THE EFFECT OF MORE FREQUENT TESTING ON A HIGH SCHOOL CHEMISTRY CLASS

A CAPSTONE PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF EDUCATION
IN
SECONDARY EDUCATION
COLLEGE OF PROFESSIONAL STUDIES
HAWAIʻI PACIFIC UNIVERSITY
HONOLULU, HAWAIʻI
SPRING 2020

By
MATTHEW L. CAPPS

Advisors:
Valentina M. Abordonado, Ph.D
Eugene Guillian, Ph.D

ABSTRACT

In today's educational atmosphere, standardized testing is the metric by which a teacher's ability to teach is measured. Teachers need an effective method for ensuring their lowest-performing students improve on these standardized exams. Thirty-eight Maryknoll High School chemistry students participated in a study to evaluate the effects of more frequent testing on student exam grades. The findings suggest a possible link between the frequency of testing and improved mean test scores on academic exams. Although the overall results of the study were inconclusive, the implication for using repetition and frequent testing as an effective method of improving student grades seems promising.

The Effect of More Frequent Testing on a High School Chemistry Class

Imagine developing a tool for teachers, which would allow them to improve student learning, student knowledge retention, and student opinions of their teacher. This tool has been the focus of thousands of studies, due to its potential for improving the effectiveness of time spent in the classroom. As impressive as this tool may sound, it also requires no expensive equipment or special training before use. Variations of this tool exist and have been utilized by teachers in one form or another for over one hundred years. The tool is more-frequent testing.

Although most teachers today understand testing as an integral part of education as a system, it has not always been this way. According to Harvard researcher William J. Reece, the vast majority of testing before the 1800s was predominantly oral, and pupils in the early 1800s did not compete in standardized, written examinations (Reece, 2013). Between 1800 and 1840, testing as a measure of student and school performance increased dramatically, at least in part, due to the infatuation with a new field of science called statistics and the desire to use testing.

Today, testing is an integral part of our education framework, so much so that many significant consequences are tied to student performance on what has become known as "high stakes testing." Under the No Child Left Behind program, schools that did not meet required annual yearly progress or "AYP" might lose their accreditation or be forced to undergo complete school restructuring (Ginsburg & Kanter, 2002). Dylan William, a British educator, and researcher from the University of London states that one of the most troubling features of "high stakes testing" in the United States is that the stakes for teachers and schools are much higher than for any students who performed badly (Wiliam, 2010). While teachers and schools had significant consequences, students who performed poorly on standardized exams typically suffered little to no repercussions at all (Wiliam, 2010). Since student results on high-stakes testing seem to have so much weight in our current educational system, finding an effective method for improving student grades on these exams seemed like a worthwhile goal.

Purpose of the Study

This study sought to find an easy to implement, inexpensive, and efficient method or tool which teachers could use to help improve student performance on standardized exams. The guiding questions were, "Would a different type of testing help the students retain knowledge better and perform better on exams, or was it simply the frequency and/or amount of knowledge tested that affected their grades?"

This study builds upon the previous research of Turney (1931), Keys (1934), Dustin (1971), and Deck (1998), who found that more frequent testing improved the grades and attitudes of students in their classes. If a similar study here in Hawaii were also to suggest that more frequent testing resulted in an improvement in student exam grades, this technique might be a viable way of improving student test scores on the Hawaii State Assessment (HSA) exam. Since the frequency of testing seemed like an easy-to-implement strategy, the improvement in test scores could conceivably be realized with less stress and preparation time by teachers and students, allowing for more time spent on non-HSA related curriculum.

The following question guided a thorough literature review of the subject: Would more frequent testing positively impact test scores in a High School chemistry class?

A review of the research relating exam frequency to test score improvement has shown an interest in testing frequency spanning back almost 100 years. Studies touting the positive effect of frequent testing on student performance can be found as far back as the 1923 dissertation by Harold Jones, who noticed frequent testing of students improved short term recall of certain specific topics by up to 200% (Jones, 1923). Other early studies included Schutte, 1925; Deputy, 1929; Maloney & Ruch,1929; Turney, 1931, and Keys in 1934 (Johansen, 2009). Many of these very early studies found that large effect sizes in student exam score improvement could be made with more frequent testing. For example, Deputy (1929) found that increased exam frequency of college students resulted in a 0.96 standard deviation improvement in unannounced quiz grades, and the study by Maoloney and Ruch (1929) found a 0.59 standard deviation improvement infrequently tested students (Bangert-Drowns, Kulik, Kulik, 1991). The early works of Turney (1931) and Keys (1934) reported increases in student grades of 12% and 23%, respectively, for frequently tested groups. Kika, McLaughlin, & Dixon (1992) showed an 8% increase in the mean test scores of High School algebra students when the students were tested weekly instead of bi-weekly. Realizing even a meager 5% improvement in Hawaii State Assessment scores simply by quizzing or testing students more frequently seemed plausible.

Further review showed several aspects of the more frequent testing had been studied over the years as possible reasons for student score improvements. The main aspects studied have been the frequency of tests per unit time, test item overlap, student opinion, the use of mastery testing, and the uses of adjunct questions. This study chose to only look at test item overlap and frequency of testing.

Frequency of Testing. Most of the studies made an effort to use the frequency of testing as one of the variables tested. Unfortunately, there is no agreed-upon standard for what is considered frequent versus infrequent testing in the studies reviewed. Several meta-analyses chose to codify frequent and infrequent use in their analysis (Bangert-Drowns, Kulik, Kulik, 1991; Johansen, 2009), but I simply chose what was my typical time between tests (20 class periods) and used that as my infrequent testing period then considered double that (every 10 class periods) to be frequent testing.

In his 1981 study, Frank Fulkerson found that the frequent, shorter tests over smaller amounts of material led to significantly better test-by-test performance than the less frequent, longer tests over larger amounts of material (Fulkerson, 1981). An interesting aspect of several of the studies reporting a positive increase associated with more frequent testing was the effect disappeared when the final examination date was announced. Keys found that more frequent testing increased examination scores by 7% on unannounced final exams but had no effect on final examination scores when the exam date was known (Keys, 1934). The studies by Fulkerson (1981), and Kika (1992) both support this finding, with the researchers concluding that frequent testing helps students to perform better on a test-by-test basis, but this advantage disappears on an announced final exam when, presumably, cramming can equalize knowledge of course material.

Test Item Overlap. Test item overlap occurs when studies use the same or similar questions on more than one of the examinations during the testing period; an example would be using quiz questions on a later examination. Dustin(1971) showed that more frequent testing led to better long-term performance on tests when there is a significant amount of question overlap between the questions on the frequent tests and the final exam (Dutin, 1971). In another study, researchers examined the effect of test item overlap by administering quizzes that covered material that overlapped completely, partially, or not at all with the material covered on the tests (Burns, 1972). Burns went on to conclude that both the direction and magnitude of the effect of evaluative quizzes on test performance is dependent on the degree of quiz/test overlap (Burns, 1972). In yet another study of test item overlap, Jeffrey Kling (2005) found that quiz/exam overlap of questions resulted in an improved student exam score of half of a letter grade. However, in classes with low content overlap, the results indicated that students did the same regardless of testing frequency, as no statistically significant differences were found.

Student Opinion. The final aspect that I considered for my study was the impact that more frequent testing had on student opinions of the teacher and the subject being taught. Study results would seem to suggest that students prefer more frequent testing to less frequent testing. Deck (1998) found that students overwhelmingly preferred frequent testing over less frequent testing. The researcher went on to say that frequent testing can promote a positive attitude toward the class, the material, and perhaps a more positive attitude about the instructor as well, making improved student evaluations an added benefit for instructors who choose to use an increased testing frequency. Bangert-Drowns, Kulik, and Kulik (1991)found that frequent testing promoted student learning and encouraged the development of extrinsic characteristics such as more frequent study, increased interest in the subject area, and positive attitudes toward the subject. Kika (1992) found that 88% of students preferred more frequent to less frequent testing because they preferred the smaller units of study, the shorter tests, and less required review. Similar results were found in a study of students in a marketing class as well (Miller, 1987)

Most studies did not attempt to determine why there was an increase in the student opinions of the class and teacher. However, Murphy & Stanga (1994)hypothesized that more frequent testing gives the students and instructors more time to interact, allowing the students to become more accustomed to the instructor and providing the general feeling that the instructor truly cares about their well-being.

The Critics. The opponents of more frequent testing point to inconsistencies in studies showing positive results and cast doubt on the validity of the effect of frequency on test performance. For instance, studies performed by Proger and Mann (1973) and Burns and Vinchur (1992) both found that past streams of research had produced inconsistent results and could not reliably show that more frequent testing improved student retention. They contended that the studies conducted by Keys (1934) and Turney (1931) were considered questionable because the researchers had grouped the students based upon initial test scores, which may have biased their results. Kling, McCorkle, Miller, & Reardon (2005) believed that the amount of improvement reported in most of these studies was less affected by the frequency of testing and more by the effect of exam question overlap. When the overlap between frequent tests and the infrequent test was the same, no improvement was gained by more frequent testing (Fulkerson, 1981). Fulkerson and Martin (1981) went on to conclude that the superior test results and higher instructor evaluations were likely the results of more frequent feedback and lower test anxiety brought about by becoming more accustomed to the expectations of the professor. In their meta-analysis of over 40 studies related to the frequency of testing, Bangert-Drowns, Kulik, & Kulik (1991) reviewed and codified many studies but only used ones that met specific criteria for procedures, instruments, and statistical procedures. The strict criteria for inclusion of studies in their meta-analysis were designed to promote consistency in data collection and statistical analysis, as well as remove previous errors that may have been introduced by what they considered improper testing procedures. Factors used for inclusion in their study were: 1) the studies had to be in actual classrooms; any past studies performed on paid volunteers were considered irrelevant;. 2) the study test groups had to receive identical instruction, except for testing frequency; 3) The tests involved had to be conventional classroom tests and not mastery testing or adjunct questioning techniques, and 4) studies had to be free of significant flaws such as pre-treatment or using biased statistical analysis.

The results of the meta-analysis performed by Bangert-Drowns, Kulik, and Kulik (1991) did suggest that more frequent testing improved student performance on exams, but at a diminishing rate of return for each doubling of test frequency. Teachers who initially doubled the amount of testing in their classes saw an initial increase of 0.54 standard deviations in the mean score, whereas subsequent doubling only resulted in 0.08 standard deviation or less improvement, and the third doubling saw an even smaller increase in the test. There was also the concern that the advantages of more frequent testing disappeared on announced exams, as found in the studies by Keys (1934), Fulkerson (1981), and Kika (1992).

Interpretation and Implications for Professional Practice. In the eighty-five years since Turney first published his positive findings on the effect of more frequent testing, much research and debate have surrounded the topic. Proponents think the studies show a correlation between more frequent testing and improvement in test scores, as well as improved affective evaluations of the teacher and the subject, whereas opponents point to the inconsistencies in the data and call the positive results unreliable. The results of all the studies warrant further study and refinement to determine if, indeed, there is a positive correlation between testing frequency and student grades. If the correct balance between testing and teaching is found, teachers could improve student achievement and interest by providing just the right amount of testing to promote an optimal student-learning environment by simply varying the frequency of the tests and quizzes they provide.

Hypothesis

Students who are subjected to testing every 10 class periods will show a greater increase in their score on a post-test than students who are subjected to testing once every 20 class periods. The outcome of this hypothesis is expected to mirror the results of studies conducted by

Keys (1934), Dustin (1971), and Turney (1931).

Method

The participants in this study were 38 tenth and eleventh-grade students in a Maryknoll High School chemistry class. Maryknoll High School is a private, Catholic school located in Honolulu, Hawaii, with a total student enrollment of approximately 400 students. The annual tuition per student at Maryknoll is $18,500 per year, making it likely that the socioeconomic status of the students in the study is at or above the mean for Honolulu. The sample contained approximately the same ethnic and cultural diversity of the school (68% Asian or Pacific Islander and 32% Caucasian) and was 39.5% male and 60.5% female. Ten of 38 students (26%) had perfect 4.0 GPAs, another 10 of 38 students (26%) had GPAs from 3.5 to 3.99, Eight of 38 students (21%) had GPAs from 3.0 to 3.49, Nine of 38 students (24%) had GPAs between 2.50 and 2.99, and only one of 38 students (2.6%) had a GPA less than 2.5. The students were separated into a control group (N=17) and an experimental group (N=21)based on their original chemistry class period, with the control group having a combined average GPA of 3.31 and the experimental group having an average GPA of 3.50. A single factor ANOVA on the control group and experimental group GPAs did not suggest that one group was academically superior to the other with an almost 20% probability of the difference being due to random error (f, 1.73; f-critical 4.12, p-value 0.197)

Sampling Procedures, Participants were selected for the control group or experimental group based on the period of their chemistry class. A purely experimental design would have produced the best results for this study, but a quasi-experimental design was chosen due to the inability to randomly assign the students to groups

Maryknoll high school includes grades nine through twelve, with approximately 100 students per grade for a total school population of approximately 400. Participants of this study were either in the 10^{th} or 11^{th} grade during their participation. Of the 39 students asked to participate in the study, only one declined, and one was absent on the day of the post-test resulting in 37 total students in the study.

Hawaii Pacific University Institutional Review Board (IRB) and Maryknoll School administration reviewed the proposed plan of the study in Appendix A and B for student safety and ethical concerns, and permission was granted before beginning the study. The risks of any ill effects due to the study were extremely slim, but all students were required to have parental consent forms signed before being allowed to participate in the study (See Appendix C). Parental consent was requested for 39 students to participate in the study and was given for 38 of the 39 students.

Student assent was gained each time a test was taken, and students were reminded before each test that they could opt-out of having their test scores included in the study data. Student test scores were part of the standard class procedure and were included in student grades regardless of their assent decision.

Possible sources of sample bias. Differences in the type of training or level of instruction between the experimental and control groups might have resulted in confounding the results of the study. To ensure this did not happen, both classes received identical instruction except for testing frequency.

Incorrect statistical analysis would have also introduced errors in the results. For this reason, all statistical analysis techniques were chosen after discussion with a Hawaii Pacific University professor knowledgeable in statistical methods typically employed in educational research.

Measures

Data collection began with both the control group and experimental groups taking the pre-test during periods one and two (See Appendix D).The pre-test, mid-test, and post-test were all scheduled on days, which had the test groups scheduled for morning classes to reduce the chance that time of day might affect the outcome of the testing.

Both the control and experimental group completed identical pre-tests) on March 31st, 2008. After completing the pre-test, both classes began a chemistry unit in chapters one through four of the Holt Chemistry book (Holt, 2004). Every five days of instruction, both classes would complete practice reviews over the material covered. These practice reviews were developed, using the chapter quizzes from the Holt Chemistry textbook and were included in the student grade but only accounted for a tiny percentage of the student's overall grade.

On April 11th, after nine instructional days and approximately halfway through the unit plan, the experimental group completed a “mid-test” covering material from chapters one and two of the Holt chemistry book (See Appendix E). Questions on the mid-test were aligned to similar topics as the pre and post-test but were not identical to the pre/post-test questions.

Both classes completed the post-test on April 25th, 2008, after 19 instructional days in chapter one to four chemistry unit (See Appendix D).

The instruments used in the study were a pre-test (Appendix D), a mid-test (Appendix E), and a post-test (Appendix D). Parental consent was obtained for all students whose data was used for the study, and student assent was obtained each time an exam was given. Absent student assent, the exams would have still counted as part of the student's grade but would not have been used in the study results. Student names were only used for the classroom grade tabulation, and any identifying student names were removed from the record.

The pre-test and post-test (Appendix D) were identical, while the mid-test (Appendix E) included questions that were similar to the pre and post-test but not identical. The pre-test and post-test were developed by sampling questions from the chapter 1-4 exams included in the Teacher's Edition of Holt Chemistry (Davis et al., 2006), while the mid-test was created using different questions than the pre-test and post-test sampled from chapter 1 and 2 exams of the same book.

The questions for the pre-test, post-test, and mid-test were developed from the textbook, much like the study completed by Deck (1998). Using questions derived from the book, ensured that test questions matched what the students were learning from the chemistry unit for chapters one through four and were written to the correct reading level for a high school chemistry class.

Research Design. A purely experimental design would have been the pinnacle of research for this study, but a quasi-experimental design was chosen due to the inability to randomly assign the students to the control and experimental groups. The students were assigned to the control group or the experimental group, based upon which period their chemistry class occurred. Differences in mean GPA for the control group and experimental group were not statistically significant, as discussed in the introduction.

Threats to Validity. The threats to internal and external validity included but were not completely limited to Testing Effect, Instrumentation Effect, Statistical Regression, Time of Day, and Low Student Motivation to Participate.

Testing Effect. Due to the short amount of time (three weeks) between the pre-test and post-test, the students may have performed better on the post-test simply since they had become "test-wise." One solution to this threat would have been to extend the amount of time between the pre and post-test; unfortunately, this could not be done due to the timeline stipulated by the course for which the study is being completed.

Instrumentation Effect. If the pre-test were harder than the post-test, a positive effect could be realized simply due to the easier post-test. The mentor teacher involved in the study verified that the pre-test and post-test are of equal difficulty.

Statistical Regression. Statistical regression describes the phenomenon where a student who scores either high or low on a pre-test will test closer to the median score on the post-test. This is typically due to factors outside the control of the experiment (how the student feels, amount of sleep, unlucky guesses on test items) and could not be accounted for in the study.

Time of Day. Only one of the classes was before lunch, and the other two were after lunch. The study used the two classes after lunch for experimental groups to avoid any bias introduced by testing after lunch.

Low Student Motivation to Participate. The students may not have been motivated to apply themselves to the examinations given if the tests were not part of their overall grade. The grades achieved for testing throughout this study were included in the student grades for the quarter.

Variables. The independent variable in the study was the frequency of testing, and the total number of questions asked. The dependent variable was student grades as measured on the pre-test/post-test (Appendix D) and mid-test (Appendix E). Frequent testing was defined as one exam every ten days or less of class time. Non-Frequent testing was defined as one exam every twenty days of classroom instruction. Student scores were percentage scores from 0% to 100% on all exams.

Procedures. Data collection began with both the control group and experimental groups taking the pre-test (Appendix D) during periods one and two, which were both before lunch. Pre-test, mid-test, and post-test were all scheduled on days during which both classes occurred before lunch to reduce the chance that time of day might affect the outcome of the testing.

Both the control and experimental group completed identical pre-tests(Appendix D) on March 31st, 2008, and immediately began a unit on matter and energy derived from the Holt Chemistry Textbook (Davis et al., 2006) chapters one through four. The same instructor taught both classes and ensured that lessons on each day were as identical as possible with the classes averaging one chapter every 5 class periods. After each chapter, both classes would have short, 15, or 20 questions "quizzes" for review, which were based on chapter quizzes from the book. On April 11th, after nine instructional days and approximately halfway through the unit plan, the experimental group completed a “mid-test” (Appendix E) covering material from chapters one and two of the Holt chemistry book. While the control group was required to review chapter one and two notes for the 30 minutes of class time, the experimental group used this time to take the mid-test. Questions on the mid-test were taken from chapters one and two chapter exam questions, which were not used on the pre-test/post-test (Appendix D). These questions were not identical to pre-test/post-test questions but did provide question overlap between questions on the mid-test and questions on the pre-test/post-test. After the mid-test on April 11th, both classes had ten more days of instruction with review quizzes after each chapter, as with chapters one and two. On April 25th, after 19 total instructional days of class time, both classes completed the post-test during periods one and two of the school day. For the duration of the study, the experimental group was tested at what was considered the more-frequent testing periodicity of about once every ten days. In comparison, the control group was only tested at the non-frequent testing frequency of about once every 20 days.

Potentially confounding variables arose when the student grades for the mid-test and practice quizzes were entered into the school's online grade report. Many of the students requested to re-take quiz and exam grades in order to improve their scores. This policy was consistent with the teacher's posted quiz policy in the class syllabus and resulted in most students retaking the chapter quizzes in order to improve their grades. Students retaking the same quizzes to improve their grades may have provided much more question overlap for both the experimental and control groups than was initially intended for the study. It would have been much better for the teacher to obtain permission from Maryknoll's administration to allow the posting of grades to be delayed until completion of the study. This would have allowed a clean collection of data without the confounding re-quizzing of students from both groups.

Results

The study lasted 19 total class periods from March 31st, 2008, until April 25th, 2008. Data collected from the students included GPA, sex, pre-test score, post-test score, and mid-test score for the more-frequently tested group. Upon completion of the study, results were analyzed for gain between pre-test and post-test scores with both groups showing gains in post-test scores when compared to pre-test scores. The gains were expected following a chemistry unit covering topics tested on the post-test, but the study is seeking to determine if the more-frequently-tested group had a larger increase in test scores than the less-frequently-tested group?

The difference in pre-test and post-test scores, or delta-score, was calculated for each student and then used to calculate the mean score for each group. The mean delta scores and standard deviations, shown in table 1 below, were (M=16.7, SD=6.6) for the frequently-tested group and (M=16.5, SD= 10.2) for the less-frequently-tested group.

Table 1
Comparison of the test means between experimental (N=20) and control group (N=17)

	Mean Test Scores			Standard Deviation
	Pre-test	Post-test	Delta-Score	Delta-Score
Less-Frequent Testing	43.18	59.65	16.47	10.2
More-Frequent Testing	47.40	64.10	16.7	6.6

Pre-test and Post-Test examinations were identical 52 question exams derived from textbook questions

On the surface, there appears to be no significant difference between the frequently-tested and less-frequently-tested groups, but analysis for statistical significance of the results was undertaken to determine if results were statistically significant.

Data Analysis Procedures. An independent, two-tailed t-test was completed to a confidence interval of 95% on the delta-score results for both classes. The outcome of the t-test was $t(35) = -0.083$, $p= .065$, meaning I fail to reject the null hypothesis similar means between the two samples. The difference between the control group and experimental group means for the study is not statistically significant.

Discussion

On the face of the findings, the study does not seem to support the hypothesis that more frequent testing results in improved grades for more frequently tested students. According to the previous findings of Keys (1934) and Turney (1931), a more pronounced and statistically significant improvement in student grades should have been noted between study groups; instead, both groups showed similar, statistically insignificant improvements in mean post-test scores.

Summary of findings. The study of more frequent testing performed on a Maryknoll High School chemistry class did not support the findings of many other studies, which showed an improvement in test grades for more frequently tested groups. These results can most likely be attributed to the findings of Keys (1934) and Kika (1992), which show that more-frequent testing improvements disappear when the final examination date is known, and the effect of student cramming reduces any improvement provided by more frequent testing.

Another possible reason the results of the study were different than expected was the ability of students to complete chapter quiz reviews during the time allocated for completion of the study. It was the normal classroom procedure prior to the study for students to be allowed to retake quizzes for grade improvement. In an oversight by the teacher completing the study, this procedure was not altered for the duration of the study resulting in many students from both groups being allowed to retake review quizzes for each chapter. This re-quizzing resulted in much more review and repetition of the material and questions than the study was originally designed to include.

Overall Interpretation. Overall the study was inconclusive; however, the lower variance of the more-frequently-tested class (N=20, M=16.7, SD=6.6) over the less-frequently-tested class (N=17, M-16.5, SD=10.2) may suggest a more predictable outcome in classes where more frequent testing is required. If teachers could improve student grades by simply asking repetitive questions or covering the same curriculum several times per testing period, the same technique should be able to be used to improve statewide Hawaii State Assessment exam results.

Limitations. The limitations of this study are numerous but not insurmountable. Although the study had hoped to be in agreement with the findings of previous studies, no effect was actually attained. Study size and previously established classroom procedures tended to be the critical detractor from positive results or being able to determine the statistical significance of more frequent testing. The length of the study, being only three weeks total, was also a significant roadblock to obtaining any useful data; the short time frame did notallow any significant trends to be noted in the results of the testing. Other limitations of the study include the lack of original data able to be retrieved from the original study. This analysis is being completed twelve years after the original study was completed. The original study also included subject and teacher rating questionnaires, which were not retained by the researcher.

Recommendations for Further Study. Additional research is needed to explore the possible positive effects of exam frequency; if possible, the studies should include multiple versions of test frequency as well as exam question overlap to determine which has a larger effect on student achievement results. The study should also be of a longer duration and across many differenttypes of classes with various teachers and teaching styles.

References

Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. L. (1991). Effects of frequent classroom testing. *Journal of Educational Research*, 85, pp. 89-99.

Burns, D. J., & Vinchur, A. J. (1992, Sept). Effects of evaluative quizzes on test performance. *Journal of Instructional Psychology*, 19(3), pp. 45-52.

Dustin, D. S. (1971). Some effects of exam frequency. *The Psychological Record*, 21, pp. 409-414.

Fulkerson, F. E., & Martin, G. (1981, Apr). The effects of exam frequency on student performance, evaluations of the instructor, and test anxiety. *The teaching of Psychology*, 8(2), pp. 90-93.

Davis, R., & Holt, Rinehart, and Winston, Inc. (2006). *Modern Chemistry Teacher's Edition*. Orlando: Holt, Rinehart, and Winston. (2006).

Ginsburg, A., & Kanter, A. No Child Left Behind. A Desktop Reference 2002, No Child Left Behind. A Desktop Reference 2002 (2002). Jessup, MD: Education Publication Center.

Johanson, B. G. (2009). Effectiveness of frequent testing over achievement: a meta-analysis study *International Journal of Human Sciences* [Online]. 6:2. Available: http://www.insanbilimleri.com/en

Jones, H. (1923). *Experimental studies of college teaching: The effect of examination on the permanence of learning* (Unpublished doctoral dissertation).

Keys, N. (1934). The influence on learning and retention of weekly as opposed to monthly tests. *Journal of Educational Psychology, 25*(6), 427-436. doi:10.1037/h0074468

Kika, F. M., McLaughlin, T. F., & Dixon, J. (1992, Jan/Feb). The effects of frequent testing of secondary algebra students. *Journal of Educational Research*, 85(3), pp 159- 162.

Kling, N., McCorkle, D., Miller, C., & Reardon, J. (2005, Nov/Dec). The impact of testing frequency on student performance in a marketing course. *Journal of Education for Business,* 81(2), pp. 67-72.

Murphy, D. P., & Stanga, K. G. (1994). The effects of frequent testing in an income tax course: an experiment. *Journal of Accounting Education*, 12(1), pp. 27-41.

Miller, F. (1987). Test frequency, student performance, and teacher evaluation in the basic marketing class. *Journal of Marketing Education*, 9(2), pp. 14-19

Turney, A. H. (1931). The effects of frequent and short objective tests upon the achievement of college students in educational psychology. *School and Society*, 33, pp. 760-762.

Wiliam, D. (2010) Standardized Testing and School Accountability, *Educational Psychologist*, 45:2, 107-122, DOI: 10.1080/00461521003703060

Hawai‘i Pacific University
Institutional Review Board
Project Application

Please complete and submit the form to the IRB chair via email: to irbchair@hpu.edu

Study title: The Effect of More Frequent Testing on a High School Chemistry Class

Investigator:

Name: Matt Capps
(Please check one)
☐ Faculty x Student ☐ Outside Investigator

Phone: (808) 428-3116

Email: cappsmatt@gmail.com

Sponsoring HPU Faculty Member: **Valentina M. Abordonado, Ph. D**
Eugene Gillian, Ph. D

Please attach a brief summary of the project. This should include an explicit statement of methods, data collection, and how confidentially of subjects/data will be protected, including a consent form.

Category for Review:

Check one level of review (Exempt, Expedited, Full) for which you believe the project qualifies, and each criterion that your project meets.

X **Exempt from review (nil or minimal risk study, or already reviewed by an IRB)**

__x__ Research involves ONLY investigation into or comparison of normal instructional strategies.

__x__ Tests, interviews, and surveys are unlikely to elicit emotion or place subjects at risk of civil/criminal liability or damage to their reputation, financial standing, employability, etc. AND information will not be recorded in such a way that subjects can be identified.

____ Research involves only the study or analysis of existing data, documents, records, or specimens that are publicly available or recorded in such a way that subjects cannot be identified.

____ If the study involves ingestion of food: only wholesome food without additives in excess of USDA recommended levels are consumed.

__x__ Brief informed consent will be done (except in the case of existing data, etc.).

____ No use of vulnerable subjects (children, prisoners, pregnant women, mentally ill, etc.).

__x__ Has already been approved by IRB at __Hawaii Pacific University 2008 _________.
(Original Copy from 2008 not available)

____ **Expedited review (minor risk study)**

__x__ Research and data collection methods are unlikely to elicit strong emotion, and deception is not involved.

____ Research involves only noninvasive, painless, and non-disfiguring collection of physical samples, such as hair, sweat, excreta.

____ No use of vulnerable subjects (children, prisoners, pregnant women, mentally ill, disabled, etc.).

____ Data are recorded using noninvasive, painless, and non-disfiguring sensors or equipment, such as EKG, weighing scales, voice/video recording.

____ Research involves only moderate levels of exercise in healthy volunteers.

__x__ Research does not involve the ingestion of drugs or the use of hazardous devices.

__x__ If existing data, documents, records, or specimens with identifiers are used, procedures are in place to ensure confidentiality.

__x__ Informed consent process will be done (attach a copy of the informed consent form).

__x__ Data will be kept confidential and not reported in identifiable fashion.

____ **Full review required (more than minor risk)**
Attach a statement that describes the use of vulnerable subjects or the study procedures and conditions that place subjects at risk. Describe the precautions that will be taken to minimize these risks. Attach a copy of the informed consent form that will be used.

Certification by Principal Investigator: The above represents a fair estimate of risks to human subjects.

Matt Capps / Student / February 18, 2020
____
Name/ Title/ Date
- -

FOR IRB USE ONLY

Certification by IRB Chair: I have read this application and believe this research qualifies as:
____ Exempt from IRB review
____ Appropriate for expedited review, and
____ approved
____ disapproved
____ Appropriate for review by the full IRB

____ _____________
IRB Chair Date

INFORMED CONSENT DOCUMENT

Project Title: The Effect of More Frequent Testing in a High School Chemistry Classroom.

investigator(s): Matthew Capps

PURPOSE

This study involved research that took place in two Maryknoll high school chemistry classrooms during the fall of 2008. The purpose of the research was to determine if more frequent testing of students would result in an improvement in student test scores. The two classrooms were given a pre-test followed by three weeks of the chemistry curriculum, and then a post-test covering the material learned during the month. The control group was given a pre-test and post-test, while the experimental group was given a pre-test, post-test, and a mid-test.

RISKS

The risks to test subjects were minimal. IRB approval was required due to the involvement of minors.

BENEFITS

There was no personal benefit to the participants of the study. It was hoped that the knowledge gained during the study might provide improved testing schedules.

COSTS AND COMPENSATION

There was no cost or compensation for participants in the study.

CONFIDENTIALITY

Records of participation were maintained and kept confidential to the extent permitted by law. However, Hawai'i Pacific University IRS may inspect and copy a subject's records pertaining to the research, and these records may contain personal identifiers. In the event of any report or publication from this study, the identity of subjects will not be disclosed. Results will be reported in a summarized manner in such a way that subjects cannot be identified.

VOLUNTARY PARTICIPATION

All participation was voluntary, and there was no penalty for withdrawal.

QUESTIONS

Questions are encouraged. Questions about this research project and questions about the rights of research subjects or research-related injury may be addressed to the IRB Chair at irbchair@hpu.edu.

Subject's name (print): ____________________ **Date:** _______________

Signature of Subject: _____________________

INVESTIGATOR STATEMENT

I have discussed the above points with the subject or the legally authorized representative, using a translator when necessary. It is my opinion that the subject understands the risks, benefits, and obligations involved in participation in this project.

[signature] 2/18/20

(Signature of Investigator) (Date)

STANDARD LANGUAGE FOR SPECIFIC ISSUES

RISKS

NEW INFORMATION

If new information related to a subject's willingness to continue to participate develops during the course of this study, subjects will be promptly informed.

CONFIDENTIALITY

CERTIFICATE OF CONFIDENTIALITY Please contact the IRB Chair, irbchair@hpu.edu, for information on how to obtain this Certificate. Subjects may receive a copy of this certificate upon request.

Hawaii Pacific University

Teacher Education Program

Dear Parents:

My name is Matt Capps. I am currently in the Masters of Education program at Hawaii Pacific University, working on a research project titled "The Effects of More Frequent Testing on a High School Chemistry Class." The purpose of this research is to determine if more frequent testing improves student test scores and the opinions of the teacher.

I am inviting students to participate in this research because they are students at Maryknoll High School where I am completing field experience courses for my degree. This project will last for four (4) weeks and will involve approximately forty students. Students who agree to participate in this project can expect the following to occur:

If the student is part of the control group, he/she will be administered:

- A short test and questionnaire at the beginning of the study.
- A short test and questionnaire at the end of the study.

If the student is part of the experimental group, he/she will be administered:

- A short test and questionnaire at the beginning of the study.
- A short quiz at the end of the two-week point over the topics covered during the period.
- A short test and questionnaire at the end of the study.

There are a few foreseeable risks to participating in this project. There may be no personal benefit for participating in this study; however, it is hoped that in the future, society could benefit from this study by determining if frequent tests help students perform better on examinations.

There will be no cost to your child for participating in this research project, nor will your child be compensated. Records of participation in this research project will be maintained and kept confidential to the extent possible. However, regulatory agencies and the Hawai'i Pacific University IRB may inspect and copy records pertaining to the research, and these records may contain personal identifiers. In the event of any report or publication from this study, the identity of research participants will not be disclosed. Results will be reported in such a way that the research participants cannot be identified.

All participation is voluntary. There is no penalty to any child who decides not to participate, nor will any child be penalized if he or she decides to stop participation at any time during the research project.

You may address any questions you have about this research project or the protection of human subjects to the IRB Chairs, Dr. Mary Sheridan at 566-2489 or at msheridan@hpu.edu.)

I have discussed the above points with the student or the legally authorized representative, using a translator when necessary. It is my opinion that the student understands the risks, benefits, and obligations involved in participation in this project. I understand and give my permission for my child ____________________________________to participate in this study.

Parent or Legal Guardian's signature: ____

Date:____

Thank you for your consideration of my request for your child's participation in this research project.

Sincerely,

(Matt Capps)

Pre-Test and Post-Test

______ **1.** A good definition of a chemical is any substance that
a. causes corrosion.
b. has a definite composition.
c. is hazardous.
d. is used in industrial processes.

______ **2.** Matter that is free to move and fills its available volume is in the ____ state.
a. liquid
b. solid
c. gaseous
d. elemental

______ **3.** The state of matter in which a material is most likely to resist compression is the
a. solid state.
b. liquid state.
c. gaseous state.
d. vaporous state.

______ **4.** A chemical change occurs when
a. dissolved minerals solidify to form a crystal.
b. ethanol is purified through distillation.
c. salt deposits form from evaporated seawater.
d. a leaf changes color.

______ **5.** Which of the following shows a physical change occurring?
a. A peach spoils.
b. A copper bowl tarnishes.
c. A piece of jewelry turns your skin green.
d. A hot-glue gun melts a glue stick.

______ **6.** Mercury(II) oxide decomposes into mercury and oxygen. In this reaction, mercury is
a. a reactant.
b. a product.
c. a foul-smelling substance.
d. sometimes a reactant and sometimes a product.

______ **7.** Which of the following observations does *not* indicate that a chemical change has occurred?
a. change of state
b. formation of a precipitate
c. absorption of energy
d. release of a gas

Use the following diagram to answer items 8–10.

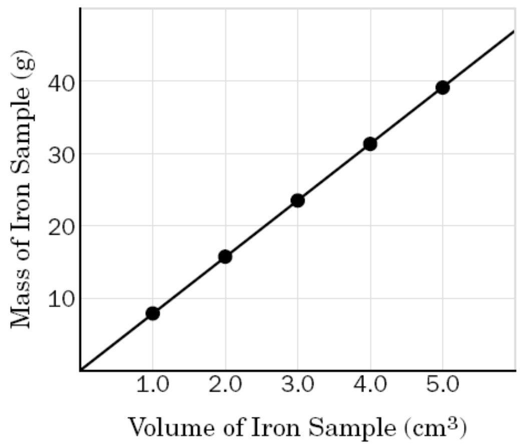


Figure 1

______ **8.** What property is illustrated by the slope of the line?
a. density
b. length
c. mass
d. volume

______**13.** To convert from mL to L, what conversion factor is used?

a. $\frac{1000\ \text{mL}}{1\ \text{L}}$

b. $\frac{1\ \text{L}}{1000\ \text{mL}}$

c. $\frac{1000\ \text{L}}{1\ \text{mL}}$

d. $\frac{1\ \text{mL}}{1000\ \text{L}}$

______**14.** Which of the following statements about density is true?

a. Density depends on the size of the sample.
b. Density is a chemical property.
c. Density is a physical property.
d. Density depends on location.

______**15.** A sample of aluminum has a mass of 20.6 g and a volume of 7.63 cm^3. What is the density of aluminum?

a. 0.370 g/cm^3
b. 2.70 g/cm^3
c. 28.2 g/cm^3
d. 157 g/cm^3

______**16.** One chemical property of matter is

a. boiling point.
b. texture.
c. reactivity.
d. density.

______18. Which of the following statements is false?
a. Nitrogen is an element that exists as a diatomic gas.
b. Atoms are elements combined in a definite ratio.
c. Water is a molecule.
d. Helium is an element.

______19. Which of the following mixtures is homogeneous?
a. iced tea
b. a cheeseburger
c. pizza
d. vegetable soup

______20. What type of matter is formed when two or more elements chemically join together?
a. compound
b. element
c. heterogeneous mixture
d. homogeneous mixture

______ 1. The capacity to do work or generate light describes
a. energy.
b. specific gravity.
c. physical changes.
d. chemical changes.

______ 2. An energy change is involved in
a. an exothermic process but not an endothermic process.
b. an endothermic process but not an exothermic process.
c. both an exothermic process and an endothermic process.
d. neither an exothermic process nor an endothermic process.

______ 6. Two hundred degrees Celsius is equal to
a. −73 K.
b. 200 K.
c. 273 K.
d. 473 K.

______ **8.** The scientific method is a

a. series of exact steps.

b. theory.

c. strategy for drawing sound conclusions.

d. lengthy experiment.

______**10.** Plants were raised for a science project. Each plant was given an equal amount of sunlight and kept at the same temperature. Different amounts of fertilizer dissolved in different amounts of water were added to each plant. Why are the results of this experiment not valid?

a. The amount of sunlight should have been varied.

b. The temperature should have been varied.

c. Either the amount of fertilizer or the amount of water should have been constant.

d. All the variables should have remained constant.

______**14.** All of the following are steps in the scientific method *except*

a. observing and recording data.

b. forming a hypothesis.

c. discarding data inconsistent with the hypothesis.

d. developing a model.

______**20.** What is 976.23 expressed in scientific notation?

a. 976.23

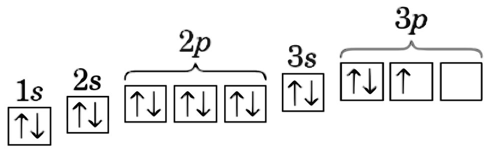
b. 9.7623×10^{2}

c. 9.7623×10^{-2}

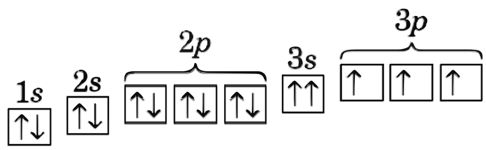
d. 9.7623×10^{4}

______ **1.** Which of the following orbital notations for phosphorus is correct?

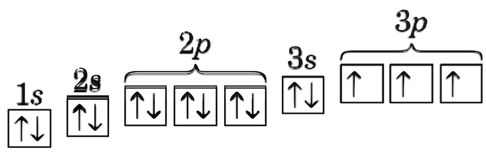
a.



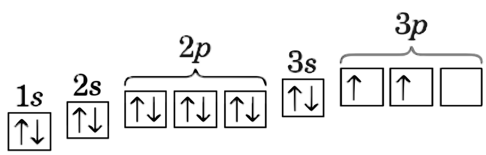
b.



c.



d.



______ **2.** Atoms contain equal numbers of
- **a.** electrons and neutrons.
- **b.** protons and neutrons.
- **c.** protons and electrons.
- **d.** protons, electrons, and neutrons.

______ **3.** The atomic symbol for beryllium, $^{9}_{4}\mathbf{Be}$, indicates that the
- **a.** atomic number is 4.
- **b.** atomic number is 9.
- **c.** mass number is 4.
- **d.** atomic number is equal to $9 - 4$.

______**16.** What is $1.245\ 633\ 501 \times 10^8$, rounded to four significant figures?
- **a.** 1246
- **b.** 1.2456×10^8
- **c.** 1.246×10^8
- **d.** 1.246×10^4

______**17.** Which of the following numbers has five significant figures?
- **a.** 23 410
- **b.** 0.006 52
- **c.** 0.017 83
- **d.** 10.292

______ **7.** Rutherford's gold foil experiment led him to conclude that
a. Thomson's plum pudding model of the atom was accurate.
b. alpha particles were a poor choice for a bombardment material.
c. a concentrated positive charge existed somewhere within the atom.
d. light was emitted by electrons returning to ground state.

______ **10.** The diagram [↑↓] shows two electrons with
a. opposite spins.
b. the same spin.
c. different energies.
d. the same energy.

______ **4.** Which of the following pairs of atomic symbols represent isotopes?
a. ^{235}U and ^{238}U
b. P_4 and P_8
c. ^{32}P and ^{83}Pb
d. ^{50}Sn and ^{51}Sb

______ **15.** The electron configuration in **Figure 1** violates

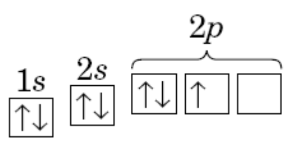


Figure 1

a. the Pauli exclusion principle.
b. the aufbau principle.
c. Hund's rule.
d. Both (a) and (c)

______**12.** The deflection of cathode rays in Thomson's experiments gave evidence of the ____ nature of electrons.

a. wave
b. charged
c. particle
d. spinning

______**13.** An atom of potassium has 19 protons and 20 neutrons. Its mass number is

a. 9.
b. 19.
c. 20.
d. 39.

______**17.** One mole of iron atoms

a. is more massive than one mole of aluminum atoms.
b. is less massive than one mole of aluminum atoms.
c. is a larger quantity of atoms than one mole of aluminum atoms.
d. is a smaller quantity of atoms than one mole of aluminum atoms.

______**18.** Avogadro's number is

a. 6.022×10^{23}.
b. 1.602×10^{-24}.
c. 3.1416.
d. 3.0×10^{-8}.

______**19.** The molar mass of an element is numerically equal to the

a. element's atomic mass.
b. element's atomic number.
c. number of neutrons found in the element.
d. mass of one atom of the element.

The Periodic Table

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10	Group 11	Group 12	Group 13	Group 14	Group 15	Group 16	Group 17	Group 18
1	1 H Hydrogen 1.01																	2 He Helium 4.00
2	3 Li Lithium 6.94	4 Be Beryllium 9.01											5 B Boron 10.81	6 C Carbon 12.01	7 N Nitrogen 14.01	8 O Oxygen 16.00	9 F Fluorine 19.00	10 Ne Neon 20.18
3	11 Na Sodium 22.99	12 Mg Magnesium 24.30											13 Al Aluminum 26.98	14 Si Silicon 28.09	15 P Phosphorus 30.97	16 S Sulfur 32.07	17 Cl Chlorine 35.45	18 Ar Argon 39.95
4	19 K Potassium 39.10	20 Ca Calcium 40.08	21 Sc Scandium 44.96	22 Ti Titanium 47.88	23 V Vanadium 50.94	24 Cr Chromium 52.00	25 Mn Manganese 54.94	26 Fe Iron 55.85	27 Co Cobalt 58.93	28 Ni Nickel 58.69	29 Cu Copper 63.55	30 Zn Zinc 65.39	31 Ga Gallium 69.72	32 Ge Germanium 72.61	33 As Arsenic 74.92	34 Se Selenium 78.96	35 Br Bromine 79.90	36 Kr Krypton 83.80
5	37 Rb Rubidium 85.47	38 Sr Strontium 87.62	39 Y Yttrium 88.91	40 Zr Zirconium 91.22	41 Nb Niobium 92.91	42 Mo Molybdenum 95.94	43 Tc Technetium 97.91	44 Ru Ruthenium 101.07	45 Rh Rhodium 102.91	46 Pd Palladium 106.42	47 Ag Silver 107.87	48 Cd Cadmium 112.41	49 In Indium 114.82	50 Sn Tin 118.71	51 Sb Antimony 121.76	52 Te Tellurium 127.60	53 I Iodine 126.90	54 Xe Xenon 131.29
6	55 Cs Cesium 132.90	56 Ba Barium 137.33	57 La Lanthanum 138.91	72 Hf Hafnium 178.49	73 Ta Tantalum 180.95	74 W Tungsten 183.84	75 Re Rhenium 186.21	76 Os Osmium 190.23	77 Ir Iridium 192.22	78 Pt Platinum 195.08	79 Au Gold 196.97	80 Hg Mercury 200.59	81 Tl Thallium 204.38	82 Pb Lead 207.2†	83 Bi Bismuth 208.98	84 Po Polonium (209)	85 At Astatine (210)	86 Rn Radon (222)
7	87 Fr Francium (223)	88 Ra Radium (226)	89 Ac Actinium (227)	104 Rf Rutherfordium (261)	105 Db Dubnium (262)	106 Sg Seaborgium (263)	107 Bh Bohrium (264)	108 Hs Hassium (265)	109 Mt Meitnerium (268)	110 Uun Ununnilium (269)	111 Uuu Unununium (272)	112 Uub Ununbium (277)		114 Uuq Ununquadium (285)				

58 Ce Cerium 140.12	59 Pr Praseodymium 141.91	60 Nd Neodymium 144.24	61 Pm Promethium 144.91	62 Sm Samarium 150.36	63 Eu Europium 151.97	64 Gd Gadolinium 157.25	65 Tb Terbium 158.92	66 Dy Dysprosium 162.50	67 Ho Holmium 164.93	68 Er Erbium 167.26	69 Tm Thulium 168.93	70 Yb Ytterbium 173.04	71 Lu Lutetium 174.97
90 Th Thorium 232.04	91 Pa Protactinium 231.04	92 U Uranium 238.05	93 Np Neptunium (237)	94 Pu Plutonium (244)	95 Am Americium (243)	96 Cm Curium (247)	97 Bk Berkelium (247)	98 Cf Californium (251)	99 Es Einsteinium (252)	100 Fm Fermium (257)	101 Md Mendelevium (258)	102 No Nobelium (259)	103 Lr Lawrencium (262)

Use the periodic table above to answer the questions in this Chapter Test.

In the space provided, write the letter of the term or phrase that best completes each statement or best answers each question.

______ **1.** Mendeleev attempted to organize the chemical elements based on their
- **a.** symbols.
- **b.** properties.
- **c.** atomic numbers.
- **d.** electron configurations.

______ **2.** A horizontal row in the periodic table is called a(n)
- **a.** family.
- **b.** group.
- **c.** octet.
- **d.** period.

______ **5.** How many electrons does carbon have in its outermost shell?
a. 1
b. 2
c. 3
d. 4

______ **6.** Which of the following elements behaves similarly to calcium?
a. magnesium
b. sodium
c. sulfur
d. chlorine

______ **7.** Which periodic group or series of elements is not correctly matched with its common family name?

a. alkaline-earth metals	Group 2
b. transition metals	Group 3
c. halogens	Group 17
d. noble gases	Group 18

______ **8.** Highly reactive metallic elements that react with water to form alkaline solutions are called
a. actinides.
b. alkali metals.
c. halogens.
d. noble gases.

______ **9.** The electron configurations of main-group elements end in

a. *d* and *f* orbitals.
b. *s* and *p* orbitals.
c. *s* and *d* orbitals.
d. *p* and *d* orbitals.

______**10.** Which of the following elements is a transition metal?

a. calcium
b. iron
c. sodium
d. sulfur

______**11.** Which property is not characteristic of a metal?

a. conductivity
b. brittleness
c. luster
d. ductility

______**12.** The alkali metal elements are found in ________________ of the periodic table.

a. Group 1
b. Group 2
c. Period 1
d. Period 2

______**14.** A measure of the ability of an atom in a chemical compound to attract electrons is called

a. electron affinity.
b. electron configuration.
c. electronegativity.
d. ionization potential.

____**13.** The energy required to remove an electron from an atom is called the atom's

a. electron affinity.
b. electron energy.
c. electronegativity.
d. ionization energy.

______**16.** Which of the following elements has the lowest electronegativity?
a. C
b. F
c. Li
d. O

______**17.** Which of the following elements has the greatest ionization energy?
a. Ga
b. K
c. Bi
d. As

______**18.** Which of the following elements has the greatest electron affinity (largest positive value)?
a. Br
b. As
c. Ar
d. I

______**19.** Naturally occurring elements are created in stars by a process known as
a. electrolysis.
b. fission.
c. transmutation.
d. fusion.

Pre-Test / Post-Test derived from a sample of Myers, R. T., Oldham, K. B. and Tocci, S. (2004) *Holt Chemistry*. 4th edn. Austin, TX: Holt, Rinehart and Winston. (*Pages 219-238)*

Mid-Test

The Science of Chemistry

In the space provided, write the letter of the term or phrase that best completes each statement or best answers each question.

______ 1. A good definition of a chemical is any substance that
- a. causes corrosion.
- b. has a definite composition.
- c. is hazardous.
- d. is used in industrial processes.

______ 2. Matter that is free to move and fills its available volume is in the ______________________________ state.
- a. liquid
- b. solid
- c. gaseous
- d. elemental

______ 3. The state of matter in which a material is most likely to resist compression is the
- a. solid state.
- b. liquid state.
- c. gaseous state.
- d. vaporous state.

______ 4. A chemical change occurs when
- a. dissolved minerals solidify to form a crystal.
- b. ethanol is purified through distillation.
- c. salt deposits form from evaporated seawater.
- d. a leaf changes color.

______ 5. Which of the following shows a physical change occurring?
- a. A peach spoils.
- b. A copper bowl tarnishes.
- c. A piece of jewelry turns your skin green.
- d. A hot-glue gun melts a glue stick.

______ 6. Mercury(II) oxide decomposes into mercury and oxygen. In this reaction, mercury is
- a. a reactant.
- b. a product.
- c. a foul-smelling substance.
- d. sometimes a reactant and sometimes a product.

______ 7. Which of the following observations does *not* indicate that a chemical change has occurred?
a. change of state
b. formation of a precipitate
c. absorption of energy
d. release of a gas

Use the following diagram to answer items 8–10.

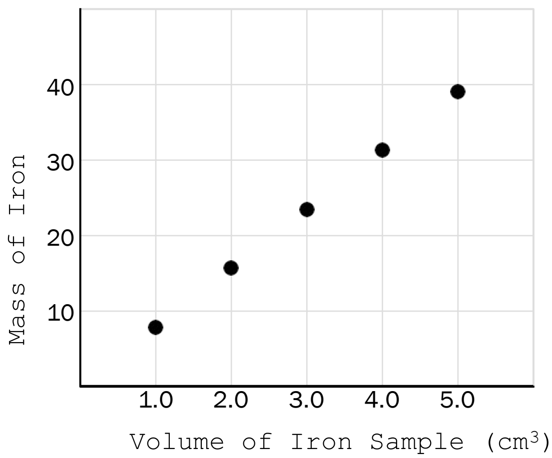


Figure 1

______ 8. What property is illustrated by the slope of the line?
a. density
b. length
c. mass
d. volume

______ 9. How can the property identified in item 8 be determined from the graph?
a. Divide the mass by the volume for any point on the line.
b. Divide the volume by the mass for any point on the line.
c. Multiply the mass by the volume for any point on the line.
d. None of the above

______10. The property shown by the graph can be used to identify a pure substance because each pure substance has its own unique
a. color.
b. density.
c. mass.
d. volume.

______11. What is the unit of measure for the answer to the following calculation?

$$\frac{19.3\ g}{cm^3} \times \frac{10\ cm^3}{1\ cm^3} \times \frac{1\ mol}{196.97\ g}$$

a. mol
b. mol/cm^3
c. cm^3
d. g/mol

______12. The SI base unit for length is the
a. meter.
b. kilometer.
c. centimeter.
d. inch.

______13. To convert from mL to L, what conversion factor is used?
a. $\frac{1000\ mL}{1\ L}$
b. $\frac{1\ L}{1000\ mL}$
c. $\frac{1\ mL}{1000\ L}$
d. $\frac{1\ mL}{1000\ L}$

______14. Which of the following statements about density is true?
a. Density depends on the size of the sample.
b. Density is a chemical property.
c. Density is a physical property.
d. Density depends on location.

______15. A sample of aluminum has a mass of 20.6 g and a volume of 7.63 cm^3.
What is the density of aluminum?
a. 0.370 g/cm^3
b. 2.70 g/cm^3
c. 28.2 g/cm^3
d. 157 g/cm^3

Matter and Energy

In the space provided, write the letter of the term or phrase that best completes each statement or best answers each question.

______ 1. The capacity to do work or generate light describes
- a. energy.
- b. specific gravity.
- c. physical changes.
- d. chemical changes.

______ 2. An energy change is involved in
- a. an exothermic process but not an endothermic process.
- b. an endothermic process but not an exothermic process.
- c. both an exothermic process and an endothermic process.
- d. neither an exothermic process nor an endothermic process.

______ 3. Recharging a heat pack by placing it in a pot of boiling water involves
- a. a transfer of energy as heat from system to surroundings.
- b. a transfer of energy as heat from surroundings to system.
- c. a transformation of chemical energy into energy as heat.
- d. no transfer of energy as heat between system and surroundings.

______ 4. Which of the following statements about mass and energy is true?
- a. Mass is conserved, but energy is used up.
- b. Under ordinary conditions, both are conserved.
- c. Mass is unrelated to energy.
- d. Energy is conserved, but mass is used up.

______ 5. Energy is transferred from hot water to a cup. This energy is in the form of
- a. chemical energy.
- b. heat.
- c. potential energy.
- d. temperature.

______ 6. Two hundred degrees Celsius is equal to
- a. −73 K.
- b. 200 K.
- c. 273 K.
- d. 473 K.

______ 7. A temperature of 590 K is equal to
- a. −317°C.
- b. 273°C.
- c. 317°C.
- d. 863°C.

______ 8. The scientific method is a
a. series of exact steps.
b. theory.
c. strategy for drawing sound conclusions.
d. lengthy experiment.

______ 9. An experiment is designed to measure the effects of increasing sulfur dioxide concentrations in the local atmosphere on the pH of rainwater. The pH of naturally occurring rainwater in the area can be considered
a. a model.
b. a variable.
c. irrelevant data.
d. the control.

______10. Plants were raised for a science project. Each plant was given an equal amount of sunlight and kept at the same temperature. Different amounts of fertilizer dissolved in different amounts of water were added to each plant. Why are the results of this experiment not valid?
a. The amount of sunlight should have been varied.
b. The temperature should have been varied.
c. Either the amount of fertilizer or the amount of water should have been constant.
d. All the variables should have remained constant.

______11. A statement or mathematical equation that describes a fact or relationship found in the universe is a(n)
a. law.
b. theory.
c. assumption.
d. hypothesis.

______12. A theory is an accepted explanation of an observed phenomenon until
a. one study conflicts with the theory.
b. repeated data and observation conflict with the theory.
c. scientists disagree about the methods used to gather the data.
d. an eminent scientist feels that the theory is inadequate.

______13. All measurements are subject to error. If a certain instrumental error recurs in each measurement, the________________________________of the experimental results is affected.
a. identification
b. accuracy
c. reportability
d. precision

______14. All of the following are steps in the scientific method *except*

a. observing and recording data.
b. forming a hypothesis.
c. discarding data inconsistent with the hypothesis.
d. developing a model.

______15. During heating, a 5-g solid gives off oxygen gas. The mass of the residue after heating is 3.5 g. To the correct number of significant figures, what is the mass of the oxygen?

a. 1.5 g
b. 2 g
c. 3.5 g
d. 8.5 g

Mid-Test derived from sample of Myers, R. T., Oldham, K. B. and Tocci, S. (2004) *Holt Chemistry*. 4th edn. Austin, TX: Holt, Rinehart and Winston.

Control Group Anonymized Data

Control Student Number	GPA	M(0) F(1)	Group Con(0) Exp(1)	Pre Test	Post Test	Diff Post Pre
8	4.00	0	0	34	72	38
3	4.00	0	0	52	82	30
5	4.00	0	0	62	90	28
2	4.00	0	0	50	76	26
6	3.63	0	0	36	60	24
4	3.53	0	0	24	44	20
15	3.88	0	0	52	70	18
16	3.19	1	0	46	62	16
12	4.00	0	0	52	66	14
13	2.94	0	0	28	42	14
9	2.81	1	0	38	50	12
10	3.71	0	0	34	46	12
1	3.41	1	0	54	64	10
17	4.00	0	0	50	58	8
7	3.47	1	0	48	54	6
11	2.81	1	0	36	42	6
14	2.65	1	0	38	36	-2

N		17	17	17
mean		43.17647	59.64706	16.4706
sd		10.46563	15.31723	10.211

Experimental Group Anonymized Data

Exp Student Number	GPA	M(0) F(1)	Group Con(0) Exp(1)	Pre Test	Post Test	Diff Post Pre
19	3.25	0	1	34	68	34
24	3.71	0	1	38	62	24
28	2.13	1	1	44	68	24
29	4.00	0	1	54	76	22
33	3.82	0	1	36	56	20
35	3.53	1	1	38	58	20
20	2.88	0	1	34	52	18
26	3.06	1	1	34	52	18
31	4.00	0	1	62	80	18
32	3.33	1	1	62	80	18
27	3.50	1	1	52	68	16
36	2.73	1	1	38	54	16
22	4.00	0	1	72	86	14
38	4.00	0	1	64	78	14
18	2.50	1	1	42	54	12
34	2.94	1	1	44	56	12
23	3.63	0	1	58	68	10
37	2.69	0	1	48	58	10
39	3.29	0	1	50	60	10
21	3.11	1	1	44	48	4
		N		20	20	20
		mean		47.4	64.1	16.7
		SD		11.44506	11.15395	6.562253