *The learner-centred curriculum: A study in second language teaching*. Cambridge: Cambridge University Press.
- Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and applications. *Language Testing*, 6(1), 1-13.
- Oskarsson, M. (1978). *Approaches to self-assessment in foreign language learning*. Strasbourg: Council of Europe.
- Pallier, G. (2003). Gender differences in the self-assessment of accuracy on cognitive tasks. *Sex Roles*, 48(5-6), 265-276.
- Peirce, B. N., Swain, M., & Hart, D. (1993). Self-assessment, French immersion, and locus of control. *Applied Linguistics*, 14(1), 25-42.

- Rolfe, T. (1990). Self- and peer assessment in the ESL curriculum. In G. Brindley (Ed.), *The second language curriculum in action* (pp. 163-86). Sydney: National Centre for English Language Teaching and Research, Macquarie University.
- Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing*, 15(1), 1-20.
- Saito, H., & Fujita, T. (2004). Characteristics and user acceptance of peer rating in EFL writing classrooms. *Language Teaching Research*, 8(1), 31-54.
- Strong-Krause, D. (2000). Exploring the effectiveness of self-assessment strategies in ESL placement. In G. V. Ekbatani & H. D. Pierson (Eds.), *Learner-directed assessment in ESL* (pp. 49-73). Mahwah, New Jersey: Lawrence Erlbaum.