THE EFFECT OF A FORMAL LAB PRACTICAL PERFORMANCE ACTIVITY BEFORE WRITTEN EXAMS ON TEST SCORES IN HIGH SCHOOL PHYSICS CLASSES

By

Matthew Andringa

A Master's Paper

Submitted to the Graduate Faculty

In partial fulfillment of the requirements for the Degree of

Master of Science in Education – Physics

____

Advisor's Signature

____

Date

University of Wisconsin – River Falls

November 25, 2013

Abstract

Mastering a physical concept is and should always be the goal of any lesson in a physics classroom. How to prove the understanding of that concept is difficult to assess. Written exams only give us a small indication of our students' competency. However, if we ask them to demonstrate their understanding of the concept with some sort of task, they can prove it. This study examines whether the use of an individual lab practical performance activity the day before their written exam increased student test scores. Test scores from three separate semesters of physics classes were compared. In the first two semesters, the classes had a traditional review day before the exam and the lab practical was not used at all. In the second year, we introduced the lab practical on the review day as a short lab-based assessment. The results of this study indicate that the students who were given the lab practical scored the same statistically as the previous semesters.

Table of Contents

Literature Review

In my experience teachers are used to traditional assessments consisting of paper-and-pencil problem-solving worksheets, quizzes and exams. What may be lacking are performance-based assessments. Performance-based assessments use hands-on experiment-type activities to demonstrate students' understanding of principles. In order to understand the purpose and effect of performance based assessments, we must first look at others' work on the topic. The goal of education is not just to provide our students with information, but to make sure those students can take that information and apply it to real life situations. "21st century skills" is a term that is used frequently in education, yet it is hard to determine what these skills are or how they will be applicable to a person's job (Silva, 2009). Critical thinking and reasoning skills are more important than simple rote memorization. "An emphasis on what students can do with knowledge, rather than units of knowledge they have is the essence of 21st century skills" (Silva, 2009). The International Baccalaureate (IB) Diploma Programme, for example, is a curriculum that requires students to both master basic subject matter as well as critical thinking skills through a variety of internal and external assessments. As an IB teacher, I have seen that these classes do increase the level of rigor and give students a better idea of the application of knowledge as shown in assessments than do traditional courses.

The purpose and authenticity of assessments becomes a key ingredient in the ability to measure a student's critical thinking and reasoning skills. "The term authenticity refers less to the particular challenge or question and more to the realism of the setting" (Wiggins, 2011). This definition provides a clear message for how students should be assessed. "The primary purpose of assessment in school should be to educate (and motivate) students about the real world of adult challenges. Assessment should better replicate or simulate what mathematicians, scientists, and historians do, not just what they know" (Wiggins, 2011). Thus, the goal of any assessment should be to test the application of students' knowledge.

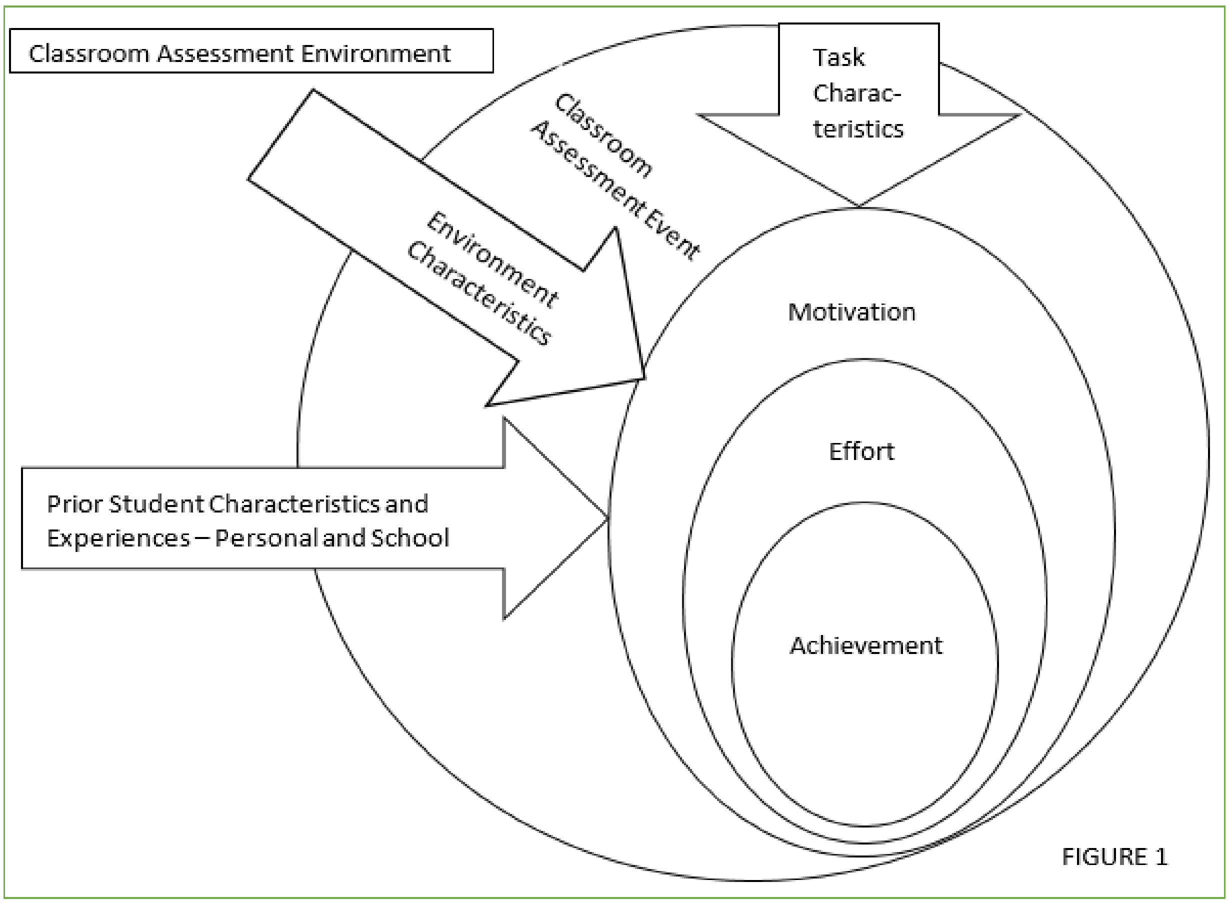
Educators are frequently reminded of Bloom's (1956) taxonomy when designing assessments. We aim to get away from simple remembering of information and move toward higher level skills like applying and analyzing. Wiliam (2011) discusses that learners must engage in actions that improve learning. With this idea in mind, assessment must support the learning of an idea. When we look at assessing students with a simple multiple choice exam, we must ask ourselves if this type of assessment is truly aiding in the learning of an idea, or is it just rote memorization. Marzano (1994) says performance tasks "provide information about students' abilities to analyze and apply information – their ability to think – whereas more traditional forms that employ forced-choice response formats (multiple-choice, fill-in-the-blank, true/false) assess only students' recall or recognition of information."

In a study done at Lincoln Memorial University, 791 graduate level teachers were surveyed about how assessment affects classroom instruction and students' learning (Cole, 2009). The study indicated that student learning is frequently put aside to simply prepare them for a standardized test. Most teachers in the study said that they were able to implement higher level assessments, but these assessments significantly reduced the time to cover the required content standards. "What is best for students is performance-based assessments – essays, research projects, science experiments, or work that shows student progress over time such as portfolios" (Cole, 2009). Perhaps educators are putting aside authentic assessments in favor of assessments that match up to the standardized tests that are very often used to rank and compare our students.

More importantly, it is worth noting that not all students learn or are able to prove their proficiency the same way. "Despite three decades of advances in information and communication technology and a generation of research on cognition and new pedagogical strategies, the field of assessment has not progressed much beyond paper-and-pencil item-based tests" (Clarke, 2010). Clarke's study shows that true higher-order thinking skills are very hard to measure with standard multiple-choice or paper-and-pencil test. This idea has led many educators to incorporate differentiation in their assessments. This practice allows students a variety of media to indicate that they have learned the objectives and standards. “Summative assessment is used to determine whether the student has successfully learned what was taught. These assessments can look as different from one another as our students do” (Levy, 2008). Giving students multiple assessments, performance-based and standardized, can at least provide more students with chances to show what they know.

The word assessment alone can be troubling for some students. In a study done by Brown, *et al.* (2009), a survey of 705 New Zealand secondary students answer questions about their conceptions and preferences of assessment. Students were given a list of twelve different assessment practices and asked to provide what they normally think of assessment. This study classified assessments into two types, Interactive-Informal and Teacher-Controlled. Interactive-Informal assessment examples include in-class questioning, teacher-scored portfolios, or the use of a rubric. Teacher-Controlled assessment examples include written tests made by the teacher, in class written essays, and long one to three hour examinations. A problem arose from this study, as the methods that educators have been pushed to implement (Interactive-Informal), were commonly seen as less important than standard tests by students. This could mean that what research says about assessment does not match what our students believe. “It may be that the students have potentially legitimate concerns about the reliability and validity of this kind of assessment practice." As educators we need to find a way to use interactive assessment to aid in their learning.

The investment of our students in their own learning and assessment might then come down to a measure of motivation. Figure 1 is a visual organizer for the factors that go into a student's achievement (after Brookhart, 2006).



"Figure 1. Factors in Student Achievement. Many factors are involved in student achievement with motivation being very significant."

The criteria that ultimately lead to achievement are many; however, motivation is a key ingredient.   In the study by Brookhart, *et al.* (2006) they found that "self-efficacy is the strongest motivational predictor of achievement." This is the belief that students can accomplish the task they are presented with. It also found that the most important factors in achievement were the same no matter what kind of assessment was presented, written pen-and-paper tests or performance assessment. Therefore, if we want to use better practice and implement performance assessments, which show a higher degree of critical thinking, we must develop performance assessments that are challenging yet attainable to increasing a student's desire to master a concept.

We must also remember that we cannot completely abandon traditional means of assessment as they are still key ingredients in the college placement of our students. "Regardless of the value of performance assessments in the classroom, a measurement-driven reform strategy that relies on performance assessments to drive curriculum and instruction seems bound to fail" (Haertel, 1999). While Haertel seems down on the idea of performance assessments, he does recognize that there is "real value" for performance assessments for instructional purposes and that there is a lot of good that can come from a good mix of assessment methods. This point is further echoed more recently by Clarke (2010): "multiple assessments are needed to make adequate observations of student performance." It is important to expose students to a variety of assessments and allow them to use multiple modes to measure their competency.

For years educators focused on teaching content and standards as opposed to teaching skills that utilize the content. "In the end, there is no real choice between content and skills. For those dedicated to improving day-to-day learning and longer-term student outcomes, designing standards, curricula, and assessments that reflect this reality is paramount" (Silva, 2004). It is this statement that drove the idea behind my study.

Purpose / Hypothesis

The purpose of this study was to find if students who completed a lab practical assessment the day prior to their written exam were able to increase their exam scores. Hypothesis: Students who receive the lab practical activity treatment will score higher than those students from previous years who did not complete the lab practical activity.

Methods Used

This study was performed at Oconomowoc High School in Oconomowoc, Wisconsin. Oconomowoc is a community with a population of approximately 16,000. The high school draws students from the city of Oconomowoc and surrounding rural areas such as Ixonia, Okauchee, and Ashippun. The total enrollment at the high school is approximately 1500 students. The high school runs a four 93 minute block schedule and all students in the study were in semester long blocks. The control group was taught in the fall of 2010 and the spring of 2011 and was made up of five sections of general physics all taught by the same teacher. There were 115 students in the control group with 59 male and 56 female. The test group was taught in the fall of 2011 and was made up of five sections of general physics taught by two teachers who worked closely together on the development of the curriculum and assessments. There were 94 students in the test group with 50 male and 44 female. The controls in this study were the length of the units, the formative assessments (worksheets, labs, activities), and the final written unit exams. The independent variable in the study was the use of the lab practical performance assessments given on the day before five different unit exams.

The lab practical performance activities were created by myself and the other instructor teaching physics during the test group semester. These activities were designed to test the fundamental concepts and objectives for each unit. Each activity included a calculation portion and a manipulative portion. The students were given some information to solve a problem and shown the apparatus, but were also tested on their ability to show that they truly understood the concept by performing a task. Students were given a summative grade based on the two parts, generally 10 points for the calculation, and 10 points for the performance. Although a grade was given on the lab practical, this study only looked at the results of the written exams after the completion of the lab assessment. Refer to Appendix A for student assessment sheets and pictures of the lab set-ups.

During Unit 2 – Constant Velocity, we created a lab that tested students on key terms like displacement, initial position, final position, average and constant velocity. In this particular lab, students were given the constant velocity of a battery operated car, a given time, and an initial position (marked with a piece of tape). The students were then directed to place a piece of tape at a position along a meter stick so when the car was started and moved between the pieces of tape, the time of travel matched the time that was given.

The Unit 3 – Constant Acceleration lab practical required students to understand the key ideas of initial velocity, final velocity, displacement, and constant acceleration. The students were given a piece of clear plastic that needed to be dropped through a photo gate. The goal was to cover the piece of plastic with tape and paper so when the plastic was dropped through the photo gate, it would produce a given time during which the gate was blocked. To test the understanding of the key ideas, we asked that the students start their taped section at either 5 or 10 cm from the bottom of the plastic so it had an initial velocity when it reached the photo gate.

In the Unit 4 – Static Forces lab practical, students were tested on the key concepts of equilibrium, mass, gravitational force, tension force, and basic trigonometry. An angle to horizontal and a tension force were given to the students. They then had to calculate a mass that, when hung in the center of the string, produced the given tension force which was read by a force probe.

Unit 6 – Kinematics in 2D was an extension of Unit 2 and Unit 3. This unit involved key concepts used in previous assessments, however, students then had to apply those concepts to two dimensional motion. A Vernier® Projectile Launcher was used to launch a steel ball off a lab station. Students had to use their knowledge of constant horizontal velocity, constant vertical acceleration, displacement, and time in the air to place a target at the appropriate range along a meter stick. Students were given a launch angle, velocity at that angle, and launch height. They then placed their target at the location they predicted, through calculation that the ball would land. A piece of carbon paper was laid on top of the target, face down, to ensure an accurate measurement.

For the Unit 7 – Energy lab practical, students had to apply the concept of conservation of energy and displacement. Students were given a velocity of a falling ball and were asked to determine the height at which the ball needed to be dropped in order to produce the given velocity through a photo gate. The students then needed to drop the ball from a height above the photo gate. This lab proved to be a bit tricky due to the ball needing to be dropped with no horizontal component.

Results

In the tables that follow, the average grades for each exam from the control and test groups are given. A two-sample t-test and probability value was used to demonstrate the statistical significance of the difference between the groups.

Table 1 – Unit 2 Constant Velocity 1D – Battery Car Lab Practical

Unit 2	Average	Standard Deviation	Number of students	2 sample t-test	p-value
Control	84.4	12.9	115	-1.47	.928
Group	81.7	13.5	95		

Table 2 – Unit 3 Constant Acceleration 1D – Free Falling Object Lab Practical

Unit 3	Average	Standard Deviation	Number of students	2 sample t-test	p-value
Control	77.3	13.7	115	-2.63	.995
Group	72.1	14.6	95		

Table 3 – Unit 4 Static Forces – Hanging Mass Tension Force Lab Practical

Unit 4	Average	Standard Deviation	Number of students	2 sample t-test	p-value
Control	76.3	17.9	115	-2.99	.998
Group	69.7	14.0	95		

** Unit 5 Unbalanced forces was not assessed due to time restraints and difficulty developing an appropriate lab practical. **

Table 4 – Unit 6 Projectile Motion (2-D Kinematics) – Projectile Lab Practical

Unit 6	Average	Standard Deviation	Number of students	2 sample t-test	p-value
Control	77.1	14.1	115	.170	.432
Group	77.4	11.2	95		

Table 5 – Unit 7 Energy, Work, and Power – Transfer of Energy Lab Practical

Unit 7	Average	Standard Deviation	Number of students	2 sample t-test	p-value
Control	78.5	16.5	115	-3.29	.999
Group	71.2	15.5	95		

** Unit 8 Circular Motion and Gravitation, and Unit 9 Momentum and Impulse were not assessed due to difficulty developing an appropriate lab practical. **

Analysis

The results of the study show that there was little difference in either group of students and therefore no statistically significant change in test scores. When analyzing the data, all units had t scores of 3.29 or less and all had probability values greater than .05. The meaning of this analysis shows that there is no significant difference in test scores between the two groups, contrary to my hypothesis. Indications are that the implementation of the lab practical performance assessment did not result in the effect of increasing written test scores that I expected.

The two groups of students had the same formative assessments (i.e. worksheets, quizzes, labs) and written exams. Also, the time spent on each unit was virtually identical, with the exception of time adjustments around holidays. The major difference between the control and treatment groups was the lab practical performance assessment. Since both groups had essentially the same class experience and scored statistically the same on written exams, one could conclude with a fair amount of certainty that the treatment had no effect on the test group.

As for possible reasons for the small apparent discrepancies in average exam scores, there is one minor difference in the groups. The test group did receive the treatment from two different instructors. Although my colleague and I worked very closely when planning the units and implementing the lab practical assessments, there could have easily been some differences in our presentation of the material.

In terms of unexpected observation, I was expecting that scores would improve. I truly expected that scores would go up with the use of these activities. I felt that they were constructed in a way to really address the main ideas of each unit and hoped that they would lead to better test performance. The only test that had an increase in average, although only .3%, was the unit on projectiles. During this unit, we invested in a Vernier® Projectile Launcher and used it in our lab practical. One wonders if the use of this device simply increased students' interest in the topic and gave them more incentive to understand the problem. While some students said they enjoyed the activities, there were those who said they dreaded them. In hindsight, the lab activities may have increased test anxiety and reduced the amount of time that students had to study on the review day for the upcoming test. Even if these ideas are true, I still feel that the performance assessments are better indications of the understanding and application of the ideas discussed in class. Though no statistical improvement was made from control to test group, the value of performance assessments still remain a priority for the curriculum since, anecdotally, many students expressed admiration for the activities as they provided them with some concrete use for the concepts they were learning indicating some attitude improvement.

If I were to repeat this study, I would likely perform the study over a longer period of time using only my students. This would eliminate the possible differences in content delivery and increase my sample size. I would also try to incorporate more lab practical assessments at the end of each unit. New assessments have been added to the remaining units since the inception of the study. Another possible change would be to give a couple informal performance activities prior to prepare them for the summative assessment.

Since the implementation of these performance based lab activities, we have continued to use them at the end of our units. In addition, since the study was concluded, another instructor has been introduced to these activities. He and I have worked at adjusting some of the lab practical assessments to make them better and more comprehensive. We believe, as a physics department, that these activities give us the best indication of conceptual comprehension and are committed to using them to gauge our students understanding.

Summary

In summary, the use of lab practical performance assessments did not produce the results I was expecting. Statistically speaking, there was no difference between the two groups of students as neither group scored significantly higher than the other. As a professional educator, I see the value of these activities and will continue to work towards a more comprehensive curriculum that challenges students to apply their learning as opposed to recalling information.

Bibliography

Bloom, B.S. (1956) Taxonomy of Educational Objectives, Handbook 1: *The Cognitive Domain.* New York: David McKay Co. Inc.

Brookhart, S. M.; Walsh, J. M.; & Zientarski, W. A. "The Dynamics of Motivation and Effort for Classroom Assessments in Middle School Science and Social Studies."*Applied Measurement in Education* 19.2 (2006):151-184

Brown, G.T.L.; Irving, S.E.; Peterson, E.R.; & Hirschfeld, G.H.F. "Use of interactive–informal assessment practices: New Zealand secondary students' conceptions of assessment." *Learning and Instruction* 19.2 (2009):97-111

Clarke - Midura, J.; & Dede, C. "Assessment, technology, and change." *Journal of Research on Technology in Education* 42.3 (2010):309-328

Cole, H.; Hulley, K.; & Quarles, P. "Does assessment have to drive the curriculum?" *Forum on Public Policy: A Journal of the Oxford Round Table*, (Spring 2009)

Haertel, E. H. "Performance Assessment and Education Reform" *Phi Delta Kappan* 80.9 (1999): 662

Levy, H.M. "Meeting the needs of all students through differentiated instruction: helping every child reach and exceed standards." *The Clearing House* 81.4 (2008):161-164

Marzano, R.J. "Lessons from the field about outcome-based performance assessments. (The Challenge of Outcome-Based Education)" *Educational Leadership* 51.6 (1994):44-50

Silva, E. "Measuring skills for 21st-century learning: designing assessments that measure newly important skills presents challenges, but that should not be an excuse for failing to evaluate what students know and are able to do." *Phi Delta Kappan*, 90.9 (2009):630-634

Wiggins, G. "Moving to modern assessments." *Phi Delta Kappan*, 92.7 (2011):63

Wiliam, D. "What is assessment for learning?" *Studies in Educational Evaluation*, 37.1(2011):3-14

Appendix A – Lab Practical Assessments

Unit 2 Lab Practical Assessment **Name:___________________________**

Given Time:_______________

Average Velocity of Car:_______________

Show work to determine the predicted displacement (10 points) including proper units:

Actual Time:________________ (+/- 20%)

Success?_______________ (10 points)

Grade: ______/20

Name:____

Distance to start of taped region - .05m or .10m

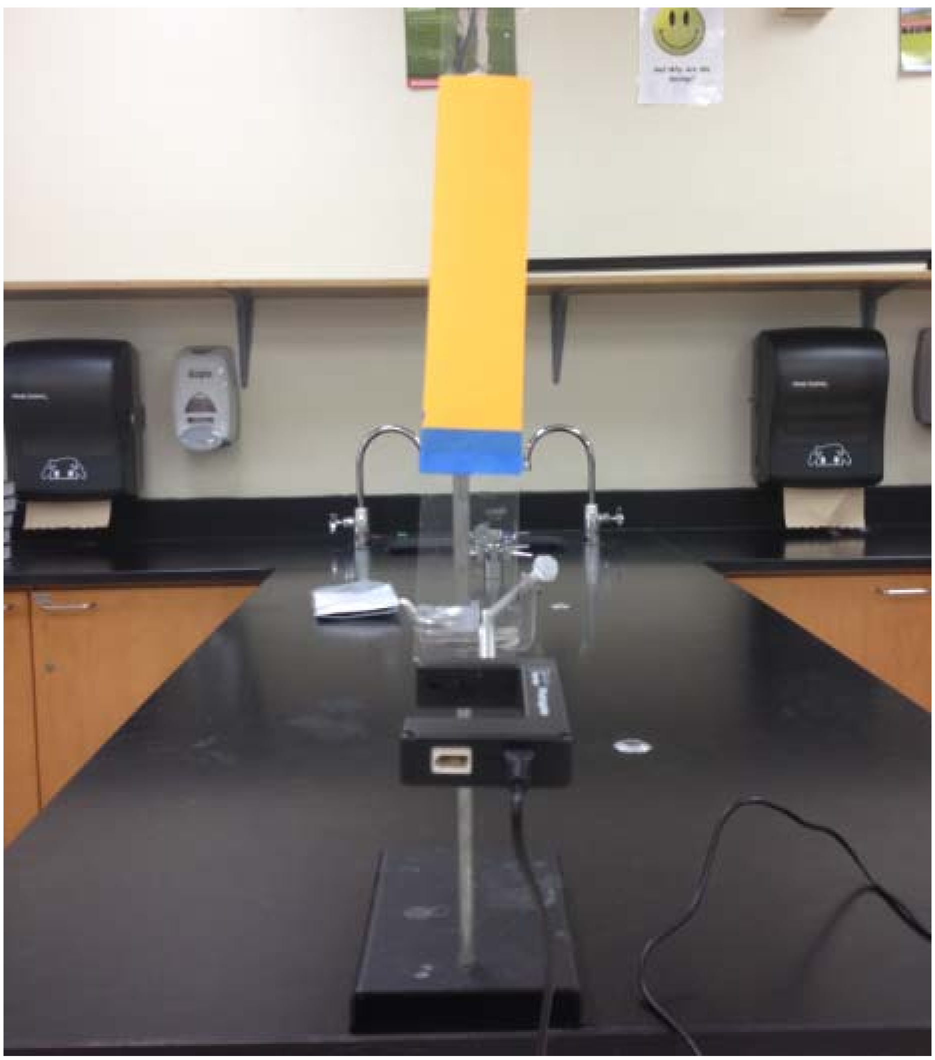
Velocity at first tape (show your work and proper units) (5 points)

Given Time - .100 s .110 .125 s .135 .150 s

Calculate the distance the tape and paper covers to produce the time given. Don't forget that the initial velocity of the plastic will not be zero when the tape covered section enters the photo gate! (5 points)

Success? _________________ /10 points

Total Grade _________________ / 20 points



Name:____

Angle of string to horizontal: ____________

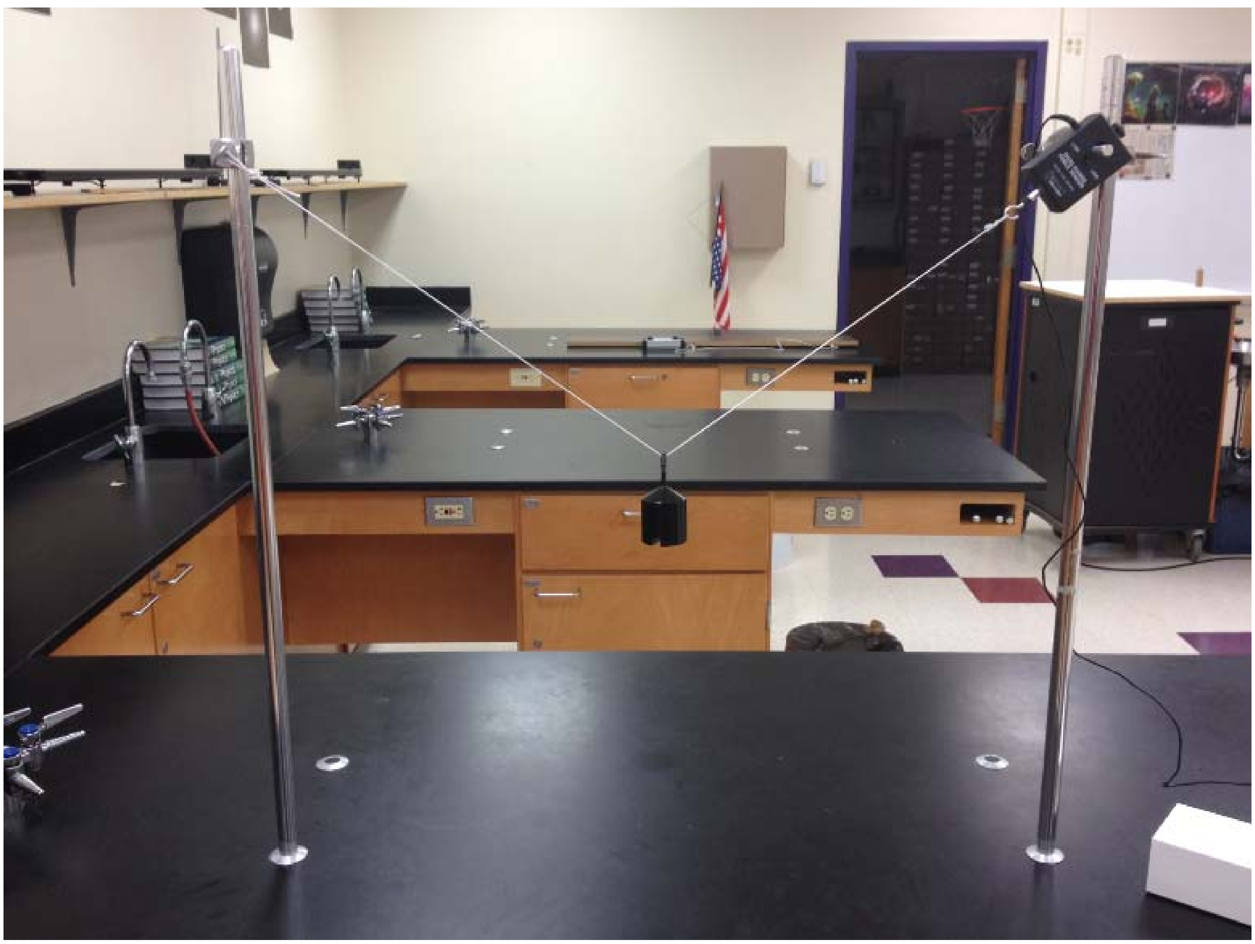
Given Tension Force - ____________

Show all work required to determine the mass needed to be hung from the string to produce the given force value in the force probe (10 points)

Hang the appropriate mass on the string to see if you did it correctly!

Success? _________________ /10 points

Total Grade _________________ / 20 points



Name: ______________________________

Unit 6 Lab Practical

Given Angle:	15^o	25^o	35^o	45^o
Given Velocity (m/s):	3.79	3.77	3.73	3.72

Given Height of Launch: 1.05 m

Show the work to find horizontal range to the projectile below. The velocity given is at the given angle.

Target Used for Projectile Lab Practical

4 cm – 100%
8 cm – 90 %
12 cm – 80 %
16 cm – 70 %

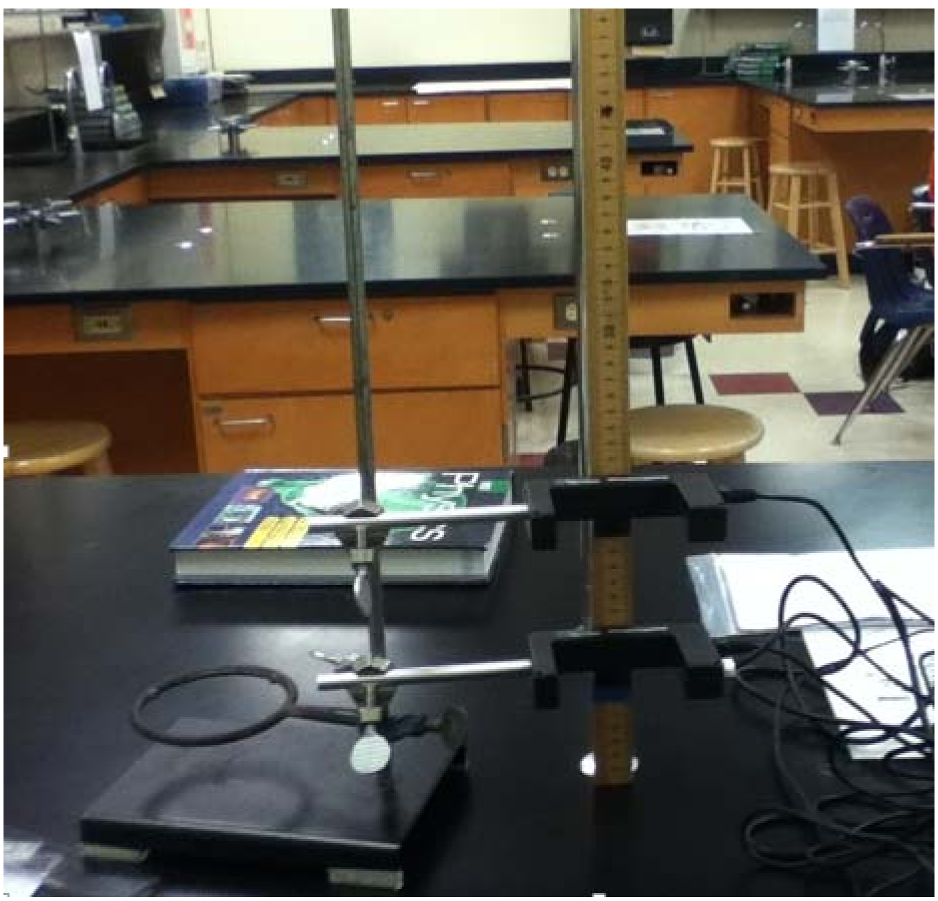
70 %
80 %
90 %
100%

Name: ______________________________

Unit 7 Lab Practical

Given Velocity (m/s): 1.98 2.42 2.80 3.13 3.43

Given the velocity above, determine the height from which a ball must be dropped in order to produce that velocity through the bottom photo gate. Use concepts involving energy (no kinematics) to solve (Your work will be graded). Note that the bottom photo gate is at a position of 10 cm.



Appendix B – Data

Test Group

Gender	Unit 2 /46	Unit 3 /48	Unit 4 /60	Unit 6 /44	Unit 7 /58
M	72.8	58.3	58.3	88.6	78.4
M	91.3	95.8	98.3	95.5	89.7
F	77.2	69.8	55.8	64.8	57.8
F	50.0	64.6	48.3	72.7	53.4
F	52.2	54.2	46.7	68.2	56.9
M	98.9	79.2	70.8	86.4	93.1
F	78.3	58.3	65.0	65.9	81.0
F	87.0	68.8	66.7	79.5	62.1
F	91.3	66.7	70.8	58.0	74.1
M	88.0	68.8	79.2	72.7	72.4
F	93.5	97.9	93.3	90.9	75.9
F	82.6	83.3	83.3	59.1	70.7
F	97.8	91.7	88.3	84.1	94.8
M	65.2	47.9	55.0	72.7	44.8
M	84.8	85.4	58.3	75.0	81.0
F	69.6	62.5	73.3	72.7	62.1
F	82.6	85.4	68.3	79.5	63.8
M	95.7	81.3	78.3	90.9	81.0
M	82.6	62.5	73.3	70.5	60.3
M	82.6	75.0	85.0	90.9	81.0
F	69.6	64.6	60.0	75.0	69.0
M	91.3	83.3	96.7	95.5	77.6
F	90.2	54.2	60.8	67.0	51.7
F	91.3	87.5	81.7	81.8	90.5
M	84.8	87.5	68.3	86.4	86.2
F	97.8	83.3	73.3	79.5	86.2
F	71.7	68.8	88.3	84.1	65.5
F	80.4	64.6	76.7	63.6	60.3
M	89.1	79.2	86.7	88.6	89.7
F	97.8	95.8	95.0	90.9	89.7
M	82.6	81.3	78.3	90.9	91.4
M	95.7	91.7	75.0	90.9	86.2

F	94.6	79.2	66.7	72.7	75.9
F	84.8	79.2	85.0	68.2	82.8
F	90.2	79.2	70.0	75.0	55.2
M	97.8	91.7	91.7	86.4	81.0
M	93.5	72.9	80.0	95.5	91.4
F	80.4	64.6	71.7	77.3	53.4
M	94.6	68.8	43.3	93.2	78.4
M	89.1	87.5	61.7	95.5	87.9
M	72.8	54.2	64.2	68.2	66.4
M	80.4	79.2	86.7	79.5	63.8
F	71.7	60.4	61.7	59.1	56.9
M	68.5	66.7	74.2	78.4	62.9
F	70.7	66.7	56.7	81.8	60.3
M	81.5	58.3	55.0	75.0	48.3
M	91.3	81.3	75.0	72.7	77.6
M	84.8	50.0	60.0	63.6	72.4
M	94.6	93.8	85.8	88.6	85.3
F	85.9	77.1	80.0	77.3	71.6
M	100.0	95.8	93.3	90.9	86.2
M	82.6	62.5	59.2	75.0	75.0
M	82.6	81.3	71.7	72.7	63.8
F	76.1	72.9	81.7	72.7	93.1
M	78.3	68.8	68.3	72.7	63.8
F	73.9	81.3	81.7	72.7	84.5
M	76.1	75.0	71.7	77.3	79.3
M	80.4	62.5	36.7	84.1	58.6
M	60.9	35.4	43.3	56.8	32.8
F	30.4	47.9	51.7	65.9	46.6
M	82.6	77.1	80.0	63.6	70.7
M	65.2	70.8	56.7	72.7	81.0
F	34.8	27.1	45.0	47.7	22.4
F	56.5	56.3	40.0	75.0	51.7
M	83.7	52.1	48.3	78.4	66.4
F	97.8	95.8	93.3	93.2	86.2
F	84.8	56.3	64.2	72.7	75.9
F	82.6	70.8	78.3	68.2	89.7
F	73.9	75.0	71.7	77.3	81.0
M	89.1	81.3	69.2	98.9	80.2

M	78.3	62.5	68.3	84.1	69.0
F	95.7	93.8	86.7	81.8	93.1
F	84.8	79.2	58.3	63.6	55.2
F	89.1	56.3	63.3	68.2	58.6
F	95.7	75.0	85.0	79.5	87.9
M	93.5	87.5	65.8	90.9	87.9
F	87.0	79.2	55.8	61.4	44.8
M	95.7	72.9	72.5	75.0	75.9
F	82.6	81.3	83.3	84.1	87.9
M	67.4	68.8	58.3	79.5	86.2
M	84.8	83.3	65.0	70.5	67.2
F	80.4	72.9	68.3	63.6	56.9
M	67.4	50.0	61.7	71.6	58.6
M	76.1	58.3	46.7	54.5	48.3
M	94.6	58.3	80.0	100.0	79.3
M	89.1	77.1	70.0	93.2	79.3
F	87.0	66.7	70.8	68.2	72.4
M	90.2	77.1	58.3	77.3	82.8
M	82.6	87.5	67.5	81.8	75.9
M	80.4	72.9	60.0	90.9	75.9
M	65.2	60.4	63.3	65.9	55.2
M	93.5	85.4	83.3	90.9	74.1
M	90.2	81.3	77.5	81.8	57.8
F	43.5	31.3	48.3	65.9	25.9
Mean %	81.7	72.1	69.7	77.4	71.2
SD	13.5	14.6	14.0	11.2	15.5

Control Group

Gender	Unit 2 – /46	Unit 3 – /48	Unit - 4 /60	Unit 6 - /44	Unit 7 - /58
M	89.1	83.3	78.3	84.1	87.9
M	95.7	87.5	136.7	100.0	93.1
M	95.7	93.8	90.0	79.5	82.8
F	95.7	91.7	95.0	86.4	82.8
M	84.8	89.6	85.0	97.7	93.1
M	93.5	83.3	88.3	86.4	94.8
M	89.1	75.0	93.3	97.7	84.5
M	87.0	85.4	81.7	97.7	91.4
F	100.0	89.6	80.0	90.9	93.1
M	82.6	56.3	60.0	68.2	77.6
M	97.8	97.9	93.3	86.4	91.4
F	89.1	77.1	83.3	97.7	87.9
M	97.8	85.4	93.3	97.7	94.8
F	45.7	60.4	46.7	59.1	25.9
M	76.1	62.5	51.7	84.1	53.4
M	67.4	41.7	51.7	61.4	36.2
F	91.3	91.7	91.7	81.8	86.2
M	76.1	50.0	40.0	63.6	53.4
F	93.5	91.7	100.0	86.4	94.8
M	58.7	58.3	71.7	72.7	48.3
F	89.1	91.7	75.0	93.2	70.7
F	73.9	75.0	53.3	68.2	67.2
F	43.5	72.9	66.7	77.3	91.4
M	91.3	56.3	45.0	88.6	75.9
F	87.0	85.4	68.3	86.4	93.1
F	91.3	85.4	95.0	84.1	84.5
F	69.6	64.6	80.0	90.9	81.0
M	78.3	83.3	71.7	75.0	62.1
F	87.0	79.2	100.0	95.5	96.6
M	100.0	81.3	91.7	100.0	79.3
F	82.6	66.7	75.0	70.5	84.5
M	91.3	91.7	53.3	90.9	94.8
F	84.8	79.2	98.3	79.5	89.7
F	39.1	47.9	11.7	15.9	25.9

F	56.5	41.7	53.3	40.9	37.9
M	82.6	68.8	75.0	70.5	72.4
M	93.5	81.3	90.0	97.7	87.9
F	80.4	91.7	93.3	75.0	82.8
F	95.7	81.3	75.0	93.2	82.8
F	82.6	68.8	76.7	77.3	82.8
F	93.5	91.7	98.3	100.0	89.7
F	93.5	83.3	85.0	81.8	96.6
M	73.9	70.8	48.3	88.6	62.1
F	82.6	60.4	75.0	79.5	60.3
F	71.7	77.1	91.7	81.8	84.5
F	58.7	64.6	83.3	63.6	63.8
M	93.5	83.3	81.7	97.7	84.5
M	95.7	97.9	88.3	93.2	87.9
F	97.8	85.4	80.0	86.4	82.8
F	87.0	81.3	65.0	54.5	77.6
F	76.1	83.3	71.7	86.4	81.0
M	82.6	70.8	68.3	50.0	72.4
F	91.3	79.2	75.0	50.0	87.9
M	91.3	81.3	81.7	84.1	82.8
M	71.7	89.6	53.3	75.0	63.8
M	87.0	81.3	86.7	97.7	94.8
F	93.5	81.3	88.3	81.8	89.7
F	93.5	81.3	65.0	84.1	89.7
M	84.8	77.1	66.7	79.5	91.4
M	93.5	91.7	88.3	86.4	96.6
M	84.8	77.1	76.7	79.5	81.0
F	80.4	87.5	61.7	81.8	87.9
M	60.9	87.5	56.7	77.3	63.8
M	100.0	91.7	98.3	97.7	94.8
F	80.4	70.8	83.3	79.5	82.8
F	76.1	83.3	68.3	77.3	84.5
F	95.7	75.0	83.3	81.8	77.6
F	91.3	85.4	80.0	79.5	93.1
M	100.0	79.2	90.0	88.6	86.2
M	95.7	75.0	95.0	79.5	93.1
M	97.8	93.8	98.3	100.0	96.6
M	97.8	97.9	86.7	100.0	96.6
M	84.8	72.9	76.7	88.6	63.8
M	60.9	37.5	71.7	70.5	74.1

M	84.8	66.7	61.7	84.1	65.5
F	93.5	70.8	88.3	70.5	89.7
M	43.5	60.4	15.0	68.2	31.0
F	95.7	56.3	98.3	77.3	96.6
M	87.0	91.7	65.0	95.5	91.4
F	76.1	70.8	68.3	56.8	82.8
M	87.0	62.5	83.3	90.9	69.0
M	76.1	77.1	83.3	79.5	74.1
F	82.6	79.2	78.3	90.9	86.2
M	80.4	47.9	45.0	88.6	65.5
F	73.9	50.0	65.0	65.9	48.3
F	76.1	70.8	66.7	65.9	84.5
M	80.4	75.0	75.0	77.3	81.0
F	80.4	58.3	53.3	70.5	46.6
M	93.5	93.8	81.7	95.5	89.7
M	87.0	81.3	61.7	59.1	60.3
M	97.8	89.6	90.0	77.3	91.4
M	97.8	89.6	70.0	81.8	89.7
M	95.7	89.6	76.7	81.8	82.8
F	97.8	93.8	78.3	84.1	91.4
F	71.7	77.1	90.0	77.3	58.6
M	71.7	62.5	55.0	56.8	79.3
F	80.4	81.3	78.3	81.8	79.3
F	97.8	95.8	96.7	86.4	96.6
F	80.4	85.4	76.7	75.0	69.0
F	84.8	77.1	80.0	81.8	84.5
M	91.3	81.3	76.7	86.4	75.9
M	91.3	93.8	78.3	84.1	86.2
F	95.7	91.7	91.7	86.4	98.3
M	89.1	56.3	80.0	79.5	60.3
M	65.2	58.3	38.3	59.1	53.4
F	67.4	66.7	83.3	59.1	58.6
M	89.1	95.8	90.0	84.1	91.4
F	95.7	64.6	88.3	100.0	82.8
F	93.5	77.1	85.0	90.9	72.4
M	95.7	85.4	83.3	93.2	84.5
F	84.8	81.3	60.0	86.4	81.0
F	93.5	85.4	86.7	93.2	87.9
M	89.1	85.4	98.3	84.1	93.1
M	73.9	83.3	71.7	72.7	44.8

F	93.5	56.3	51.7	50.0	58.6
mean %	84.4	77.3	76.3	80.6	78.5
SD	12.9	13.7	17.9	14.1	16.5