Journal of Laboratory Chemical Education 2017, 5(5): 95-107
DOI: 10.5923/j.jlce.20170505.01

# The Effectiveness of General Chemistry Lab Experiments on Student Exam Performance

**Junyang Xian, Daniel B. King***

Department of Chemistry, Drexel University, Philadelphia, PA, USA

**Abstract** Verification labs are widely used, despite the belief by many that other types of labs are more effective. For those who use verification labs it is important to understand how these experiments contribute to student learning. The main purpose of this study is to determine if there is evidence that completing a verification experiment leads to improved performance on a corresponding exam question related to the lab content. Student performance on exam questions that related to lab content was compared to student performance on exam questions that were not related to lab content in two general chemistry courses. Student performance on quantitative exam questions that related to lab content was also compared to student performance on conceptual exam questions that related to lab content for a few experiments. Overall results were mixed. Students performed better on lab-related questions for some topics, such as kinetics and electrochemistry. The graphing component of the kinetics experiment seemed to benefit science majors and female students more than engineering majors and male students, respectively. In contrast, similar performance was observed on lab-related quantitative and lab-related conceptual questions related to pH titration and electrochemistry.

**Keywords** General chemistry, Verification laboratory, Exam performance

## 1. Introduction

Laboratories have been a standard educational technique in science classes at all levels for more than a hundred years [1, 2]. From a sample of 40 public universities in the United States 58% offer general chemistry laboratories as a required portion in the students' curriculum [3]. Science laboratory activities provide students a chance to observe and understand the natural world by their interactions with the environment [2]. Students gain more benefit in practical manipulations from the lab than they gain from other types of instruction. In addition to practical skills, some of the studies concluded that laboratory experiments could improve the learners' favorable attitudes toward science [4, 5]. Labs have been reported to help students develop their creative thinking and problem-solving skills [6]. In addition, labs can help students develop scientific thinking [7, 8], scientific abilities [9], intellectual abilities [10] and conceptual understanding [11, 12]. Students' interest in specific activities, such as experimenting, dissecting and working with microscopes) are significantly higher if they have lab experience [13]. However, even though laboratories have been viewed as a fundamental part of a science course and they are assumed to be important and necessary, it is sometimes hard to justify the inclusion of lab courses due to the high cost in terms of time, finances and personnel required [14]. Therefore, more study is needed to determine whether or not the high laboratory costs are a necessary component of science education.

The majority of colleges and universities that require lab (70%) have their students complete verification chemistry experiments [15], in which students attempt to reproduce a result or measurement they have already been provided. Verification labs are beneficial when there is need to minimize resources, including time, space, equipment and personnel [16]. Verification labs provide well stated procedures, which the students can easily follow. They are well designed for a large number of students to complete the experiment simultaneously. It has been reported that verification labs can improve students' lab experiment skills and teach them factual information [1]. Verification experiments are often preferred when lab time is limited, due to the fact that other lab types, e.g., inquiry-based labs, can require that students spend up to four times longer in the lab. The additional lab time associated with the inquiry-based experiments didn't produce increased content understanding, as there was no significant difference in student exam scores between students enrolled in verification labs and students enrolled in inquiry-based labs [9]. In a review of a group of studies conducted over a 32-year period, it has been found that the test scores on post-tests of students who enrolled in verification labs are about 10 points higher than those of their pre-tests [17].

* Corresponding author:
dk68@drexel.edu (Daniel B. King)
Published online at http://journal.sapub.org/jlce

However, despite their benefits, verification experiments are heavily criticized by some faculty [18]. The first reason is that the outcome of a verification experiment is predetermined, so students are trying to reproduce the expected result. This means that students who complete verification labs spend more time determining if their outcome agrees with the predicted result than they spend thinking about or interpreting the data [19-21]. The second reason is that when the students are doing verification experiments, they often focus more on mathematical skills required to complete the corresponding calculations than on conceptual understanding of the associated chemistry [9]. Students usually don't understand the purpose of the experiments, because their sole focus is on the procedure and there is no time or effort spent thinking about the experiment. Researchers also pointed out that students failed to be engaged in discussion and analysis of the science phenomena in lab courses that used verification experiments [1, 2].

Understanding the benefits that verification labs might contribute to general chemistry teaching is important for those who use these labs. At our institution, all general chemistry labs use verification experiments, and students are required to write individual lab reports in which they summarize their results. The goals of the lab reports are to increase the time that students spend thinking about the lab content, to improve their conceptual understanding of the lab content, to improve their ability to do the experimental calculations, and to improve their ability to graph data and write scientifically. The main purpose of this study is to investigate the relationship between the content in the general chemistry verification experiments and student performance on corresponding content questions on general chemistry exams (these exams are primarily based on the lectures; there are no exams or quizzes given in lab). Student performance on exam questions that related to lab content was compared to student performance on exam questions that were not related to lab content. Student answers on quantitative questions and conceptual questions were analyzed. Student exam performance was also compared as a function of major and gender. In addition, the lab-related exam performance of students with higher average exam scores was compared to the exam performance of students with lower average exam scores.

## 2. Methodology

Student performance has been evaluated during two terms of a general chemistry sequence: General Chemistry I (CHEM101) and General Chemistry II (CHEM 102). During each 10-week term students complete four verification labs, three in-term exams, and one (cumulative) final exam, along with other course components. The lab is part of the general chemistry course, so all students are required to take the labs and lecture together. Individual lab reports, which include an introduction, a discussion and data analysis section, and conclusions, are submitted one week after the experiment is completed. Students are provided access to the rubric used to grade the lab reports before they write the report. This study used data collected during academic years 2012-2013 and 2013-2014. In the first year, 1177 students were enrolled in CHEM 101 and 1005 students were enrolled in CHEM 102, 915 of whom were enrolled in both 101 and 102. In the second year, 985 students were enrolled in CHEM 101 and 882 students were enrolled in CHEM 102, with 760 enrolled in both courses. Most of these students are first-year science and engineering majors. Students in the honors section of the course were not included in this study, as their performance is statistically different than the rest of the sections.

Student exam performance on individual lab-related questions was analyzed. Student in-class performance and student exam performance on the same lab topic were also compared. In-class performance was determined for a subset of the students who were enrolled in a lecture section where personal response devices (i.e., clickers) were used. To help determine the effectiveness of lab for students as a function of average chemistry ability, students were sorted into different groups by their average exam scores. Each group is assigned to represent a corresponding "ability level": high-performing, medium-performing, and low-performing. The "difficulty level" of a question is determined from the percent correct of all the test takers on that question [22], where a lower percent correct corresponds to higher difficulty. Item characteristic curves have been used as a graphical method to show the range of student performance (percent correct) at different "ability levels". The same technique was used to analyze student performance on American Chemical Society (ACS) exams [22]. In this study, students were grouped by their average exam scores over the whole term, which was considered as their "ability level" or "student proficiency". All students who took each exam were evenly divided into ten groups. Each group contains 70-90 students, depending on the number of students who took the exam. Students' standard z scores are calculated with equation (1).

$$z = (\text{score-average score})/(\text{standard deviation}) \tag{1}$$

where score is the average of all exam scores for one student; average score is the average of all exam scores for the whole class; and standard deviation is calculated from the student average exam scores of all the first and second term exams. Z scores represent the student average exam performance over the whole term compared to the rest of the class. The scales for each set of z scores are slightly different due to the different number of students who took each exam. An example of grouping the students by z score is shown in Table 1.

Students who had the highest average exam scores compared to the rest of the students were considered as high-performing students (for example, z scores higher than 0.60, as shown in Table 1). Students who had the lowest average exam scores were regarded as low performing students (for example, z scores lower than -0.33, as displayed in Table 1).

**Table 1.** Z score ranges and sizes of different student groups. K1 and K2 are kinetics questions related to the lab from CHEM 102 exams in 2014 and 2013, respectively

| Z score increment | Number of students answered (K1) | Number of students answered (K2) | Student performance |
|---|---|---|---|
| | | | |
| less than -1.30 | 88 | 88 | Low |
| -1.31 to -0.73 | 92 | 92 | |
| -0.73 to -0.34 | 88 | 88 | |
| -0.33 to -0.07 | 87 | 87 | Medium |
| -0.06 to 0.12 | 86 | 86 | |
| 0.13 to 0.36 | 87 | 87 | |
| 0.37 to 0.60 | 89 | 89 | |
| 0.61 to 0.83 | 87 | 87 | High |
| 0.84 to 1.12 | 87 | 87 | |
| more than 1.12 | 93 | 93 | |

Each item characteristic curve contains results from a pair of questions: a lab-related and a non-lab related question or a quantitative and a conceptual question. Lab-related questions refer to exam questions that have content directly related to the experimental data or lab report calculations; non-lab related questions cover the same general chemistry concept as the corresponding lab-related exam questions in a broader or tangential way. Paired questions are taken from the same exam. To compare the exam performance of students at different levels of proficiency, chi square values were calculated. This test was chosen because the student performance of all the groups does not correspond to a normal distribution, which is required for an ANOVA test. The average student performance on lab-related exam questions and non-lab related exam questions and the average student performance on lab-related quantitative questions and lab-related conceptual questions were compared by chi square test on the performance of all levels of students, with the significance level set at 0.05. The chi square test was performed on all ten groups of students, which were grouped by their average exam scores of CHEM 101 and CHEM 102. To further compare the performance at difference levels of student proficiency, chi-square tests of high-performing, medium-performing, and low-performing students are reported.

There are some limitations associated with this study. First, there were a small number of questions directly related to lab content on the in-term exams. This limited analysis for some of the experiments. Second, some of the experiments were not tested at all on the exams, which meant that they could not be included in the study. Finally, student performance on exam questions is not a direct measure of student learning associated with the experiments. Exam scores are a function of many parameters, including lecture content, textbook, online homework practice and overall study skills. It is assumed that comparisons of lab-related and non-lab-related questions related to the same general topic on the same exam help to minimize the impact of these other factors.

# 3. Results and Discussion

Student performance on exam questions related to several different lab topics has been analyzed. The analysis includes lab-related and non-lab related questions; lab-related quantitative and lab-related conceptual questions; and the impact of student major and gender on their performance on lab-related and non-lab related questions. Table 2 shows a list of topics that were covered in CHEM 101 and 102 labs. Some of these topics were not included in this study because there were no questions about these topics on the general chemistry exams or because the exam questions didn't match the criteria of this study (e.g., exams only included questions that were not lab related).

**Table 2.** List of general chemistry lab topics in CHEM 101 (Experiments 1-4) and CHEM 102 (Experiments 5-8)

| Experiment | Topic | In this study? |
|---|---|---|
| 1 | spectroscopy | no |
| 2 | conductivity of solutions | yes |
| 3 | stoichiometry and limiting reagents | no |
| 4 | freezing point depression | no |
| 5 | preparation of esters | no |
| 6 | kinetics of alcohol oxidation | yes |
| 7 | acids and bases (pH titration) | yes |
| 8 | electrochemical cells (voltaic cells) | yes |

## 3.1. Lab-related and Non-lab Related Questions

In the kinetics experiment (CHEM 102), students recorded times and calculated the rate of a pseudo-first order reaction. The data were recorded by computer using the Vernier LoggerPro System (Beaverton, OR). In the lab report, each student plotted three kinetics curves (zero order, first order and second order) and chose the graph that had the best linear relationship to determine the reaction order [23, 24]. Figure 1 shows a lab-related question (K1) from the CHEM 102 final exam (2014) and the corresponding distribution of students' answers. Statements I and III in the question refer to kinetics curves of zero and first order, respectively, which were directly related to the lab content. Many students could identify that statement III was correct (79% chose c or e), but many of those students failed to identify that statements I and II were incorrect (52% chose e). The students who chose answer e failed to show complete conceptual understanding of first order reactions; they thought the concentration of the reactant should decrease linearly with time and the rate of reaction is constant. This result indicates that this experiment might not help those students improve their conceptual understanding of first order reactions. This may be due to the fact that students cannot connect the graphs they made in the kinetics lab report to the fundamental concepts associated with this experiment.

**K1.** Which of the following statements is or are correct for the first-order reaction below?

I. The concentration of A decreases linearly with time.
II. The rate of reaction is constant.
III. A plot of ln [A] versus time is a straight line with the slope equal to -k.

a. I only
b. II only
**c. III only**
d. I and II
e. I, II, and III

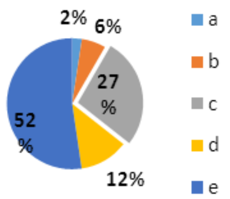



**Figure 1.** Student performance on a lab-related question about reaction orders from CHEM 102 final exam (2014, n=754). The correct answer is indicated with bold font

Student in-class performance and exam performance on the same content was also compared. Students' answers during classroom lectures (CHEM 102, 2013) were collected using personal response systems. Students in one lecture section were provided a device (a clicker) to use in class, but no points were associated with the student answers. The content of the "clicker questions" and the exam questions were similar, and both clicker and exam questions were multiple choice. The clicker results for questions asked in lecture before the content was discussed were used as an indicator of prior knowledge related to that content. Student performance on an in-class kinetics question was compared to student performance on a corresponding kinetics question on the second in-term exam. The sample population for this comparison was limited to students from the lecture section that used "clicker questions" (n = 45). In the lab-related question students were asked to identify the characteristic curve of a second order reaction. In the non-lab related question students were asked to determine the change of the reaction rate based on a change of the reactant concentration. There was a much greater improvement associated with the lab-related questions than with the non-lab related questions (Table 3). While much of the larger difference for lab-related questions is due to the lower in-class performance (i.e., prior knowledge) on the lab-related question (13.0% correct), the score on the lab-related exam question (76.9%) is also significantly higher than the score on the non-lab related question (23.0%, Table 3).

**Table 3.** In-class student performance (% correct) and in-term exam student performance (% correct) on lab-related and non-lab related questions about reaction orders (CHEM 102, 2014, n=45)

| Non-lab-related | | | Lab-related | | |
|---|---|---|---|---|---|
| In-class | Exam | Difference | In-class | Exam | Difference |
| 50.0 | 73.0 | 23.0 | 13.0 | 89.9 | 76.9 |

Figure 2 shows sample lab-related and non-lab related questions from years 2013 and 2014. Student performance on the kinetics questions appeared to vary from 2013 to 2014 (Figure 3). Students performed better on the lab-related question (K2) than on the non-lab related question (K3) in 2013 (Figure 3A). The differences were most evident for students with medium and low average exam scores (medium performing and low performing students). A chi square test of all ten groups of students who answered K2 and K3 shows that on average students performed significantly better on the lab-related kinetics question (K2, 90% correct) than on the kinetics question not related to the lab (K3, 80% correct), p=0.01. As shown in Table 3, low-performing and medium-performing students have significantly different scores on questions K2 and K3. However, in 2014 there did not appear to be a difference between student performance on the lab-related question (K4) and the non-lab related question (K5, Figure 3B). A chi square test of the ten groups of students who answered K4 and K5 shows that on average, student performance on the lab-related questions (76% correct) and non-lab related questions (78% correct) are not significantly different, p=0.73. The lack of difference in 2014 could be due to a change in the level of difficulty associated with the lab-related questions.

The content of the lab-related questions chosen from the different years was very similar. However, the format of these two questions was different, which could explain the lower score during the second year. In the lab-related question K2, students were required to identify the reaction order by remembering the shape and/or key features of the characteristic kinetics plots. This question was a straightforward application of the content incorporated into the lab experiment and report, with a visual cue (graph) that they would have created as part of the lab report. The lab-related question K4 might have been harder for the students because it asked them to interpret descriptions of the characteristic kinetics plot of a second order reaction without the visual cue. Even though both the graphs and the descriptions of the kinetics curves were included in the lab content, it might have been easier for the weaker students to use the graph to answer the question than to use the description. Both the content and the format of K3 and K5 were similar, which might help to explain why the scores on the non-lab related questions were comparable both years. Since the performance on the lab-related question decreased from 2013 to 2014 while the performance on the non-lab related question remained constant, the results suggest that the lack of observed difference between lab-related and non-lab-related questions in 2014 is real. As, shown in Table 4, the high-performing students (higher than 0.61 in 2013 and higher than 0.59 in 2014) performed similarly on the two types of questions (p=0.71 and p=0.98 for 2013 and 2014, respectively). These results suggest that the high-performing students were better able to interpret the non-lab related question.

Lab-related and non-lab related questions from the electrochemical cells experiment (CHEM 102) were also analyzed. In the electrochemical cells experiment, students measured the cell potential of different pairs of half-cells and used their measurements to determine the positive probe and negative probe. The concept of spontaneous oxidation/reduction reactions was presented in the lab manual [23, 24] and also discussed in the lecture class. Figure 4 shows a lab-related question (E1) and a non-lab-related question (E2) on the topic of electrochemical cells. In E1 the students were asked to identify the correct half-cell reaction for a given cell; in E2 the students were required to identify the oxidant and reductant based on which reaction had a higher standard potential.

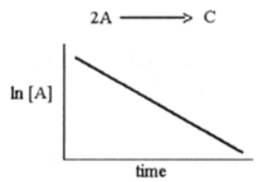
| Lab-related questions | Non-lab related questions |
|---|---|
| **K2.** The graph shown below depicts the relationship between concentration and time for the following chemical reaction. Determine the order of this reaction.<br><br><br>a. zero order<br>**b. first order**<br>c. second order<br>d. pseudo-second order<br>e. The order cannot be determined from this information | **K3.** A reaction was found to be second order in carbon monoxide concentration. The rate of the reaction__________if the [CO] is doubled, with everything else kept the same.<br><br>a. doubles<br>b. does not change<br>c. triples<br>**d. increases by a factor of 4**<br>e. decreases by a factor of 2 |
| **K4.** What statement below best describes the graph representing the integrated second–order rate law?<br>a. plot of [A]t vs t having a straight line negative slope equal to -*k*<br>b. plot of 1/[A]t vs *t* having a straight line negative slope equal to -*k*<br>c. plot of ln[A]t vs *t* having a straight line positive slope equal to *k*<br>d. plot of ln[A]t vs *t* having a straight line negative slope equal to -*k*<br>**e. plot of 1/[A]t vs *t* having a straight line positive slope equal to *k*** | **K5.** The reaction A + 2B → products<br>was found to have the rate law below: rate = $k[A][B]^2$<br>If the concentration of B increases from x to 3x while that of A does not change, the rate of reaction will<br>a. triple.<br>b. increase by a factor of 6.<br>c. stay the same.<br>d. double.<br>**e. increase by a factor of 9.** |

**Figure 2.** Lab-related questions and non-lab related questions about reaction orders from CHEM 102 second in-term exams of 2013 (K2, K3) and 2014 (K4, K5). Correct answers are highlighted in bold font

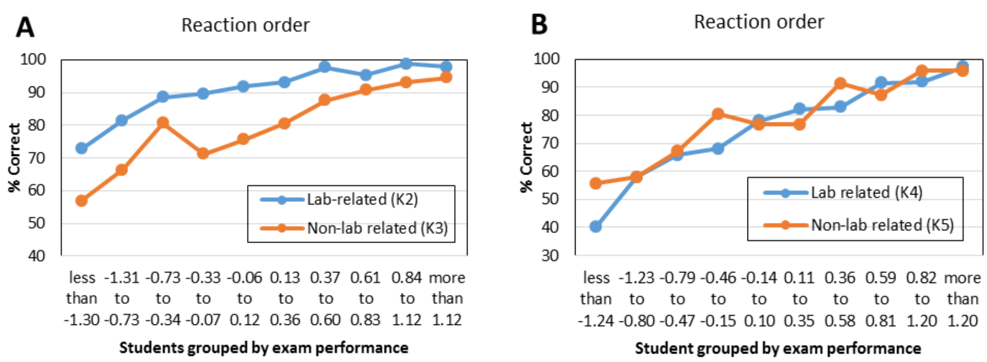



**Figure 3.** Item characteristic curves for kinetics questions from CHEM 102 second in-term exam in 2013 (A) and 2014 (B), using standard z scores (which represent student average exam performance of the whole term compared to the rest of the class)

**Table 4.** Chi square tests of student performance (%) on lab-related and non-lab related questions about reaction orders from CHEM 102 second in-term exams in 2013 (K2 and K3) and 2014 (K4 and K5)

| Reaction order | Number of students | Lab-related | Non-lab-related | Difference | P value (chi-square) |
|---|---|---|---|---|---|
| **Ability** | | **K2** | **K3** | | |
| High<br>(z score ≥ 0.61) | 267 | 97.4 | 92.8 | 4.6 | 0.71 |
| Medium<br>(-0.33 ≤ z score ≤ 0.60) | 349 | 93.1 | 78.7 | 14.4 | 0.01 |
| Low<br>(z score ≤ -0.34) | 268 | 81.0 | 67.9 | 13.1 | 0.01 |
| | | | | | |
| **Ability** | | **K4** | **K5** | | |
| High<br>(z score ≥ 0.59) | 217 | 94.3 | 93.7 | 0.6 | 0.98 |
| Medium<br>(-0.46 ≤ z score ≤ 0.58) | 288 | 79.0 | 81.4 | -2.4 | 0.61 |
| Low<br>(z score ≤ -0.47) | 219 | 54.1 | 60.7 | -6.6 | 0.12 |

| **Lab-related question** | **Non-lab-related question** |
|---|---|
| **E1**. The cell notation for a voltaic cell constructed with copper and silver is shown below<br>$Cu(s) \mid Cu^{2+}(aq) \parallel Ag^{+}(aq) \mid Ag(s)$<br><br>Which of the following processes occurs at the cathode?<br>**a.** $Cu(s) \rightarrow Cu^{2+}(aq) + 2e^{-}$<br>**b.** $Cu^{2+}(aq) + 2e^{-} \rightarrow Cu(s)$<br>**c.** $Ag(s) \rightarrow Ag^{+}(aq) + e^{-}$<br>**d. $Ag^{+}(aq) + e^{-} \rightarrow Ag(s)$**<br>**e.** $Cu(s) + 2\,Ag^{+}(aq) \rightarrow Cu^{2+}(aq) + 2\,Ag(s)$ | **E2.** Given that E° for $X + e^{-} \rightarrow Y$ is greater than E° for $A + 2e^{-} \rightarrow B$, it is correct to say that, under standard conditions,<br>______<br>a. X will oxidize A<br>b. Y will oxidize A<br>c. Y will reduce A<br>**d. X will oxidize B**<br>e. X will reduce B |

**Figure 4.** Lab-related question and non-lab-related question about electrochemical cells from CHEM 102 final exam of 2014. Correct answers are highlighted in bold font

On average, students scored better on the lab-related question (56%) than on the non-lab related question (39%) (Figure 5A). Students at all levels (high-performing, medium performing, and low-performing) did better on the lab-related question (Figure 5B). A chi square test of the ten groups of students indicates that all of these differences between lab-related and non-lab-related questions are statistically significant (p<0.01, Table 5). The content of the E1 and E2 questions was similar. E1 was more closely related to the lab content in that students wrote half-cell equations in the lab report. However, the content in E1 wasn't completely covered in the experiment; the concepts of cathode and cell notation were not covered in the lab manual or the lab report. E2 might have been more difficult for the students because it required complete understanding of how to identify the oxidant and reductant of an unknown chemical reaction based on the comparison of half-cell potentials. In addition, without actual potential values or chemicals given in this question (E2), students needed to identify the oxidant and reductant by qualitative methods that were not taught through the lab. Students at the low performance level may have struggled on both the lab-related and non-lab-related questions, yielding a smaller difference on these two questions than observed for the other students.

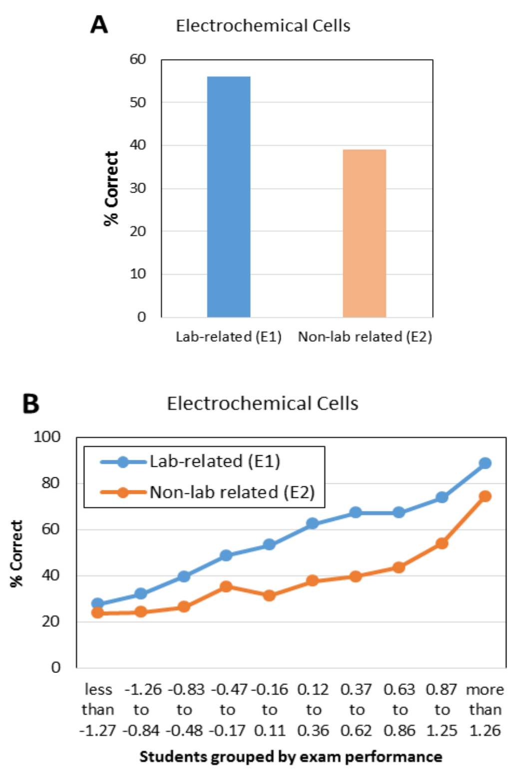



**Figure 5.** Average student performance (A) and item characteristic curve (B) on lab-related and non-lab related exam questions about electrochemical cells from CHEM 102 final exam in 2014

Lab-related and non-lab related questions from the conductivity experiment (CHEM 101) were also compared. In this experiment, students measured the conductivity of different solutions (nonelectrolyte, strong electrolyte, and weak electrolyte). Students evaluated the effect of ionic charge/concentration on conductivity by graphing their data. The dissociation equations of ionic compounds that generate ions with different charges (NaCl, $CaCl_2$, and $AlCl_3$) were included in the lab manual [23, 24]. Figure 6 lists two sample exam questions related to this topic.

Students had similar performance on the lab-related exam question (C1) and the non-lab related exam question (C2) (Figure 7A). There was a difference with high-performing students, who scored better on the lab-related question than on the non-lab related question (Figure 7B). A chi square test of the ten groups of students who answered questions C1 and C2 shows that on average students performed slightly better on the non-lab related solutions question (C2, 66% correct) than on the solutions question related to the lab (C1, 64% correct), p=0.04. However, this difference is primarily due to the lowest performing students (z score less than -1.44) (Figure 7B). This group of students had a relatively low performance on this lab-related question. As displayed in Table 6, high-performing students received higher scores (10% higher) on this lab-related question compared to the corresponding non-lab related question, p=0.09 (at the 90% confidence level). Low-performing students received lower scores on C1 compared to C2, p=0.01, although much of this lower score can be attributed to the poor performance of the weakest students on C1.

**Table 5.** Chi square tests of student performance (%) on lab-related and non-lab related questions about electrochemical cells from CHEM 102 final exams in 2014

| Voltaic Cells | Number of Students | Lab | Voltaic Cells | Number of Students | Lab |
|---|---|---|---|---|---|
| **Proficiency** | | **E1** | **E2** | | |
| High (z score ≥ 0.61) | 222 | 76.5 | 57.2 | -19.3 | 0.00 |
| Medium (-0.33 ≤ z score ≤ 0.60) | 304 | 57.8 | 35.9 | -21.9 | 0.00 |
| Low (z score ≤ -0.34) | 227 | 33.0 | 24.7 | -8.3 | 0.01 |

| **Lab-related question** | **Non-lab-related question** |
|---|---|
| **C1.** What are the products (with correct stoichiometric coefficients) of the dissociation of magnesium nitrate in water?<br>a. $Mg^+ + NO_3^-$<br>b. $Mg^{2+} + NO_3^-$<br>**c. $Mg^{2+} + 2NO_3^-$**<br>d. $2Mg^{2+} + NO_3^-$<br>e. $Mg^+ + 2NO_3^-$ | **C2.** Which of the following statements is false?<br>a. Ice in water is a heterogeneous mixture.<br>b. Formation of cations from atoms is an oxidation process.<br>**c. Formulating a law to explain experimental observations is the first step of the scientific method.**<br>d. Sugar dissolving in boiling water illustrates a physical change.<br>e. A horizontal row in the periodic table shows how many units of atomic number, Z, are in between elements with similar chemical and physical properties. |

**Figure 6.** Lab-related question and non-lab- related question about solutions from CHEM 101 first in-term exam of 2013. Correct answers are highlighted in bold font

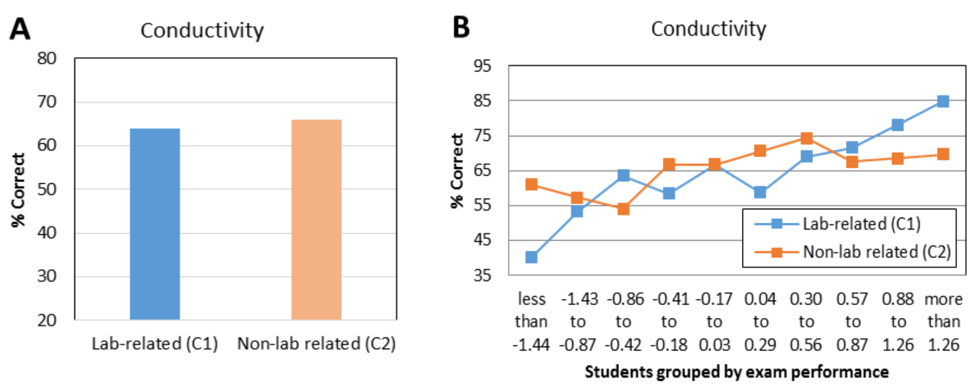



**Figure 7.** Average student performance (A) and item characteristic curves (B) on lab-related and non-lab-related exam questions about solutions from CHEM 101 first in-term exam in 2013

**Table 6.** Chi square tests of student performance (%) on lab-related and non-lab-related questions about solutions from CHEM 101 first in-term exam of 2013

| Conductivity | Number of Students | Lab-related | Non-lab related | Difference | P value (chi square) |
|---|---|---|---|---|---|
| **Proficiency** | | **C1** | **C2** | | |
| High ($z$ score $\geq 0.57$) | 226 | 78.2 | 68.6 | 9.6 | 0.09 |
| Medium ($-0.41 \leq z$ score $\leq 0.56$) | 296 | 63.1 | 69.6 | -6.5 | 0.32 |
| Low ($z$ score $\leq -0.42$) | 226 | 52.4 | 57.5 | -5.1 | 0.01 |

In the lab-related question C1, students were required to choose the correct dissociation equation for magnesium nitrate. Students needed to know the number and charge of the anions and cations. C1 was partially an application of the content incorporated into the lab experiment and report due to the fact that dissociation equations were included in the lab manual, and the students had to use them to interpret their data; however, students were not required to write dissociation equations as part of the lab report. The high-performing students might have been more familiar with dissociation equations than the low-performing students, i.e., better able to balance the charges of the ions, and more likely to choose the correct answer. C2 was a broader conceptual question that contained a range of topics, including dissolving and melting, formation of ions, fundamental steps of scientific method and periodic trends. Student performance on the non-lab related question was fairly consistent across all performance levels. This is likely due in part to the fact that students did not need to understand solutions to answer this question correctly.

## 3.2. Lab-related Quantitative and Lab-related Conceptual Questions

To gain more information about the effectiveness of verification experiments, student performance on quantitative and conceptual questions related to several experiments was investigated. One example of paired lab-related quantitative and lab-related conceptual questions related to a pH titration experiment (CHEM 102). In this experiment, students performed a titration of a vinegar solution with NaOH standard solution and observed the color change of an indicator at the endpoint. Students used the equivalence point volume to calculate the concentration of acetic acid in vinegar. In the lab report, students used the pH value of the half equivalence point to calculate the pKa value. The conversion between pH values and $[H^+]$ was provided in the lab manual [23, 24]. Figure 8 shows two pH titration questions (T1 and T2) from the third in-term exam of CHEM 102 in 2014. To answer T1, students had to convert pH values to $[H_3O^+]$ and $[OH^-]$. In addition, students had to know the pH ranges that correspond to acidic and basic solutions, which was also related to the lab content. To answer T2, students were asked to identify the pH range of a solution using the behavior of two indicators, which also directly relates to the lab content.

| **Lab-related quantitative question** | **Lab-related conceptual question** |
|---|---|
| **T1.** If the pH of a solution is raised from 6.27 to 7.57, ________<br>a. the solution has gone from acidic to basic<br>b. $[H_3O^+]$ has decreased by a factor of approximately 20<br>c. $[OH^-]$ has increased by a factor of approximately 20<br>d. the new $[H_3O^+] = 2.7 \times 10^{-8}$<br>**e. All of the above.** | **T2.** Phenolphthalein is an acid-base indicator that is colorless in its acid form and pink in its basic form, changing color at pH = 8.5. Bromocresol green is yellow in its acidic form and blue in its basic form, changing color at pH = 4.8. An aqueous solution is colorless in phenolphthalein and blue in bromocresol green. Therefore, we can conclude that the pH of the solution is _____<br>a. exactly 7.0<br>b. greater than 8.5<br>c. between 7.0 and 8.5<br>**d. between 4.8 and 8.5**<br>e. less than 4.8 |

**Figure 8.** A lab-related quantitative question and a lab-related conceptual question about pH values from CHEM 102 third in-term exam of 2014. Correct answers are highlighted in bold font

In general, students performed well on both the lab-related quantitative question (T1) and the lab-related conceptual question (T2) (Figure 9A). Student performance on the lab-related quantitative question was slightly higher than student performance on the lab-related conceptual question. While the item characteristic curve appears to show a difference between the quantitative and conceptual questions for the high-performing students (Figure 9B), this difference is not statistically significant ($p=0.14$, Table 7). The performance of all groups of students is similar for both questions ($p>0.05$ for all groups), which may be due to the low "difficulty level" of this pair of questions.

The electrochemical cells experiment provided another opportunity to compare student performance on a lab-related quantitative question (E3) and a lab-related conceptual question (E1) (Figure 10). The procedure and content associated with this experiment were described in a previous section. To answer E3, students were required to calculate the standard cell potential from given standard half-cell reduction potentials; to answer E1, students were asked to identify the correct half-cell reaction for a given cell. Figure 11A shows that students performed slightly better on the lab-related conceptual question E1 (56%) than the lab-related quantitative question E3 (51%). In Figure 11B, it appears that this difference could be attributed to higher scores of the medium-performing students on the conceptual question. However, Table 8 suggests that the difference between quantitative and conceptual questions is not statistically significant for this experiment (p=0.17 for medium-performing students). It should also be noted that student performance on the conceptual question was significantly better than student performance on the non-lab-related question (E2) presented earlier.

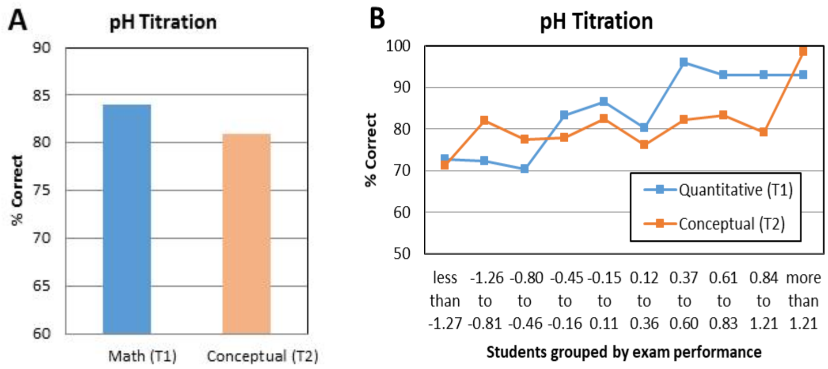



**Figure 9.** Average student performance (A) and item characteristic curve (B) on a lab-related quantitative question and a lab-related conceptual exam question about pH titration from CHEM 102 third in-term exam in 2014

**Table 7.** Chi square tests of student performance (%) on a lab-related quantitative question and a lab-related conceptual question about pH titration from CHEM 102 third in-term exam of 2014

| pH Titration | Number of Students | Quantitative | Conceptual | Difference | P value (Chi-square) |
|---|---|---|---|---|---|
| **Proficiency** | | **T1** | **T2** | | |
| High (z score ≥ 0.61) | 214 | 93.0 | 87.0 | 6.0 | 0.14 |
| Medium (-0.33 ≤ z score ≤ 0.60) | 290 | 86.5 | 79.6 | 6.9 | 0.37 |
| Low (z score ≤ -0.34) | 216 | 71.7 | 76.9 | -5.2 | 0.40 |

| **Quantitative Question** | **Conceptual Question** |
|---|---|
| **E3.** You are asked to construct an electrochemical cell using the information given below<br>$Cu^{2+}(aq) + 2e^- \rightarrow Cu(s)$<br>$Sn^{4+}(aq) + 2e^- \rightarrow Sn^{2+}(aq)$<br><br>What would be the standard potential for this cell?<br>a. 0.47 V<br>**b. 0.21 V**<br>c. – 0.21 V<br>d. – 0.47 V<br>e. 0.42 V | **E1**. The cell notation for a voltaic cell constructed with copper and silver is shown below<br>$Cu(s) \mid Cu^{2+}(aq) \parallel Ag^{+}(aq) \mid Ag(s)$<br><br>Which of the following processes occurs at the cathode?<br>**a.** $Cu(s) \rightarrow Cu^{2+}(aq) + 2e^-$<br>**b.** $Cu^{2+}(aq) + 2e^- \rightarrow Cu(s)$<br>**c.** $Ag(s) \rightarrow Ag^{+}(aq) + e^-$<br>**d. $Ag^{+}(aq) + e^- \rightarrow Ag(s)$**<br>**e.** $Cu(s) + 2\,Ag^{+}(aq) \rightarrow Cu^{2+}(aq) + 2\,Ag(s)$ |

**Figure 10.** Lab-related quantitative question and lab-related conceptual question about electrochemical cells from CHEM 102 final exam in 2014. Correct answers are highlighted in bold font

These results show that these two experiments (pH titration and electrochemical cells) are equally effective at helping students to answer quantitative and conceptual questions relating to the lab content. All students performed well on both pH titration-related questions. In contrast, low-performing students scored much lower on the electrochemical cells-related questions. However, as with the pH titration experiment, the electrochemical cells experiment did not provide any added benefit with respect to the quantitative or conceptual questions.

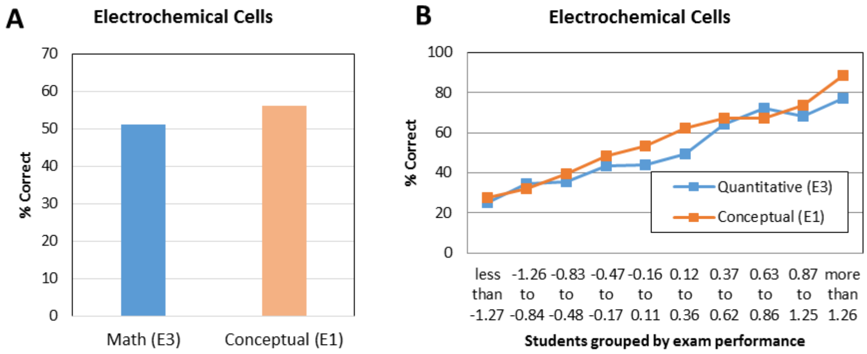



**Figure 11.** Average student performance (A) and item characteristic curves (B) on a lab-related quantitative question and a lab-related conceptual question about electrochemical cells from CHEM 102 final exam in 2014

**Table 8.** Chi square tests of student performance (%) on lab-related math questions and lab-related conceptual questions about voltaic cells from CHEM 102 final exam of 2014

| Voltaic Cells | Number of Students | Quantitative | Conceptual | Difference | P value |
|---|---|---|---|---|---|
| **Proficiency** | | **E2** | **E1** | | |
| High<br>(z score ≥ 0.61) | 222 | 72.6 | 76.5 | -3.9 | 0.65 |
| Medium<br>(-0.33 ≤ z score ≤ 0.60) | 304 | 50.3 | 57.8 | -7.5 | 0.17 |
| Low<br>(z score ≤ -0.34) | 227 | 31.7 | 33.0 | -1.3 | 0.32 |

## 3.3. The Impact of Major and Gender

To further investigate potential benefits of verification experiments, results of the lab-related and non-lab related questions were also analyzed as a function of student major and gender. Since the student population in CHEM 101 and CHEM 102 is primarily composed of engineering majors (about two-thirds of the students) and biology majors (about one-sixth of the students), comparisons by major were made using only these two groups of students. Comparisons by gender, however, used all students enrolled in the courses. The first results presented relate to the kinetics experiment. The average performance of biological science students on the lab-related kinetics questions (K2, 93.5%, and K4, 78.6%) was higher than the performance of engineering students on the same questions (88.6% and 75.5%, respectively, Table 9). The average performance of female students on the lab-related kinetics questions (K2, 93.0%, and K4, 81.2%) was higher than that of male students on the same questions (89.3% and 72.7%, respectively). These differences are not due to an inherent difference in performance between these groups, as demonstrated by the fact that the average exam scores are the same for each group in a given year (Table 9). Since engineering majors are predominantly male and biology majors are more evenly mixed with respect to gender, it is not surprising that similar results (i.e., lower scores) are observed for both engineering majors and male students. When comparing student performance on lab-related and non-lab related questions, again, engineering majors and male students performed in a similar way. Engineering students performed 7.9% higher on question K2 (lab-related) than K3 (non-lab related), whereas biological science majors showed a bigger difference, scoring 18% higher on question K2 than K3, with all groups scoring higher on the lab-related questions during 2013. In 2014, however, some of the groups scored higher on the non-lab-related question. Engineering students performed 5.7% lower on question K4 (lab-related) than K5 (non-lab-related), while biological science majors performed 8.5% higher on question K4 than K5. As was the case in 2013, engineering majors and male students showed similar results relative to biology majors and female students, respectively. It is not known why the results differed between the two years in the response to lab-related and non-lab related questions.

The second results presented relate to the electrochemical cells experiment. The results for this experiment were similar to those observed for the kinetics experiment. Table 10 shows that biology majors performed better (63.0%) on question E1 (lab-related) than the engineering majors (53.7%), and they had lower scores on question E2 (29.4% correct) compared to the engineering majors (43.2% correct). Biology majors scored much higher on question E1 than E2 (33.6% difference) compared to the engineering majors (10.5% difference). As on the kinetics questions, the behavior of male and female students paralleled the behavior of engineering and biology majors for this pair of questions.

**Table 9.** The comparison of the average performance (%) of different student majors and genders on lab-related (K2 and K4) and non-lab related questions (K3 and K5) about reaction orders from CHEM 102 second in-term exams of 2013 and 2014

| Major | Number of Students | Average Exam Scores | 2013 K2 | 2013 K3 | 2013 Difference |
|---|---|---|---|---|---|
| Engineering | 564 | 84.0 | 88.6 | 80.7 | 7.9 |
| Biological Science | 155 | 83.4 | 93.5 | 75.5 | 18 |
| **Gender** | | | | | |
| Male | 591 | 84.2 | 89.3 | 80.0 | 9.3 |
| Female | 286 | 83.5 | 93.0 | 80.1 | 12.9 |
| **Major** | **Average Exam Scores** | **2014 K4** | **2014 K5** | **2014 Difference** | **Average Exam Scores** |
| Engineering | 473 | 78.8 | 75.5 | 81.2 | -5.7 |
| Biological Science | 117 | 78.4 | 78.6 | 70.1 | 8.5 |
| **Gender** | | | | | |
| Male | 479 | 78.5 | 72.7 | 80.0 | -7.3 |
| Female | 245 | 78.5 | 81.2 | 75.5 | 5.7 |

**Table 10.** Average performance (%) of different student majors and genders on lab-related (E1) and non-lab-related (E2) questions about electrochemical cells from CHEM 102 final exam of 2014

| Major | Number of Students | E1 | E2 | Difference |
|---|---|---|---|---|
| Engineering | 495 | 53.7 | 43.2 | 10.5 |
| Biological Science | 119 | 63.0 | 29.4 | 33.6 |
| **Gender** | | | | |
| Male | 502 | 52.8 | 40.6 | 12.2 |
| Female | 251 | 61.8 | 34.7 | 27.1 |

Results of the solutions questions related to the conductivity experiment were also analyzed as a function of student major and gender. As was observed with the kinetics questions, engineering majors and male students showed similar results relative to biology majors and female students, respectively. The average performance of biology majors on the lab-related solutions question (C1, 52.5% correct) was lower than the average performance of engineering students on the same question (70.5% correct) (Table 11). The average performance of female students on the lab-related question (C1, 56.1% correct) was lower than that of male students on the same question (68.7% correct). The results of lab-related and non-lab-related questions were different than either of the results for the kinetics questions. Engineering majors performed 4.1% higher on the lab-related question (C1) than on the non-lab related question (C2), whereas biology majors showed the opposite result, 10% lower on C1 than C2. This might be because biology majors were better at remembering or understanding the concepts listed in question C2.

In addition, student performance on quantitative questions (T1, E3) and conceptual questions (T2, E1) were analyzed as a function of student major. It was found that the average performance of biology majors on quantitative questions (T1, 73.7% correct; E3, 43.7% correct) was lower than the average performance of engineering majors on the same questions (T1, 86.1% correct; E3, 53.5% correct). Biology majors, however, scored higher on conceptual questions (T2, 86.0% correct; E1, 63.0% correct) than engineering majors (T2, 79.7% correct; E1, 53.7% correct). These results may be because biology majors received more training in answering conceptual questions in other courses, while engineering majors have more experience answering quantitative questions and stronger math training.

**Table 11.** Average performance (%) of different student majors and genders on lab-related (C1) and non-lab-related (C2) questions about solutions from CHEM 101 first in-term exam of 2013

| Major | Number of Students | C1 | C2 | Difference |
|---|---|---|---|---|
| Engineering | 495 | 70.5 | 66.4 | 4.1 |
| Biological Science | 119 | 52.5 | 62.5 | -10 |
| **Gender** | | | | |
| Male | 502 | 68.7 | 67.1 | 1.6 |
| Female | 251 | 56.1 | 62.8 | -6.7 |

The variable results observed in these comparisons suggest that major and gender are unlikely to have a reproducible impact on student learning of the content presented in verification experiments. While some students will benefit more from these types of experiments than others, there is insufficient evidence from this study to draw a general conclusion about the effectiveness of verification experiments with respect to a specific group of students.

# 4. Summary and Conclusions

Evidence from this study suggests that students tend to perform better on lab-related exam questions associated with some verification experiments, such as kinetics and electrochemical cells. Results for the kinetics experiment were mixed, however. While students appeared to benefit from the practice gained generating graphs, they weren't always able to translate that learning to different questions. In 2013, an item characteristic curve and chi square tests showed that low-performing and medium-performing students scored higher on a lab-related kinetics question than a corresponding non-lab related kinetics question. However, in 2014, there was no difference in student scores at all performance levels. Further investigations are needed to determine why the experiments help some students answer some questions better than others. While students at all levels performed better on an electrochemical cell lab-related question than on an electrochemical cell question that wasn't related to the experiment, results are only available for one year. Therefore, it is unknown if this result is reproducible. It doesn't appear that the conductivity experiment helped most students to improve their performance on corresponding exam questions, as the results on lab-related and non-lab-related questions were very similar.

When lab-related quantitative questions were compared to lab-related conceptual questions, the results also were variable. For the pH titration experiment, students scored slightly higher on the quantitative question. For the electrochemical cells experiment, students scored slightly higher on the conceptual question. It is likely that the effect of verification experiments is content-specific. Different experiments will provide benefits that correspond to the particular design of that experiment and the skills and/or techniques that are reinforced in that experiment.

Student performance on lab-related and non-lab related questions does not appear to be a function of student major or gender. As with quantitative and conceptual questions, the results vary by experiment. For the kinetics experiment, biology majors and female students scored higher on the lab-related question. For the conductivity experiment, engineering majors and male students scored higher on the lab-related question. Rather than observing that verification experiments are generally more effective for a particular group of students (e.g., engineering majors), the effectiveness of a particular experiment should be determined for each population.

Student performance on quantitative and conceptual questions also was compared as a function of major. Analysis of two pairs of questions related to pH titration and electrochemical cell experiments indicates that engineering majors were better at answering quantitative questions and biological science majors tended to score better on conceptual questions. These results might be due to their relative prior experience with these types of questions.

The results of this study suggest that there is a relationship between student exam performance and some verification experiments. Future studies should include investigations of experiments related to chemistry topics not included in this study. Additional studies should also include a larger number of paired exam questions. Since this study only focused on the impact of verification experiments on exam performance, there is still a need to investigate the effectiveness of this type of experiment on the development of other skills, such as lab work and scientific writing.

# REFERENCES

[1] Hofstein, A.; Lunetta, V. N., Rev. Educ. Res. 1982, 52 (2), 201-217.

[2] Hofstein, A.; Lunetta, V. N., Science Education 2004, 88 (1), 28-54.

[3] Matz, R. L.; Rothman, E. D.; Krajcik, J. S.; Holl, M. M. B., J. Res. Sci. Teach. 2012, 49 (5), 659-682.

[4] Bybee, R. W., Science Education 1970, 54 (2), 157-161.

[5] Ben-Zvi, R.; Hofstein, A.; Samuel, D.; Kempa, R. F., J. Chem. Educ. 1976, 53 (9), 575-577.

[6] Hill, B. W., J. Res. Sci. Teach. 1976, 13 (1), 71-77.

[7] Wheatley, J. H., J. Res. Sci. Teach. 1975, 12 (2), 101-109.

[8] Raghubir, K. P., J. Res. Sci. Teach. 1979, 16 (1), 13-17.

[9] Etkina, E.; Karelina, A.; Ruibal-Villasenor, M.; Rosengrant, D.; Jordan, R.; Hmelo-Silver, C. E., J. Learn. Sci. 2010, 19 (1), 54-98.

[10] Renner, J. W.; Fix, W. T., J. Chem. Educ. 1979, 56 (11), 737-740.

[11] Kösem, Ş. D.; Özdemir, Ö. F., Science & Education 2014, 23 (4), 865-895.

[12] von Aufschnaiter, C.; von Aufschnaiter, S., Eur. J. Phys. 2007, 28 (3), S51-S60.

[13] Holstermann, N.; Grube, D.; Bögeholz, S., Research in Science Education 2010, 40 (5), 743-757.

[14] John Carnduff, N. R., Enhancing Undergraduate Chemistry Laboratories, Pre-Laboratory and Post-Laboratory Exercises, Examples and Advice. Royal Society of Chemistry: London, 2003; 32 pages.

[15] Bopegedera, A. M. R. P., J. Chem. Educ. 2011, 88 (4), 443-448.

[16] Lagowski, J. J., J. Chem. Educ. 1990, 67 (3), 185.

[17] Abraham, M. R., J. Chem. Educ. 2011, 88 (8), 1020-1025.

[18] Domin, D. S., J. Chem. Educ. 1999, 76 (4), 543-547.

[19] Pickering, M., J. Chem. Educ. 1987, 64 (6), 521-523.

[20] Louis E. Raths, S. W., Arthur Jonas, and Arnold M. Rothstein, Teaching for Thinking: Theory, Strategies, and Activities for the Classroom. Teachers College Press: New York, 1986; 264 pages.

[21] Tobin, K., Tippins, D. J., & Gallard, A. J., Research on Instructional Strategies for Teaching Science. In Handbook of Research on Science Teaching and Learning, Gabel, D. L., Ed. Macmillan: New York, 1994; pp 45-93.

[22] Holme, T.; Murphy, K., J. Chem. Educ. 2011, 88 (9), 1217-1222.

[23] Thorne, E. J., Laboratory Manual for General Chemistry. John Wiley & Sons, Inc: Hoboken, NJ, 2012; 101 pages.

[24] Thorne, E. J., Laboratory Manual for General Chemistry. John Wiley & Sons, Inc: Hoboken, NJ, 2013; 103 pages.