

Development and use of a rubric to assess undergraduates' problem solutions in Physics

Aysel KOCAKÜLAH*

Mathematics and Science Education Department, Balikesir University, Balikesir, Turkey

ORCID: 0000-0002-3472-4707

Article history

Received:
09.09.2021

Received in revised form:
18.12.2021

Accepted:
30.12.2021

Key words:

electromagnetic induction;
rubric development;
university physics;
assessment; problem solving

The aim of this study is to develop and apply a rubric to evaluate the solutions proposed for questions about electromagnetic induction belonging to university second year pre-service teachers. In this study which has pretest-posttest quasi-experimental design with control group, teaching of the topic of electromagnetic induction was applied to both groups with the same teaching method and a test consisting of four questions was applied before and after the teaching. 73 students in the experimental group were informed about the properties and usage of rubrics and asked to create a rubric. The effect of rubric on the success of students was examined by applying descriptive statistics and t-tests to the scores obtained from both tests. The validity and reliability of the scoring with rubric were analyzed by calculating the linear regression, t-test statistics, Pearson correlation, Intraclass correlation and Cronbach-alpha correlation coefficients. The results of the analyses show that the developed rubric was used consistently by the researcher and an independent coder and there was a high and significant ($p=.000$) relationship between the scores for all questions. In the inter-rater reliability analysis for each question, the lowest ICC coefficient was specified as .826. In light of the findings obtained from the study, it was concluded that the developed rubric helped to make consistent and stable ratings independent of the scorer, to determine the characteristics of the problem solutions and to increase the level of achievement of students. Another type of analytical rubric can be created and used for other topics of the area of Physics and the results can be compared with the findings of this study.

Introduction

A rubric is a tool that evaluates students' work according to the criteria defined by predefined performance levels. Rubrics with well-defined scoring criteria and definitions of performance levels can help to assess students' understanding. Moreover, rubrics are used to provide teacher and peer feedback, and to review what students should do to improve their understanding (Brookhart & Chen, 2015).

There are many case studies on the formative use of rubrics to improve student performance

* Correspondency: ayselko@balikesir.edu.tr

(Panadero & Jonsson, 2013). In these studies, it was reported that the students had 'increased transparency' (Jonsson, 2014; Reynolds-Keefer, 2010), higher 'self-confidence' (Andrade & Du, 2005; Andrade, Wang, Du & Akawi, 2009) and 'increased performances' (Auxtero & Callaman, 2021; Shadle, Brown, Towns & Warner, 2012) by informing them about what was expected in advance. Lee and Cherner (2015), Vercellotti and McCormick (2021) consider the explicit definition of evaluation criteria of rubrics and scoring levels as crucial. They emphasize that the educational potential of rubrics in teaching practices may not be analyzed by teachers when evaluation criteria and scoring levels are not clearly defined. From the researcher's perspective, a carefully designed rubric will help students and teachers define what the criteria for a successful product or process are before starting a given task and during the completion of the task. In this sense, besides the evaluative role of the rubrics, they can play an instructive role.

The role of rubrics in students' learning

There are few experimental studies on the value of rubrics to increase pupils' learning. Howell (2014) reports that only 10 out of 75 studies conducted in the past 40 years have investigated the development of students' learning through rubric use. Additionally, Reddy and Andrade (2010) point out that most studies did not define the process of rubric development. Cheng and Chan (2019) also concluded that students should actively participate in the development of the rubrics and learn how to use them for supporting students' learning and increasing their success. Menendez-Varela and Gregori-Giralt (2018) found that participation in the rubric design and moderation discussions promoted the development of assessment skills and rubrics may support classroom debate if they are considered as instructional resources. It is understood from the results of these studies that students should actively use rubrics during teaching and know how to use rubrics to support learning.

When learning is considered in the conceptual understanding dimension, it means that scientific facts and algorithms are understood beyond memorization and a dynamic process between what is desired and achieved in the students is structured. In this sense, it is expected students to move from answers containing scientific terms and formula-based solutions to answers integrating new concepts and making connections between those concepts to achieve transitions to higher-order thinking in explaining scientific truth (Claesgens, Scalise, Wilson & Stacy, 2009). However, it is difficult to explain concepts using the content knowledge for students, especially the concepts related to electromagnetism, regardless of age group (Beer, 2010). Additionally, studies in physics education focus on the topic of mechanics but electromagnetism attracts less attention (Dori & Belcher, 2005; Zuza, De Cock, van Kampen, Kelly & Guisasola, 2020).

Findings reported in studies related to electromagnetic induction

The abstract structure of the concepts such as magnetic field, magnetic flux, magnetic force in electromagnetism is shown as the most important reason of students' difficulties (Kocakulah, 2002; Thong & Gunstone, 2008). Additionally, the topic of Electromagnetic Induction (EI) requires students to use and associate the basic concepts of electricity and magnetism. In fact, the students often confuse the concepts of field and flux (Karam, 2014; Kocakulah, 2003). Studies conducted at the university level show that many students could not explain the change in flux and how the induced current would be formed in a conductor without physical interaction with the field source (Härtel, 2018; Zuza, Almudi, Leniz & Guisasola, 2014). These problems cause students to use Faraday's law without understanding and to mix the area of conductor and its area in a magnetic field.

It has been revealed that some students use primitive concepts that contain naive explanations when they encounter more abstract and mathematically structured electromagnetic concepts at the university. These explanations include observable properties such as 'number of magnetic field lines passing through the coil', 'proximity to the magnetic field source', 'contact with field' and 'the direction of the current in the coil' (Zuza et al., 2014). Studies in the literature also indicate that students prefer writing the formula and reaching the result without considering the conceptual basis of the event in the solutions of problems related to EI (Albe, Venturini & Lascours, 2001). Thong and Gunstone (2008) emphasize that the relationships between concepts should be specified with qualitative explanations as well as the mathematical representations because this is the indicator of a good knowledge structure. In this sense, it is necessary to pay attention to the students' qualitative explanations as well as their mathematical representations and how they associate these with each other in the solutions of problems related to EI.

In many studies conducted in physics education, the focus has been on improving students' problem-solving skills and step-by-step problem-solving strategies are used to achieve this goal (Hull, Kuo, Gupta & Elby, 2013). Moreover, there are few studies conducted with university students on the problems of learning EI. However, there is no study in the literature about rubric use in solving EI problems. It has therefore emerged that an experimental study should be conducted showing that step-by-step problem solutions can be carried out using rubrics in teaching EI in physics. In addition to the rubric created by the researchers, there are few studies in which students are included in the process of creating rubrics, and the role of rubric in students' learning is a point that needs to be investigated. Such a study will prove that rubric can be an instructional resource besides its evaluative role.

Considering the above points, this study aims to develop and apply a rubric to evaluate the performance of students related to EI by considering their conceptual difficulties in EI and solving the problems. The students were asked to design a rubric as it is useful in seeing their learning goals and evaluating their solutions to a given problem. In particular, the researcher aimed to reveal whether the students who solve the questions using the rubric are more successful in problem solving than those of others. The effectiveness of a rubric depends on making a valid and reliable assessment of students' problem solutions (Reddy & Andrade, 2010). Therefore, this study also aims to evaluate the effectiveness of the rubric by assessing consistency between raters in coding students' solutions.

The concept of EI was selected considering students' difficulty in solving the problems related to EI in the researcher's previous teaching experiences in electromagnetism course. These challenges include failure to determine magnetic field source and the direction of the magnetic field, the inability to determine the direction or magnitude of the magnetic force, the inability to determine whether there is a change in the magnetic flux, the inability to understand the meaning of the Faraday's law and the inability to determine the direction of the induced current using the Lenz's law. These challenges and poor problem-solving performances of the students as well as the structured problem-solving strategy proposed by Gaigher, Rogan and Braun (2007) to improve students' conceptual understanding were considered to develop a rubric. The problem-solving strategy of Gaigher et al. (2007) is preferred because it gives importance to the qualitative analysis of problems and improves the conceptual understanding of the conservation of energy, which is similar to EI in terms of difficulty.

The main purpose of this research is to develop and apply a rubric to examine whether the use



of rubrics may lead students to perform better in solving problems related to EI in physics. To analyze the students' problem-solving performance, the researcher sought answers to the following questions:

- (1) To what extent does the designed rubric yield a valid and reliable assessment of students' problem solutions?
- (2) Do students in the process of developing and using rubrics perform better in solving problems than students who have never used rubrics but instructed with the same teaching methods?

Method

Research design

In this study, a quasi-experimental pretest-posttest control group design was used. Having determined that the four existing classes were equivalent in terms of their abilities in physics, each class was assigned randomly to either the experimental or control group. The same content and teaching strategy were used, and the same tests were applied in both groups except the development and use of rubric in the experiment group during the academic term of 2018-2019. Maximum effort was shown by providing the same teaching conditions (dealing with the same problems, studying the same simulations or visuals, etc.) and environment (using the same classroom in different days of the week) in both groups of students to minimize the Hawthorne effect indicated by Caleon and Subramaniam (2005). It should be noted that experiment group students are not subjected to extra activities or problem-solving practices related to EI during the development of the rubric. The control group students strived at the same rate as the experiment group to attend the activities related to EI.

Context and participants

A total of 142 second-year undergraduate students who were studying in four intact classes in the primary mathematics education department at a state university participated in this study. Ethical principles taken into account for the participants (Vanclay, Baines & Taylor, 2013) in the study are as follows. First, the students were briefed about the purpose of the research and consequences of them of taking part to gain informed consent. Secondly, students were also told that participation was voluntary and that they would not be subjected to any coercion or threat of harm for non-participation. Moreover, participants were reminded that they could withdraw at any time and that their already collected data could be removed from the analysis whenever possible. Furthermore, students were asked to respond in confidence by offering them protection of their privacy.

The rubric designed in this research (see Appendix A) has been developed to be a supportive tool during the solution of the problems of EI in the general physics course which is compulsory for the second-year undergraduates (18-19 years old), and to be an evaluation tool after the completion of the solutions. Students had been taught electromagnetic induction in the 11th grade (16 years old) of the high school and in the university entrance examination preparation stage.

The students were divided into classes by the Council of Higher Education according to the university entrance exam scores from the higher to the lower scores. To check for differences between classes, the university entrance exam scores of students were compared with the independent samples one-way analysis of variance (ANOVA) test. The analysis results

showed that there were no statistically significant differences between the averages of four classes ($F=2.48$; $p=.06$). In addition to the scores of the university entrance exam, an achievement test was administered to compare the students in terms of their abilities in physics.

There were 69 students in the control and 73 students in the experiment group. In both groups, the gender ratio was in favor of girls. In the control group, this rate was 3.928:1 with 55 girls and 14 boys, while in the experiment group it was 2.842:1 with 54 girls, 19 boys. By doing power analysis, when group sizes were selected as 69 and 73, $\beta:\alpha$ rate was selected as 1 and $d=.5$ medium effect size was selected, the power value was calculated as $P=.90$. Since this value is greater than the conventional value of .80 (Dattalo, 2008), it is considered that the study has sufficient power with the composed sample size.

Data collection tools

In this study, an achievement test and a test involving physics problems about EI were used as data collection tools. The achievement test includes questions about the topics of the magnetic field of magnets (4 questions), the magnetic fields of current carrying a straight and a circular conductor (9 questions), the magnetic force acting on a charged particle (8 questions) and a current carrying conductor (5 questions), the induced current (4 questions), the self-induction (2 questions) and induced emf (5 questions) in high school physics courses.

The first form of the achievement test, devised by the researcher, consists of 40 multiple choice questions. The questions were reviewed with the opinions of three experts in the field of physics education regarding their face validity, comprehensibility of questions and suitability for the age level concerned. The test was applied to the group consisting of 143 students, who took general physics course, in the science education department of the same university and a pilot study was performed in spring 2018. A 22-item test was developed as a result of the item analysis of the data obtained from this pilot study. The final test had a Cronbach alpha of .82 and an average test difficulty index of .55. The final form of the achievement test was applied at the same time to the sample of this study in four classes in spring 2019. Achievement test data were analyzed with ANOVA test which also showed no significant difference between the four classes of the real study ($F=1.11$; $p=.35$, see Table 1).

Table 1. Comparison of the differences between the classes according to the achievement test scores.

Variance	Sum of squares	df	Mean square	F	p
Between groups	361.55	3	120.52	1.11	.35
Within groups	14978.33	138	108.54		
Total	15339.88	141			

The test comprised of four questions related to EI (see Appendix B) was administered to all four classes as a pre-test and post-test. After examining the physics textbooks to determine the conceptual areas of EI topic to be covered, the questions were designed by the researcher. The content validity of the questions was determined by taking the opinion of three experts in the field of physics education who carried out the teaching of general physics courses at the university for more than 10 years. After testing the appropriateness of the questions and the suitability of their use for the general physics course, small changes were made in the figures and question roots in line with the feedback given by the experts. The test was administered after the experts approved the final version of the test in terms of coverage of EI topic and comprehensibility of the questions.



Experimental process

Development of rubric

The teaching of Faraday's induction law unit was completed in two weeks and eight hours of lessons in both groups. In the next week, the students of the experiment group were educated on what a rubric was, what types of rubrics were, and how they could be developed and used. After the detailed introduction of rubrics, the students were asked to develop a rubric in small groups to evaluate the solutions to the problems related to EI. The students were divided into 15 groups in two classes, with approximately 4-5 students in each group, and were asked to form their rubrics as a group within one week. In the following week, students were asked to explain how they formed the rubrics and how useful and efficient their rubrics were in evaluating the problem solutions. While the students were free to decide on the type of rubric (holistic or analytical), it was stated that the developed rubric would be an effective tool to be able to make a reasonable and valid evaluation about learning of a student who solved the problem considering the evaluation criteria.

The following week, the groups started to present their rubrics. The most practical rubric was determined in two stages, using the sum of the points given by the researcher, the three experts in the field of physics education and spokesperson of a group that reflect the opinion of each group. Firstly, groups were allowed to make presentations in their classes and the highest scored rubric was determined. In the second stage, two classes were gathered in the conference hall and the rubrics, which received the highest score in both classes, were presented once again and the best rubric was chosen. Thus, instead of using a ready-made rubric, it was tried to ensure that all students participated in the rubric development process.

To score the solutions to the questions about EI, three groups chose to develop holistic rubrics while 12 groups developed analytical rubrics. The analytical rubric, which had six criteria topics, became the first in the final elimination. The selected rubric was then reviewed with the class and discussed whether it is a valid tool for performance assessment. In this discussion, six criteria topics in the rubric, the ratio of the criteria topics in the total points, the sub-evaluation criteria that form the criteria topics and the point values of these sub-criteria were finalized by reaching consensus on the views of the students. The correct order of the criteria titles and the percentage of scoring and the validity of expressions of the sub-evaluation criteria were verified by taking the approval of four experts in the field of physics education and two experts in the field of measurement and evaluation.

It should be noted that studies in both experimental and control groups were conducted by the same researcher. While the control group solved the selected questions without using the rubric, the experiment group discussed the feasibility of the rubric on these questions for two weeks period. The evaluation criteria of the rubric can be explained in detail by looking the solution of the first problem in Figure 1.

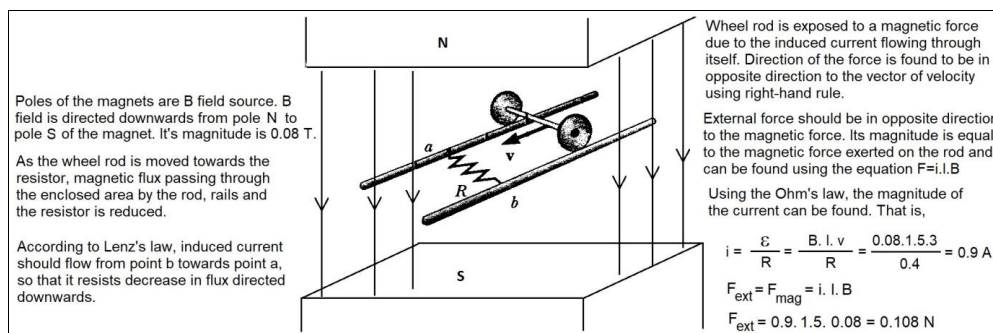


Figure 1. Solution of the first question using the designed rubric.

First, the source of the magnetic field should be determined, and the direction of this field should be indicated correctly. In the first question, the magnetic field source is a magnet, and the magnetic field is directed downwards from the north to the south pole. When the wheel rod is moved between the poles of the magnet, students must indicate that the flux passing through the closed circuit consisting of rails, rod and resistance has been reduced. The correct explanation of the 'flux change', which is the key concept in the formation of induced emf, is a prerequisite for the correct application of Lenz's law. Thus, students can determine the direction of the induced current using Lenz's law and the direction of the induced current for the first question will be from point b to point a. After the direction of the induced current is determined, the direction of the magnetic force can be found by using the right-hand rule. To find the magnitude of the magnetic force generated in the next stage, the value of the induced current is obtained with the $i = \mathcal{E}/R = B l v/R$ equation from the Ohm's law. Finally, the equation of $F = i l B$ is solved and the result is calculated. The exact solution of the problem is completed by checking whether each physical quantity is written in the correct system of units.

Use of the rubric

The experiment group was informed that they would use the rubric as an auxiliary tool while solving the EI problems and the problem solutions would be scored according to the evaluation criteria in the rubric by the researcher. When the teaching of the subject was completed, the students were asked to solve the problems in the post test via the rubric. Then, the researcher scored the solutions of the experiment and control group students' pre-test and post-test questions using the rubric. To ensure confidentiality in the scoring of both tests, students were asked to write their chosen nickname on the tests. Once the researcher completed the scoring of students' solutions, the data analysis was started.

Data analysis

Rubric-based scores given by the researcher to the solutions of the students to pre-test and post-test questions in the experiment and control groups were analyzed to explore quantitative data. Descriptive statistics were used to determine the overall mean, standard deviation, skewness and kurtosis values and their standard errors, and the range of the researcher's scores for all questions.

Whether the analyzed variables had a normal distribution was determined by using Liliefors correction in Kolmogorov-Smirnov test and by calculating skewness and kurtosis values. Both skewness and kurtosis values were found to be ranging between -1 and $+1$. Additionally, histograms with normal curves, normal Q-Q plots, detrended normal Q-Q plots were also examined to be sure that assumptions for normality are fulfilled (Tabachnick &

Fidell, 2015). It was found that the scores of the students in the experiment and control groups in the pre- and post-tests were normally distributed. For this reason, parametric statistics (t-tests, ANOVA, Pearson correlation, etc.) were used in the analysis of the data.

The qualitative definitions of the responses given to the pre-test and post-test questions and the conceptual features were presented to make comparisons of these features. First, it was described generally how students solve the pre-test questions in both groups, and then, the solution of a pre-test question of one student from each group was presented. Moreover, the same two students' solutions were given by outlining the general characteristics of solutions in the post test.

During qualitative analyses, six dimensions of the rubric (see Appendix A), which show the correct order of problem solutions in EI, were considered. The first dimension involves the definition of the source, direction and magnitude of the magnetic field. The next dimension includes whether there is a change in magnetic flux and weights 25% of total performance in the solutions. The third-dimension weights 30% of the solutions and requires identifying the direction of the induced current. In the fourth dimension, direction, and magnitude of magnetic force, which weights 20% of the performance, should be defined. In the next two dimensions, equations should be solved and written in correct units and symbols that weight 10% and 5% respectively in the complete solution to the related question.

To determine the validity of the scoring to the students' solutions, linear regression analyses were performed to check the congruence between the scores of the researcher and the second coder. Another way to examine the validity of the scoring was to check whether the inferences based on the scores collected with the rubric were supported by evidence collected by another instrument that was used as a criterion (Wu, Wu & Hsu, 2014). In this study, criterion-related validity was calculated by examining Pearson's correlations between the pre-, post-test rubric scores and achievement and final exam scores respectively. The final exam test of the general physics course includes 24 multiple-choice questions which are similar to the questions solved by students using the rubric. The average difficulty index of the final exam test was .60. Based on the item analyses, the difficulty level of the items ranged from .27 to .92 values and the discrimination index of the items ranged between .31 and .68 values. The Cronbach-alpha reliability coefficient of the test was .82.

Several methods were used to determine the reliability of coding. Firstly, an independent coder in the physics education department was asked to score the solutions of the students to the four questions by using the rubric. The second coder was one of four physics education specialists who followed the development process of the rubric from beginning to end and examined the criteria topics and the scoring percentages of these topics and the statements of the sub-evaluation criteria. When the scoring of the second coder was completed, independent samples t-tests were conducted to determine the difference between the scores of the researcher and the second coder.

Secondly, the Pearson's correlation analysis between the scores of the researcher and the second coder was performed to determine the agreement between the two raters. However, Weir (2005) argues that Intraclass Correlation Coefficient (ICC) should be preferred because the correlation analysis does not determine the systematic errors made. Therefore, ICC values were calculated for each of pre- and post-test question scores of two coders.

Lastly, the reliability of the rubric as a scoring tool can be evaluated by using Cronbach's

alpha coefficient, since the same student's solution to four different questions is scored using the rubric by each coder. This coefficient is an indicator of the internal consistency of the rubric and generally Cronbach's alpha values over .70 are considered as acceptable (Angell, 2015; Kocakulah, 2010). Thus, the Cronbach-alpha reliability coefficient was also calculated to investigate the single coder reliability of the researcher for both tests.

Results

Investigation of the assessment scores

Each question in the pre- and post-tests is evaluated on a maximum of 100 points using the rubric. Descriptive statistics are presented in Table 2. In the control group, the mean score given by the researcher in the pre-test was 8.03 and the standard deviation was 3.57, after the instruction, it reached the value of 25.18 ± 7.14 by increasing in the post test. While the average score before the instruction in the experimental group was 8.19 ± 3.83 , with a notable increase in the post test, it reached the value of 80.77 ± 7.45 .

Table 2. Descriptive statistics of the researcher scores in two groups.

Researcher scores	Experiment group		Control group	
	Pre-test	Post-test	Pre-test	Post-test
Mean	8.19	80.77	8.03	25.18
SD	3.83	7.45	3.57	7.14
Skewness (p*)	0.45 (0.20)	-0.55 (0.20)	0.23 (0.20)	0.14 (0.20)
Range	0.50-19.00	59.50-96.25	1.00-17.75	12.25-42.75

* Shows the value of significance obtained from the Kolmogorov-Smirnov test for normality

Paired samples t-test was used to investigate the changes in the test scores of the groups. The pretest and posttest mean of the experimental group was statistically significant ($t=73.04$; $p=.000$; $d_{\text{rep-means}}=8.79$) while a considerable increase was observed in the test averages of the control group ($t=22.99$; $p=.000$; $d_{\text{rep-means}}=2.77$). On the other hand, difference between the pre-test means of the experiment ($\bar{X}_e = 8.19$) and control ($\bar{X}_c = 8.03$) groups was analyzed with independent samples t-test and no statistically significant difference was found ($t=.26$; $p=.80$). However, there was a significant difference between the post-test mean scores of the two groups ($\bar{X}_e = 80.77$; $\bar{X}_c = 25.18$; $t=45.35$; $p=.000$) and Cohen's d was calculated as $d=7.62$ which corresponds to a large effect (Cohen, 1988).

Qualitative descriptions about problem solutions

A qualitative description of the answers to the questions is presented below. Before the instruction, the students in both groups intuitively responded to the question by writing emf or magnetic force formula without finding the source of the magnetic field and indicating the direction of the magnetic field. In doing so, they mixed the $\varepsilon = B l v \sin \alpha$ formula with the formula of magnetic force and mostly wrote $F = B l v \sin \alpha$. This confusion exists not only in the formula but also in the application of the right-hand rule. The students, who try to apply the right-hand rule, use magnetic field and current direction and try to define the direction of emf, instead of defining the direction of magnetic force.

In all questions, flux change should be analyzed thoroughly and induced current or emf should be interpreted accordingly, but students cannot go beyond specifying the direction of magnetic field and/or magnetic force. This distress is manifested by their use of the concept of

current instead of flux. Indeed, students answered the questions as, for example, "to increase the current, we should pull the rod" and "when the switch is closed, the current passing through the ring increases".

The main problem for the students in the pre-test is the inability to define how the induced current is produced. In this case, the students find the way out by not questioning whether there is a change in the magnetic flux through the area that is enclosed by the conductor but by considering the direction of the current generating the magnetic field (question 4) or the direction of the current supplied from the power source (question 2), and by determining the direction of induced current in the same or opposite direction to that current supplied from the power source.

Figure 2 shows the solutions of the students to the first question before the instruction. In both groups, the students did not draw the direction of magnetic field, force and emf. While the experiment group explained that "magnetic field is directed from the pole N to the pole S", the control group wrote unsure statements such as "the direction of F from N to S may be correct". Control group tried to calculate the magnitude of the magnetic force with the equation of $F=B\Im l$ by confusing the emf formula with magnetic force formula. The experimental group student, however, was affected by the electric force formula in electrostatics and derived the $F=Bi$ equation for magnetic force but could not go any further. In both groups, the students tried to calculate the magnetic force without mentioning the formation of the induced current and the change in magnetic flux.

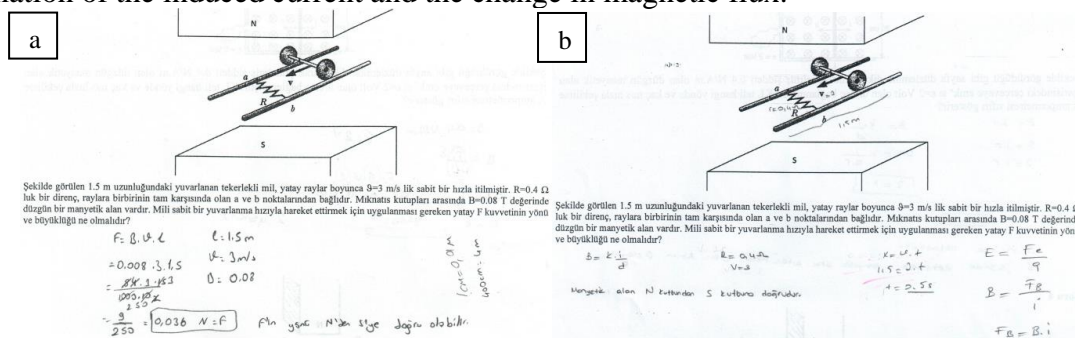


Figure 2. Solutions of the selected students in the a-control and b-experiment groups to the first question in the pre-test.

Figure 3 presents the same two students' responses to the first question in the post test. The student in the experiment group started by defining the source of the magnetic field, and then indicated the direction of the magnetic field. After defining the closed circuit in which the induced current will flow, it was emphasized that there is a decrease in the flux as the closed area decreases with the movement of the wheel rod. The student stated with a reason that the decrease in flux would create an induced current which should flow from point b to a. Although the student did not associate the direction of the magnetic force with the induced current by using the right-hand rule, she found it in the correct direction according to the direction of the external force. In the last stage, emf and current values were miscalculated, and the magnitude of the magnetic force was not found.

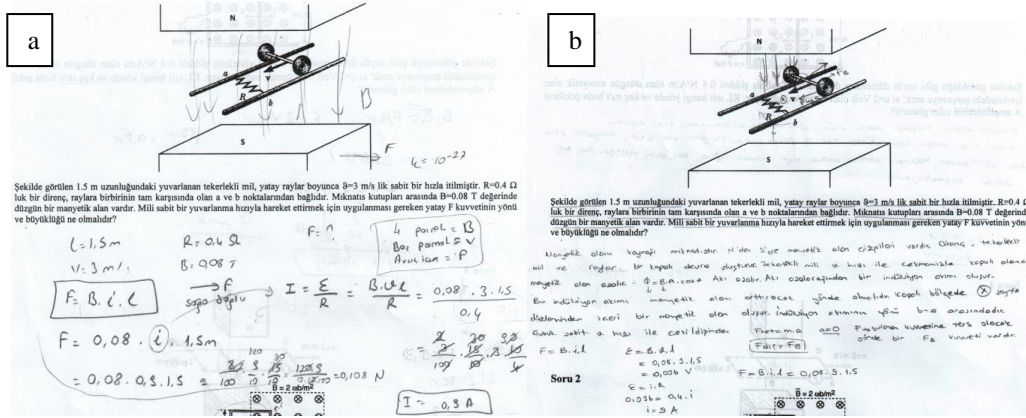


Figure 3. Solutions of the selected students in the a-control and b-experiment groups to the first question in the post-test.

After the instruction, the student in the control group showed the direction of the magnetic field lines between the magnet poles correctly. He showed how he found the magnetic force direction by indicating how he applied the right-hand rule for the conductor frame. However, he drew magnetic force vector incorrectly on the figure instead of the direction of the induced current and confused these two concepts together. It was enough for the student that the conductor should be in the magnetic field and move to form an induced current. The formation of the induced current and its effect on the movement of the conductor is ignored. Finally, the student, who wrote the equations for emf and magnetic force as correct, reached the correct result with the right system of units.

Testing the validity of the rubric

A regression analysis was performed to check the validity for each of the four questions in the rubric. In this way, the accuracy of the scoring of the researcher was checked by looking at the agreement level between the scores of the researcher and the second coder. As shown in Table 3, it was found that there were significant positive functional relations (see values R and p) related to the scores of two coders for the questions in both tests.

Table 3. Regression analyses of the researcher's rating on the independent coder's rating for each question in the pre- and post-tests.

Group Type	Question Number	Pre-test					Post-test				
		b	R	R ²	F	p	b	R	R ²	F	p
Control	1	.974	.983	.966	1913.024	.000	.988	.988	.976	2750.079	.000
	2	.958	.986	.972	2347.377	.000	.933	.983	.967	1935.978	.000
	3	.857	.827	.684	144.853	.000	.948	.970	.942	1080.880	.000
	4	.929	.957	.917	737.576	.000	.981	.992	.985	4388.370	.000
Experiment	1	.959	.966	.933	983.610	.000	.979	.986	.972	2455.985	.000
	2	.915	.905	.819	320.536	.000	.994	.970	.942	1148.907	.000
	3	.925	.897	.805	293.552	.000	.921	.942	.886	554.246	.000
	4	.938	.938	.880	521.260	.000	.938	.939	.881	527.903	.000

Secondly, the correlation analysis was made to calculate criterion-related validity of the scoring by rubric and the relationship between the rubric scores and the selected test scores that were taken as criteria was investigated. When the achievement test scores were taken as criteria for the pre-test scores, the correlation coefficient was calculated as r=.661 for the control group and r=.684 for the experimental group. On the other hand, when the final exam



scores were accepted as criteria for the post-test rubric scores, the correlation values of $r=.761$ and $r=.784$ were obtained for the control and experiment groups respectively. As the correlation values were found to be significant at .01 level ($p=.000$), it can be assumed that the rubric scores were consistent with the scores of achievement test and final exam and were valid to determine the cognitive level of the students.

Reliability of the rubric in scoring students' solutions

Firstly, independent samples t-test analysis was performed to determine the level of agreement between the researcher and the second coder in assessing the scores for the problem solutions in both groups. The t-test results showed no statistically significant differences in the pre- and post-tests as shown in Table 4.

Table 4. Results of t-test analyses showing a comparison of researcher and independent-coder scores.

Question number	Coder	Pre-test						Post-test					
		Control			Experimental			Control			Experimental		
		\bar{X}	SD	t (p)	\bar{X}	SD	t (p)	\bar{X}	SD	t (p)	\bar{X}	SD	t (p)
1	A	8.23	5.13	.07	8.64	5.64	.06	18.12	6.86	.26	83.53	10.26	.20
	B	8.29	5.18	(.95)	8.70	5.67	(.95)	17.81	6.86	(.79)	83.19	10.33	(.84)
2	A	9.58	6.80	.01	9.62	6.54	.11	27.55	10.13	.02	81.26	10.29	.27
	B	9.56	6.99	(.99)	9.74	6.46	(.91)	27.58	10.67	(.98)	81.71	10.04	(.79)
3	A	7.10	5.00	.19	7.30	4.38	.04	25.96	8.20	.02	80.44	7.99	.19
	B	6.94	4.82	(.85)	7.33	4.25	(.97)	25.98	8.39	(.98)	80.70	8.17	(.85)
4	A	7.20	4.57	.07	7.19	3.36	.12	29.22	13.72	.05	77.84	7.91	.13
	B	7.26	4.71	(.94)	7.26	3.36	(.90)	29.33	13.88	(.96)	77.67	7.92	(.90)

* A denotes the researcher and B denotes the independent coder

Secondly, the inter-rater reliability was calculated by using the two-way mixed model and the consistency procedure for each question and the average of four questions in the pre-test and the post-test. As seen in Table 5, the ICC coefficients calculated for the average scores of raters are above .90, which indicates an excellent agreement among the raters. When the raters are compared based on each question, the ICC coefficients yield excellent agreement among coders except for the third question which indicates good level of agreement with the .897 and .826 values in the pre-test scores for both the experiment and the control groups, respectively. The 95% confidence interval of the third question for the experiment group ranges between .841 and .934 that indicates the level of reliability changes from good to excellent. On the other hand, the lower and upper limits of 95% confidence interval for the third question of the control group range between .734 and .889, which indicates the change in the level of reliability from moderate to good. Although the ICC value of the second question for the experiment group in the pre-test is above .90, its 95% confidence interval ranges between .852 and .939 with a changing reliability level from good to excellent. However, the ICC values for the scores of the rest of the questions in the pre-test and each question in the post test for both groups are over .90, including the lower limit of the 95% confidence interval that indicates an excellent level of reliability.



Table 5. ICC estimates with 95% confidence interval for rubric scores.

Test type	Group type	Question number	ICC	95% confidence interval (lower-upper bounds)	Test type	Group type	Question number	ICC	95% confidence interval (lower-upper bounds)
Pre-test	Control	1	.983	(.973-.989)	Post-test	Control	1	.988	(.981-.993)
		2	.986	(.977-.991)			2	.982	(.971-.989)
		3	.826	(.734-.889)			3	.970	(.952-.981)
		4	.957	(.931-.973)			4	.992	(.988-.995)
		1-4	.969	(.951-.981)			1-4	.961	(.938-.976)
	Experimental	1	.966	(.946-.978)		Experimental	1	.986	(.978-.991)
		2	.905	(.852-.939)			2	.970	(.953-.981)
		3	.897	(.841-.934)			3	.941	(.908-.963)
		4	.938	(.903-.961)			4	.939	(.904-.961)
		1-4	.963	(.941-.976)			1-4	.986	(.977-.991)

Finally, Cronbach-alpha coefficient was calculated to determine the reliability of the researcher's scores for four questions. The reliability of coding made by a single rater is close to the threshold value of .70 and takes values above .70 as can be seen in Table 6.

Table 6. The reliability analysis results of the researcher's scores for four questions in the pre- and post-tests.

Type of analysis	Control group		Experimental group	
	Pre-test	Post-test	Pre-test	Post-test
Cronbach-alpha coefficient	.71	.73	.74	.78

Discussion and conclusions

In this study, the effect of a rubric developed with students on their problem-solving performances related to the EI was investigated. It was found that the average of the post-test scores of the experimental group in which the rubrics were used in problem solving process was significantly different from the average scores of the control group. Such a result suggests that if students, who use rubrics, know what they need to focus on and what steps they need to follow, their success in problem solving is significantly improved. Similarly, an increase in students' expertise of problem solving in kinematics (Hull et al., 2013) and calculus (Auxtero & Callaman. 2021) and improvement in students' chemical problem-solving skills (Shadle et al., 2012) were reported in other studies in which the students were able to establish a relationship more easily between the qualitative definition of the related concept and its mathematical definitions in the form of equations. These findings indicate that students can provide the correct answer to the problem and present the rationale for the solution indicating the problem is understood correctly using the rubrics in which step-by-step problem-solving stages are determined.

It was tried to take some precautions against the threats to internal validity in this study. For instance, students were randomly assigned to groups and groups were trained by the same person against the history threat. Ariel, Bland and Sutherland (2021) suggest inclusion of a control group as a design consideration which is applied in this study to alleviate the threat of maturation. The comparison of post test scores between the experiment and control group allowed the researcher testing the effect of rubric use beyond the maturational changes within the subjects. Employing a control group that did not use the rubric was also helped to avoid the potential threat of testing and to clarify whether changes in achievement of groups were caused by the development and use of the rubric or by the testing itself. The same strategy is used to reduce the potential threat of regression. Regression effect may be experienced by both groups and can be controlled for by the control group change. A second method to



manage regression is considered to use stratified randomization. It is thought that university entrance examination scores could provide stratified randomization because equivalent numbers of students with high, middle, and low values are randomized to each class in this study. Such randomization to groups may prevent selection threat to internal validity. However, the researcher decided to assess both groups' equivalency prior to study on their physics abilities. Therefore, physics achievement test was used for this purpose.

Reinforcement of motivation strategy (Ariel et al., 2021) was used to contribute to the students' desire to participate and to minimize mortality threat in this study. This was basically achieved by engaging students to the design process of the rubric. Group contamination threat was tried to be managed by limiting communication between the students in both groups. Courses were set in different days to prevent control group students to learn rubric use in the experiment group and somehow apply it themselves that could reduce post-test differences between both groups. Statistical powers of the tests were provided to make valid interpretations and normal distribution of data from each group was checked to ensure that certain assumptions of parametric statistics were not violated against the threat of statistical conclusion validity. Moreover, accuracy in measurement over time is considered as vital in reducing instrumentation threat which is reported to be posed when the test does not have adequate reliability (Ariel et al., 2021). Thus, a rubric with accepted levels of reliability and validity was developed and used to reveal any difference between groups as a result of experimental manipulation.

The validity of the rubric was examined by applying regression analysis to the researcher's and the second coder's scores for each question. The high correlation values between the scores of the coders for both tests (the lowest value is $r = .827$ for the scores of the third question of the pre-test of the control group) show that both coders used the rubric in the same way and similar scorings of the raters can be attributed to a clear and the same interpretation of the scoring criteria of the rubric designed in this study. This finding also suggests that the developed rubric in this respect can be used to make valid evaluation of students' performance independent of the scorer.

In this study, not only the content validity of the rubric is ensured by providing how the rubric is constructed, but also the criterion-related validity is presented by comparing the scores obtained from the rubric with external criteria such as achievement test and final exam scores. On visiting the literature, external validity was reported with modest correlation values between .40 and .60, in many studies based on the correlations between the rubric scores and other measurements such as post course evaluations or tests of prior knowledge (Jonsson & Svingby, 2007). In this study, the correlation values between the pre- and post-test rubric scores and achievement and final exam scores were found to be lowest .661 and highest .784, respectively. These values are higher than the values obtained in other studies in the literature, and it could be argued that the scores obtained from the rubric developed in this study can be a good predictor of the students' pre-knowledge and post-teaching status.

To reveal the reliability of the scoring by rubric for the problem solutions of the students, ICC coefficients were also calculated. ICC coefficient, which is deemed to be a more desirable measure of inter-rater reliability and reflects together the level of consistency by presenting correlation between the scores of the raters and the level of consensus by indicating whether the raters have the same score, should have the value above .90 that indicates excellent reliability (Gray, Connolly & Brown, 2019; Koo & Li, 2016). Yet, in the literature, no study gives the ICC values of inter-rater reliability for the rubric scores of the students' solutions to

the physics problems. Furthermore, Brookhart and Chen (2015) argue that appropriate reliability values would change according to the purpose of the study and that the ICC value of .80 and above is considered acceptable in the studies. It may therefore be stated that the ICC values in this study are acceptable and student performances can be interpreted very similar by raters using the developed rubric.

The Cronbach-alpha reliability analysis of the same coder scores for four different questions in the same content yielded values above .70 for the pre- and post-test questions in both groups. This finding is consistent with the results of Kocakulah's (2010) study in which the values of .77 and .75 obtained for the pre- and post-test data from the analytical rubric that was developed to score undergraduates' solutions of problems in mechanics. Obviously, distinct feature of the evaluation criteria and performance level definitions allow the raters how to use the rubric. Therefore, coding with a well-designed rubric will not change according to the scorer type and even in questions with different context in the same topic, consistent coding can be made.

In this study, it has been demonstrated that undergraduate students' problem-solving performances on a topic of EI is improved considerably and there is a need for studies using a developed rubric to make more evaluations to students' problem solutions in a selected science topic. The reason for the increase in problem solving performance of the students with rubric is thought to be the preparation of the rubric together with the students instead of giving them a rubric prepared beforehand. Indeed, studies that report no significant increase in students' performance used rubrics with clearly designed criteria and attainment levels but distributed them without instructing students (Panadero & Jonsson, 2013). In this study, the participation of the students during the preparation of the rubric was not only related to how they would use the rubric, but also to understand what the content of the rubric was and what was being evaluated. Such an activity is considered as 'motivational instruction method' by Zheng, Ding, Lu and Branch (2020). In addition, the application of the rubric's criteria on different physics problems is thought to have a significant effect on the development of the students' conceptual understanding. The most important indicator of this was that the experimental group students correctly defined flux changes, determined the direction of the induced current correctly and verified the direction of the current by indicating the direction of the magnetic force during the analysis of solutions to the post test questions. Moreover, it could be argued that the use of the rubric allowed students to consider the concepts of field and flux as separate criteria and prevented the confusion of these concepts as indicated by Karam (2014).

It has been observed that the control group students who do not consider flux change, have tried to use the mathematical equations. This finding suggests that the rubric allows students to review their decisions by focusing on the conceptual basis underlying the physical phenomenon during problem solving and, in the meantime, it eases in what aspect they should consider. Indeed, the control group students have concluded that the conductor which does not carry current will be under the influence of magnetic force when it is in the magnetic field instead of focusing on flux change and its effect on the conductor that forms a closed path. Thus, if students internalize the evaluation criteria of the rubric, they may learn what to consider in problem solving even if they have not been accompanied by a rubric.

In this study, it was found that students' problem solutions could be evaluated in a valid and reliable way with a rubric. In addition to being an assessment tool, it has also been demonstrated in this study that it increases student achievement significantly compared to



students who do not use rubrics, by reinforcing the concepts of the topic to be taught. Therefore, it is thought that the study will contribute to the literature by exemplifying the use of an original rubric that develops students' problem-solving skills in EI and evaluates their problem solutions in a valid and reliable way.

The idea of teaching physics topics using rubrics can help devise a teaching plan that will support students in developing a problem-solving strategy without memorizing the solution to the problem, but by considering which points and in what order they need to be addressed. It is thought that the use of such a strategy by teachers as well as the students during teaching of the topic will be important in terms of creating an effective teaching atmosphere and meaningful learning outcomes.

In this study, students' views and affective characteristics about the teaching process were not taken into account. These points, which are thought to be the limitations of the study, can be investigated in future studies. The sample of the study consists of university students. A similar study can be conducted with students of different age groups and education levels, and the results obtained from this study can be compared. Furthermore, the findings of this study are limited to data obtained from students at a university. Therefore, it is thought that the comparison of the data to be obtained from different university students and the results of the current study will increase the external validity of this research.

The rubric developed in this study can be used in the teaching of EI to deepen students' conceptual understanding and to improve their problem-solving skills by making qualitative and quantitative associations. The criteria in the rubric may allow students to self-diagnose their solutions (Safadi and Saadi, 2021), to analyze the problem conceptually before starting to solve the problem and may provide students a working forward strategy (Larkin, 1985) by turning them into purposeful problem solvers. The method of analyzing the problem conceptually was studied by Zajchowski and Martin (1993) with introductory college students on mechanical questions and they determined that students use the 'working forward' strategy in the same manner as expert problem solvers. Therefore, students can take the opportunity to work on solutions of problems and to see what they can do in solving problems with the rubrics before requesting help from teachers.

Although the validity of the rubric developed in determining the success of problem solving of the students in EI and the reliability values of scoring with the rubric were presented, developed rubric can be examined by using similar sample groups in different cultures by other researchers. This would increase the external validity of this study and demonstrate the usefulness of the rubric developed on problem solving skills. Additional research can be designed to examine gender neutrality of the rubric in assessing problem solving.

More research can be designed to analyze the distribution of the scores obtained from the rubric to compare the scores obtained from different student groups, who took the same course for several years, as Hafner and Hafner (2003) used this approach to examine the consistency of outcomes. In addition, students may be asked to evaluate their peers' problem solutions by using the rubric in future studies. Apart from this implementation, the rubric can be used as a self-assessment form and students can use the rubric to evaluate their solutions in pairs. Selecting a friend during self-assessment would increase the validity of a student's scoring by checking the problem solution with his peer and allow reviewing and improving his learning about the subject by discussing the points he did not understand or misunderstood with his peer. Finally, clues for teaching can be obtained to develop concepts from which the

students have gained low scores within evaluation criteria during the use of the rubric developed in this study.

References

- Albe, V., Venturini, P., & Lascours, J. (2001). Electromagnetic concepts in mathematical representation of physics. *Journal of Science Education and Technology*, *10*, 197-203. doi:10.1023/A:1009429400105
- Andrade, H., Wang, X. L., Du, Y., & Akawi, R. L. (2009). Rubric-referenced self-assessment and self-efficacy for writing. *Journal of Educational Research*, *102*(4), 287-301. doi:10.3200/JOER.102.4.287-302
- Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research, and Evaluation*, *10*(3). doi:10.7275/g367-ye94
- Angell, K. (2015). The application of reliability and validity measures to assess the effectiveness of an undergraduate citation rubric. *Behavioral and Social Sciences Librarian*, *34*(1), 2-15. doi:10.1080/01639269.2015.996481
- Ariel, B., Bland, M., & Sutherland, A. (2021). *Experimental designs*. London: Sage.
- Auxtero, L. C., & Callaman, R. A. (2021). Rubric as a learning tool in teaching application of derivatives in basic calculus. *Journal of Research and Advances in Mathematics Education*, *6*(1), 46-58. doi:10.23917/jramathedu.v6i1.11449
- Beer, C. P. (2010). *How do pre-service teachers picture various electromagnetic phenomenon? A qualitative study of pre-service teachers' conceptual understanding of fundamental electromagnetic interaction*. (Doctoral dissertation). Ball State University.
- Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, *67*(3), 343-368. doi:10.1080/00131911.2014.929565
- Caleon, I., & Subramaniam, R. (2005). The impact of a cryogenics-based enrichment programme on attitude towards science and the learning of science concepts. *International Journal of Science Education*, *27*(6), 679-704. doi:10.1080/09500690500038306
- Cheng, M. W. T., & Chan, C. K. Y. (2019). An experimental test: Using rubrics for reflective writing to develop reflection. *Studies in Educational Evaluation*, *61*, 176-182. doi:10.1016/j.stueduc.2019.04.001
- Claesgens, C., Scalise, K., Wilson, M., & Stacy, A. (2009). Mapping student understanding in chemistry: The perspectives of chemists. *Science Education*, *93*(1), 56-85. doi:10.1002/sce.20292
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Lawrence Erlbaum Associates.
- Dattalo, P. (2008). *Determining sample size: Balancing power, precision and practicality*. Oxford: Oxford University Press.
- Dori, Y. J., & Belcher, J. (2005). How does technology-enabled active learning affect undergraduate students' understanding of electromagnetism concepts? *Journal of the Learning Sciences*, *14*(2), 243-279. doi:10.1207/s15327809jls1402_3
- Gaigher, E., Rogan, J. M., & Braun, M. W. H. (2007). Exploring the development of conceptual understanding through structured problem-solving in physics. *International Journal of Science Education*, *29*(9), 1089-1110. doi:10.1080/09500690600930972
- Gray, J. S., Connolly, J. P., & Brown, M. A. (2019). Measuring intercultural knowledge and competence in college essays: Does a performance-based rubric have construct



- validity? *Studies in Educational Evaluation*, 62, 142-148. doi:10.1016/j.stueduc.2019.05.007
- Hafner, J. C., & Hafner, P. M. (2003). Quantitative analysis of the rubric as an assessment tool: An empirical study of student peer-group rating. *International Journal of Science Education*, 25(12), 1509-1528. doi:10.1080/0950069022000038268
- Härtel, H. (2018). Electromagnetic induction: An alternative for teaching and understanding. *European Journal of Physics Education*, 9(2), 1-13. doi:10.20308/ejpe.v9i2.210
- Howell, R. J. (2014). Grading rubrics: Hoopla or help? *Innovations in Education and Teaching International*, 51(4), 400-410. doi:10.1080/14703297.2013.785252
- Hull, M. M., Kuo, E., Gupta, A., & Elby, A. (2013). Problem-solving rubrics revisited: Attending to the blending of informal conceptual and formal mathematical reasoning. *Physical Review Special Topics-Physics Education Research*, 9(1), 010105-1-16. doi:10.1103/PhysRevSTPER.9.010105
- Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment and Evaluation in Higher Education*, 39(7), 840-852. doi:10.1080/02602938.2013.875117
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2, 130-144. doi:10.1016/j.edurev.2007.05.002
- Karam, R. (2014). Framing the structural role of mathematics in physics lectures: A case study on electromagnetism. *Physical Review Special Topics-Physics Education Research*, 10(1), 010119-1-23. doi:10.1103/PhysRevSTPER.10.010119
- Kocakülâh, M. S. (2002). An investigation of first year university students' understanding of magnetic force relations between two current carrying conductors: A case study. *Hacettepe University Journal of Education*, 23, 155-166.
- Kocakülâh, M. S. (2003). The assessment of university students' conceptual understanding and interpretation abilities on electromagnetic induction. *Journal of Balikesir University Institute of Science and Technology*, 5(2), 101-111.
- Kocakülâh, M. S. (2010). Development and application of a rubric for evaluating students' performance on Newton's laws of motion. *Journal of Science Education and Technology*, 19(2), 146-164. doi:10.1007/s10956-009-9188-9
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155-163. doi:10.1016/j.jcm.2016.02.012
- Larkin, J. H. (1985). Understanding problem representations and skill in physics. In S. F. Chipman, J. W. Segal & R. Glaser (Eds.), *Thinking and learning skills: Research and open questions* (pp.141-159). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lee, C.-Y., & Cherner, T. S. (2015). A comprehensive evaluation rubric for assessing instructional apps. *Journal of Information Technology Education: Research*, 14, 21-53. doi:10.28945/2097
- Menendez-Varela, J.-L., & Gregori-Giralt, E. (2018). Rubrics for developing students' professional judgement: A study of sustainable assessment in arts education. *Studies in Educational Evaluation*, 58, 70-79. doi:10.1016/j.stueduc.2020.100971
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129-144. doi:10.1016/j.edurev.2013.01.002
- Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment and Evaluation in Higher Education*, 35(4), 435-448. doi:10.1080/02602930902862859

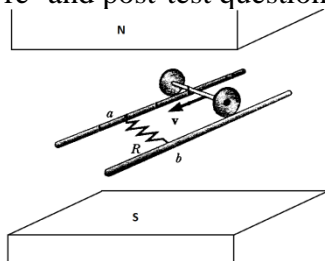
- Reynolds-Keefer, L. (2010). Rubric referenced assessment in teacher preparation: An opportunity to learn by using. *Practical Assessment, Research and Evaluation*, 15(1). doi:10.7275/psk5-mf68
- Safadi, R., & Saadi, S. (2021). Learning from self-diagnosis activities when contrasting students' own solutions with worked examples: The case of 10th graders studying geometric optics. *Research in Science Education*, 51, 523-546. doi:10.1007/s11165-018-9806-8
- Shadle, S. E., Brown, E. C., Towns, M. H., & Warner, D. L. (2012). A rubric for assessing students' experimental problem-solving ability. *Journal of Chemical Education*, 89, 319-325. doi:10.1021/ed2000704
- Tabachnick, B. G., & Fidell, L. S. (2015). Çok değişkenli istatistiklerin kullanımı [Using multivariate statistics]. (M. Baloğlu, çev. ed.). Ankara: Nobel Publishing.
- Thong, W. M., & Gunstone, R. (2008). Some student conceptions of electromagnetic induction. *Research in Science Education*, 38, 31-44. doi:10.1007/s11165-007-9038-9
- Vanclay, F., Baines, J. T., & Taylor, C. N. (2013). Principles for ethical research involving humans: Ethical professional practice in impact assessment Part I. *Impact Assessment and Project Appraisal*, 31(4), 243-253. doi:10.1080/14615517.2013.850307
- Vercellotti, M. L., & McCormick, D. E. (2021). Constructing analytic rubrics for assessing open-ended tasks in the language classroom. *The Electronic Journal for English as a Second Language*, 24(4), 1-19.
- Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research*, 19(1), 231-240.
- Wu, P.-H., Wu, H.-K., & Hsu, Y.-S. (2014). Establishing the criterion-related, construct and content validities of a simulation-based assessment of inquiry abilities. *International Journal of Science Education*, 36(10), 1630-1650. doi:10.1080/09500693.2013.871660
- Zajchowski, R., & Martin, J. (1993). Differences in the problem solving of stronger and weaker novices in physics: Knowledge strategies or knowledge structure. *Journal of Research in Science Teaching*, 30, 459-470. doi:10.1002/tea.3660300505
- Zheng, H., Ding, L., Lu, Z., & Branch, R. M. (2020). The motivational effects of involving students in rubric development on animation instruction. *TechTrends*, 64(1), 137-149. doi:10.1007/s11528-019-00443-w
- Zuza, K., Almudi, J.-M., Leniz, A., & Guisasola, J. (2014). Addressing students' difficulties with Faraday's law: A guided problem solving approach. *Physical Review Special Topics-Physics Education Research*, 10(1), 010122-1-16. doi:10.1103/PhysRevSTPER.10.010122
- Zuza, K., De Cock, M., van Kampen, P., Kelly, T., & Guisasola, J. (2020). Guiding students towards an understanding of the electromotive force concept in electromagnetic phenomena through a teaching-learning sequence. *Physical Review Physics Education Research*, 16(2), 020110-1-14. doi:10.1103/PhysRevPhysEducRes.16.020110

APPENDIX A**The rubric for the assessment of problem solutions to Electromagnetic Induction questions**

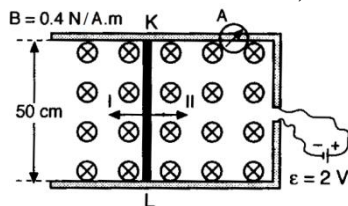
Weight: 1. Defining the source, direction and magnitude of magnetic field source	
10%	
10	Source of magnetic field was defined and its direction and magnitude were correctly determined
8	Source of magnetic field was defined and its direction or magnitude was correctly determined
6	Source of magnetic field was defined and its direction or magnitude was incorrectly determined
4	Direction and/or magnitude of magnetic field were/was correctly determined without indicating its source
3	Direction and magnitude of magnetic field were incorrectly determined without indicating its source
2	Direction or magnitude of magnetic field was incorrectly determined without indicating its source
0	No work done
Weight: 2. Defining the change in magnetic flux	
25%	
25	Whether there was a change in magnetic flux was correctly defined with a valid reason
20	Whether there was a change in magnetic flux was correctly defined with an invalid reason
15	Whether there was a change in magnetic flux was correctly defined without a reason
10	Whether there was a change in magnetic flux was incorrectly defined with a valid reason
5	Whether there was a change in magnetic flux was incorrectly defined with an invalid reason
2	Whether there was a change in magnetic flux was incorrectly defined without a reason
0	No work done
Weight: 3. Identifying the direction of the induced current/emf	
30%	
30	Direction of the induced current/emf was correctly identified by applying Lenz's law
20	Direction of the induced current/emf was incorrectly identified by applying Lenz's law
7	Direction of the induced current/emf was correctly identified without applying Lenz's law
2	Direction of the induced current/emf was incorrectly identified without applying Lenz's law
0	No work done
Weight: 4. Defining the direction and magnitude of magnetic force	
20%	
20	Source of magnetic force was defined and its direction and magnitude were correctly determined
15	Source of magnetic force was defined and its direction or magnitude was correctly determined
10	Source of magnetic force was defined and its direction or magnitude was incorrectly determined
5	Direction and/or magnitude of magnetic force were/was correctly determined without indicating its source
3	Direction and magnitude of magnetic field were incorrectly determined without indicating its source
2	Direction or magnitude of magnetic field was incorrectly determined without indicating its source
0	No work done
Weight: 5. Solutions for equations	
10%	
10	Necessary mathematical equations were correctly written and a correct result was obtained
7	Necessary mathematical equations were correctly written and an incorrect result was obtained
4	Necessary mathematical equations were correctly written but a result was not obtained
2	Necessary mathematical equations were incorrectly written and an incorrect result or nothing was obtained
0	No work done
Weight: 6. Writing Units and Symbols Correctly	
5%	
5	Each term was represented by correct symbols in a correct system of units
4	Some terms were represented by correct symbols in a correct system of units
3	Some terms were represented by incorrect symbols in a correct system of units
2	Some terms were represented by correct symbols in an incorrect system of units
1	Some terms were represented by incorrect symbols in an incorrect system of units
0	No work done

APPENDIX B

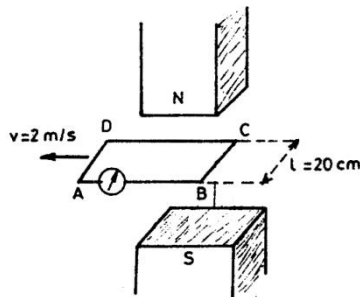
Pre- and post-test questions



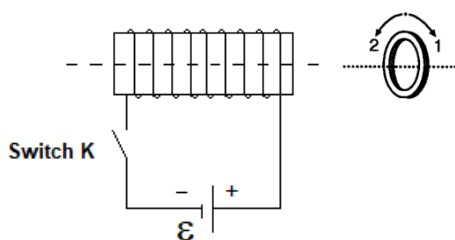
1. The 1.5 m long rolling wheel rod is pushed along the horizontal rails at a constant speed of $v = 3 \text{ m/s}$. A resistance of $R = 0.4 \Omega$ is connected to the rails at points a and b which are parallel to each other. There is a uniform magnetic field of $B = 0.08 \text{ T}$ between the poles of the magnet. What should be the direction and magnitude of the external force F , which must be applied to move the rod at a constant rolling speed?



2. As shown in the figure, the power supply is connected to the conductor frame, which is placed in a uniform magnetic field perpendicular to the plane of the page and has a magnitude of 0.4 N/A.m . In which direction and what speed should be the wire KL pulled so that the ammeter A indicates zero?



3. A wire frame with a resistance of 10Ω is pulled with a constant speed of 2 m/s out of the magnetic field. What is the magnitude and direction of the induced emf in the frame, if the magnetic field is uniform and has a magnitude of 2 mT inside but decreases to a value of zero outside the poles of the magnet?



4. A coil is placed next a conductor ring with an area of 0.5 m^2 so that their axes coincide. When the switch is turned off, the magnitude of magnetic field around the coil changes 2 N/A.m per second. Find the direction and magnitude of the induced emf formed in the ring.