TÜRK FEN EĞİTİMİ DERGİSİ
Yıl 12, Sayı 2, Mart 2015

Journal of
TURKISH SCIENCE EDUCATION
Volume 12, Issue 1, March 2015

http://www.tused.org

The Effect of Performance Based Evaulation on Preservice Biology Teachers' Achievement and Laboratory Report Writing Skills

Perihan GUNES[1], Hikmet KATIRCIOGLU[2], Mirac YILMAZ[3]

[1] Assist. Prof. Dr., Aksaray University, Faculty of Education, Aksaray-TURKEY
[2] Assoc. Prof. Dr., Gazi University, Faculty of Education, Ankara-TURKEY
[3] Assist. Prof. Dr., Hacettepe University, Faculty of Education, Ankara-TURKEY

Received: 06.06.2014 **Revised:** 03.11.2014 **Accepted:** 22.01.2015

The original language of article is English (v.12, n.1, March 2015, pp.71-83, doi: 10.12973/tused.10134a)

ABSTRACT

The aim of this study is to determine the effect of performance based evulation on preservice biology teachers' cognitive achievements and laboratory report writing skills about DNA isolation. In the study, nonequivalent control-group design were used to determine the effect of performance based evaluation on pre-service biology teachers' achievement and laboratory report skills about DNA isolation. The sample of the study was 70 pre-service teachers. A rubric and DNA isolation achievement test were developed for the data collection. Data was collected by pre-and post-administration of achievement test and administration of rubric. Before the beginning of the study, a pre-test was applied in order to determine the cognitive field levels of students on DNA isolation. Then, a DNA isolation test was applied to the experiment and control groups. While the experiment group was given a rubric to be used during writing their laboratory reports, the control group was not. At the end of the study, a post-test was applied in order to determine students' cognitive field levels. To detect the differences between the experimental and the control groups, the independent samples t-test was used. At the end of the study, it was determined that preservice teachers who use rubrics display a higher skill in writing laboratory reports and have a higher cognitive field level compared to those who do not.

Keywords: Rubric; Analytic Rubric; Holistic Rubric; DNA Analysis; Gene Technology; Preservice Teacher; Report Writing Skill.

INTRODUCTION

Recently, many educators have emphasized that the evaluation methods used today do not provide knowledge for the individuals' self-assessment nor the state of his development. These kind of evaluation methods should be used less frequently for evaluating student achievement (Kutlu, 2004; Bahar et al., 2006). As such evaluations fall short in measuring the higher order cognitive skills (such as designing an experiment, writing a new story, or presenting a paper) of students that are required to be graded (Kutlu, 2004). However, throughout their education, students must be able to use advanced mental skills in addition to knowledge that is based on recall to participate in meaningful learning. Kutlu, Doğan & Karakaya (2008) described the higher order cognitive skills as the entire cognitive, affective

Corresponding author e-mail: perihanguness@gmail.com

and psychomotor features used by individuals while displaying their skills. Zoller (2000) approached the higher order cognitive skills as asking questions, critical and systematic thinking, problem solving, analyzing, evaluating, and synthesizing new information and decision-making. To educate individuals who can meet the needs of modern society, eduaction systems should set the development of students' higher order as high priority (Kutlu et al., 2008). This necessity requires educators to also measure and evaluate higher order cognitive skills. However, the current tests used in science teaching give little information about how students use the knowledge they have gained. That situation makes an emergent call for the use of new evaluation approaches that check the product and the process, which is globally referred to as performance-based evaluation, combining with pre-existing measurement and evaluation approaches.

Performance-based evaluation can be applied to real-life situations. Moreover, it can measure an activity with multiple solutions or strategies that includes higher order cognitive skills, which can be measured over a wide-ranging period of time, from a couple of minutes to a couple of days (such as modeling), or a situation that can produce original answers (such as explaining the solution to a mathematical problem) (Aschbacher, 1991; Baron, 1991; Madaus and O'Dwyer 1999; Stiggins, 1987). Kubiszyn and Borich (1996), on the other hand, underlined the necessity for propounding a product, having an observable performance while propounding that product and enabling it to involve the process of high order thinking in order to ground an evaluation on performance. Besides, the researchers indicated that the performance-based evaluations had to enable the social skills and group studies, as well as the interdisciplinary transition and information exchange (as cited in Berberoğlu, 2006). Performance-based evaluation is composed of two important parts. One is the performance task, and the other is the rubric (Popham, 2007). Miller (2005) defines performance tasks as activities where students are required to develop their own answers rather than choosing among the options presented to them. Students' knowledge can be structured by giving them performance tasks, and this is important for effective learning (as cited in Marzano, Pickering, and Mctighe, 1993). Thus, performance-appropriate evaluation tools are necessary to evaluate performance. Rubrics are said to be one of the most widely used tools to complete performance evaluations (Kutlu et al., 2008)

Rubrics are documents where the criteria taken into consideration for a given study are listed and the quality of each criterion is provided with detailed definitions (Andrade, 2000; 2001; 2005; Andrade and Du, 2005; Andrade, Wang, Du & Akawi, 2009; Goodrich, 1997). Having various forms and levels, the rubric could be used for a good number of course fields (Moskal, 2000). Rubric is especially used by teachers due to their dissatisfaction in giving a mark to practice-based fields like projects and oral presentations (Reddy, 2007). According to Popham (2007), rubrics are composed of three parts, namely, evaluation criteria, criterion definitions, and grading strategies.

1. Evaluation criteria: These are the factors that an evaluator uses when deciding on the quality of a student's performance. In other words, these are the requirements for a student to be considered successful (Wiggins, 1991).

2. Quality definitions: These are the detailed definitions of what a student has to do to achieve a certain performance level (Popham, 2007).

3. Grading strategies: These include whether the grading will be conducted according to the process or the result (Moskal, 2000).

Rubrics are descriptive grading schemes, including grading requisites. They are important because they include the use of higher order cognitive skills, which are the direct result of the emphasis on performance (Hafner and Hafner, 2003). While there is no standardized method to develop rubrics, certain criteria should be considered during the rubric development process. These are continuity, parallelism, consistency, even distribution, reliability, and validity (WNCP, 2006).

Recently, researchers have discussed the benefits of using rubrics in education, indicating that the application of rubrics in classes is appealing for both teachers and students. The reasons for this can be better understood when the reasons for using rubrics are examined (McCollister, 2002; Halonen et al., 2003; Andrade & Du, 2005). One of the most important goals of rubric use is to enable students to openly express their expectations about learning (Luft, 1997; 1999). They feel more responsibility for the evaluation of their learning (Phillip, 2002), and they actively participate in the evaluation process by partaking in self-evaluation (Wittaker, Spencer, and Duhaney, 2001). Students are able to determine their own needs and evaluate their own performance with rubrics (Andrade and Du, 2005). When used for the right reasons, rubrics make each teacher's grading system transparent so that students better understand the standards they are expected to meet (McCollister, 2002). In addition to enabling students to develop higher order cognitive skills and meta-cognitive strategies (Halonen et al., 2003), the application of rubrics supports and enhances learning (Andrade, 2000; Andrade and Boulay, 2003). The use of rubrics improves student achievement and learning by increasing each student's belief in their own self-efficiency (Quinlan, 2006). Moreover, it helps students improve their self-regulation skills by supporting meta-cognitive strategies such as planning, observation, and regulation (Saddler & Andrade, 2004). Reddy (2007) states that effective, valid, and reliable rubrics that provide clear, satisfactory and detailed feedback increase student satisfaction with the evaluation. They also augment learning by influencing self-evaluation, where students observe their own processes, such as evaluation, renewal, and performance grading. This implies that, in addition to increasing academic achievement, rubrics can augment other aspects of a student's learning, such as interest, personal proficiency and individual regulation.

As tools, rubrics are also appealing to teachers because they are strong in both teaching and evaluation (Andrade, Du & Wang, 2008; Goodrich, 1997). With the help of rubrics, teachers can have a more objective and consistent evaluation of their grading practices (Andrade, 2005; Wittaker, Spencer, and Duhaney, 2001). A valid rubric decreases the possibility of a biased decision about a student's performance by preventing the evaluator from focusing on factors such as a student's gender, race, age, appearance, ethnicity, or previous academic achievement. This quality of rubrics makes it possible to evaluate students objectively, and it becomes good supporting documentation for the teacher when he or she meets with the student and his or her parents (Andrade, 2000; Whittaker, Spencer, and Duhaney, 2001). At this point, the need for high quality rubrics that are to be used in teaching and learning processes comes to the fore. This implies that there is a necessity for a good comprehension of the rubric preparation process and that learning processes should be supported by well-prepared rubrics. For well-prepared rubrics to be used, the teachers, i.e., the evaluators should be well-educated in the use of rubrics. Additionally, the rubrics that are implemented should be developed by experts and be well tested.

In today's world, where our expectations of school and education consistently increase, science and the biological sciences, which facilitate our life and increase its quality, have an important role in supporting our decisions and enabling us to make the right ones. Therefore, the benefits of rubrics and their expansion in terms of science, technology, society, and environment, should be discussed by the field educators. When considering the transference of knowledge on the topics of molecular genetics in the biological sciences to individuals, the importance of multifaceted learning that will be realized on an advanced mental level must be better understood.

"Modern biotechnology" (or "gene technology"), in terms of its advantages and disadvantages related to health, economics, industry, environment, ecology, social and ethical topics, should be carefully discussed. The importance of communicating the ethical, economic, social, medical, and ecological results of gene technology to the individual and carefully examining the development of personal judgment and evaluations (approving/not

approving, risk perception, and the ability to evaluate dangers), as well as the fact that these activities increase the importance of gene technology education, have been emphasized by several scientists (Schallies and Wellensiek, 1995; Harms and Bayrhuber, 1999; Harms, 2002). Discussing the teaching and the learning level of DNA isolation, which includes purification of the DNA by separating it from the live cell, using a rubric, is also important.

In the light of previous literature, the current study aimed to determine the effect of performance based evaluation on pre-service biology teachers' cognitive achievements and test report writing skills about DNA isolation. The following questions were addressed.

I. Is there a difference between the skills of students writing reports on DNA isolation levels when rubrics are used vs. when they are not used?

II. Is there a difference between the cognitive field levels of students studying DNA isolation when rubrics are used vs. when they are not used?

METHODOLOGY

a) Type of the Study

In this study, quantitative research designs patterns are used. The data are gathered using a nonequivalent control-group design, which is a quasi-experimental design. The most commonly used quasi-experimental design in educational research is the nonequivalent control-group design. In this design, research participants are not randomly assigned to the experimental and control groups, and both groups take a pretest and a posttest (Gall, Gall, Borg, 2007)

b) Study Group

The study group consists of 70 pre-service biology teachers who are sophomores at a state university. They were selected using cluster sampling, and they were appointed as one experiment and one control group in class level (Karasar, 2006). There were 31 students in the experimental group and 39 students in the control group. Ninety percent of the study group were female participants (N=63), and 10% were male participants (N=7). Second grade preservice teachers did not have an experience about rubrics. Besides, they did not have theoretical and practical studies about the DNA isolation, either.

c) Data Gathering Tools

DNA Isolation Cognitive Field Achievement Test: To determine students' cognitive field achievement levels with respect to DNA isolation, a 5-choice multiple-choice test was prepared. For this test, 5 objectives of student achievement were determined by the researcher. At least two questions that measure each objective were written, and a multiple-choice achievement test with a total of 22 questions was prepared. The achievement test was presented to two biology experts and two biology education experts. In accordance with the views and suggestions of the experts, necessary corrections were made and a pre-trial form was constructed. The constructed pre-trial form was administered to 174 university students studying in the biology education department during the spring semester of the 2010-2011 academic year. After this application, the article difficulty and differentiation indexes of the test were calculated. According to the results obtained from these calculations, 3 articles with an article difficulty lower than 0.20 were discarded. The remaining articles had article difficulties ranging from 0.22 to 0.59. The average article difficulty was calculated as 0.61, and it was determined that the test was of medium difficulty. The inner consistency reliability coefficient of the final test with 19 articles was determined to have a Cronbach's α=0.73.

DNA Isolation Rubric: To determine each student's laboratory report writing skills with respect to DNA isolation, including their weaknesses and strengths within the scope of general biology classes, a rubric was prepared. During the development process of the analytical rubric, which was intended to determine students' levels in reaching the objectives of laboratory report writing, four quantifiable objectives were determined. A performance item was established for each objective. These performance items became the evaluation criteria for the rubric. The resulting evaluation criteria were as follows: being able to write down the materials of the experiment, being able to write down the actual realization of the experiment, being able to write down the observations of the experiment, and being able to write down the results of the experiment. Each evaluation criterion was graded using 1-3 points. The summary of the scoring is as follows:

1 point: The skill needs to be developed

2 points: Acceptable

3 points: Highly successful

Then, each evaluation criterion was individually defined. For each evaluation criterion, for the spot where the scoring levels intersect, the performance expected of the student for that level is defined in detail from 3 to 1 (from good to bad). The ability to make observations was divided into two sub-categories; likewise, the ability to interpret was divided into 5 sub-categories, and all these sub-categories were separately defined. For example, for the experiment materials evaluation criterion, "All materials used in the experiment were documented in full" was written down for 3 points; "Many of the materials used in the experiment were documented" was written down for 2 points; and "Materials used in the experiment were not documented" was written down for 1 point (Appendix 1: DNA Isolation rubric). 4 experts examined the score adjustment levels between the evaluators for the reliability of the rubric. And while doing so, the evaluators were required to score a laboratory report that was prepared with the help of the rubric separately through the rubric. And then the reliability coefficients between the scores given by the evaluators were calculated with the Kendall conformation coefficient. As a result of the study, the conformation coefficient between the evaluators was determined as Kendall W=0, 81 (Kendall, Babington-Smith, 1993)

d) Data Analysis

The SPSS 20.00 statistics program was used to analyze the data. Item difficulty and discrimination indexes were measured by ITEMAN, which is an article analysis test, and the internal consistency reliability coefficient was calculated with Cronbach's alpha. To detect the differences between the experimental and the control groups, the independent samples t-test was used. In the interpretation of the results, the level of meaningfulness of the p value was determined as 0.05.

e) Intervention

The study was conducted by performing the DNA isolation experimentation within the context of the lesson of general biology laboratory II. The lesson of general biology laboratory II is conducted in two sections. While the 1st section is separated as the experimental group, the 2nd section is separated as the control group. We studied with the experimental group in the first two hours of the 4-hour lesson of general biology laboratory II and with the control group in the following two hours. At the beginning of the study, the pre-service teachers in the experimental and control group were informed about the DNA isolation and they were required to apply the experimentation of DNA isolation on their own.

The experimental and control group realized the experimentation in 4 groups according to the following order.

1. A half tomato was peeled and then diced. It was smashed until it became a puree within the press.
2. The solution A was prepared by putting 3 gr salt, 3 ml detergent and 24 ml distillate water in a beaker. This solution was put in a press and properly smashed with the pulped tomato.
3. The thickened puree was distilled with the help of a gauze or a thin strainer and put in a beaker.
4. 15 ml pineapple juice was added to the distilled part and then mixed.
5. 5 ml was taken from that mixture, placed within a tube and 10 ml cold alcohol was added to it.
6. We waited for 4 – 5 minutes and then observed DNA on the alcohol layer.

As a result of the intervention, the pre-service teachers in the experimental and control group were required to discuss about the results of the experiment with their friends in the group. As a result of the discussions, they were all expected to write a report regarding the experiment. They were required to write the experiment report by considering the materials being used in the experiment, as well as the performance of the experiment, observations of the experiment and the order of the results/interpretation of the experiment. While the experimental group was given the DNA isolation rubric that was developed to determine how to evaluate the reports while writing their reports, the control group was not given such a rubric. The control group only received oral explanations about how to write the report. The experiment reports that were obtained at the end of the study were evaluated by an expert through considering the criteria in the rubric.

FINDINGS

Test Report Writing Skill Acquisition Levels of Preservice Teachers

Table 1 shows the t-test findings. The results indicate that there is a statistically significant difference between the report writing achievement scores of the experiment and the control groups in favor of the experiment group.

Table 1. *Preservice Teachers' Rubric Total Point Averages Independent Samples T-Test Results*

Laboratory Report Writing Skills	Group	N	$\overline{X}$	SS	sd	t	p
Total Points	Experiment	31	23.5	3.265	68	13.96	0.000*
	Control	39	14.4	2.137			

*p<0.05

Table 2 shows the t-test findings. The data indicate that while there is not a statistically significant difference in the material writing skills of the control and the experiment group, there is a meaningful difference in favor of the experiment group in terms of the skills of being able to realize, observe and interpret the experiment.

Table 2. *Preservice Teachers' Rubric Averages Independent Samples t-Test Results*

Laboratory Report Writing Skills	Group	N	$\overline{X}$	SS	sd	t	P
Materials used in the Experiment	Experiment	31	2.58	0.502	68	0.330	0.743
	Control	39	2.54	0.555			
Construction of Experiment	Experiment	31	2.90	0.301	68	2.003	0.049
	Control	39	2.69	0.521			
Observation of Experiment	Experiment	31	5.10	1.270	68	5.469	0.000*
	Control	39	3.43	1.250			
Interpretation of Experiment	Experiment	31	12.90	2.330	68	16.497	0.000*
	Control	39	5.77	1.220			

*p<0.05

The data in Table 3 show the rubric skills averages of the experiment and the control groups with the t-test findings. In Table 2, observations and interpretations of the experiment were taken into consideration as a whole, whereas in Table 3, they are examined separately. Data obtained from the rubric show that there is a statistically significant difference between the experiment and control groups in favor of the experiment group when observation 1, observation 2, interpretation 1, interpretation 2, interpretation 3, and interpretation 5 achievement scores are considered.

Table 3. *Preservice Teachers' Observation and İnterpretation Rubric Averages Independent Samples T-Test Results*

Laboratory Report Writing Skills	Group	N	$\overline{X}$	SS	sd	t	p
Observation 1	Experiment	31	2.77	0.617	68	4.661	0.000*
	Control	39	1.82	0.997			
Observation 2	Experiment	31	2.32	0.945	68	3.361	0.001*
	Control	39	1.62	0.815			
Interpretation 1	Experiment	31	2.61	0.715	68	11.460	0.000*
	Control	39	1.13	0.339			
Interpretation 2	Experiment	31	2.35	0.661	68	8.307	0.000*
	Control	39	1.26	0.442			
Interpretation 3	Experiment	31	2.55	0.810	68	11.960	0.000*
	Control	39	1.00	0.000			
Interpretation 4	Experiment	31	2.77	0.617	68	13.140	0.000*
	Control	39	1.10	0.447			
Interpretation	Experiment	31	2.61	0.715	68	8.158	0.000*
	Control	39	1.28	0.647			

*p<0.05

Cognitive Field Levels of Preservice Teachers

The experiment and control groups' cognitive field achievement test pre-test averages independent samples t-test results are shown in Table 4. The results indicate that there is no statistically significant difference between the cognitive field achievement test pre-test averages of the experiment and control groups.

Table 4. *Pre-Application Cognitive Field Levels of Students' Independent Samples T-Test Results*

Group	N	$\overline{X}$	SS	sd	t	p
Experiment	31	14.26	3.44	68	1.150	0.254
Control	39	13.28	3.59			

The experiment and control groups cognitive field achievement post-test averages independent groups t-test results are shown in Table 5. The data indicate that there is a statistically significant difference between the cognitive field achievement post-test averages of the experiment and control groups in favor of the experimental groups.

Table 5. *Post-Application Cognitive Field Levels of Students*

Group	N	$\overline{X}$	SS	sd	t	p
Experiment	31	15.52	3.48	68	2.015	0.048*
Control	39	13.82	3.51			

*p<0.05

The results shown in Table 6 show that there is a statistically significant difference between the cognitive field achievement pre-test and the post-test averages of the experiment group and that there is no statistically significant difference between the cognitive field achievement pre-test and post-test averages of the control group. It was detected that the experiment groups achieved a higher success level in the post-test compared to the pre-test.

Table 6. *Experiment Group Pre-test and Post-test Average Points Matched t-Test Results*

Group	Test Type	N	$\overline{X}$	SS	Sd	t	p
Experiment	Pre-test	31	14.26	3.44	30	-3.198	0.003*
	Post-test	31	15.52	3.48			
Control	Pre-test	39	13.28	3.59	38	-1.007	0.320
	Post-test	39	13.82	3.51			

*p<0.05

DISCUSSION AND SUGGESTIONS

As a result of the study, it was determined that the rubric and achievement test successes of students showed a difference according to the educational method. As a result of the study, when the total points students earned on the DNA isolation rubric were considered, the experiment group was more successful compared with the control group in terms of acquiring report writing skills (Table 1). In similar studies on this topic, in terms of skills

acquisition levels, students who use rubrics are more successful compared with those who do not (Andrade, 2001; Sefer, 2006; Gunes, 2011). In their respective studies, Andrade (2001) proves that rubrics are influential in enabling students to acquire effective writing skills; Sefer (2006) proves that they are influential in enabling students to acquire problem solving skills, and Gunes (2011) proves that they are influential in enabling students to acquire research skills.

Using rubrics to develop and acquire skills can be given as one of the reasons affecting student achievement. The reason students who use DNA isolation rubrics are more successful is that learning outcomes and the required achievement levels are clearly indicated in rubrics.

When the experiment report writing skills of the experiment and control groups were examined individually, it was observed that the experiment group was more successful compared to the control group (Table 2).

When the report writing skills acquisition levels of the preservice teachers were examined individually, the experimental group achieved a higher success level in terms of observation and interpretation of experiment results skills compared with the control group (Table 3). In his research on students' skills in interpreting experiment results, Rutherford (2007) found that students who use rubrics can interpret experimental results more successfully. In their study about learning studies with groups, Cohen, Lotan, Scarloss, Schultz and Abram (2002) suggested that individuals who were informed about the evaluation criteria had better-quality group studies and discussions. Andrade and Du (2005), on the other hand, indicated that using a rubric could enable students to do quality homeworks, have better grades and decrease their anxieties about what they would learn. According to the results of our study, it could be asserted that the use of a rubric in the DNA isolation experimentation will be effective upon the development of the skills of preservice teachers such as writing an experimentation report involving higher order cognitive skills like making accurate observations, as well as interpreting and writing the experimentation results, which could signify that the rubric would increase the higher order cognitive skills. Interpreting the study results in a broader term, on the other hand, it could be thought that the success of preservice teachers who were informed about the evaluation criteria by means of the rubric was affected by the decrease of their anxieties about how to write the experimentation report and the betterment of their perceptions regarding this subject.

The reason for the high success level in DNA isolation is that the perceptions of the preservice teachers regarding experimental report writing improves, and using rubrics is therefore effective in developing experimental report writing skills.

It was found that there is no meaningful difference in any class level between the cognitive field pre-test points of preservice teachers who use or do not use rubrics (Table 4). The points taken from the pre-test indicated that the control and the experimental group have a certain level of presumption about the topic.

There was a meaningful difference between the cognitive field post-test points of preservice teachers who use DNA isolation rubrics and those who do not, in favor of the experiment group (Table 5). The fact that the experimental group used rubrics is shown as the reason for this.

In this study, it was found that compared to the pre-test, the experimental group had a higher success level in the post-test (Table 6). The reuslts showed similarities with the research findings of Gunes' study (2011) related to the effect of rubrics on the research skills and cognitive field levels of primary school students. Gunes (2011) found that the cognitive level of the experiment group that used rubrics was higher in the pre-test compared to the post-test. Moreover, the cognitive level of the control group did not show any difference from the pre-test to the post-test. In the qualitative study of Gunes (2011) that was conducted with primary school students, the students indicated that the rubrics increased their success, which signifies that the results of study are supported by qualitative data. In our study, the fact that

the experimental group used rubrics in research activities may have been effective in the occurrence of this difference. Rubrics may have a positive effect on the improvement of preservice teachers' perceptions and skills and on the improvement of their academic achievements.

The findings of the study allowed to make several suggestions. First, expressing what pre-service teachers should take into consideration when preparing an experimental report can have a positive effect both on the enhancement of their perceptions and skills and on the enhancement of their academic achievement. Second, the importance of using rubrics during the evaluation process, expanding the use of rubrics, and improving preservice teachers by informing them about rubrics can contribute to their education. Third, a discussion of developing new rubrics for molecular genetics, gene technology and other biology subjects and applying these rubrics to wide and various groups, thereby increasing student success and advancing cognitive skills and the effects of complimentary evaluation tools can be suggested. Finally, It could be suggested to also conduct studies that investigate the factors affecting the success and are supported by qualitative data.

REFERENCES

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership, 57* (5), 13-18.

Andrade, H. G. (2001). The effects of instructional rubrics on learning to write. *Current Issues in Education [On-line],* 4 (4). Accessed October 12. 2010 http://cie.asu.edu/volume4/number4/.

Andrade, H.G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching, 53* (1), 27-31.

Andrade, H.G., & Boulay, B. A. (2003). Role of rubric-referenced self-assessmentin learning to write. *The Journal of Educational Research, 97* (1), 21-30.

Andrade, H. G., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research and Evaluation 10* (3), 1-11.

Andrade, H. G., Du, Y., & Wang, X. (2008). Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educational Measurement: Issues and Practice, 27* (2), 3–13.

Andrade, H. G., Wang, X., Du, Y., & Akawi, R. L. (2009). Rubric-referenced self assessment and self efficacy for writing. *The Journal of Educational Research, 102* (4), 287-301.

Aschbacher, P. R. (1991). Performance assessment: State activity, interest and concerns. *Applied Measurement in Education, 4* (4), 275-288.

Bahar, M., Nartgün, Z., Durmuş, S., & Bıçak, B. (2006). *Geleneksel alternatif ölçme ve değerlendirme öğretmen el kitabı*, Ankara: Pegem A Yayıncılık.

Baron, J. B. (1991). Strategies for development of effective performance exercises. *Applied Measurement in Education, 4* (4), 305-318.

Berberoğlu, G., (2006). *Sınıf içi ölçme ve değerlendirme teknikleri*. İstanbul: Morpha.

Cohen, E., Lotan, R., Scarloss, B., Schultz, S. & Abram, P. (2002). Can group learn? *Teacher College Record, 104*, 1045-1068.

Gall, M. D., Gall, J. P., & Borg, W., (2007). *Educational research an introduction*, USA:Pearson İnternational Edition.

Goodrich , H. G. (1997). Understanding rubrics. *Educational Leadership, 54* (4). 14-17.

Gunes, P. (2011*). Dereceli puanlama anahtarının ilköğretim öğrencilerinin araştırma becerisi ve bilişsel alan düzeyine etkisi*. Yayınlanmamış doktora tezi, Hacettepe Üniversitesi Fen Bilimler Enstitüsü, Ankara.

Hafner, J. C., & Hafner, P. M. (2003). Quantitative analysis of the rubric as an assessment tool: An Emprical Study of Student Peer-Group Rating. *International of Science Education, 25* (12), 1509-15298.

Halonen, J. S., Bosack, T., Clay, S., McCarthy, M., Dunn, D. S., Hill IV, G., W., McEntarffer, R., Mehrotra, C., Nesmith, R., Weaver, K. A., & Whitlock, K. (2003). A rubric for learning, teaching, and assessing scientific ınquiry in psychology. *Teaching of Psychology, 30* (3), 196-208.

Harms, U. & Bayrhuber, H. (1999). Biotechnologie im Unterricht, Biotechnologie und Gentechnik. In M. Schallies, und K.D. Wachlin eds. Biotechnologie und Gentechnik, Springer Verlag Berlin, Heidelberg, Newyork, 87-98.

Harms, U. (2002). Biotechnology Education in schools. *Electronic Journal of Biotechnology, 5*(3). http://www.ejbiotechnology.info/content/vol5/issue3/ teaching/01/ (Accessed August 8. 2004).

Karasar, N. (2006). Sosyal bilimlerde nitel araştirma yöntemleri (16th Ed.), Ankara: Nobel Yayın Dağıtım.

Kendall M. G., Babington-Smith B. (1993). The Problem of m Rankings. *The Annals of Mathematical Statistics, 10* (3), 275- 287.

Kutlu, Ö. (2004). *Tek Soru İle Öğrenci Performansının Belirlemesi*. Accessed November 13. 2010. http://www.erg.sabanciuniv.edu/iok2004/bildiriler.

Kutlu, Ö., Doğan, C. D., & Karakaya, İ. (2008). *Öğrenci Başarısının Belirlenmesi: Performansa ve Portfolyaya Dayalı Durum Belirleme.* Ankara: Pegem Akademi.

Luft, J. A. (1997). Design your own rubric. *Science Scope*, *20* (5), 25-27.

Luft, J. A. (1999). Rubrics: Design and use in science teacher education. *Journal of Science Teacher Education, 10* (2), 107-121.

Madaus, G. F., & O'Dwyer, L. M. (1999). A short history of performance assessments. *Phi Delta Kappan, 80* (9), 688-696.

Marzano, R. J., Pickering, D., & McTighe, J. (1993). *Assessing student outcomes: Performance assessment using the dimensions of learning model*. Virginia: Association for Supervision and Curriculum Development.

McCollister, S. (2002). Developing criteria rubrics in the art classroom. *Art Education, 55* (4), 46-52.

Moskal, B. M. (2000). Scoring rubrics: What, when and how? *Practical, Assessment, Research and Evaluation,* 7(3). Retrieved October 24, 2013 from http://PAREonline.net/getvn.asp?v=7&n=3.

Phillip, C. (2002). Clear expectations: Rubrics and scoring guides. *Knowledge Quest, 31*(2), 26-27.

Popham, J. W. (2007). *Classroom Assessment: What Teachers Need to Know (5th Ed.)*. USA: Pearson Education.

Ouinlan, A. (2006). *Assessment made easy: Scoring rubrics for teachers from K-college*. Lanham, MD: Rowman and littlefield education.

Reddy, M. (2007). Effect of rubrics on enhancement of student learning. *Educate 7* (1), 3-17.

Rutherford, S. (2007). Using a Laboratory Conclusion Rubric. *Science Activities: Classroom Projects and Curriculum Ideas, 43* (4), 9-14, DOI: 10.3200/SATS.43.4.9-14.

Saddler, B. & Andrade, H. G. (2004). The writing rubric. *Educational Leadership, 62*(2), 48-52.

Schallies, M. & Wellensiek, A. (1995). Biotechnologie/Gentechnik Implikationen für das Bildungswesen. Arbeitsbericht, Akademie für Technikfolgenabschätzung in Baden-Württemberg, Stuttgart.

Sefer, D. G. (2006). Dereceli puanlama anahtarının (rubrik) problem çözme becerisinin değerlendirilmesinde kullanılması. Yayınlanmamış yüksek lisans tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.

Stiggins, R. J. (1987). Design and development of performance assessments. *Educational measurement: Issues and Practice, 6* (1), 33-42.

Whittaker, R. C., Spencer, S.J., & Duhaney, D. (2001). Creating instructional rubrics for inclusive classroom. *Teaching Exceptional Children, 34* (2), 8-13.

Wiggins, G. (1991). Standart, not standardization: Evoking quality student work. *Educational Leadership, 48* (5), 18-25.

WNCP, (2006). *Rethinking classroom assessment with purpose in mind: Assessment for learning, assessment as learning, assessment of learning*. Canadian. Accessed August 22. 2010 http://www.wncp.ca/media/40539/rethink.pdf.

Zoller, U. (2000). Teaching college science towards the next millennium: Are we getting it right? *Journal of College Science Teaching, 29*, 409-414

Appendix 1: DNA Isolation Rubric

Score / Criterion		3 (Highly successful)	2 (Acceptable)	1 (Needs to be developed)	Score
Materials used in the experiment		All materials used in the experiment were documented in full.	Many of the materials used in the experiment were documented.	Materials used in the experiment were not documented.	
Realization of the experiment		All phases related to the realization of the experiment were correctly listed.	Phases related to the realization of the experiment were somewhat correctly listed.	Phases related to the realization of the experiment were listed completely incorrectly.	
Observation of the experiment	**Observation 1**	Completely correct observations were made about the place of the structure we observe in the experiment.	Somewhat correct observations were made about the place of the structure we observe in the experiment.	Completely incorrect observations were made about the place of the structure we observe in the experiment.	
	Observation 2	Correct observations were made as to what the structure we observe in the experiment looks like.	Somewhat correct observations were made as to what the structure we observe in the experiment looks like.	Completely incorrect observations were made as to what the structure we observe in the experiment looks like.	
Results/Interpretation of the Experiment	**Interpretation 1**	Completely correct results were arrived at as to why mechanical breaking is performed.	Partially correct results were arrived at as to why mechanical breaking is performed.	Completely incorrect results were arrived at as to why mechanical breaking is performed.	
	Interpretation 2	Completely correct results were arrived at as to why Solution A is used.	Somewhat correct results were arrived at as to why Solution A is used.	Completely incorrect results were arrived at as to why Solution A is used.	
	Interpretation 3	Completely correct results were arrived at as to what function filtration has in the experiment.	Somewhat correct results were arrived at as to what function filtration has in the experiment.	Completely incorrect results were arrived at as to what function filtration has in the experiment.	
	Interpretation 4	Completely correct results were arrived at as to what function the pineapple juice has in the experiment.	Somewhat correct results were arrived at as to what function the pineapple juice has in the experiment.	Completely incorrect results were arrived at as to what function the pineapple juice has in the experiment.	
	Interpretation 5	Completely correct results were arrived at as to what function the alcohol has in the experiment.	Somewhat correct results were arrived at as to what function the alcohol has in the experiment.	Completely incorrect results were arrived at as to what function the alcohol has in the experiment.	
				TOTAL SCORE	