

The Ability of Non-Music Majors to Self-Evaluate at the End of a Music Course

Dan Keast, The University of Texas of the Permian Basin, Odessa, Texas, USA

Larke Tapper, The University of Texas of the Permian Basin, Odessa, Texas, USA

Abstract

The researchers of this study investigated the participants' (N = 177) use of a self-evaluation tool employed at the end of an online undergraduate music course that fulfilled the Texas general education requirement for the creative arts. Participants' use of the two aspects of the tool correlated at $r = .5548$ – interpreted as a high positive relationship. The Pearson coefficient for instructor final grade and the student desired grade was $r = .4456$ – a moderate positive correlation. A t-test of the instructor grade to the student desire grade was -2.814 with a $p = .002$. Free-response justifications for the desired grade generated a mean of 3.53 sentences each that were qualitatively coded into themes: Effort/Completed 22%, Generalized Statements 20%, and Positive Affirmations 11% as the larger categories. Participants consistently used the form to justify a desired grade, evaluate their work, and suggest a final grade. However, that grade was statistically significantly lower than the final grade issued by the professor.

Keywords: Online learning, self-evaluation, online capstone courses, case-based learning, problem-based learning

INTRODUCTION

University educators have utilized self-evaluation – the ability to look at one’s own work and know what met the standard and what fell short of the standard – for decades as a form of formative assessment. However, including the self-evaluation into the final course grade is something rare in higher education. As a capstone event in an online undergraduate music course, a self-evaluation helps a student to understand what was learned in the course and why they should expect a particular grade. The process also teaches students to analyze what was expected of them and their performance.

As musicians, the researchers of this study rely on, and teach, metacognitive skills to students. Student musicians enter a practice room each day to practice material for an upcoming lesson. Applied Music teachers instruct students not to waste time practicing material that is easy or ready for performance, but to isolate the shorter passages that challenge the student the most. This is a metacognitive skill – knowing what is known – and relies heavily upon self-evaluation. Music is not alone in its use of this metacognitive skill (Lynch, Mannix McNamara, & Seery, 2012).

Several careers require staff to routinely self-assess their skills and/or knowledge to perform a task (Mahlberg, 2015). Take, for instance, a medical surgeon. If the surgeon is not properly trained to perform a procedure, they are likely to seek the training or postpone a procedure until confident about their ability. In other fields, a mentor often observes new staff to guide them during an initial/probationary period. At some point, the new staff member either demonstrates the ability to perform solo tasks, or indicates to their mentor the confidence and ability to do so. In education, this could be the student teaching process, the internship, or even the first year of teaching.

In higher education, a self-evaluation by the student aids the student in understanding what grade they are likely to receive based on their performance in the course. Self-evaluation helps the student express what they did in concert with what was expected. The ability to express their opinion is valuable to the instructor in case there is a misconception or disputed work that could be cleared up to prevent a final grade appeal or academic grievance.

Teaching the ability to self-assess is not particularly challenging, just time consuming when faculty struggle to cover increasing amounts of content in courses. With a trend in online learning favoring shorter courses than the traditional 16-week or 12-week semester, time is a premium and self-assessment is, perhaps, a casualty to the accelerated course programs.

Modern careers are fraught with educational demands on employees who need to keep abreast of the advances within the field. Medical breakthroughs occur every day and the information is widely disseminated through conferences, journals, lectures, and pharmaceutical companies. Telecommunication companies incorporating new technologies must train their staff to install or use the new systems. Those working in the field of automotive repair are not immune to the advances in engine or various safety devices utilized in new vehicles. Their employees are routinely trained on these mechanisms to insure proper repair when necessary.

Lifelong learning is not a fad, but a way of life in the twenty-first century for both blue- and white-collar jobs. The premise of lifelong learning requires self-evaluation (Arthur, 1995; Mattheos, Nattestad, Falk-Nilsson, & Attstrom, 2004; Dochy, Segers, & Sluijsmans, 1999). Without the ability to self-evaluate skills, a person will not stay proficient at a job for long. Computer technology, the Internet, and a host of other advances in science and math disrupted many industries in the twentieth century – and continue to enact change in many fields.

Employees are continuously self-evaluating their ability, seeking information, and altering their habits in response to the new information.

A trip to the doctor is now a self-evaluation when a medical professional performs a medical history and current symptoms assessment. Automobile drivers perform a self-evaluation prior to operating their vehicle to determine if they are coherent, sober, or physically able to maneuver the vehicle safely. Self-evaluation permeates society, yet is rarely taught in our educational system.

LITERATURE REVIEW

The goals of higher education changed from domain-specific knowledge proficiency, to a more liberal arts minded individual (Dochy, Segers, and Aluijsmans, 1999). Professors converted from a lecture-test model of education, popular in the mid-twentieth century, in favor of case-based scenarios and problem-based learning with ill-structured problems (Jonassen, 1997). The demand for new assessment methods surfaced to evaluate the students' constructed knowledge instead of measuring their reproduction of knowledge. The use of rubrics surged as faculty adapted new evaluation methods and incorporated self-evaluations in the courses.

Self-evaluation research exists primarily in education and medical education literature concerning college and universities students (Lynch, Mannix McNamara, and Seery, 2012). Professors in higher education have used the process of self-evaluation as formative and summative assessment. "Good' students have always been effective self-assessors, but it is becoming increasingly recognized that in order to develop this skill more widely among students, explicit attempts need to be made to develop the capability, and opportunities need to be given for it to be openly practiced" (Boud & Falchikov, 530).

The process of evaluation is a higher-order thinking skill which requires synthesis and analysis to create a judgement. The revised taxonomy by Anderson and Krathwohl (2000) still has evaluation as a higher-order thinking skill of the cognitive domain originally developed by Benjamin Bloom (1956). Evaluating is described in the taxonomy as making judgements about the value of ideas or materials. Such activities may include summarizing, justifying, critiquing, concluding, and appraising. The goal of higher education is for students to create an understanding of knowledge using a set of criteria (Boud and Falchikov, 1989). The student's participation in assessing that knowledge is perceived by researchers as fair, reliable, and a contribution to the growth of competence (Dochy, Segers, and Sluijsmans, 1999; Dunning, Heath, and Suls, 2004). Mahlberg found that self-evaluation linked to better school performance: "Students exposed to self-assessment in their classes report using significantly more self-regulated processes such as coming to class prepared, setting goals, reflecting on learning objectives, and modifying study strategies to increase understanding than students exposed only to traditional assessment" (p. 779).

The upsurge of interest in studying and implementing self-evaluation can be attributed to the publication of benefits. Boud and Falchikov (1989) found that "...there has been a principled desire on the part of teachers for learners to take greater responsibility for their own learning through involvement in a crucial act of learning: assessing one's own competence" (p.530). Faculty cannot simply rely on tests to measure a learner's competence, but need to include the learner in the assessment process (Kurt, 2014). A self-evaluation is not to replace other methods, but learners deserve the opportunity to be a part of the process. Similarly, self-evaluation should not replace an existing process, but become a part of an existing one (Lopez-Pastor, Fernandez-Balboa, Santos Pastor, and Fraile Aranda, 2012; Kurt, 2014).

The literature of self-evaluation includes concerns of studies generalizing correlations between students who used a particular scale and a faculty member using a different scale. The use of a five-point scale was espoused by Boud and Falchikov (1989) as being more effective than a percentage. Bergee (1997) found an increased coefficient between instructor grade to student self-evaluation as the structure of the assessment increased. Students expecting higher grades often underrate while students expecting lower grades tend to overrate in comparison to instructor evaluations (Weimer, 2014; Bergee, 1997; Brew and Boud, 1995). However, students can become better self-evaluators with continued practice (Sadler & Good, 2006; Dochy, Segers, & Sluijsmans, 1999; Boud and Falchikov, 1989; Carrigan and Hardham, 2011).

The quantitative literature for self-evaluation illustrates the diversity of methods for the process itself. Falchikov and Boud (1989) discovered coefficients from -0.05 to 0.82 with a mean of $r = 0.39$. Cohen (1977) set the definitions of $r = 0.10$ as small, $r = 0.30$ as moderate, and $r = 0.50$ as high. Another meta-analysis by Mabe and West (1982) reported coefficients from -0.26 to 0.80 with a mean of 0.29. In Bryan, Krych, Carmichael, Viggiano, and Pawlina, (2005), peer evaluations were found to be significantly higher than self-evaluations where $r = 0.22$ ($p = .0001$). A study of suggested self-grade, professor grade, and negotiated final grade completed by Lopez-Pastor, Fernandez-Balboa, Santos Pastor, and Fraile Aranda (2012) found significant difference between each of the grades ($p = .001$). A study by Arthur (1995) used a test-retest method for self-evaluation of knowledge and found a $r = 0.68$. And in music, the field of the authors, Bergee (1993) reported the self-evaluation ability by musicians to be poor as compared to peer and master teacher evaluations.

Palloff and Pratt (2007) propose a set of questions to assist students in a self-evaluation:

What was most useful to me in my learning process?

Did I achieve my learning objectives in this course?

What did I learn about my own learning process by taking this course?

How did I change as a learner through my involvement in this course?

Do I feel what I learned in and through this course will have application in other areas of my life?

How well did I participate in this course?

How would I evaluate my performance in this course overall? (p. 225).

This study measured the participants' use of six stimulus questions, suggested desired grade, free-response justification, and instructor grade to answer four guiding questions:

Research Question 1: Did the mean of the questions equate to the student's suggested overall grade?

Research Question 2: What themes surfaced in the free-response question as to why they deserve the grade?

Research Question 3: Do students who desired an "A" in the course write more justification in the free response area than those who expected a "B" or "C?"

Research Question 4: Were students' desired grades comparable to the earned grade provided by the professor?

METHODOLOGY

The participants in this study (N=176) were selected from students enrolled in an online undergraduate general education music course and who had completed the prescribed self-evaluation form in its entirety. The task to self-evaluate was given during finals week in the course and used a form supplied by the instructor. Students were motivated to complete the form

as it comprised 5% of the student's final grade in the course. The identity of the participant was not coded, only the data on the form.

Participants in the study were registered students completing the 16-week course at a mid-sized Texas university with a distance education initiative. Students in the course ranged from 15 to 65+ years old, although the majority aged from 16-20 years of age. The course is designated as a dual-credit course for high school students, so younger students are common. As a university, the student body is 59% female, 40% part-time students, 45% online only, 67% minority, and 40% who receive government aid to offset tuition.

Mahlberg's study (2015) informed the design of this study to including questions and the Likert-type scale to quantify the student responses. Mahlberg's design utilized a five point Likert-type scale asking students to assess their work from 0 (unacceptable) to 5 (excellent). Questions included the following:

-I contributed meaningfully to every classroom discussion by sharing examples and observations.

-I read the assigned chapters in the textbook.

-I completed all of the assignments included in the 'A' assignments.

Participants in this study were presented with six stimulus questions in which they were asked to use the following scale to rate themselves: 1 (needs improvement), 2 (fair), 3 (average), 4 (good), and 5 (excellent). The questions were:

- 1) Attendance and participation in group meetings.
- 2) Was prepared and accomplished tasks on time.
- 3) Helped keep focused on goals.
- 4) Contributed quality ideas and information to each part of the project.

- 5) Supported and encouraged other group members.
- 6) Overall rating I would give myself.

The six questions are the same questions that the students used to evaluate each other, so the choice was made to allow the student to evaluate themselves using the same form. The form also asked two open-ended questions: “What grade do you deserve for this course?” and “Why do you deserve that grade?” The form was downloaded from the Learning Management System (LMS) by the student and completed offline. Once completed, the student would send the form, via an attachment to a private email, to the professor for credit.

The students normally reported the grade they expected as a letter grade, some with plus and minuses, which were coded as percentages using the existing grading scale published in the syllabus. Other responses were either percentages or use of the Likert-type scale from the questions. In the rare instances of the Likert-type scale, the number was equated to a percentage to maintain integrity of the student’s desired score.

The student’s desired grade is entered into the gradebook as a percentage and weighted as five percent of the final course grade. Thus, the student is highly motivated to complete the assignment as it is worth half the value of a quiz in this course. Additionally, the student is instructed to make a convincing case for their desired grade. The assignment could be viewed as a short persuasive essay – an extension of the curricular goals.

RESULTS

Research Question 1: Did the mean of the questions equate to the student’s suggested overall grade?

A Pearson correlation coefficient was employed because the normally distributed data were intervallic. The Pearson coefficient determined the relationship between the mean of the six

stimulus questions and the student's desired grade. The $r = .5548$ which indicated the relationship between the variables had a tendency for high grades to correlate with a high mean on the stimulus questions. The strength of Pearson's coefficients (r) was large/high from .5 to 1.0 (Cohen, 1977). Since the coefficient was positive, the relationship was interpreted by the researchers as a high positive relationship.

The Pearson coefficient for earned grade given by the instructor to the self-reported mean of stimulus questions was $r = .5518$. The relationship between the variables indicated a tendency for high earned grades to correlate with a higher mean on the stimulus questions.

Additional Pearson tests revealed that the relationship of a participant's desired grade to the earned grade provided by the professor was a moderate relationship at $r = .4456$. Pearson correlation coefficients are considered moderate from .3 to .5, whether positive or negative (Cohen, 1977).

The mean for each stimulus question was calculated to compare to the overall desired score. The data for the means of the stimulus questions are reported in Table #1. The mean for the average suggested grade was 89% and students used an average of 3.54 sentences to justify their grade.

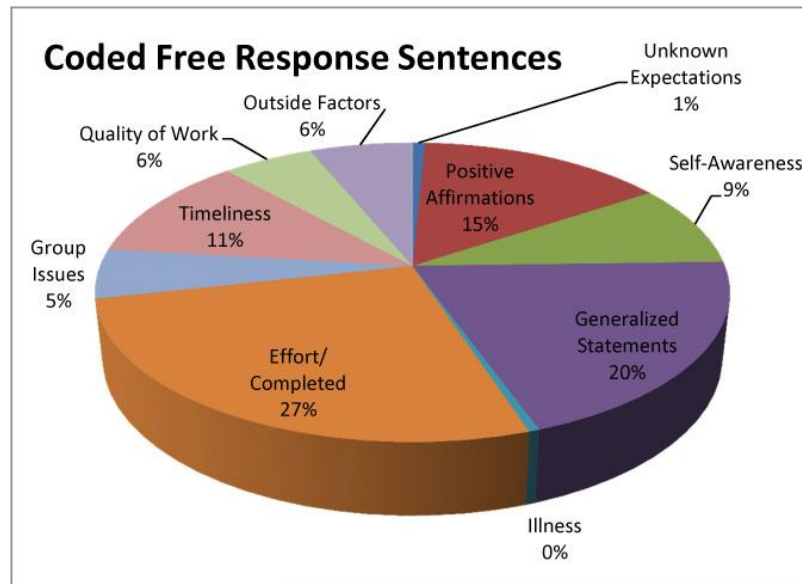
Table 1
Mean for Each Stimulus Question and Open Question on the Survey

4.39 – Attendance and participation in group meetings.
3.92 – Was prepared and accomplished tasks on time.
4.21 – Helped keep focused on goals.
4.35 – Contributed quality ideas and information to each part of the project.
4.21 – Supported and encouraged other group members.
4.30 – Overall rating I would give this member.
89% – Overall suggested percentage
3.54 – Number of sentences used to justify grade

Research Question 2: What themes surfaced in the free-response question as to why they deserve the grade?

Six hundred and twenty-four sentences were included in the study from 176 participants over the course of three years (six semesters). The justifications were removed from the form and separated into individual sentences that were divided into thematic categories related to generalized topics.

The free response data (N=624 sentences) were qualitatively analyzed using the grounded theory method of triangulation. Participants authored sentences justifying a desired grade. Those sentences were coded by the researchers then verified by an independent third party with an 84% accuracy. The themes are displayed in Graph #1 with Effort/Completed (27%), Generalized Statements (20%), and Positive Affirmations (15%) garnering the majority of sentences.

Graph #1. Coded Free Response Sentences

The themes were generated by using key words and recurring phrases. The Effort/Completed category (27%) consisted of sentences with references to “I worked hard,” “I learned a lot,” “I tried hard,” and “I completed all of the assignments.” The Generalized Statements category (20%) used a wide variety of statements that did not fit well into other categories. These statements included a reference to a newfound appreciation of a specific artists studied, amazement about the number of jazz styles, attentiveness to the class website or email from group members. The Positive Affirmations category (15%) contained words describing the course overall as fun, enjoyed the course, interesting course, enjoyed the professor, and informative. The Timeliness category (11%) included references to time such as “turned all of my assignments in on time,” “I did the work that was assigned to me by the due date,” and “I collaborated with my group to turn in activities before the deadline.” The Self-Awareness category (9%) was interesting as it mirrors some college-readiness skills ascertained in the

prerequisites. “I know that I could have done better if I had taken more time to do the coursework,” “I work better in face-to-face courses,” or “I had a lot on my plate this semester and now know that I need to divide my time to do the work more evenly.” Quality of Work (6%) referenced the individual’s work quality as “my work was outstanding,” “all my work was up to par, and good work all-in-all,” or “I know it [the work] was all of great quality because I worked very hard on it.” Outside Factors (6%) tied with Quality of Work, resonated with the older students in the course because of references to “I have a job and work all day,” “I have a 1-year old and a full time job,” or “my computer broke and that really put me behind.” Group Issues (5%) was expected to be a popular thread as group work is often unpopular with online students. However, only a small percentage groused about group problems in their self-evaluations. Statements did address the distaste for groups with references to poor group communication, lack of a time to meet or coordinate, and “doing it alone was easier than involving others.” Unknown Expectations (1%) included statements like “this is my first online course so I didn’t know what to expect,” “I’ve never taken a music course so I didn’t know what I was getting into,” or “this was my first college course and I wasn’t sure what to expect.” The Illness (<1%) referred to long term illnesses or family emergencies such as a death or natural disaster.

Research Question 3: Do students who desired an “A” in the course write more justification in the free response area than those who expected a “B” or “C?”

The average number of sentences written by those participants desiring an “A” was 3.618 (N=76) while the participants desiring a “B” or “C” averaged 3.426 (N=100.) The overall mean of sentences authored was 3.53.

A t-test compared the number of sentences authored by the students desiring an “A” and those desiring a “B” or “C”. The t-value was .4831 and interpreted as not significant at .05

alpha. Therefore, there is no significant difference between the number of sentences authored by the two populations.

Research Question 4: Were students' desired grades comparable to the earned grade provided by the professor?

The data were considered to be within a normal distribution due to the standard deviation of desired grades = 0.1078 and earned grades = 0.111. A t-test of the desired grade to the earned grade given by the professor was $t = -2.814$ with a $p = .002$. There was a statistically significant negative difference between the desired and earned grades.

DISCUSSION

Research Question 1 – The participants self-evaluated by using the questions in a relatively close manner to their overall desired grade for the course. While the correlation was $r = .5548$, it is positive and high. The researchers' premises were unfounded that these two aspects of the self-evaluation form were disconnected, and rife with variation. The data illustrated the students were consistent in their analysis and able to link the two measures in their self-evaluation with a high degree of correlation.

The idea that the desired grade correlated to the mean of the stimulus questions prompted the researchers to test if the earned grade was correlated as well. The correlation of the earned grade to the mean of the questions was also positive and high. Student responses to the stimulus questions appeared to be consistent to the assessment by the professor.

Research Question 2 – The students used the justification section to their full advantage by discussing their struggles and triumphs in the course. Several themes were evident in the data that were coded and then verified by an independent third party.

The data were expected by the instructor to address the quality of the work completed by the student. However, the category Quality of Work was third to smallest in the coded data. Students used over 20% more statements about their effort in the course than the quality of their work. Only lamenting about Illness, Group Issues, and Unknown Expectations were lower than the amount of data in the Quality of Work category. This illustrated a disconnection between faculty who are interested in the quality of work to the student who is more focused on the completion of the work or the expended effort in the course.

Research Question 3 - Participants were thought to have written more when justifying their grade for an “A” than for a lower mark. While the mean variance between the “A” desired grades and “B/C” desired grades were only slight (3.618 to 3.426), a t-test confirmed that there was no significant difference in the use of sentences ($t = .5964$). The variation was investigated by searching for outliers in the data, yet no obvious outlier was found. The skew = 1.55 indicating an asymmetrical distribution with a longer tail to the right.

This finding was interpreted by the researchers that the participants were equally verbose about earning an “A” as that of a “B” or “C” in the course. The question stem on the form did not specify how many sentences – or even that a justification was required. Students appear to have replied with a similar number of sentences to justify a grade regardless of their desired course grade.

Research Question 4 - The coefficient of earned grade to desired grade was only a moderate correlation at $r = .4456$. The standard deviations were minimal and suggested that the data were similar. However, a t-test of the desired grades to the earned grades determined statistical significance and lead the researchers to believe that the participants were not able to effectively evaluate themselves. These data echo the findings of Boud and Falchikov (1989) in

that the students' self-assessment overrated the earned score provided by the faculty. The findings in this research (mean = .4456) reinforce the findings of Mabe and West (1982) who found a range of coefficients from -.26 to .80 with a mean of .29.

A few of Palloff and Pratt's (2007) questions resemble college-readiness questions about the learning process and knowing one's desired way to learn. Since the researchers' course enrolls high school students, the recommended questions by Palloff and Pratt were informative to the redesign of the self-evaluation instrument. A revision now ties the stimulus questions to the stated learning objectives for the course. The form ascertains the student's knowledge of their learning process (what worked and didn't work for them). The form includes a question for the student's description of their participation in the course (consistency in posting, contributions to the group activities, quality of submitted work, and time spent engaged in content.) The final question still allows the students an opportunity to evaluate their overall performance in the course. The question now directs students to provide a percentage and two-three paragraphs for a justification of that expected grade.

The self-evaluation process is not as easily embedded into a course as this instructor thought. After careful consideration of the data in this study, the instructor was able to rebuild the self-assessment mechanism to assist the young adults in achieving a more successful self-evaluation. Leading students through a process is necessary as self-evaluation is a taught skill as discussed by other researchers. With consistent and clear expectations, self-evaluation skills can mature over time and lead to helpful habits throughout a lifetime of learning.

Further studies in this line of research of self-assessment of online college courses should consider investigating the age of the student, if the student has taken previous online courses, if this is his/her first college course, or a correlation to the student's overall GPA. Participating

faculty should alter questions to reflect specific course content and goals stated in the syllabus. A blended model is recommended to gain the depth of student input gained through qualitative data and breadth from the quantitative data.

REFERENCES

- Anderson, L. W. & Krathwohl, D. R. (Eds.) (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of Educational Objectives*. Boston: Allyn & Bacon.
- Arthur, H. (1995). Student self-evaluations: How useful? How valid? *International Journal of Nursing Studies*, 32(3), 271-276.
- Bergee, M. (1993). A comparison of faculty, peer, and self-evaluation of applied brass jury performances. *Journal of Research in Music Education*, 41(1), 19-27.
- Bergee, M. (1997). Relationships among faculty, peer, and self-evaluations of applied performances. *Journal of Research in Music Education*, 45(4), 601-12.
- Bloom, B., Engelhart, M., Furst, E., Hill, W., & Krathwohl, D. (1956). *Taxonomy of educational objectives: The classification and educational goals*. Handbook I: Cognitive Domain. New York: David McKay Company.
- Boud, D., & Falchikov, N. (1989). Quantitative studies of student self-assessment in higher education: A critical analysis of findings. *Higher Education*, 18, 529-49.
- Brew, A., & Boud, D. (1995). Teaching and research: Establishing the vital link with learning. *Higher Education*, 29(3), 261-73.
- Bryan, R. E., Krych, A. J., Carmichael, S. W., Viggiano, T. R., & Pawlina, W. (2005, September.). Assessing professionalism in early medical education: Experience with

- peer evaluation and self-evaluation in the Gross Anatomy course. *Annals Academy of Medicine*, 34(8), 486-491.
- Carrigan, R. & Hardham, G. (2011). Use of technology to enhance student self evaluation and the value of feedback on teaching. *International Journal of Therapy and Rehabilitation*, 18(10), 579-90.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York: Routledge.
- Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, 24(3), 331-50.
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5(3), 69-106.
- Falchikov, N. & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Education Research*, 59(4), 395-430.
- Lopez-Pastor, V. M., Fernandez-Balboa, J. M., Santos Pastor, M. L., & Fraile Aranda, A. (2012). Students' self-grading, professor's grading and negotiated final grading at three university programmes: Analysis of reliability and grade difference ranges and tendencies. *Assessment & Evaluation in Higher Education*, 37(4), 453-64.
- Lynch, R., Mannix McNamara, P., & Seery, N. (2012). Promoting deep learning in a teacher education programme through self-and peer-assessment and feedback. *European Journal of Teacher Education*, 35(2), 179-97.
- Mahlberg, J. (2015). Formative self-assessment college classes improves self-regulation and retention in first/second year community college students. *Community College Journal of Research and Practice*, 39(8), 772-83.
- Kurt, M. (2014). Collaborative assessment: Fostering ownership in assessment. *Education*,

134(3), 332-39.

Mabe, P. A. & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, 67(3), 280-96.

Mattheos, N., Nattestad, A., Falk-Nilsson, E., & Attstrom, R. (2004). The interactive examination: Assessing students' self-assessment ability. *Medical Education*, 38, 378-89.

Palloff, R. M. & Pratt, K. (2007). *Building online learning communities: Effective strategies for the virtual classroom* (2nd ed.). San Francisco: Jossey-Bass.

Sadler, P. M. & Good, E. (2006). The impact of self- and peer-grading on student learning. *Educational Assessment*, 11(1), 1-31.

Weimer, M. (2014). Developing students' self-assessment skills. *Faculty Focus*, December 10, 2014. Retrieved from <http://www.facultyfocus.com/articles/teaching-professor-blog/developing-students-self-assessment-skills/>