

Simple Formative Assessment, High Learning Gains in College General Chemistry

Yalçın Yalaki*

Suggested Citation:

Yalaki, Y. (2010). Simple formative assessment, high learning gains in college general chemistry. *Egitim Arastirmalari - Eurasian Journal of Educational Research*, 40, 223-241.

Abstract

Problem Statement: The utility of formative assessment for improving student performance is acknowledged in the literature. For formative assessment to be useful, quality feedback, the proper use of feedback by students and revision of practice to accommodate formative assessment is necessary. However, in the contexts of university level science courses, where, in many cases, there is a heavy load of teaching and assessment on both students and instructors, providing quality feedback, using feedback properly, and changing practice may not be as forthcoming as hoped.

Purpose of Study: The purpose of this study is to show that providing opportunities for formative assessment, even at a relatively simple and manageable level, could significantly impact students' achievement and also their attitude toward university level science classes without a considerable change in practice.

Methods: This study was conducted in an undergraduate general chemistry course with an emphasis on organic chemistry taken by pre-service middle school science teachers over the period of two spring semesters in two consecutive years. In total, 163 students participated in the study. A qualitative research methodology accompanied quantitative methods for more in-depth understanding. Summative exam results, responses to a questionnaire, observations, and interview transcripts provided the data for the study.

Findings and Results: Quantitative results showed statistically significant improvement in experimental group students' exam grades, which was an indication of improvement in achievement and learning. Through qualitative data, positive student reactions toward the formative

* PhD, Hacettepe University, Faculty of Education, yyalaki@hacettepe.edu.tr

assessment process used in the study were observed. Insight into students' perceptions of the formative assessment methods and their use of feedback were obtained.

Conclusions and Recommendations: The relatively simple formative assessment process utilized in this study has the potential to improve both student achievement and learning and also students' attitudes toward the courses they take. Therefore, university level instructors, who teach science courses, should consider using formative assessment methods, at least at a simple level without a too heavy a load of extra work, to improve student achievement and attitude. However, it is not possible to claim that formative assessment methods will improve student achievement and attitude in all cases.

Keywords: Formative assessment, formative feedback, peer-feedback, self-assessment, science education, higher education

The utility of formative assessment for improving student learning is acknowledged in the literature. Widely cited publications about formative assessment such as the meta analysis by Black and Wiliam (1998) and a book published by Organisation for Economic Co-operation and Development [OECD] (2005): *Formative assessment: improving learning in secondary classrooms*, claim that formative assessment has a high potential of providing positive learning gains among other educational innovations. However, applying formative assessment in classrooms is a challenge for many teachers and instructors, especially in higher education (Carless, 2007; Yorke, 2003). The purpose of this study is to demonstrate that in classrooms without severe motivation challenges, even simple applications of formative assessment can be highly successful in terms of student attitude and learning.

According to the CERI (2005) book "Formative assessment refers to frequent, interactive assessments of student progress and understanding to identify learning needs and adjust teaching appropriately (p. 21)." Successful applications of formative assessment have several features. Nicol and Macfarlane-Dick (2006) identify the following principles of good feedback practice in formative assessment:

Good feedback practice:

1. helps clarify what good performance is (goals, criteria, expected standards);
2. facilitates the development of self-assessment (reflection) in learning;
3. delivers high quality information to students about their learning;
4. encourages teacher and peer dialogue around learning;
5. encourages positive motivational beliefs and self-esteem;
6. provides opportunities to close the gap between current and desired performance;
7. provides information to teachers that can be used to help shape teaching (p. 205)

These principles summarize what features an ideal feedback practice in formative assessment would have. However, there are many factors that affect the success of

formative assessment such as learning context, student motivation, student-teacher relations, and students' level of cognitive development (Elwood 2006; Higgins, Hartley & Skelton, 2002; Perrenoud 1998; Yorke 2003). These factors influence how many of the principles above could be utilized effectively by instructors and students. Higgins et al. (2002) ask important questions about student motivation and the practice of feedback:

For formative assessment to work in practice, feedback must 'connect' with students. But, at a time when student numbers are rising and competition for graduate jobs is growing, are students increasingly becoming instrumental consumers, driven by the extrinsic motivation of the mark? If so, will they heed written feedback which encourages them to reflect on their learning? Or will they simply pay attention to the grade, and seek feedback only when it is perceived to provide 'correct answers' to commit to memory (and only then when their grade expectation has not been met)? (p. 54)

They answer these questions by arguing that even if grades and competition for jobs are a source of motivation for students, students do use feedback in many different ways. Perrenoud (1998) expresses the importance of learners in formative assessment, reminding us that the idea of feedback and formative assessment would be incomplete without the learner and that 'feedback does not in itself guarantee better learning.

Feedback is a simple message. How can it assist the learning process? Because pupils take it into account, because it affects their cognition. Communication theory teaches us that the effectiveness of a message is measured at the level of the recipient: an intervention or a piece of information only helps a pupil learn better if their thought processes are modified. This is an abstract way of stating that no learning takes place without the learner. One can only stimulate, reinforce, reorient, readjust or accelerate the pupil's mental processes, in the hope of modifying the learning processes. (p. 86)

Yorke (2003) agrees by arguing that "the importance of the student's reception of feedback cannot be overstated." (p. 488) Students' goals, motivation, personality orientation, self-efficacy and other factors influence how they receive and use the provided feedback. For example, self-efficacy plays an important role on how individuals regulate the use of feedback, among other things, which varies from person to person (Akbaş, 2010; Erdem 2008). On the other hand, providing quality feedback is a challenge for teachers and instructors. As Black and Wiliam (2009) argue, teachers usually have only a few seconds to interpret student contributions and choose an optimum response. From these arguments, it is possible to conclude that using formative feedback is a complex process on the side of both students and teachers and does not guarantee improved learning. As Yorke (2003) argues, establishing a direct connection between formative assessment and learning and achievement warrants caution. Black, Harrison, Hogden, Marshall and Wiliam (2005) also agree that formative assessment does not guarantee improvement in student achievement in all cases.

We do not claim that formative assessment leads to improved student achievement in *all* cases, with *all* teachers on *all* occasions. ... Our claim is that formative assessment in *general* is an effective intervention, although we do not estimate the difficulties in translating theory into practice. (p. 7, original emphasis)

Carless (2007) argues that "In particular, large class sizes and heavy workloads often present a barrier to teachers' implementation of formative assessment." (p.

173). He recognizes that "...there is a need to find ways to make formative assessment more attractive and manageable for practitioners..." (p. 174). Higgins et al. (2002) and Yorke (2003) also mention the problems regarding the implementation of formative assessment. They also argue that in higher education, the challenges that instructors face in utilizing formative assessment include high student/instructor ratio, other instructor responsibilities than teaching, and emphasis on summative assessment by the institutions. Hence, practicing formative assessment by providing quality feedback, using feedback on the part of students, and adjusting teaching may not be a simple matter in higher education.

Recognizing the difficulties faced in the practice of formative assessment, many authors still agree that formative assessment is a part of good teaching practice (Ahmed & Teviotdale, 2007; Black et al., 2005; Black & Wiliam, 2009; Carless, 2007; Sadler, 1998; Yorke, 2003). This study is based on the argument that formative assessment has a potential to improve student learning. The problems faced in utilizing formative assessment could be surmounted by simple approaches. In order to keep instructor workloads relatively unchanged while implementing formative assessment practices in teaching, some choices could be made on the type of feedback to utilize. The feedback in formative assessment can take different forms such as, written feedback, oral feedback, grading, peer feedback and self-assessment (Sadler, 1998; Black, 2003; Nicol & Macfarlane-Dick, 2006). All these feedback forms have their advantages and disadvantages. For example, written feedback may be more time consuming for instructors, while oral feedback can be given more easily and in a more timely manner. Peer feedback and self-assessment provide better reflection on the part of learners (Liu & Carless, 2006). Encouraging peer dialogue during the formative assessment process promotes a sense of ownership among students while enhancing learning (Nicol, 2007). In this study, oral feedback, peer feedback and self-assessment were utilized, because they required less work on the part of instructors. This decision was important since the purpose of this study was to demonstrate that least amount of additional work on the part of instructors, who are already busy with other occupations, could still stimulate useful feedback practices. Besides, an important role of formative assessment is to improve student self-regulation in learning, which would be improved by peer feedback and self-assessment (Nicol, 2007 and 2009).

The research questions of the study were:

- 1- Could formative assessment improve pre-service middle school teachers' achievement in chemistry classes?
- 2- What are the perceptions of pre-service middle school teachers toward the utilization of formative assessment in chemistry classes?
- 3- How do pre-service middle school teachers use feedback in a formative assessment process?

A mixed qualitative and quantitative research methodology was used in this study with a strategy of case oriented action research. Even though these two methodologies have very different assumptions about the phenomena to be investigated, mixed methodologies involving both quantitative and qualitative research are advocated and used in educational research (Johnson & Onwuegbuzie,

2004). The reason for choosing this mixed methodology for this research was the richness of understanding achieved by the complementary use of the two methods at the same time. Sale, Lohfeld and Brazil (2002) argue that these two methodologies, if used together, should be used in a way that they complement each other.

In this study, the phenomena investigated by quantitative and qualitative approaches can be differentiated easily. The quantitative data provided information on the general progress and views of participants while qualitative data provided more in-depth and enhanced understanding of their perceptions of the formative assessment used in this study and how they utilized it. Based on Johnson and Onwuegbuzie's (2004, p.22) conceptualization of mixed methodology, this study used quantitative and qualitative methods concurrently and with equal emphasis.

Method

The study was conducted at a major university in Turkey in a general chemistry course, with an emphasis on organic chemistry, taken by second-year pre-service middle school science teachers over the period of two spring semesters in two consecutive years. 163 students participated in the study. The study was designed in a way to compare the achievement scores of students who took the class in 2007 with those who took it in 2008. The 2007 class received regular instruction without a formative assessment process and limited feedback while the 2008 class received instruction with a formative assessment process and deliberate feedback. The 2007 class consisted of two groups with 47 and 48 students in each group. This grouping was done to better manage classroom space and class schedules. Both groups received two midterm exams and one final exam. The 2008 class also consisted of two groups for the same reasons each containing 34 students. Instead of conducting the research with two groups in the same year, the reason for this sequential design was to prevent control and treatment groups from interacting with each other and also preventing a feeling of unfairness among students. The students were assigned to the groups based on the alphabetical order of their last name, which provided a level of randomization in group assignment of individuals. However, convenience sampling rather than random sampling was employed in this study, since the purpose of the study was reporting a case and not generalizing the results to a population.

The 2008 class received four quizzes without grading as treatment for the purpose of providing formative feedback to students on top of the required two midterm exams and one final exam. Feedbacks on the midterm exams were also provided to the 2008 class. Midterms and final exams were used as a measure of achievement for both classes. The tests applied for the midterms and final-exams provided the quantitative data for the study. Reliability and validity studies were completed for all of the tests applied. Table 1 shows the reliability statistics of the tests. As seen in the table, because of conflicting schedules, the groups in the 2007 class received different tests, but they were similar in structure and content, with same number of items. The groups in the 2008 class received the same midterm and final tests. 2007 and 2008 tests contained different but similar items and the format and structure of all the tests were the same. The test items were classical paper and pencil items, most of which required short answers. However, in order to calculate a

reliability score, all of the test items were coded as “right” or “wrong” even though partial credits were given for partially correct answers. The point worth of items varied, but, for example, for a test item worth 10 points, if a student got 6 points or more for the item, the item was coded as “right,” and if the student got less than 6 points, then the item was coded as “wrong.” Even though the results averaged out to some degree, some information was lost in this process regarding the grades of the students. However, all of the tests applied provided high reliability scores at the end of this process, as shown in Table 1.

Table 1*Reliability Statistics of the Applied Tests*

| | 2007 Class | | | | 2008 Class | |
|-----------|---------------------|------|---------------------|------|---------------------|-------------------|
| | Group A (47 Std.) | | Group B (48 Std.) | | Group A (34 Std.) | Group B (34 Std.) |
| | Number of questions | KR21 | Number of questions | KR21 | Number of questions | KR21 |
| Midterm 1 | 24 | 0.80 | 24 | 0.82 | 25 | 0.74 |
| Midterm 2 | 17 | 0.84 | 17 | 0.82 | 21 | 0.79 |
| Final | 23 | 0.79 | 23 | 0.73 | 30 | 0.71 |

In order to achieve content validity for the tests, a table of the course objectives was developed. The number of questions that needed to be asked for each objective, and at what level, was decided prior to the administration of the tests. Most of the questions in the test measured procedural knowledge, which required application of some rule, while other questions measured declarative knowledge (Oosterhof, 2001).

Another instrument that was developed for this study was a questionnaire with a Likert type scale to measure students’ interests toward organic chemistry and the class they took, as well as their reactions toward the formative assessment process utilized, which contained 26 items plus four open ended questions. Each item had five possible responses ranging from “Do not agree at all” to “Totally agree.” The students were asked to write a pseudonym rather than their names on the questionnaire papers to ensure anonymity. A frequency analysis of the items in the instrument was performed after gathering students’ responses in a pilot study; the opinions of four colleagues about the items were solicited, and some revisions were made based on these findings to ensure validity. The reliability score of the instrument was found to be Cronbach alpha = 0.896. At the end of the study, 66 students from the 2008 class had returned the instrument with their responses. The questionnaire provided both quantitative and qualitative data regarding students’ views, reactions and perceptions.

As the author and investigator of this study, I taught all of the courses, conducted the quizzes and exams and also evaluated them. As mentioned before, a case-oriented action research strategy was employed to conduct the study. I was a participant observer and investigator at the same time. During the formative assessment process, I took notes about my observations related to students' use of the formative feedback that they got from their peers and me as the instructor.

Finally, semi-structured interviews were conducted with seven volunteer students to collect further data on students' perceptions of the formative assessment process. Both the open ended questions in the questionnaire and interviews provided detailed data on the students' opinions about the formative process utilized in the class and also how they used the formative feedback generated during this process. The interviews were recorded and transcribed for data analysis.

Data Analysis

As mentioned above, quantitative and qualitative data were collected in this study. The administered test results and the scores obtained from a Likert type questionnaire served as the quantitative data. For analyzing these data, an independent sample t-test and a frequency analysis were used, respectively. Answers given to the open ended questions in the questionnaire, observation notes, and interview transcripts made up the qualitative data. To analyze the qualitative data, codes were developed in order to find themes and patterns in the data (Bogdan & Biklen, 2002), which helped with interpreting the data and also with the process of making meaning out of data.

Findings and Discussion

The quantitative data gathered through the applied tests revealed a significant difference in the scores of both classes. When the test results of the combined groups for each class were compared, it was found that the 2008 class, in which formative assessment was utilized, had significantly higher scores in all of the tests. Table 2 show the related statistics with the assumption that the homogeneity of variance is not violated for all the cases.

Table 2

Group statistics (Chemistry 4 test results are out of 100)

| | Class | N | Mean | Std. Dev. | Std. Error Mean | p (2-tailed sig.) |
|-----------|-------|----|------|-----------|-----------------|-------------------|
| Midterm 1 | 2008 | 68 | 87.3 | 10.38 | 1.26 | <.001 |
| | 2007 | 95 | 66.0 | 19.18 | 1.97 | |
| Midterm 2 | 2008 | 68 | 76.5 | 18.16 | 2.20 | <.001 |
| | 2007 | 95 | 63.9 | 23.66 | 2.43 | |
| Final | 2008 | 68 | 68.8 | 15.53 | 1.88 | .010 |
| | 2007 | 95 | 61.4 | 19.45 | 2.00 | |

No pre-test was administered in this study, because there was evidence that showed all of the participating students had similar competencies in the beginning. When their average achievement scores for the previous three chemistry classes that they took prior to this study are compared, as shown in Table 3, there is no significant difference. The first two previous chemistry courses were taught by the same instructor for both groups, with similar tests in content and design. The third previous chemistry course was taught by the author of this study to both groups, again with similar tests in content and design. The tests applied to both groups in these three classes were similar enough in content and design to allow comparison, but the questions had to be different enough to prevent 2008 class students from having an advantage over the 2007 class students by collecting information on previous tests. This information was considered enough to assume that the compared groups had similar achievements in chemistry in the beginning.

Table 3

Comparison of average achievement scores of the previous three chemistry classes

| 2007 Class | 2008 Class |
|---|---|
| average of the achievement scores for Chemistry 1, 2, and 3 classes | average of the achievement scores for Chemistry 1, 2, and 3 classes |
| 73 (out of 100) | 74 (out of 100) |

When it comes to the questionnaire data, the 26 items in the questionnaire were designed to collect information about students' perceptions about the course that they were taking, the content of the course, the instructor, and also the formative assessment practices used in the course. As noted earlier, 66 participants returned the questionnaire with their responses. The frequency analysis of the responses is shown in Table 4. The responses are given either as the total of positive responses (agree + strongly agree) with a + mark or as the total of negative responses (disagree + strongly disagree) with a - mark.

Table 4

Student responses to the items on the questionnaire (n=66)

| Questionnaire Items | Responses (%) |
|--|---------------|
| 1 This course will be useful to me in my job in the future. | +80.3 |
| 2 This course includes information that I can use in daily life. | +75.8 |
| 3 I value what I learn in this course. | +90.9 |
| 4 If we didn't take this course, we would not lose anything. | -86.4 |

| | | |
|----|---|-------|
| 5 | The taught subjects were in agreement with the course objectives | +94.0 |
| 6 | The way this course was taught was in agreement with the course objectives | +81.8 |
| 7 | I did not fully learn the taught subjects | -59.1 |
| 8 | I am interested in Organic Chemistry. | +63.6 |
| 9 | The instructor was not very competent in course subjects | -92.4 |
| 10 | Organic chemistry is over my capacity of comprehension | -84.8 |
| 11 | This course could have been taught in a better and clearer way | -57.6 |
| 12 | This course has no contribution to my job training | -80.3 |
| 13 | The course content was not suitable for my expectations | -86.4 |
| 14 | I find organic chemistry totally uninteresting | -80.3 |
| 15 | The instructor gave the course in a way that was clear and easy to understand | +93.9 |
| 16 | I did not learn anything useful in this course | -80.3 |
| 17 | I learned the subjects well that was discussed in the course | +77.3 |
| 18 | The instructor was insufficient in terms of teaching techniques | -77.3 |
| 19 | The instructor knows organic chemistry very well | +92.4 |
| 20 | Organic chemistry is one of the easy classes for me | +40.9 |
| 21 | The quizzes applied in this course were very useful. | +94.0 |
| 22 | The quizzes applied in the course helped me learn. | +95.5 |
| 23 | There should have been more quizzes. | +42.4 |
| 24 | Quizzes made no contribution to my learning. | -97.0 |
| 25 | Nothing would have changed if no quizzes were applied. | -97.0 |
| 26 | Having too many quizzes was boring. | -92.4 |

The findings in Tables 2 and 3 imply a higher level of chemistry achievement in the 2008 class. However, it is not possible to claim that the sole reason for the improvement observed in the achievement scores of the 2008 class is the formative

assessment process. There are many variables that influence a group of students' achievement that are not possible to control. For example, one factor that may have affected the achievement of students is the reduced class sizes in the 2008 group. Other factors may include the frequent changes in the central exam and student placement system for university entrance in Turkey. The content of questions asked in the national university entrance examinations, the rules of these examinations, and the number of students accepted to a particular program each year have changed over the years, which may have influenced the backgrounds of students who entered university in different years. However, the percentile range of students that enter the particular program in the university where this study was conducted did not change significantly over the years. This provides some confidence about the conclusion that at least some of the increased achievement scores of the 2008 class, when compared to the 2007 class, come from the utilized formative assessment process.

Students' very positive reactions toward the use of quizzes for formative purposes support this conclusion to some extent. As seen in Table 4, the vast majority of the students thought that the applied quizzes contributed to their learning. They also had mostly positive attitudes toward the class and its content, which is a crucial issue that influences how much they value and utilize the formative assessment process. Without this positive attitude it would be much more difficult for an intervention to have a positive affect. The reason why students had positive attitudes toward the course is because they all have a relatively strong science background, since they chose to study in the scientific field in high school. The positive reactions of students toward the formative assessment process that was used in this study were also reflected in their responses to some of the open ended questions in the questionnaire that was given to them.

As indicated before, quizzes were given to students for formative purposes. Before giving examples of students' responses, the application of quizzes should be made clearer at this point. As explained in the methods section, four quizzes were administered in the 2008 General Chemistry class during one semester. The first two quizzes were administered before the first midterm examination, while the last two were administered before the second midterm examination. Students were told that the quizzes were not going to be graded in order to focus their attention on the learning aspect of the quizzes rather than the grades. Not grading the quizzes also helped with decreasing the grade-related stress on the part of students. Also, not having to grade the quizzes kept the workload of the instructor virtually unchanged. The quizzes were given to students for a period of about 30 minutes and during this time they were encouraged to solve the quiz questions on their own to see where their weaknesses and misunderstandings were. After this time, students were asked to swap papers with a friend and compare their answers. Then those who swapped papers were encouraged to discuss with each other and try to find out the reasons for their different answers, if any. During this time, students also discussed the answers in small groups and had an opportunity to interact with their friends' ideas about quiz questions. They were able to hear their friends' answers for questions that they could not answer. After this period, a break was given, and when the students came back, the instructor answered questions during a whole class discussion. During the small group and class discussions, the instructor encouraged the students to find their weaknesses and misunderstandings. In order to aid students in improving their

learning, the instructor adjusted the teaching by making further explanations, giving different examples, and solving more problems instead of just continuing onto the next subject.

The open-ended questions related to the given quizzes in the questionnaire were: "What are your thoughts about the quizzes given in the class? Do you think that quizzes contributed to your learning? How should the quizzes have been done to be more useful?" Almost all students gave positive answers to these questions. One example of a student response was:

I think that quizzes given in the class were definitely useful, because they allowed us to see where we are lacking and our mistakes before the course advanced further. This way we could complete our knowledge in these subjects. Besides, thanks to the quizzes we were able to solve different examples. The way quizzes were conducted was nice. (Student response to open ended questions)

Another students' response was:

Students usually study from exam to exam and because of this many subjects remained uncomprehended before the exams. Thanks to quizzes, these subjects were better understood. (Student response to open ended questions)

There were many more examples similar to the quotes above. Besides quizzes, similar feedback was given about the midterm examinations as well. Students had a chance to review their exam papers, compare their answers with their friends and finally learn from their mistakes during whole class discussions. In this process, students performed self-assessment and were also exposed to peer feedback and instructor feedback. As a result of this process, some subjects, where many students had shortcomings, were taught and explained in different ways with different examples. This was also a unique experience for the instructor, who did not have a chance to see and intervene with his students understanding and learning to this extent before.

Interviews provided further insight into students' perceptions of the formative process used in this study and also how they made use of the feedback that they got from peers and the instructor. One of the findings that came out of the interview transcripts had to do with the influence of the quizzes that was used for formative assessment on how often the students studied for the class. When asked "How often do you study for this class at home?" one of the students, Ferda (pseudonym), explained:

As often as the quizzes... Quizzes were very good, in that they made this class one of the easiest in terms of preparing for the midterms and the final exam. Because we did not wait till the last moment to study. You know, in general it is student psychology to study for an exam only a week before, but this class wasn't like that. (Interview transcripts with Ferda 22.05.2008)

Another student, Tulin's (pseudonym) response to the same question was:

I was studying right before the quizzes. Other than that I was taking regular notes in the class, I have all the notes. ...[The quizzes] definitely influenced. ... Since no grade was given, nobody memorized as they do before the exams. We studied for learning really. (Interview with Tulin, 21.05.2008)

Finally, Yelda (pseudonym) had this to say about this question:

... I will generalize about all students; we leave everything to the last day. If there were no quizzes, we probably would have studied chemistry only the last day [before the exams] but since there were quizzes, we at least looked at what we have done in the quizzes, we could see where we were lacking, and we could specify which subjects to study. This, I believe, encouraged us to study gradually; we would not just go to the exam and feel disappointed, we knew what we know and what we don't. (Interview with Yelda, 26.05.2008)

These responses show that quizzes did effect how often students studied for the class. Most students usually study for a class just before the exams, but the quizzes applied in this study appeared to have increased the frequency of study and review on the part of students. Another issue that the interview findings revealed was that most students attach great importance to peer-review during the application of quizzes. One of the students, Mine (pseudonym), said,

I think reviewing the quizzes among ourselves is very useful, because, it matches human psychology, what your friend says may be more effective and longer lasting in memory than what the teacher says, because our way of thinking is similar. I think it was necessary, I mean, it was logical to review with our peers. (Interview with Mine, 22.05.2009)

Another student, Aynur (pseudonym) agreed,

I think sharing is better... It was good to discuss with friends. I think I learned better. The more I hear information from my friends, the better. No matter how much I study on my own, if I can't learn something, I can't learn it. But when there is interaction, everything is faster. (Interview transcripts with Aynur 21.05.2008)

Tulin also agreed and she argued that it is better to do peer-review than self-assessment, because this way, students cared more about what they did in the quizzes.

It was good to do exchanges among friends. ... I learned there. ... When we gave [our papers] to friends, it was more effective, because, I don't know, maybe it is this worry that my friend will think that I don't know anything, but at that time there was this psychology of trying to do more, since somebody else is going to review it; that was very effective. It is good to have friends to evaluate you, otherwise when there is only self-assessment, my thinking is "I am going to study anyway" and people may care less. (Interview with Tulin, 21.05.2008)

One student, Uhde (pseudonym), however, did have reservations about peer-review. She said,

...There were things that I learned incorrectly during the review of the first quiz. We discussed with friends, I placed it in my memory as I learned this, as if it was right. As I was studying, I looked at the exam and said, "This is how I learned it." I mean, instead of discussing with friends, well, maybe we should do that too, but it was more effective for me when you solved the questions on the board. (Interview with Uhde, 27.05.2008)

Finally, another student, Umit (pseudonym) said this about peer-review,

If we knew the subject, we could see where our friends were lacking, if we didn't, we could say "this is how my friends did this, why didn't I think about that. (Interview with Umit, 26.05.2008)

From these findings, it is obvious that most students in this study valued peer-review and they saw this process as something that aided their learning. All of the

interviewed students, except Mine, thought that it was good that no grade was given to the quizzes. They said that not giving quizzes took away the stress associated with grades and expressed the learning aspect of the quizzes. Mine, on the other hand, thought that giving at least partial credit to the quizzes could have made it more important for students and they would have taken it more seriously. Others disagreed. Mine's comments about giving grades on the quizzes came after the question related to the last two quizzes. During the last two quizzes, students seemed less interested and less engaged. When I asked the reason, Mine thought that if grades were given to the quizzes, students would have taken them more seriously. However, all other students thought that not giving grades was not the problem. The problem was that during the last two quizzes students were so busy with other homework and exams that they were just exhausted. They thought that giving grades to the quizzes would have made it a negative experience; students would have stressed out on top of being exhausted and this would not have helped their learning at all.

Another issue that was raised during the interviews was the importance of quizzes. When students were asked, "What is the most important contribution of the quizzes, helping you learn or helping you to improve your grades?" all of the students responded that both are important. Therefore, at the college level, where grades are very important for students for many good reasons, it is possible to conclude that learning is not easily detached from grades. This poses a challenge to formative assessment, since the purpose of formative assessment is to improve learning rather than giving summative grades. This is why no grades were offered during the formative assessment process during this study.

When it comes to the types of questions that were asked in the quizzes, students had mixed responses. While most agreed that the quiz questions were fine, there were suggestions to increase the types of questions or add questions that were more related to daily life. Five of the interviewed students thought that quiz questions should be similar to the exam questions, while two of them said that this was not necessary. For example, Ferda thought that if the quiz questions were very different than those on the exam, the value of the quizzes would have been diminished. This thinking is in line with the importance attached to exam grades and exam preparations. Most students thought that exam preparation was extremely important in addition learning.

Observation notes revealed that most students were engaged during the quizzes, discussing with each other while answering questions. Observation and interview data also revealed that students used feedback that came from their peers or the instructor to review their learning. There were always a few students who were disengaged during the quizzes for various reasons. The number of disengaged students increased as the semester progressed, and as the interviews revealed, this was due to the higher work load for students.

Conclusion

It can be argued from the evidence presented in this study that formative assessment, even at a relatively simple level, could significantly improve

achievement in science classes in higher education, if instructors are prepared to utilize it. Even though there may be other factors that influenced the improved achievement scores in this study, the formative assessment process seemed to have a positive contribution. Many studies confirm this positive contribution in higher education in many different fields such as health sciences (Carrillo-de-la-Penā et al., 2007), medicine (Kibble, 2007; Dobson, 2008), biology (Peat and Franklin, 2002), psychology (Costa, Mullan, Cothe & Butow, 2010), and math (Lawson, 1999). Kibble (2007) and Dobson (2008) report using online quizzes as formative assessment tools with positive results. Most of the studies above used online or computerized tests as tools for formative assessment and the point of doing so was to reduce the workload on the instructor while utilizing formative assessment.

The formative assessment process used in this study was also designed in a way to prevent adding too much workload onto the busy instructors while still providing further learning opportunities for students. From the findings of this study, it could be asserted that a formative assessment process using quizzes or short exams as a means to assess student learning and provide feedback should take the following factors into consideration:

- No grades, even partial credit, should be given to quizzes. This helps focus students' motivation on the learning aspect rather than grades.
- Peer feedback is important and provides students opportunities to interact with their peers while learning. Sometimes peers use language that is easier to understand for other students, and may aid learning.
- Quizzes increase the frequency of assessment and the frequency of study for students, which seem to benefit learning.
- Grades on exams are as important as learning, if not more, for students; therefore, one of the incentives of quizzes for students may be that they help improve grades. This could be used to the advantage of the formative assessment process, in that students could attach more value to the quizzes instead of seeing them as mere exercises.

This study is an example of a successful application of formative assessment without significant changes to teaching practices. However, it is not possible to claim that this process would be successful in any college level science class. One of the important issues that need to be considered is student motivation and self-efficacy. Without student motivation or value placed by students on the process, it is very difficult for formative assessment to make a difference. For further research, I suggest exploring other simple methods of utilizing formative assessment in undergraduate and post-graduate levels that do not place too great an extra workload on the instructors. Otherwise, formative assessment may not find widespread application at the college level.

Acknowledgments

I would like to thank Prof. Fitnat Kaptan for her valuable help and suggestions during the study and the data analysis process. I also would like to thank Dr. Ali Ekber Sahin, Dr. Serkan Yilmaz, Dr. Pinar Ozdemir, and Bilge Gök for their help during the instrument development and data analysis.

References

- Ahmed, N. & W. Teviotdale (2007). The Value of Formative Assessment in Higher Education. Education in a Changing Environment, Greater Manchester, UK, Informing Science Press.
- Akbaş, A. (2010). Attitudes, self-efficacy and science processing skills of teachign certificate master's program (ofmae) students. *Egitim Arastirmalari - Eurasian Journal of Educational Research*, 39, 1-12.
- Black, P. (2003, April). *Formative and Summative Assessment : Can They Serve Learning Together?* Paper presented at the Amercian Educational Research Association, Chicago, IL.
- Black, P., Harrison, C., Hogden, J., Marshall, B., & Wiliam, D. (2005). Dissemination and evaluation: a response to Smith and Gorard. *Research Intelligence*, 93, 7.
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-73.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.
- Bogdan, R. C., & Biklen, S. K. (2002). *Qualitative research in education*. Massachusetts: Allyn and Bacon.
- Carless, D. (2007). Conceptualizing pre-emptive formative assessment. *Assessment in Education: Principles, Policy & Practice*, 14(2), 171-184.
- Carrillo-de-la-Pena, M. T., E. Bailles, Caseras, X., Martinez, A., Ortet G. & Perez, J. (2009). "Formative assessment and academic achievement in pre-graduate students of health sciences." *Advances in Health Sciences Education* 14(1): 61-67
- CERI. (2005). *Formative assessment: Improving learning in secondary classrooms*. Paris: OECD.
- Costa, D. S. J., B. A. Mullan, et al. (2010). "A web-based formative assessment tool for Masters students: A pilot study." *Computers & Education* 54(4): 1248-1253.
- Dobson, J. L. (2008). "The use of formative online quizzes to enhance class preparation and scores on summative exams." *Advances in Physiology Education* 32(4): 297-302.
- Elwood, J. (2006). Formative assessment: possibilities, boundaries and limitations. *Assessment in Education: Principles, Policy & Practice*, 13(2), 215-232.
- Erdem, M. (2008). The effects of the blended teaching practice on prospective teachers' teaching self-efficacy and epistemological beliefs. *Eurasian Journal of Educational Research*, 30, 81-98.

- Higgins, R., Hartley, P., & Skelton, A. (2002). The Conscientious Consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, 27(1), 53-64.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, 33(14), 14-26.
- Johnson, R. B. and A. J. Onwuegbuzie (2004). "Mixed Methods Research: A Research Paradigm Whose Time Has Come." *Educational Researcher* 33(14): 14-26.
- Kibble, J. (2007). "Use of unsupervised online quizzes as formative assessment in a medical physiology course: effects of incentives on student participation and performance." *Advances in Physiology Education* 31(3): 253-260.
- Lawson, D. (1999). "Formative assessment using computer-aided assessment." *Teaching Mathematics and Its Applications* 18(4): 155-158.
- Liu, N., & Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teaching in Higher Education*, 11(3), 279-290.
- Nicol, D. J. (2007). Principles of good assessment and feedback: Theory and practice. International Online Conference sponsored by the REAP Project: Assessment design for learner responsibility. Retrieved April 05, 2010, from http://www.reap.ac.uk/public/Papers/Principles_of_good_assessment_and_feedback.pdf
- Nicol, D. J. (2009). "Assessment for learner selfregulation: Enhancing achievement in the first year using learning technologies." *Assessment & Evaluation in Higher Education* 34(3): 335-352.
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and selfregulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218.
- Oosterhof, A. (2001). Classroom applications of educational measurement. Upper Saddle River, N.J., Merrill.
- Peat, M. and S. Franklin (2002). "Supporting student learning: The use of computer-based formative assessment modules." *British Journal of Educational Technology* 33(5): 515-523.
- Perrenoud, P. (1998). From Formative Evaluation to a Controlled Regulation of Learning Processes. Towards a wider conceptual field. *Assessment in Education: Principles, Policy & Practice*, 5(1), 85-102.
- Sadler, D. R. (1998). Formative Assessment: revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, 5(1), 77-84.
- Sale, J. E. M., Lohfeld, L. H., & Brazil, K. (2002). Revisiting the Quantitative-Qualitative Debate: Implications for Mixed-Methods Research. *Quality & Quantity*, 36, 43-53.
- Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education*, 45, 477-501.

Üniversite Genel Kimya Derslerinde Basit Biçimlendirici Değerlendirme, Yüksek Öğrenme Kazanımları

(Özet)

Problem Durumu: Literatürde biçimlendirici değerlendirmenin öğrencilerin performansını artırmadaki yararı çokça tartışılmış ve genel olarak kabul görmüştür. Biçimlendirici değerlendirme öğrenci gelişiminin ve algısının sıkça, etkileşimli olarak değerlendirilmesi ve öğrencilerin ihtiyaçlarının belirlenerek öğretimin buna göre yeniden düzenlenmesi olarak tanımlanabilir. Biçimlendirici değerlendirmenin yararlı olabilmesi için kaliteli dönütler verilmesi, bu dönütlerin öğrenciler tarafından uygun bir şekilde kullanılması gereklidir. Fakat üniversite seviyesindeki fen derslerinde, çoğu zaman, hem öğretmenler hem de öğrenciler üzerinde eğitim-öğretim ve ölçme-değerlendirme açısından yoğun bir yük vardır. Bu nedenle kaliteli dönütler vermek, bu dönütleri uygun bir şekilde kullanmak ve öğretim uygulamalarını değiştirmek beklendiğinden daha zor olabilir.

Araştırmanın Amacı: Bu çalışmanın amacı biçimlendirici değerlendirme için kolay uygulanabilir bir düzeyde fırsatlar oluşturulmasının, öğrencilerin başarılarını ve üniversite seviyesindeki fen derslerine yaklaşımlarını, öğretim uygulamalarında fazla bir değişikliğe gerek kalmadan, olumlu bir şekilde etkileyebileceğini göstermektir.

Araştırmanın Yöntemi: Bu araştırma ilköğretim fen bilgisi öğretmen adaylarının aldığı lisans düzeyindeki organik kimya ağırlıklı bir genel kimya dersi kapsamında yapılmıştır. Dersi 2007 ve 2008 yıllarında alan toplam 163 kişi bu çalışmaya katılmıştır. Araştırmada hem nicel hem de nitel metotlar bir arada kullanılmıştır. Çalışmada 2007 ve 2008 yıllarında dersi alan öğrencilerin başarıları karşılaştırılmıştır. 2007 yılında dersi alan öğrencilere normal öğretim yapılırken, 2008 yılında dersi alan öğrencilere öğrenmeyi desteklemek amacıyla biçimlendirici değerlendirme uygulanmıştır. Biçimlendirici değerlendirme, dönem içerisinde not verilmeden yapılan kısa sınavlar çerçevesinde öğrencilerin kendilerini değerlendirmeleri, akranlarından ve öğretim elemanından dönüt almaları ve bu dönüt ve değerlendirmeler ışığında öğretimde düzenlemeler yapılması şeklinde gerçekleştirilmiştir. Çeşitli veriler ışığında kontrol ve deney grubu öğrencilerinin başlangıçtaki başarı düzeylerinin aynı olduğu görülmüştür. Öğrencilerin başarılarını ölçmek için güvenilirlik ve geçerlik çalışmaları yapılmış ara sınavlar ve final sınavı kullanılmıştır. Ayrıca geliştirilen Likert tipi bir anket yardımıyla uygulama grubundaki öğrencilerin derse, dersin işlenişine ve ders içeriğine karşı tutumları ve biçimlendirici değerlendirme uygulamasına karşı tepkileri ölçülmüştür. Anketin içerdiği dört adet açık uçlu soruya verilen cevaplar, yedi

öğrenciyle yapılan yarı yapılandırılmış mülakatlar ve gözlem notları çalışmanın nitel verilerini oluşturmuştur.

Araştırmanın bulguları: Nicel verilerin analizi uygulama grubu öğrencilerinin başarısında istatistiksel olarak anlamlı bir artışın olduğunu göstermiştir. Her ne kadar bu artışa etki eden ve kontrol edilemeyen başka faktörlerin olması olası ise de, bu artışta kullanılan biçimlendirici değerlendirme uygulamasının da bir katkısı olduğu değerlendirilmiştir. Anket sorularının frekans analizi öğrencilerin derse, dersin içeriğine, dersin işlenişine ve kısa sınavların kullanımına karşı oldukça pozitif tutumlar sergiledikleri görülmüştür. Bu pozitif tutumların biçimlendirici değerlendirme uygulamasının başarısına olumlu bir katkı getirdiği söylenebilir. Anketlerdeki açık uçlu sorulara verilen cevaplar ve yedi öğrenci ile yapılan mülakat çözümlenmeleri öğrencilerin uygulanan biçimlendirici değerlendirme yaklaşımına karşı olumlu tutumlarını teyit etmiş ve öğrencilerin bu süreci nasıl algıladıkları ve kullandıkları ile ilgili bilgiler vermiştir. Bu veriler ışığında öğrenciler genel olarak biçimlendirici değerlendirme amacıyla yapılan kısa sınavlara not verilmemesinin olumlu bir uygulama olduğu, kısa sınavlar sırasında arkadaşlarından dönüt almalarının önemli olduğu, kısa sınavların hem öğrenmelerine hem de notlarına olumlu katkı getirdiği ve kısa sınavların derse çalışma sıklıklarını artırdığı yönünde görüş bildirmişlerdir.

Araştırmanın Sonucu ve Önerileri: Lisans düzeyinde bağıl olarak basit ve uygulanması kolay bir şekilde tasarlanmış olan bir biçimlendirici değerlendirme yöntemi, uygun şartlarda, öğrencilerin başarısını ve aldıkları derslere karşı tutumlarını olumlu bir şekilde etkilemektedir. Bu nedenle lisans düzeyinde eğitim veren öğretim elemanlarının, kendilerine fazla bir yük getirmeden, biçimlendirici değerlendirme yaklaşımlarını kullanmayı düşünmeleri önerilmektedir. Biçimlendirici değerlendirme çok farklı şekillerde uygulanabilir, fakat şu hususlara dikkat edilmesi önerilebilir:

- Biçimlendirici değerlendirmede not kullanılmamalıdır.
- Akran dönütlerine ve öz değerlendirmeye fırsat verilmelidir.
- Kısa sınavlar öğrencilerin çalışma sıklığını ve konulara hâkimiyetlerini artırmaktadır.
- Biçimlendirici değerlendirmenin sınavlarda alınan puanlara olumlu katkı getireceği düşüncesi öğrenciler açısından motive edici olabilmektedir.

Anahtar kelimeler: Biçimlendirici değerlendirme, biçimlendirici dönüt, akran dönütü, öz değerlendirme, fen eğitimi, lisans eğitimi

To learn more about what a formative assessment is, and how you could be using one to guide your own studying, read on. As the content you learn becomes more and more advanced, having a general idea of your background knowledge will be more and more critical. After all, you can't be expected to undertake a calculus-based physics class if you have gaps in your understanding of basic algebra. Teachers at every level will use formative assessments to determine student starting points, guide instruction, and identify areas in need of review. Why should I use a formative assessment? Teachers aren't the only ones who should take advantage of these best teaching practices. Over the last 30 years a range of assessment strategies have been developed aiming to effectively capture students' learning in Higher Education and one such strategy is measuring students' learning gains. The main goal of this study was to examine whether academic performance within modules is a valid proxy for estimating students' learning gains. Yalaki, Y.: Simple formative assessment, high learning gains in college general chemistry. *Egitim Arastirmalari-Eurasian J. Educ. Res.* Using assessment for learning and learning from assessment. *Assessment & Evaluation in Higher Education*, 27, 501-510. [Google Scholar]. Teddlie, C., & Tashakkori, A. (2009). Simple formative assessment, high learning gains in college general chemistry. *Eurasian Journal of Educational Research*, 40, 223-240. [Google Scholar]. Formative assessment or assessment for learning is one of the most emphasized educational innovations around the world. Two of the common strategies that could be used in formative assessment are use of summative tests for formative purposes and comment only marking. We utilized these strategies in the form of formative quizzes in a general chemistry course and measured its effect on students' learning. Yalaki, Y. (2010). Simple Formative Assessment, High Learning Gains in College General Chemistry. *Eurasian. Journal of Educational Research*, 40, 223-241. Citations (2). References (15). Formative assessment practice is still not well developed in most schools. Much practice depends on the formative use of summative tests. In addition, the National Curriculum level descriptors provide an imprecise notion of progression through many science concepts. The study sought to use more research-based materials to improve the author's formative assessment practice by carrying out an evaluation of a single action research cycle. A backward design approach was used to generate a series of 31 diagnostic questions which could assess understanding in a KS3 solution chemistry topic in a teach