Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi (EFMED)
Cilt 7, Sayı 2, Aralık 2013, sayfa 1-17.

Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 7, Issue 2, December 2013, pp. 1-17.

Using Self-Assessment to Improve College Students' Engagement and Performance in Introductory Genetics

Mehmet Aydeniz* **& Michael, A. Gilchrist**

The University of Tennessee, Tennessee, USA

Received: 12.02.2013 Accepted: 08.10.2013

Abstract - The purpose of this study was to explore the role of structured self-assessment in receiving feedback on students' perceptions of classroom instruction, the learning strategies they use, and the type of instructional support they need to comprehend the course material in a 90-student college sophomore genetics course. The results indicate that weekly-administered structured self-assessments make a range of information accessible to the professor and engage students in self-reflection about their learning and the teaching strategies used in the classroom. Preliminary statistical analysis of participation in self-assessment and student performance on exams suggest that self-assessments have a moderately positive effect on student performance. Our discussion focuses on the challenges and opportunities presented to the course professor while administering and evaluating self-assessments. Finally, we discuss the role of technology in facilitating students' effective engagement with self-assessment.

Keywords: self-assessment, college science, instruction, genetics.

DOI No: http://dx.doi.org/10.12973/nefmed200

Özet - Bu çalışmanın amacı, öz-değerlendirmenin üniversite öğrencilerinin genetik dersindeki derse etkin katılımı ve dersteki başarıları üzerindeki etkilerini araştırmaktır. Bu çalışmaya 90 üniversite öğrencisi katılmıştır. Bu araştırmanın sonucunda öz-değerlendirmenin dersin hocasına ve öğrencilere çeşitli yararlı bilgilere ulaşmasında katkı sağladığı görülmüştür. Öz-değerlendirmeler sayesinde öğrenciler, konuyla ilgili değişik zayıflıklarının farkına varma şansı elde edebilmiş ve bu zayıflıklarını gidermek için değişik çalışma metotları hakkında düşünme aktivitelerine girişmişlerdir. Ders hocası öğrencilerden aldığı dönütler sayesinde kullandığı öğretim metotlarının etkinliği ve öğrencilere faydası hakkında değişik bilgilere sahip olma şansı yakalamıştır. Öğretim görevlisi bu bilgileri kullanarak derslerini daha etkili metotlar kullanarak anlatmak için motive olmuştur. öz-değerlendirme her ne kadar öğrencilerin basarisi üzerinde olumlu bir etki yapsa da bu etki istatiksel olarak anlamlı bir etki değil. Öğretim görevlisinin öz-değerlendirmeleri uygulama ve sonuçlarını kullanma aşamasında yaşamış olduğu sorunlar hakkında bilgi verilmektedir. Bu noktada teknolojinin bu sureci kolaylaştırma ve etkinleştirmesindeki rolü tartışılmıştır.

Anahtar kelimeler: öz-değerlendirme, üniversite fen dersi, öğretim, genetik.

* Corresponding Author: Dr. Mehmet Aydeniz, A 408 Jane and Bailey Education Complex The University of Tennessee, Knoxville, 1126 Volunteer Boulevard, Knoxville, Tennessee,37996-3442.
Email: maydeniz@utk.edu

Introduction

Assessment refers to the process of gathering information on students' or teacher's knowledge, skills, performance or attitudes in relation to some predetermined curricular or instructional goals (Berry & Adamson, 2011). Assessment can be used either for summative or formative purposes (National Research Council [NRC], 2001). Summative assessments are used to collect information on students' level of achievement of a particular instructional goal at the end of instruction (NRC, 2001; Kearney & Perkins, 2010). A growing number of educators believe these end-of-semester assessments are of limited use to teachers who need immediate feedback on the effectiveness of their instruction and to the students who need immediate feedback on the level of their learning (Aydeniz, 2007; Berry, 2006; Black, Harrison, Lee, Marshall & Wiliam, 2003; Bloxam & Boyd, 2007; Morgan & Watsin, 2002). These educators argue that if the goal of assessment is to enhance student learning and to inform the teacher of the effectiveness of his/her instruction, assessment must be used for formative purposes. (Berry, 2008; Black et al., 2003; Furtak & Ruiz-Primo, 2008).

Airasian (2001) defines formative assessment as "the process of collecting, synthesizing and interpreting information for the purpose of improving student learning while instruction is taking place" (p. 421). The key purpose of formative assessment is the use of information for improving instruction and learning rather than using information just for the purpose of documenting students' failure or success. It can be used both by the teacher and the students. Formative assessments can take place in various forms. Self-assessment and peer-assessment are the two most frequently used forms of formative assessments. In this study, we used student self-assessment (SA), which is a type of formative assessment that engages students in reflective learning.

Review of Literature on Self-Assessment

Self-Assessment (SA) refers to the process of critically analyzing factors that contribute or impede one's learning or performance and designing subsequent actions to meet the learning expectations (NRC, 2001). Research studies that examine SA in the context of college science classrooms (Aydeniz & Pabuccu, 2011; Lin, Hong, Wang & Lee, 2001), show that SA can result in various positive educational outcomes. If successfully utilized, SA can guide student thinking about what they know, how they learn, and how they can change the way they study to improve their learning (Zoller & Ben-Chaim, 1998). Studies show that SA can engage students in reflective learning, thus enhance students' metacognition and motivation during learning (Schunk, 2001; Sundstrome, 2005), engage students in mastery learning (Schunk, 2001), and increase the efficiency at which they learn (McMillan, 2004; McMillan & Hearn, 2009; Ross, 2006).

According to Schunk (1996), three processes embodied in SA help students to engage in metacognitive activity. First, SAs give students the opportunity to make observations about specific aspects of their learning. For instance, they can identify learning strategies or resources that make important contributions to their learning and those that do not make contribution to their learning. Second, SAs force students to make judgments about their performance in relation to the learning objectives of the lesson or the course. Third, it forces students to make decisions about the subsequent steps they need to take to accomplish learning goals in a course. For instance, SAs may prompt students to ask their instructor questions about the concepts they are most confused about.

While, in general, it is assumed that collectively these three processes will improve student learning, empirical studies testing this assumption in higher education science classes are limited. One reason why the numbers of SA studies in science are so limited is likely due to the fact that most university professors are unaware of their existence or utility as a pedagogical tool (Balinsky, 2006), or lack the time and motivation to develop self-assessment tools (Siebert, 2001). In order to make a contribution to the SA literature in college science teaching we designed and used a technology-based self-assessment tool to explore students' perceptions of: 1) classroom instruction, 2) the learning strategies they use, and 3) the type of support they believe will help them improve their comprehension of course materials.

The research questions guiding this inquiry are:

1) How does self-assessment engage college science students in self-reflection about:

a. their level of understanding of weekly course content,

b. instructional or learning strategies that helped them to develop conceptual understanding of weekly course content,

c. factors that impacted their limited understanding of the weekly course content,

d. instructional strategies that could have helped them to develop a better conceptual understanding of course content?

2) What impact does self-assessment have on college students' conceptual understanding of the course material?

Methodology

We used both qualitative and quantitative methods in this study. Qualitatively, we used thematic analyses to identify the factors (e.g., instructional strategies) that contributed or inhibited students' learning in the course. Quantitatively, we used descriptive statistics to calculate the frequency of different instructional and learning strategies that either helped or hindered students' learning in the course. In addition, we fitted a linear model between our dependent variables and student performance as a weighted combination of three mid-term and one final exam to determine the effects of SA participation on students' performance in the course.

Context, Participants and Intervention

The study took place at a research-intensive university with an enrollment of 27,000 in the southeastern part of the United States. Participants for this study consist of 90 university students enrolled in an introductory, sophomore level college genetics course. This course is a requirement for all biology majors and, additionally, taken by most students hoping to attend veterinarian, dental, nursing, or medical school.

This study is the result of a collaboration between a science education faculty (First author) and a university science professor (Second author) who is interested in enhancing his students' learning in his courses. Both faculty members met and discussed different strategies that could had been used to enhance student learning. After several meetings, the course professor decided to use self-assessment (SA) in his course. Next, the course professor and science education faculty collectively developed a SA protocol (see Appendix A) to be administered on a weekly basis through their university's online course management system, Blackboard.

The SA protocol was administered to the participants on an approximately weekly basis with a total of 10 SAs assigned over the course of the semester. Participation in the SAs made a 3% contribution to the students' final grades. Due to the large number of students in the course and the inefficiency of the online course management system interface, only a subset of responses of a week's SA were graded by the course instructor. Thus, the credit a student received was based on a random sample of 4 of the 10 SAs.

Data and Data Analyses

Data consisted of the weekly-administered SA results and the end of course exam reports. During the semester, the authors read the subset of self-assessments (SAs) that were to be graded. The purpose of these analyses was to receive immediate feedback on students' responses to the weekly instruction. At the conclusion of the course, the authors downloaded all of the data from the Blackboard, the course management program used by the university. In an effort to ensure confidentiality, students' names were removed and replaced with a unique, randomly generated number. Students' responses were combined across the SAs on a prompt-by-prompt basis as each prompt was designed to collect specific information related either to students' learning or teaching throughout the semester. Each prompt and example answers given to the students can be found in Table 1.

Table 1 Self-Assessment survey prompts and example answers provided to students in survey.

P#	Prompt Text	Provided Example Responses
P1	Please list the main concepts you have learned within the past week.	The role of transcription in genetics, transcription initiation, transcript elongation
P2	[On a scale of 1 to 10,] please indicate how well you think you understand the topics you listed under [P1]	None Given
P3	For each of the concepts you just listed [in P2] explain the quality of your understanding.	I understand the main concepts. I can apply my understanding to solve the simpler problems we've seen. I feel confident enough about my understanding to explain it to my classmates
P4	Explain how you ended up learning the concepts you understood well. You may list more than one reason.	The lecture slides were useful, especially the ones on initiation. Some of [Course Professor]'s explanations on elongation were too detailed for me to follow. I do remember some of this stuff from BIO130 and that helps
P5	Tell me why you did not understand the concepts you rated yourself low on. You may list more than one reason.	I don't really understand how the RNA polymerase knows where to start initiation and the whole sigma factor thing confused me. Are there more than 1? How many more?

P6	Tell me what I can do to help you improve your learning in this course.	Use more analogies in your explanations, use group work more frequently, go through more examples, slow down on the math
P7	Tell me what you can do to improve your learning in this course.	I could still spend some more time studying before class, I could ask more questions when I do not understand something in class

Briefly, the purpose of the first prompt (P1) was to have students reflect on the topics covered the previous week. Because the topics varied between weeks, we did not analyze student responses to P1 nor was there any value in analyzing such descriptive data. The purpose of the second prompt (P2) was to have students assess and quantify their understanding, on a scale of 1 to 10, of the topics they listed in response to P1. We analyzed these responses by looking at the average value students gave themselves.

The purpose of the third prompt (P3) was to have students provide a verbal description of their understanding of each topic they listed in response to P1. The purpose of the fourth prompt (P4) was to have students reflect on and explain how they developed their understanding of the topics they felt they understood well. The purpose of the fifth prompt (P5) was to have students reflect on and explain the reasons behind their lack of understanding of the topics they did not feel they understood well. The purpose of the sixth prompt (P6) was to have students reflect on how the material was presented in class and solicit constructive feedback from the students for the instructor on how he could improve his teaching of the material. The purpose of the seventh and final prompt (P7) was to have students reflect on and decide upon what steps they could take to improve their learning in the course in the future.

Responses to P3-P7 were open-ended textual in nature and, as a result, we analyzed student responses to these prompts using QDA-Miner software package. For our analyses we used QDA-Miner's frequency and phrase-finder functions to identify the most frequently used strategies or factors cited by the students. Due to the volume of data, we only report the strategies or factors that were most frequently cited by the students.

Self-Assessment and Conceptual Understanding. The course professor integrated various forms of assessments in his instruction. These include: self-assessments (SAs), class activities such as group work, quizzes, and homework assignments. Our assumption was that there would be a positive relationship between each of our dependent variables (SAs, class activities, quizzes, and homework assignments), and the response variable, the cumulative exams score. To test our hypothesis we used two-tailed t-test by fitting a linear model with all predictors first and then by systematically excluding the predictor variables with the lowest significance (i.e. the largest estimated p-value).

Results

Consistent with the order of our research questions, presentation of our findings are ordered as follows: 1) students' perceived level of understanding of the course content, 2) perceived effective instructional or learning strategies that helped students to develop conceptual understanding of course material, 3) perceived factors that impacted students' limited understanding of the course material, 4) perceived instructional strategies that could have helped the students to develop conceptual understanding of course material, 5) strategies that students could have implemented to enhance their conceptual understanding of the concepts covered in the course, and 6) the results of a statistical analysis of the impact of SA on students' conceptual understanding of course content.

1-Students' Perceived Level of Understanding of Course Content

One purpose of the SA project was to give the course professor the chance to identify students' level of understanding of the course material each week. The purpose of P1 was to receive feedback on whether the students were successful in identifying the key concepts covered each week. Since the topics varied between SAs, we did not analyze responses to P1. In contrast, responses to the remaining prompts were combined across SAs and analyzed together.

Responses to P2 and P3, which asked students to quantify their understanding of the topics covered, provided the instructor with information on whether the students felt they were learning the concepts they had identified as important. For P2, where students quantify their understanding of each topic on a scale from 1-10, the average value was 7.0533 ($n = 90$, standard deviation (SD)= 1.68829, coefficient of variation (cv)= 0.239361). These results indicate that students felt confident in their understanding (i.e. solid understanding) of core ideas at the end of instruction.

Students provided a verbal description of their understanding in response to P3. The results of our analyses are summarized in Table 2. It should be noted that we only provide the top three responses, which represented 90% of all of the student responses (Table 2).

Table 2 Students' perceived level of understanding.

Level of Understanding	Frequency	Percent
I understand main concepts	347	67.0
I can apply my understanding to solve similar problems	74	14.0
I understand it enough to explain to my classmates	98	19.0

2- Perceived Effective Instructional and Learning Strategies

A second purpose of this SA project was to identify teaching and learning strategies that contributed to students' conceptual understanding of the target concepts for each week. These most common learning strategies and the frequency at which they were listed are summarized in Table 3.

Table 3 Strategies and frequency of strategies that helped the participants to understand the course content.

Strategy	Frequency	Percent
Lecture slides	134	37.0
Reading the course material	63	17.0
Participation in in-class activities	44	12.0
Professors' explanations	35	9.5
Homework problems	28	8.0
Remembering information from previous courses	24	6.5
Asking questions	19	5.0
Studying	19	5.0

3- Perceived Factors That Impacted Students' Limited Understanding of the Course Content

A third purpose of this SA project was to identify factors that interfered with student performance in the course. This was achieved through P5. In general, students provided multiple reasons for their limited understanding of the course material. As mentioned in the previous section, the majority of the student responses stated that they had a good grasp of the concepts covered during the class. For the remaining concepts, the main causes listed by students for their limited understanding of these concepts are listed in Table 4.

Table 4 Perceived reasons for students' limited understanding of the course content.

Reason	Frequency	Percent
Confused/Confusing	104	55.0
Need more time to review the material	44	23.0
Limited understanding of formulas/equations and how to use them.	24	13.0
Lost in details	17	9.0

4- Perceived Instructional Strategies that Could Have Helped Students to Develop Conceptual Understanding

An additional purpose of this SA was to receive feedback on the perceived effectiveness of instruction and ways to improve the effectiveness of instruction. The suggestions proposed by the students for changes to the professor's instructional strategies, are summarized in Table 5.

Table 5 Perceived effective instructional strategies for helping students to understand the course material.

Suggested Strategy	Frequency	Percent
Provide more examples	116	29.0
Group work	94	23.0
Working through worksheets or problems	75	19.0
Clicker questions	61	15.2
Professor should continue to explain the lectures	38	9.4
Pictures representations	18	4.4

5- Perceived Learning Strategies That Students Could Have Helped Students to Develop Conceptual Understanding

Finally, a fifth purpose of this SA study was to help students develop learning strategies for themselves. Students were prompted in P7 to state what they could do to improve their learning of the course materials. Students suggested diverse strategies that they could have used to improve their performance in class. The nine most common strategies provided and their frequencies are provided in Table 6.

Table 6 Student learning strategies.

Strategy	Frequency	Percent
Asking questions	217	24.8
Spending more time studying the course material	168	19.2
Taking notes in class	134	15.3
Reading the textbook more frequently	108	12.3
Paying more attention to the lectures	93	10.6
Paying more attention to completing homework assignments.	63	7.2
Reviewing the course material	54	6.2
Taking more advantage of the discussion board	21	2.4
Preparing for the course	17	2.0

6- Participation in Self-Assessment as the Predictor of Conceptual Understanding

Finally, we wanted to see if participation in SAs had a positive effect on students' conceptual understanding as measured through cumulative exam scores. The course professor used multiple assessment activities in the course. Our assumption was that there would be a positive relationship between each of our dependent variables, (SAs, class activities, quizzes, and homework assignments) and the response variable, the cumulative exams score. To test our hypothesis we fitted a linear model between our dependent variables and student performance as a weighted combination of three mid-term and one final exam. Despite a relatively large sample size of 90 students, we found no significant effects between any of our dependent variables, SA, class activities, quizzes, and homework assignment scores, and the response variable, exam scores. The results of our analyses are summarized in Table 7.

Table 7 Relationship between instructional activities and cumulative exam scores.

	Estimate	Standard Error	T value	p-value
(Intercept)	0.62130	0.07868	7.897	1.61e-11 ***
SAs	0.08211	0.07693	1.067	0.289
Class Activities	0.07425	0.07776	0.955	0.343
Quizzes	0.02640	0.09139	0.289	0.773
Homework	-0.04732	0.12458	-0.380	0.705

Residual standard error: 0.1348 on 77 degrees of freedom, Multiple R-squared: 0.05793, F-statistic: 1.184 on 4 and 77 DF, p-value: 0.3247. ***= p<0.0005 for intercept.

In addition to fitting our full model, we also fitted several simpler models in which we eliminated dependent variables with the greatest p-value from our model. In none of these models were we able to detect a significant effect of any of the dependent variables on the outcome variable. Nevertheless, we do note that SAs consistently had the largest absolute effect on exam scores and, correspondingly, the lowest p-value of all of our dependent variables. So while our final and simplest model which included only SAs as a dependent variable suggests that students who did all of the SAs had an approximately 11% gain in exam performance, the effect was not significant at the $p<0.05$ level. The results of this last and simplest model are summarized in Table 8.

Table 8 Relationship between self-assessments and cumulative exam scores.

	Estimate	Std. Error	t value	p-value
(Intercept)	0.63695	0.04265	14.935	<2e-16 ***
SA	0.10733	0.05988	1.792	0.0768

Residual standard error: 0.1336 on 80 degrees of freedom, Multiple R-squared: 0.03861,F-statistic: 3.213 on 1 and 80 DF, p-value: 0.07684 , ***= p<0.0005 for intercept.

It should also be noted that since we could not directly manipulate student effort on SA, a substantial component of the positive relationship between effort on SAs and exam scores may be correlative rather than causative.

Discussion

Science educators have paid considerable attention to improving undergraduate students' learning experiences within the last two to three decades (Andrade & Valtcheva, 2009). Gopal et al, 2010; Rissing & Cogan, 2009; Taylor, Gilmer & Tobin, 2002). They have used multiple teaching strategies including the use of clickers, collaborative learning strategies, studio teaching and argumentation, and educational technologies (Beatty & Gerace, 2009; Dawson, Meadows & Haffie, 2010). However, science educators have rarely used SA in higher education science courses. In addition, there are limited SA models in science education literature. In this study, we attempted to address this issue in science education.

The results show that the SA strategy that we used is promising in making a range of information about the effectiveness of instruction and students' learning strategies accessible to the instructor. For instance, it gives the instructor the opportunity to assess students' level of understanding of key concepts covered for every unit and a chance to revisit the concepts with the students if needed. It also encourages students to reflect on their own learning: identify the concepts they are most confused about and those that are comprehended very well. For example, our analyses revealed that students referred to the complexity of the concepts covered in the course as one of the reasons for not understanding the course material. Additionally, students stated that they needed more time to review the course material before they could understand the course material. Some cited their limited understanding of the formulas and equations as the causes of their limited understanding of the course material. These comments were primarily related to the course content that required students to calculate probabilities of different possible outcomes.

In addition, these results indicate that students benefited the most from the course lectures and found explaining genetics concepts through examples contributed positively to their learning. This is not surprising because the course professor provided explanations of rather complex genetics concepts through his lectures and incorporated group-based activities and clicker questions. Similarly, it is not surprising that learning strategies such as reading the course material, attending lectures, completing homework problems, asking questions and studying contributed to student learning. What is surprising is the fact that direct, student studying of the material was towards the bottom of this list. It is not clear if this means students spent relatively little time outside of class studying or, alternatively, students spent substantial amounts of time studying the material but found it to be of little value.

The results showed that the complexity of the concepts covered in the course made it difficult for the students to learn on their own. Although the course professor's explanations supported by examples helped them to overcome such difficulty, students failed to execute the learning strategies such as asking questions that could have perceivably helped them develop a better understanding of course content before the exams. Online course management based discussion boards may be used as a venue for students to express their questions to the course professor or to the graduate teaching assistants after each session. However, in our experience and despite our repeated requests for students to do so, students rarely posted questions to the discussion boards. Possible reasons include students not wanting to appear "stupid" in front of their peers or finding it difficult to pose their questions using these tools.

The course professors can adopt several strategies to help their students to overcome the fear of asking questions. First, course professors can use anonymous communication venues to help their students. For instance, by using clickers the professor can gauge the level of students' confusion for each topic covered and provide additional explanations by using a variety of examples. Similarly, course discussion boards can be set in ways that will not identify students that are asking questions. Alternatively, the course professors can use collaborative learning strategies during class meeting times so that students can support each other to comprehend the course material.

Unfortunately we found that using group work frequently resulted in additional challenges for the course professor. First, doing the exercises in class generally takes substantial amounts of class time. As a result, the frequent use of group work interfered with instructor's ability to cover the intended course materials. Second, coming up with engaging ideas upon which to base these exercises was challenging for the instructor who has had no formal training or background in this area. Third, even when a suitable idea for an exercise was identified, the development of these ideas generally took several hours since they usually involved making figures or tables accompanying slides, and a solutions guide. The grading of the exercises also required additional instructor time. Although some of the assessment of student participation in the activities could be done via clickers or an online survey that is graded automatically, most grading could not be easily automated. Finally, after an exercise was developed there were frequently unforeseen problems that became apparent only once the exercise had begun. Although the instructor could generally address these problems in class, their existence undermined the effectiveness of the exercise and reduced the students' confidence in the instructor's abilities. Clearly, with more time and experience the number of problems would diminish over time. Nevertheless, these problems do represent substantial barriers for instructors wishing to begin using in class learning exercises.

A majority of students stated that reading the course material or studying the material in the course could have potentially helped them to develop a better understanding of the course material. Surprisingly, despite students identifying their need to study more, studying was at the bottom of the list of techniques we assembled. Further, the fact that student performance on unannounced in class quizzes was relatively poor, suggesting they did not complete their assigned textbook readings before class and the threat of an unannounced quiz was not a sufficient incentive for getting the students to prepare before class. An alternative approach would be to require students to take an online quiz or write a short summary of the assigned readings before each lecture. This strategy can be effective for several reasons. First, it will expose students to the technical terminology before they come to the class. Second, it can increase the probability of students' questions in class. Third, it can empower students with the prior knowledge needed to engage in group-based activities more effectively. However, the development of quiz questions or the assessment of student summaries would lead to a substantial increase in time the course professor spends on student assessment. The grading issue could be partly solved by randomly choosing and grading a subset of the essays every other week. This strategy can place pressure on the students to invest sufficient time on reading and reflecting on the course content before they come to class and thus more effectively engage in the lectures and other in class learning activities.

All of these suggest that SAs made a range of useful information accessible to the course professor. However, the course professor was not able to use this feedback right away to make substantial changes in his instruction to address his students' learning needs. Although the instructor was not able to use this information to make large changes to the instruction the year these SAs were implemented, they did prompt the instructor to redesign the course activities the following year.

SAs produce substantial amount of feedback for the instructor of a course with 90 students, especially when SA prompts are open-ended. The technology we used, Blackboard, did not have the functionality to summarize the results for the course instructor on a weekly basis. These summaries had to be done manually, which is very time consuming. Using a technology that can provide the instructor with more timely and precise summaries of students' understanding after each SA is implemented could provide valuable information for the course professor to gauge his or her students' level of understanding for each concept and respond accordingly and timely by making modifications to his teaching and planning. We plan to use the lesson we learned through this study to develop such a computer-based system in the near future, implement it in the classroom and monitor its effect on student learning.

Limitations

As is the case with many education studies, there are several limitations to this study. We recognize these limitations and caution our readers to keep these limitations in mind as they consider the application of the methods and results reported in this paper for their specific contexts. First, we believe that the study design could have been improved by having a control treatment where students do some kind of 'sham' activity or, perhaps, an alternative activity to SA. We could then randomly assign students to these treatments and then directly test for the effect of SAs. Second, although our analysis indicates that the completion of SAs provided useful information to the course instructor, the value of these reflections to the course professor was limited. This was primarily due to the fact that the extensive amount of data made it difficult for the course professor to process the students' responses in a timely manner. This illustrates a need for the development of the necessary software tools to analyze and present SA data in an efficient manner.

References

Aydeniz, M. (2007). *Understanding the challenges of enacting assessment reform in public high school science classrooms.* Unpublished doctoral dissertation, Florida State University, Tallahassee, FL.

Aydeniz, M & Pabuccu, A. (2011). Understanding the impact of formative assessment strategies on first year university students' conceptual understanding of chemical concepts, *Necatibey Faculty of Education, Electronic Journal of Science and Mathematics Education, 5*(2), 18-41.

Airasian, P. W. (2001). *Classroom assessment: Concepts and applications* (4th ed.). Boston: McGraw Hill.

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory into Practice, 48(*1), 12-19.

Balinsky, M. G. (2007). *Forging* an *identity*: *Four science doctoral students* in a *collaborative partnership* with *K-12 science teachers*. Unpublished doctoral dissertation, Tallahassee, FL.

Beatty, I. D., & Gerace, W. J. (2009). Technology-enhanced formative assessment: A research-based pedagogy for teaching science with classroom response technology. *Journal of Science Education & Technology*, *18*(2), 146-162.

Berry, R. & Adamson, B. (Eds.) (2011). *Assessment Reform in Education: Policy and Practice*. Dordrecht, Netherlands: Springer.

Berry, R. (2008). *Assessment for learning*. Hong Kong: Hong Kong University Press.

Berry, R. (2006). Activating learners using the learner autonomy approach: An action research on the relevance of teaching to classroom practice. *Curriculum Perspectives, 26*(3), 34–43.

Black, P., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*, 139-148.

Black, P., Harrison, C., Lee, C., Marshall, B., &William, D. (2003). *Assessment for learning: Putting it into practice*. Buckingham, UK: Open University Press.

Bloxam, S. & Boyd, P. (2007). *Developing effective assessment in higher education: a practical guide.* Berkshire: Open Link Press.

Dawson, D., Meadows, K., & Haffie T. (2010). The effect of performance feedback on student help-seeking and learning strategy use: Do clickers make a difference? *The Canadian Journal for the Scholarship of Teaching and Learning*, 1(1).Retrieved from http://ir.lib.uwo.ca/cjsotl_rcacea/vol1/iss1/6/

Furtak, E. M., & Ruiz-Primo, M. A. (2008). Making students' thinking explicit in writing and discussion: An analysis of formative assessment prompts. *Science Education*, 92, 799-824.

Gopal, T., Herron, S. S., Mohn, R. S., Hartsell, T., Jawor, J. M., & Blickenstaff, J. S. (2010). Effect of an interactive web-based instruction in the performance of undergraduate anatomy and physiology lab students. *Computers & Education, 55*, 500 - 512. Retrieved from http://dx.doi.org/10.1016/j.compedu.2010.02.013.

Kearney, S. P., & Perkins T. (2010, Nov). *Developing students' capacity for innovation, creativity and critical thinking through contemporary forms of assessment: a case study in progress*. Paper presented at ATN Assessment Conference, University of Technology, Sydney.

Lin, H.-S., Hong, Z.-R., Wang, H.-H., & Lee, S.-T. (2011). Using reflective peer assessment to promote students' conceptual understanding through asynchronous discussions. *Educational Technology & Society*, *14* (3), 178–189.

McMillan, J. H., &. Hearn, J. (2009). Student self-assessment: the key to stronger student motivation and higher achievement. *The Education Digest*, *74*, 39-44.

National Research Council [NRC]. (2001). *Knowing what students know: The science and design of educational assessment.* Washington, DC: National Academy Press.

Rissing, S.W. & Cogan, J.G. (2009). Can an inquiry approach improve college student learning in a teaching laboratory? *CBE Life Science Education,* 8, 55-61.

Ross, J.A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment*, *Research*, and *Evaluation*, *11(*10*),* 1-13

Schunk, D. H. (2001). Social cognitive theory and self-regulated learning. *In* B. J. Zimmerman & D. H. Schunk *(Eds*.), *Self-regulated learning academic achievement: Theoretical perspectives*(*2nd ed.*, pp. *125–151).* Mahwah, NJ*: Erlbaum.*

Siebert, E. D. (2001). Science education program standards*. In E.D.Siebert & W.J .McIntosh (Eds.), College pathways to the science education standards (pp. 115–138).* Arlington, VA*:* National Science Teachers Association Press.

Stiggins, R. J. (2004). *Classroom assessment for student learning*. Portland, OR: Assessment Training Institute.

Sundström, A. (2004). *Self-assessment of knowledge and abilities* – a *literature study*. (Em no 54). Umeå university: Department of Educational Measurement.

Schunk, D. H. (1996). Goal and self-evaluative influences during children's cognitive skill learning. *American Educational Research Journal, 33*(2), 359-382.

Taylor, P., Gilmer, P., & Tobin, K. (Eds) (2002). *Transforming undergraduate science teaching: Social constructivist perspectives*. New York, NY: Peter Lang Publishing

Zoller, U., Tsapaiiis, G., Fatsow, M., & Lubezky, A. (1997).Student self-assessment of higher-order cognitive skills *in* college science teaching*. Journal of College Science Teaching, 27*(2), 99-101.