Student self-assessment in higher education: Alone or plus?

Paper accepted for presentation at the CPLA Conference, October 29-30,

2010, Lebanese American University, Lebanon

Abdel Salam Abdel Khalek El-Koumy
Professor of Curricula and Instruction of English as a Foreign Language at Suez Canal University (Egypt) and Amman Arab University for Graduate Studies (Jordan)

Abstract

Higher education in Egypt has recently shifted towards the constructivist theory that involves students in the learning process in all of its aspects. Therefore, the need for self-assessment has grown to engage learners in constructing their own knowledge and developing their thinking skills to succeed academically and in life. In response to this need, the researcher conducted two experimental studies with English majors at Suez Canal University. The first study took place during the 2000/2001 academic year and examined the effect of student self-assessment versus no assessment on English-major seniors' knowledge achievement and academic thinking. During the 2004/2005 academic year, he conducted the second study which examined the effect of self-plus-teacher assessment versus self-assessment on the same dependent variables that were investigated in the previous study with other English-major seniors at the same school. The results of the two studies provided evidence that statistically significant improvement in knowledge achievement and academic thinking can occur only when the teacher assesses students self-assessments. This paper is a report on these two studies. The researcher started the paper by defining student self-assessment and explaining the theory behind it. Then, he showed the advantages and shortcomings of this form of assessment. Next, he surveyed the research related to it. After that, he displayed the findings of the two studies and explained the reasons behind these findings. Finally, he drew conclusions and offered recommendations based on the results of the two studies.

What is student self-assessment?

Student self-assessment is a form of authentic assessment in which each student reflects on her/his strengths and weaknesses in order to identify learning needs and reinforce weaknesses with the aim of improving achievement and/or performance (Fitzpatrick, 2006). This form of assessment can be used for formative and summative purposes. However, in the two studies conducted by the researcher, it was only used for formative purposes because most researchers agree that it can be useful for formative purposes and disagree about its role for summative purposes.

Theoretical background to student self-assessment

Student self-assessment is compatible with the constructivist theory which advocates the full engagement of learners in the construction of their own knowledge. In order to construct new sound knowledge, learners must assess this knowledge to fill gaps in it and to make sure of connections between its parts. In the absence of these revisiting processes, there can be no real construction of knowledge. Self-assessment can therefore be regarded as vital for the building of sound knowledge to occur.

Advantages of student self-assessment

Given the appropriate conditions and supports, student self assessment can have the following advantages:

(1) It enhances students awareness of their own learning and thinking processes (Zohar, 2004).

(2) It provides direction for future learning (Chamot and O'Malley, 1994).

(3) It makes teachers aware of their students' needs and provides them with an additional lens through which to view their achievement (Blanche, 1988).

(4) It helps learners see gaps in their own learning and initiate self-repair to redirect their learning toward the learning goal (Liang, 2006).

(5) It stimulates learners to consider course content critically and helps them achieve a high level of academic thinking skills (Teh, 2006).

(6) It promotes learners' autonomy and brings a sense of responsibility and accountability to them (Asadoorian and Batty, 2005).

(7) It increases learners' knowledge of their learning goals and thus enhances their motivation and goal orientation (Liang, 2006).

(8) It stimulates learners to exercise a variety of learning strategies and higher order thinking skills (Chamot and O'Malley, 1994).

(9) It develops appropriate study skills and strategies and fosters lifelong learning (Zohar, 2004).

(10) It helps learners get better marks (Harris, 1997).

Shortcomings of student self-assessment

The shortcomings of student self-assessment are mentioned in the literature as follows:

(1) Students are unaware of what they do not know (Dunning, Johnson, Ehrlinger and Kruger, 2003).

(2) Students tend to view themselves in desirable ways (Shrauger and Osberg, 1981).

(3) Many students believe that the duty of assessment is the responsibility of the teacher (Falchikov, 2003).

(4) There are doubts about the reliability of this form of assessment in formal education (Dickinson, 1987).

(5) There are doubts about the confidentiality of the learners if self-assessment is used for high-stakes decisions (Dickinson, 1987).

(6) Instructors are reluctant to lose their control over assessment (Blue, 1988).

(7) There may be a mismatch between the goals of assessment as conceived by the learner and the teacher (Oscarson, 1997).

(8) Students feel a sense of humiliation in reporting their own weaknesses (Liang, 2006).

(9) Some students perceive self-assessment as irrelevant and inappropriate (Liang, 2006).

(10) There is a confusion over the diverse functions of self-assessment (e.g., certification, diagnosis, placement, feedback, etc.) (Liang, 2006).

Research on self-assessment

Most of the research on self-assessment focused on its validity. The findings of the studies investigating the degree of validity of self-assessment (e.g., Frye, Richards, Bradley and Philp, 1991; Arthur, 1995; Zoller, Fastow, Lubezky and

Tsaparlis, 1999) showed that self-assessments have low to moderate validity compared to expert raters. The factors that were found to affect the accuracy of self-assessment include: tendency toward self enhancement, inability in the area being assessed and lack of experience in self-assessment.

With respect to tendency toward self-enhancement, many studies (e.g., Risucci, Tortolai and Ward, 1989; Kruger and Dunning, 1999; Mattheos, Nattestad, Falk-Nilsson and Attstrom, 2004) showed that most students tended to overestimate their own abilities in relation to external raters. In contrast, a minority of studies (e.g., MacIntyre, Noels and Clément, 1997; Chur-Hansen, 2000; Lind et al., 2002; Rees, 2003; Bryan et al., 2005**)** found that some students–particularly females, highly competent males and anxious students–tended to underestimate their own abilities**.**

With respect to students' inability in the area being assessed, some studies revealed that students did not have the full range of competence in the area being assessed, and therefore, they evaluated themselves on different dimensions of performance and used different benchmarks for quality than did their teachers. Woolliscroft, TenHaken, Smith and Calhoun (1993), for example, studied the

ability of second-year medical students to evaluate their physical examination skills as compared to expert faculty and concluded that students were not able to accurately assess their own performance, and added that as students skill level in the area being assessed increased, their self-assessment accuracy also increased. Similarly, Kirby and Downs (2007) found that college science students were not able to accurately self-assess in relation to the standards set by their teachers. They recommended that practice and teacher feedback can contribute to accurate self-assessment. Likewise, in their meta-analysis of seventy-seven studies about self-assessment covering under- and post-graduate students in clinical practice, Colthart et al. (2008) reached the conclusion that the least competent are also the least able to self-assess accurately. More recently, Mort and Hansen (2010) found that first-year pharmacy students had difficulties performing accurate self-assessment of counseling skills with the largest percentage of students overestimating their performance. They also found that students in the lower quartiles tended to overestimate while those in the upper quartile underestimated their skills. They recommended that teachers should provide students with frequent self-assessment opportunities along with feedback on skill development and self-assessment accuracy.

With respect to students' experience in self-assessment, a number of researchers found that experience is an important factor of self-assessment accuracy. Mabe and West's (1982) meta-analysis of self-assessment research, for example, concluded that self-assessments became more accurate as subjects gained experience in evaluating their abilities. Gordon (1991), for another example, reached a similar conclusion in his review of the validity of self-assessment and recommended that explicit feedback on self-assessment is vital to its development.

The findings of the previous studies on self-assessment in various content areas support the assertion that self-assessment alone, without experience and teacher feedback, is generally inaccurate. In accordance with this, in their review of the literature on self-assessment accuracy and utility, Breidert and Fite (2009) state:

> Self-assessment, as currently used, is generally inaccurate. Although some research has found self-assessment to be an efficient measure of one's ability or performance, the majority of research typically found it to be an under- or overestimation of one's actual performance. However, with appropriate consideration of the

> moderating variables and clarification of terminology, self-assessment could be accurately utilized. (p. iv)

Based on the results of the studies reviewed earlier, one can conclude that (1) there is only a moderate correlation between self and teacher assessment, and (2) students are not able to identify their own strengths and weaknesses as compared to their teachers. Therefore, some assessment specialists (e.g., Angelo and Cross, 1993; Hinsz and Matz, 1997; Brew, 1999) propose that instructors and researchers should go beyond the accuracy of self-assessment and consider what will make self-assessment most effective in identifying students' needs and enhancing learning. These educators believe that self-assessment validity is not important if self-assessment is used for formative purposes. Unfortunately, there is little published research on the contribution of self-assessment to achievement in various content areas and development of academic thinking skills (El-Koumy, 2001; MacDonald and Boud, 2003). Therefore, the two studies conducted by the researcher aimed to investigate the effects of self-assessment, compared to no-assessment and self-plus-teacher assessment, on students' knowledge achievement and academic thinking skills. To achieve these aims, the researcher employed the pretest-posttest quasi-experimental design. In this

design, each student knowledge achievement and academic thinking skills were measured before and after treatment by using two valid and reliable tests, one for knowledge achievement and the other for academic thinking skills **(**See Appendix I). The researcher also developed and used two valid and reliable scoring scales, one for knowledge achievement and the other for academic thinking skills (See Appendix II).

Study 1

The first study took place during the 2000/2001 academic year and examined the effect of student self-assessment versus no assessment on English-major seniors' knowledge achievement in the course of English Language Teaching Methods and their academic thinking skills in the same content area. The study involved a treatment group and a control group**.** In the treatment group (n=47) each student was asked to assess her/his knowledge before and after the lecture. In the control group (n=47) students neither assessed themselves nor received teacher evaluation. Both groups were well balanced to eliminate the possible confounding effects of gender and competence in the content area. The results of the study revealed no significant differences between the self-assessment group and the no-assessment control group in knowledge achievement or academic thinking skills ($t=0.88$, $p > 0.05$; $t=96$, $p > 0.05$, respectively). On average, the

treatment group received slightly higher scores on both achievement and thinking skills than the control group, but the differences were statistically insignificant. A possible explanation for these findings is that students' lack of previous experience in self-assessment and their inability to assess their own achievements made self-assessment difficult and time-consuming for them. Like any other skill, students were in need to practice self-assessment under the teacher's supervision to become skilful and comfortable enough to incorporate it into their own learning. The teacher (researcher) observed that when students were asked to assess their own achievement in each session, most of them did not know where to begin or how to assess. In line with this explanation, Bose, Oliveras and Edson (2001) are of the opinion that self-assessment is a skill that needs to learned and honed over time. Brookhart, Andolina and Zuza (2004) also point out that self-assessment is not a magic bullet that generates learners' learning and involvement, but like other strategies, it needs to be instilled in learners. Another explanation for the findings of study one is that students in the treatment group felt they were deserted by the teacher because no assistance was received from him before, during or after self-assessment. This led many of them not to take self-assessment seriously.

Study 2

The second study took place during the 2004/2005 academic year and examined the effect of self-plus-teacher assessment versus self-assessment on the same dependent variables that were investigated in study one with other English-major seniors at the same school. The study involved two treatment groups. In group one (n=30) each student assessed her/his own knowledge and received teacher evaluation on her/his self-assessment without assigning grades to help her/him change misconceptions and improve weaknesses in knowledge and thinking. In group two (n=32) each student assessed her/his own achievement independently. Both groups were well balanced to eliminate the possible confounding effects of gender and competence in the content area.

The results of study two revealed that self-plus-teacher assessment was more successful than self-assessment in improving students' knowledge achievement and academic thinking ($t=2.67$, $p < 0.01$; $t=2.90$, $p < 0.01$, respectively). These findings are due to five reasons. First, self-plus-teacher assessment encouraged and motivated students to take self-assessment seriously. Second, the teacher assessment of student self-assessment provided the teacher (researcher) with ongoing information to gauge student progress and adjust teaching accordingly.

Third, the assessment dialogue between the teacher and the student provided students with deep information and developed their ways of thinking as well as their assessment skills. Fourth, the teacher feedback on students self-assessments helped them know how the teacher thinks and allowed individual students to compare their own self-assessments with those of the teacher, which in turn, could improve their assessment skill and make them think more, and consequently learn more. Fifth, the teacher assessment, which focused only on identifying strengths and weaknesses in student self-assessment without assigning grades, motivated students to involve actively and comfortably in their self-assessments. In accordance with this reason, many assessment specialists believe that grades may be detrimental to formative assessment and its purposes. As Taras (2001) puts it: "Experience has shown that the grade interferes with students' judgments and prevents them from focusing on their work" (p. 609). In support of this, in their meta-analysis of forty-eight studies about student self-assessment, Falchikov and Boud (1989) found that in an entry-level required course, self-assessment did not work because students were too preoccupied with grades and consequently inflated their grades and reflections in order to pass a course.

It is evident that the findings of study two support Vygotsky's (1978) theory that the learner's ability for independent strategic functioning can evolve through social interaction with an expert within her/his "zone of proximal development" (p. 86). They also support Donato's (2000) view that self-assessment should leave room for a dialogue to occur between the student and the teacher. Instead of being unidirectional (as under the traditional paradigm), self-assessment should allow for a bi-directional flow of information in which both the teacher and the student are involved in the process.

Conclusions

Based on the results of the two studies, the researcher can conclude that (1) university teachers should not expect students to demonstrate expert assessment skills without support. In other words, it is unwise to assume that university students can assess their own knowledge or performance, particularly if they lack the experience in carrying out the assessment process; (2) self-assessment alone, without the teacher's ongoing support, does not improve students achievement or thinking skills; (3) self-plus-teacher assessment is more effective than self-assessment alone in improving students' achievement in content areas; and (4) university students can effectively achieve a high level of academic thinking skills if they are involved in self-plus-teacher assessment.

Recommendations

In light of the findings of the two studies conducted by the researcher, the following recommendations are offered:

(1) Rather than viewing self and teacher assessments as opposing strategies, it is more useful to capitalize on the advantages of both. In other words, for self-assessment to be effective, students are in need to practice it with teacher feedback. The teacher feedback can be decreased gradually, and the student can take greater responsibility for assessment as her/his self-assessment skill is developed. In view of this, the researcher proposes the following three stages for implementing self-assessment:

   (a) Self-assessment awareness-raising: In this stage learners are made aware of the purposes, techniques, advantages and disadvantages of self-assessment.

   (b) Guided self-assessment practice: In this stage students apply self-assessment techniques with teacher feedback.

   (c) Independent self-assessment: In this stage each student assesses her/his own learning independently.

(2) Guided self-assessment should be incorporated at the earliest stages of education.

(3) All university teachers should make self-assessment part of their classroom activities under their own guidance and supervision to enable students to become independent learners.

(4) Self-assessment should not free university teachers from the duty of classroom assessment because most students do not develop this skill independently.

(5) University teachers should change students negative attitudes towards self-assessment by raising their own awareness of the benefits of this type of assessment, especially when it is a new experience to them.

(6) The ultimate purpose of student self-assessment in higher education should be to improve students' learning not prove it.

References

Angelo, T. and Cross, K. (1993). *Classroom assessment techniques* (2nd ed.). San Francisco: Jossey-Bass.

Arthur, H. (1995). Student self-evaluations: How useful? How valid*? International Journal of Nursing Studies*, 32(3), 271-276.

Asadoorian, J. and Batty, H. (2005). An evidence-based model of effective self-assessment for directing professional learning. *Journal of Dental Education*, 69(12), 1315-1323.

Blanche, P. (1988). Self-assessment of foreign language skills: Implications for teachers and researchers. *RELC Journal*, 19 (1), 75-93.

Blue, G. (1988). Self-assessment: The limits of learner independence. In A. Brookes and P. Grundy (eds.), *Individualization and autonomy in language learning* (pp. 100-118). London: Modern English Publications and the British Council.

Bose, S., Oliveras, E. and Edson, W. (2001). How can self-assessment improve the quality of healthcare? *Quality Assurance Project*, 2(4), 1-19.

Breidert, J. and Fite, J. (2009). *Self assessment: Review and implications for training.* Arlington, Virginia: U.S. Army Research Institute for the Behavioral and Social Sciences.

Brew, A. (1999). Towards autonomous assessment: Using self-assessment and peer-assessment. In S. Brown and A. Glasner (eds.), *Assessment matters in higher education: Choosing and using diverse approaches* (pp. 157-159). Buckingham, U. K.: Society for Research into Higher Education.

Brookhart, S., Andolina, M. and Zuza, M. (2004). Minute math: An action research study of student self-assessment. *Educational Studies in Mathematics*, 57, 213-227.

Bryan, R., Krych, A., Carmichael, S., Viggiano, T. and Pawlina, W. (2005). Assessing professionalism in early medical education: Experience with peer evaluation and self-evaluation in the gross anatomy course. *Annals of Academy of Medicine in Singapore*, 34(8), 486-491.

Chamot, A. and O'Malley, J. (1994). *The CALLA handbook: Implementing the cognitive language learning approach.* Reading, MA: Addison Wesley.

Chur-Hansen, A. (2000). Medical students' essay-writing skills: Criteria-based self- and tutor-evaluation and the role of language background. *Medical Education*, 34, 194-198.

Colthart, I., Bagnall, G., Evans, A., Allbutt, H., Haig, A., Illing, J. and McKinstry, B. (2008). The effectiveness of self-assessment on the identification of learner needs, learner activity, and impact on clinical practice. *Medical Teacher*, 30(2), 124-145.

Dickinson, L. (1987). *Self-instruction in language learning*. Cambridge: Cambridge University Press.

Donato, R. (2000). Sociocultural contributions to understanding the foreign and second language classroom. In J. P. Lantolf (ed.), *Sociocultural theory and second language learning* (pp. 27-50). Oxford: Oxford University Press.

Dunning, D., Johnson, K., Ehrlinger, J. and Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science,* 12 (3), 83-87.

El-Koumy, A. (2001). *Effects of student self-assessment on knowledge achievement and academic thinking*. Washington DC: ERIC Clearinghouse on Languages and Linguistics. [ERIC Document No. ED452731]

El-Koumy, A. (2005). *Effects of self-plus-teacher assessment versus student self-assessment on English-major seniors' knowledge achievement and academic thinking*. Unpublished Paper Presented at the National Conference on Assessment, May 23-24, 2005.

Falchikov, N. (2003). Involving students in assessment. *Psychology, Learning and Teaching*, 3(2), 102-108.

Falchikov, N. and Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Higher Education Research*, 59 (4), 395-430.

Fitzpatrick, J. (2006). Self-assessment as a strategy to provide interactive learning within a professional degree programme. *Learning in Health and Social Care*, 5(1), 23-34.

Frye, A., Richards, B., Bradley, E. and Philp, J. (1991). The consistency of students' self-assessments in short-essay subject matter examinations. *Medical Education,* 25, 310-316.

Gordon, M. (1991). A review of the validity and accuracy of self-assessments in health professions training. *Academic Medicine*, 66, 762-769.

Harris, M. (1997). Self-assessment of language learning in formal settings. *English Language Teaching Journal*, 51(1), 12-20.

Hinsz, V. and Matz, D. (1997). Self-evaluations involved in goal setting and task performance. *Social Behavior and Personality*, 25(2), 177-182.

Kirby, N. and Downs, C. (2007). Self-assessment and the disadvantaged student: Potential for encouraging self-regulated learning? *Assessment and Evaluation in Higher Education*, 32, (4), 475-494.

Kruger, J. and Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Psychology*, 77(6), 1121-1134.

Liang, J. (2006). *Overview of self-assessment in the second language writing classroom*. Paper Presented at the 2006 TESOL Convention, Tampa,

Florida. Available at: secondlanguagewriting.com/documents/Overview.Doc. Accessed May 10, 2006.

Lind, D., Rekkas, S., Bui, V., Lam, T., Beierle, E., and Copeland, E. (2002). Competency-based student self-assessment on a surgery rotation. *Journal of Surgical Research*, 105, 31-34.

Mabe, P. and West, S. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, 67(3), 280-296.

MacDonald, B. and Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education: Principles, Policy and Practice*, 10(2), 209–220.

MacIntyre, P., Noels, K., and Clément, R. (1997). Biases in self-ratings of second language proficiency: The role of language anxiety. *Language Learning*, 47 (2), 265-287.

Mattheos, N., Nattestad, A., Falk-Nilsson, E. and Attstrom, R. (2004). The interactive examination: Assessing students' self-assessment ability. *Medical Education*, 38, 378-389.

Mort, J. and Hansen, D. (2010). First-year pharmacy students' self-assessment of communication skills and the impact of video review. *American Journal of Pharmaceutical Education*, 74(5), 78.

Oscarson, M. (1997). Self-assessment of foreign and second language proficiency. In C. Clapham and D. Corson (eds.), *Encyclopedia of language and education, Volume 7: Language testing and assessment* (pp. 175-187). Dordrecht: Kluwer Academic Publishers.

Rees, C. (2003). Self-assessment scores and gender. *Medical Education*, 37, 571-573.

Risucci, D., Tortolai, A. and Ward, R. (1989). Ratings of surgical residents by self, supervisors and peers. *Surgical Gynecology and Obstetrics*, 169(6), 519-526.

Shrauger, J. and Osberg, T. (1981). The relative accuracy of self-predictions and judgments by others in psychological assessment. *Psychological Bulletin*, 90(2), 322-351

Stiggins, R., Rubel, E. and Quellmalz, E. (1988). *Measuring thinking skills in the classroom* (Revised Edition). Washington, D.C.: Northwest Regional Educational Laboratory.

Taras, M. (2001). The use of tutor feedback and student self-assessment in summative assessment tasks: Towards transparency for students and for tutors. *Assessment and Evaluation in Higher Education*, 26, 605-614.

Teh, S. (2006). Implementing student self-assessment–tools for learning and reflection. *EARB and ELTC*, 1-7.

Vygotsky, L. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.

Woolliscroft, J., TenHaken, J., Smith, J. and Calhoun, J. (1993). Medical students' clinical self-assessments: Comparisons with external measures of performance and the students' overall self-assessments of overall performance and effort. *Academic Medicine*, 68(4), 285-294.

Zohar, A. (2004). *Higher order thinking in science classrooms: Students' learning and teachers' professional development.* Dordrecht: Kluwer Academic Publishers.

Zoller, U., Fastow, M., Lubezky, A., and Tsaparlis, G. (1999). Students' self-assessment in chemistry examinations requiring higher- and lower order cognitive skills, *Journal of Chemical Education*, 76 (1), 112-113.

Appendix I

(A) Knowledge achievement test

English Language Teaching Methods

Instructions to test takers: This is a test of your ELT knowledge. It consists of four tasks and takes three hours. Three marks will be given to each response in accordance with the attached scoring scale.

Test tasks

Answer the following questions:

(1) Outline the components of the following:

(a) Listening, and

(b) Reading.

(2) What are the similarities and differences between spoken and written language?

(3) What are the advantages and disadvantages of guided composition?

(4) What are the differences between formative and summative evaluation?

(B) Academic thinking skills test

English Language Teaching Methods

Instructions to test takers: This is a test of your thinking skills in the area of English language teaching. It consists of four tasks and takes three hours. Twelve marks will be given to each response in accordance with the attached scoring scale.

Test tasks

Answer the following questions:

(1) Which do you think is the most effective in our context, a whole-language approach, a skills-based approach, or a compromise between the two? Why?

(2) Do you think grammar needs to be deliberately taught in the EFL classroom? Why? Why not?

(3) Do you think that overemphasis on language mistakes can distract the learner's attention from meaning? Why? Why not?

(4) What do you think of breaking down language into components for the purpose of testing?

Appendix II

(A) Scale for scoring knowledge achievement

0 No evidence of knowledge about the topic in question

1 mark Limited knowledge of the topic in question

2 marks Moderate knowledge of the topic in question

3 marks Evidence of complete knowledge about the topic in question

(B) Scale for scoring academic thinking skills*

Thinking skills	Very poor	Poor	Good	Very good
	0	**1**	**2**	**3**
Examinee analyzes the elements of the topic in question and addresses their relationships to each other.				
Examinee compares/contrasts multiple views on the topic in question.				
Examinee creates a personal opinion.				
Examinee gives reasons to support her/his point of view.				
TOTAL				

* Adapted from R. Stiggins, E. Rubel, and E. Quellmalz (1988), *Measuring thinking skills in the classroom* (Revised Edition), Washington, D.C.: Northwest Regional Educational Laboratory, p. 12.