

April 2014

Physics Education Research: The effects of context-rich problem solving in groups in introductory electricity and magnetism courses

Jack Richard O'Brien
Worcester Polytechnic Institute

Follow this and additional works at: <https://digitalcommons.wpi.edu/mqp-all>

Repository Citation

O'Brien, J. R. (2014). *Physics Education Research: The effects of context-rich problem solving in groups in introductory electricity and magnetism courses*. Retrieved from <https://digitalcommons.wpi.edu/mqp-all/3533>

This Unrestricted is brought to you for free and open access by the Major Qualifying Projects at Digital WPI. It has been accepted for inclusion in Major Qualifying Projects (All Years) by an authorized administrator of Digital WPI. For more information, please contact digitalwpi@wpi.edu.

Physics Education Research

The effects of context-rich problem solving in groups in
introductory electricity and magnetism courses

A Major Qualifying Project report
submitted to the Faculty
of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the
Degree of Bachelor of Science.

Submitted by:

Jack O'Brien

Submitted to:

Germano Iannacchione

April 9, 2014

Abstract

This project investigated whether or not interactive-engagement through context-rich problem solving in groups can augment the comprehension of concepts in introductory electricity and magnetism courses. The study took place during the B13 occurrence of the physics course Principles of Physics: Electricity & Magnetism (PH 1121). Students in four sections of the course were studied.

Half the students were given context-rich problems to solve in groups, and half were given traditional textbook style problems that covered the same topics. These sessions occurred between the second and final exams. Following the sessions, students' second and final exam grades were compared, and the gains (positive or negative) between the two exams were recorded, and averaged.

The complete average gain for this project was $\langle \bar{G} \rangle = 4.26 \pm 14.5$, which indicates that, on average, students exposed to context-rich problem solving in groups improved by five more points on their final exams than students that were not. This suggests that context-rich problem solving is helpful to students but, due to the large standard deviation, is not statistically conclusive. In addition to this general result, several subgroups were analyzed (declared major, year of graduation, and gender). It was found that while context-rich problem solving is useful overall, some groups may benefit more than others, and some groups may in fact be hindered by exposure to these methods.

This educational method appears to be, on average, beneficial to students learning concepts in PH 1121. These results, however, must be tempered with the fact that while the *average* student may benefit from this technique, there are certain *individual* students that may not. Ergo, educational techniques must be adapted to students' particular learning styles to maximize student success.

Acknowledgements

Completing an MQP is a colossal undertaking, regardless of the size of one's group. I would like to thank those, without whom, this project would not have succeeded: Professor Qi Wen, who so graciously welcomed my project and so readily volunteered assistance, as well as Klaida Kashuri and Cai Waegell, both of whom were kind enough to assist with the conference sessions and offer feedback on the problems themselves.

I would be remiss if I did not devote an entire paragraph to my project advisor, Professor Germano Iannacchione. Without his constant guidance and encouragement this project would not have seen the light of day. There were a few bumps along the way (both small and large) and without his support I would not have been able to complete my research. To him I extend my eternal gratitude.

To all the students who participated, to Jackie Malone who made sure that Professor Iannacchione and I could meet, and to those I have not included here, please accept my heartfelt thanks and warmest regards.

-Jack

Table of Contents

Abstract	2
Acknowledgements	3
Table of Contents	4
Table of Figures	6
Introduction	7
Physics education research	7
<i>History</i>	7
<i>Interactive-engagement</i>	8
<i>Context-rich problem solving in groups</i>	8
<i>Future research</i>	9
Methodology overview	10
Results	11
Structure of report	12
Literature Review	13
A brief history of education research.....	13
<i>Humble beginnings</i>	13
<i>World War I and beyond</i>	14
<i>Research in the modern era</i>	14
Current state of physics education research.....	15
<i>Necessity of physics education research</i>	15
<i>First year courses and interactive-engagement</i>	16
<i>Problem-solving pathways in context-rich problems</i>	17
<i>Solving context-rich problems in groups</i>	19
Goals of future research	20
Methodology	22
Hypothesis and experiment overview	22
Part 1: Crafting the problems	22
<i>How to craft a context-rich problem</i>	22
<i>Topic selection</i>	25
Part 2: Group structure and session flow.....	26
Part 3: Analysis and demographic data.....	28
Implementation	29
Results	31
Overview of results	31
Explanation of data analysis	32
Overall results	33
Gender subgroup.....	35
Class year subgroup.....	36
Declared major subgroup.....	38
Statistical conclusions.....	40
Analysis	41

Sources of error	41
Overall results	43
Gender subgroup.....	44
Class year subgroup.....	47
Declared major subgroup.....	48
Conclusions	51
Results of the experiment	51
Future research goals.....	52
<i>Corroboration</i>	52
<i>Other variables to explore</i>	52
Time of day	53
Subgroups.....	53
Group structure.....	53
Problem presentation.....	54
Instructors	54
Particular topics in the curriculum.....	54
Problem type variance	55
Other physics courses.....	55
Final comments	56
Works Cited	57
Appendix A – Conference problem sets	59
Conference Session 1 – Experimental problem	59
Conference Session 1 – Control problem.....	59
Conference Session 2 – Experimental problem	60
Conference Session 2 – Experimental problem	60
Appendix B – Survey	61
Appendix C – Table of results	62

Table of Figures

Fig. 1 - An example control problem statement. This is what one might find in a textbook.	23
Fig. 2 - A context-rich problem encourages critical thinking.....	24
Figure 3 - A bar graph of the individual average gain for all students.	34
Figure 4 - A bar graph displaying the complete average gain for the gender subgroup.....	36
Figure 5 - A bar graph of the complete average gain by class year.....	37
Figure 6 - A bar graph of the complete average gain by declared major.....	39
Figure 7 - Histogram displaying the average individual gain of the control group versus the individual average gain of the experimental group.....	44
Figure 8 - A histogram comparing the male control average individual gain against the experimental male average individual gain.....	46

Introduction

The purpose of this project was to investigate whether or not interactive-engagement through context-rich problem solving in groups can augment the comprehension of difficult concepts in introductory electricity and magnetism courses. This particular line of inquiry was arrived upon after a thorough review of both historical data in educational research as well as current research initiatives. This project was carried out during the B13 iteration of Professor Qi Wen's PH 1121 (Principles of Physics: Electricity & Magnetism) course, a calculus-based introductory physics course that covers fundamental electricity and magnetism concepts. Students in several conference sections were studied, and the benefits of context-rich problem solving were analyzed.

Physics education research

History

Education research in the early 20th century was heavily influenced by earlier efforts to systematically document and study teaching methods (Comp). With the appointment, in 1909, of Charles Judd as the chair of the School of Education at the University of Chicago, a new era of rigorous education research began.

Philosophical theory was limited in the field, and educators began to look to performance evaluations (standardized tests, etc.) for insight into the effectiveness of their teaching methods.

During the years following World War II, several circles of thought began to emerge as the field matured. Constructivism (learning through experience), functionalism

(a focus on the functional social benefits of education), and postmodernism (contextual learning based on culture, as opposed to universal truths) challenged previously dominant ideologies, such as positivism (focus on facts derived only from sensory experience and experimentation). This competition, spurred on by the creation of academic journals, fostered a thriving research atmosphere. One such academic journal, the American Journal of Physics (published by the American Association of Physics Teachers and the American Institute of Physics) contains a plethora of Physics Education Research and is an invaluable resource.

Interactive-engagement

During the modern era, more attention is being given to “action research”; a kind of applied research intended to solve an immediate problem. These studies have provided interesting data, particularly with regards to “interactive-engagement” in first year students. Evidently, when using the traditional lecture method of teaching “the basic knowledge gain... is essentially independent of the professor” (Halloun and Hestenes 1043-1055). Conversely, courses which “made substantial use of [interactive-engagement] methods achieved an average gain [of] almost two standard deviations” above courses taught in the traditional lecture style (Harke 64-74).

Context-rich problem solving in groups

In addition to the methods of teaching being employed, recent studies have investigated the type of practice problems given to students. A study conducted by Antonenko et al. discovered that students have become accustomed to well-structured problems, and will struggle when confronted with multi-step, complex problems. When faced with these “real-world” scenarios, students often searched

for an “algorithm that might work (‘plug-and-chug’), immediately ask for direct help, or flounder along doing considerable busy-work with no real planning or direction” (Antonenko et al. 323-342). This is somewhat alarming, and is in stark contrast to how an expert problem solver may approach a similar problem.

This study (Antonenko et al. 323-342) showed that “context-rich” problems increased problem-solving skill, and shifted the problem solving pathways of students to align more closely with that of an expert problem solver. Context-rich problems are designed specifically to disabuse students of the notion that “plug-and-chug” methods are the only effective means of solving problems. Research by Antonenko et al. found that student groups exposed to context-rich problems decreased their length of time to complete a problem by almost 20% over the course of a semester.

It has also been shown that group exercises are beneficial to each individual student of the group, and thus can be effective means of introducing students to context-rich problem solving (Heller, Keith, and Anderson 1992). In the course of this study, it was found that all students exposed to context-rich problem solving in groups that successfully solved a particular problem on an exam did so by using force diagrams, whereas only 57% of students in control sections (not exposed to context-rich problem solving in groups) that solved the problem used force diagrams. Ergo, students exposed to context-rich problem solving in groups exhibited problem solving that much more closely aligned with that of an expert.

Future research

Up to this point, much research has been geared towards introductory mechanics courses. This project hopes to apply these methods, proven successful at conveying mechanics concepts, to an introductory electricity and magnetism course to determine if context-rich problem solving is an effective technique for these

concepts as well. Electricity and magnetism concepts can be particularly difficult for students to grasp, as they likely have little opportunity in their lives to develop an intuitive sense of, for example, acceleration due to an electric field, whereas most individuals will have a strong sense of the acceleration due to gravity. If these methods can be proven beneficial to students taking introductory electricity and magnetism courses, instructors can begin using these methods to augment their teaching styles.

Methodology overview

This project was carried out during the conference sessions of the aforementioned PH 1121 course. First, through coordination with Professor Qi Wen, a lesson plan was obtained and analyzed. Then, conference sessions were selected such that they occurred in between the second and third exams. Next, problems were written by this experimenter to test the students' knowledge and reinforce recently learned concepts from lecture. Context-rich problems were distributed to half of the conference sections (the experimental sections), and traditional, textbook style problems were given to the other half (the control sections). The problems used are listed in Appendix A. Both the control and experimental sections were broken into groups of three (when possible) and given 10-15 minutes to complete the problem. The papers were then graded, and returned.

Success of the experiment was determined by comparing the gain (positive or negative) from the second to the third exam. These gains were averaged (for control and experimental), and then subtracted (control from experimental). This number is the "complete average gain" (denoted here as $\langle \bar{G} \rangle$) for that particular subgroup, and is a measure of the average difference in improvement from the second exam to the third exam between the experimental and control groups. A

positive number indicates that the experimental groups improved more than the control groups, while a negative number indicates the opposite.

In addition to this overall result, subgroups were analyzed. Students were surveyed at the culmination of the course to gain demographic data (class year, declared major, and gender). The exam grades for students that fell into particular subgroups were analyzed in the same fashion as the overall exam grades and $\langle \bar{G} \rangle$ was calculated for these various subgroups as well.

Results

The results of the project suggest that students, overall, benefit from context-rich problem solving in groups. The complete average gain for the experiment was $\langle \bar{G} \rangle = 4.26 \pm 14.5$, which means that, on average, students in the exposed to context-rich problem solving improved by four more points on their exams than students in control groups. Due to the large standard deviation, this result cannot be considered conclusive, however.

In addition to the overall result, several subgroups were also analyzed. Some groups performed quite well. Particularly, males earned a complete average gain of $\langle \bar{G} \rangle = 13.1 \pm 13.0$, AE majors earned a complete average gain of $\langle \bar{G} \rangle = 17.7 \pm 12.1$, and ECE majors earned a complete average gain of $\langle \bar{G} \rangle = 26.3 \pm 15.1$. Other groups suffered when exposed to context-rich problem solving in groups. For example, BME majors had the largest negative complete average gain with $\langle \bar{G} \rangle = -15.0 \pm 8.66$, and females had the second largest negative complete average gain ($\langle \bar{G} \rangle = -10.4 \pm 16.0$).

Evidently, while context-rich problem solving in groups is beneficial overall, some groups of students may benefit from a different approach. Thus, a major conclusion to draw from this project is that educational techniques may need to be adapted to particular learning styles in order to maximize learning.

Structure of report

This report is divided into five major sections, and several appendices. First, a brief literature review will introduce vocabulary, concepts, and research initiatives that lead to the development of this project. Next, the methodology of the project is described in detail. This section will cover the execution of the experiment as well as the development of the context-rich problems. Following this, the results themselves are conveyed both in full and across the experimental subgroups. Then, an analysis of these results is presented. This analysis will offer insight into the results, and discuss possible sources of error in the experiment. The penultimate section contains conclusions drawn and final comments with regards to the experiment are offered, including suggestions for future projects and research. Finally, works cited for this project are listed, and various appendices containing pieces, which do not fit elsewhere in the report, are listed.

Literature Review

A brief history of education research

Humble beginnings

The purpose of this project was to investigate whether or not interactive-engagement through context-rich problem solving experiences in groups can augment the comprehension of difficult concepts in introductory electricity and magnetism courses. Education research in general, however, is a much broader and varied field. From its humble beginnings in the 1830s, educational research has changed dramatically in both its intended goals, and research techniques (Comp).

With the revival of the “common school” in the late 1830s and early 1840s, curricula and teaching methods were, for the first time, documented, studied, and were influenced by systematic experimentation and data compilation (Travers 7). These new applications were, according to Travers, “an examination of the ideas on which education was based, an intellectual crystallization of the function of education in a democracy, and the development of a literature on education that attempted to make available to teachers and educators important new ideas related to education that had emerged in various countries.” Research tendencies during this time were largely philosophical, choosing to focus more on logical argument than on scientifically obtained data. However, in 1909 when Charles Judd became the Chair of the School of Education at the University of Chicago, he ushered in a new, rigorous, era of education research. Judd was a major proponent of the scientific method, and pushed to apply it to the understanding of education, thus limiting the use of philosophical theory in the field. His work helped form and shape the field of educational psychology (“Charles Hubbard Judd”), and was crucial to defining future research.

World War I and beyond

By 1915, education research was flourishing (Comp). It had attracted the attention of major schools in the United States, such as Harvard and the University of Chicago. The United States was growing at a rapid pace due to a large influx of immigrants, the arrival of which coincided with the testing movement that had emerged during World War I when the United States Army began issuing standardized tests to its recruits. Educators called upon education researchers to assist in the development and oversight of curricula to accommodate a wide range of racially and socially diverse students. As a result of this need, the growth in schools of education, and the appearance of academic journals with a focus on the topic, education research continued to mature in the years after World War II ended.

In the years following World War II, the positivist movement that had been largely dominant since the early 1900s began to be challenged by newer factions of thought such as constructivism (focused on learning through experience), functionalism (focused on the functional social benefits of education), and postmodernism (focused on contextual learning based on culture, as opposed to universal truths) (Pring 90). While the positivist movement has its base almost entirely within the confines of the scientific method (by utilizing experimental result as the only true rhetoric), these other movements had their bases in social science, philosophy, and psychology. Their inception offered criticism of positivism, and debates as to which method stands supreme continue today (Comp).

Research in the modern era

During the 1950s and 1960s, the United States government became more heavily involved in the regulation of education research. The government imposed a new

“evidence-based” method, as institutions strove to provide proof that funding allocations for education research were justified. According to Travers, since the 1970s, “virtually every bill authorizing particular educational programs has included a requirement that the particular program be evaluated to determine whether the program was worth the money spent upon it.” (Travers 539). As such, a shift in focus from “pure research” to “applied research” occurred largely due to the potential profitability and utilitarianism of the latter (Greenwood, and Levin 143-44). Ultimately, in recent years, education research has taken the form of “action research”.

The objective of action research is not to provide new knowledge, but to solve an immediate concern or apparent problem within a specific school of thought. The goal of any action research is to progressively improve how an institution addresses and solves problems (“Action Research”). This is a type of research that is thriving in present day education research, and the type that was conducted during this project.

Current state of physics education research

Necessity of physics education research

Physics education research (PER) is a sub-genre of education research, and shares many goals and methods with the field at large. The focus of PER, however, is much more direct and is inherently action research-oriented. Because so many elementary concepts in physics are abstract and/or counter-intuitive, the lecture method often fails to correct many misconceptions that students might have developed prior to having had any formal education in the subject (“Physics education research”).

Teaching unintuitive concepts through analogy, as is often the case in the lecture method, can lead to dangerous confusion, and can actually reinforce misconceptions as opposed to correcting them. A prominent example of this is Newton's First Law of inertia. A careless analogy may not completely express the subtle notion that friction is a force that acts on an object in motion, and students may retain the Aristotelian misconception that an object in motion requires a constant "push" to remain in motion. This subtle fallacy is one of any number of examples of the need for PER, particularly in modifying and improving curricula and teaching methods in introductory courses.

First year courses and interactive-engagement

A great deal of effort has been applied to improving introductory physics courses. In 1985, Halloun and Hestenes concluded a large study of mechanics concepts in introductory physics courses. Through the use of pre- and post-testing, they determined that "the basic knowledge gain [of a student] under conventional instruction is essentially independent of the professor" (Halloun and Hestenes 1043-1055). Thus, regardless of the instructor, conventional teaching methods (classical lecture method) are equally ineffective at imparting a comprehension of fundamental physics concepts to first year students.

New research suggests that, in sharp contrast to the results of the conventional method of teaching, methods favoring "interactive-engagement" (IE) methods are actually quite successful in augmenting conceptual understanding. IE learning involves the use of active problem solving, and encourages the application of concepts discussed in lectures. Courses which "made substantial use of IE methods achieved an average gain [of] almost two standard deviations above that of the traditional courses" (Harke 64-74), indicating that there is absolutely room for improvement over traditional instructional methods in introductory courses.

These results are promising. Harke cautions, however, that of all the courses studied, none were in a region deemed to be “High-*g*” and thus all were insufficiently instructing students. Also, “67% of [IE] courses were taught at least in part by individuals who had devoted considerable attention to PER” (Harke 64-74). This result helps to exemplify the pressing need to advance and study PER, and reminds that IE methods are rendered inadequate when used by instructors without a complete understanding of their means of implementation. The primary conclusion from Harke’s research is that IE courses have significant improvements over traditional methods, but counsels that there is space to improve upon IE’s instructional methods, and a more widespread adoption of PER among instructors is needed.

Problem-solving pathways in context-rich problems

A logical extrapolation of IE methods is to supplement traditional lecturing with problem-solving exercises to apply concepts covered in lectures. This method requires that careful consideration must be given to the style of problem that students are given, and can often illustrate weak areas in understanding that require improvement. Students that have become accustomed to well-structured problems will often struggle when confronted with a multi-step, complex, “real-world” problem, and will likely search for an “algorithm that might work (‘plug-and-chug’), immediately ask for direct help, or flounder along doing considerable busy-work with no real planning or direction” (Antonenko et al. 323-342).

Antonenko et al. studied the difference between the ways in which “novices” and “experts” approached an ill-structured problem. Experts tend to have extensive knowledge in the domain, and great confidence solving complex problems. This allows them to qualitatively analyze a problem, and avoid distractions in what might be a complex presentation. Experts will also understand when their chosen pathway fails to remain productive, and the importance of testing the validity of

their solutions. This is in sharp contrast to the novice problem solver, who will often resort to a somewhat mechanical solving style, likely involving equations recently learned in lecture.

One can also conclude that:

“...strong problem-solvers...recognize the benefits of incorporating experiences gained from each problem into their knowledge structure that can be useful to draw on when confronted with new ill-structured problems” (Antonenko et al. 2010).

This emphasizes the importance of introducing students to complex problems at an initial stage, so that they can begin to develop methods with which to solve these problems as early as possible.

“Context-rich” problems have been shown to increase problem-solving skill and, with repeated exposure over time, shift the pathways of a novice to align more closely with that of an expert. A context-rich problem, as compared against a “standard” or “text-book” problem, encourages critical thinking on the part of the solver. These problems are typically more conversational in tone, and motivate students to find a solution due by exploring the relationship of physical principles to real-world scenarios. Context-rich problems are specifically crafted to disabuse students of the notion that “plug-and-chug” methods are the only effective means of solving problems. They may appear insurmountable at the outset, but are imminently solvable when first approached with the qualitative methods of an expert problem-solver, and become simpler over time as students gain the proper problem solving mindset.

Research by Antonenko et al. found that student groups exposed to context-rich problems throughout the course of a semester completed qualitative analysis of the problem, a crucial weapon in the expert problem solver’s arsenal, 20% more quickly during the last problem of the semester than during the first problem of the

semester. This result confirms that students, when given the opportunity, will begin to utilize the problem-solving pathways of an expert and do in fact benefit from context-rich problem solving.

Solving context-rich problems in groups

Context-rich problems have been shown to increase students' comprehension of fundamental physical concepts in introductory courses. The question of how to deliver these problems, however, still remains. Two methods are logically possible, presenting each individual student with a problem to be completed during lecture (or for homework), or utilizing a recitation session to distribute the problems to a group of students. Heller, Keith, and Anderson (1992) studied the possible benefits of group work compared against working individually on problem sets of similar objective difficulty, covering the same topics. Problem solving in groups that are "effective" will invariably yield higher scores on individual assignments as there are simply more resources available to the individual (his or her partners), "effective" here of course refers to a group of motivated students focused on solving a problem.

A major difficulty of group work is balancing groups effectively so that weaker students receive benefits of assistance, but not to the detriment of stronger students. Group work is only worthwhile if *all* students can benefit. It has been shown that this in fact the case, as the "highest-ability students improved at approximately the same rate as the other students" (Heller, Keith, and Anderson 1992) when solving context-rich problems in groups. Additionally, the groups consistently scored better on problem sets than the highest-performing students that were working independently.

It was found that stronger students often provided leadership during problem solving, while weaker students would help to check and monitor the work to ensure

steps were not being missed. Thus, stronger students that might normally skip steps, or perform work in their head, can benefit from learning to introduce these more disciplined behaviors into their methods.

Problem solving cannot always be done in groups. During an exam, for example, students typically work independently. Ergo, successful group work will prepare students to solve problems by augmenting not only the group's performance, but *each* member of the group's performance as well. Interestingly, in their study, Heller, Keith, and Anderson discovered that individual students benefitted greatly from solving context-rich problems in groups.

On an exam, all students in the experimental sections of a course (the sections exposed to context-rich problem solving in groups) who successfully solved a particular problem did so by using force diagrams, whereas only 57% of students in control sections (not exposed to context-rich problem solving in groups) drew force diagrams. Also, the students in the experimental sections scored significantly higher than the students in the control section. Most notably, however, was the revelation that students exposed to context-rich problem solving in groups displayed methodology more in line with that of an expert problem solver than those students in control sections.

Goals of future research

It has been shown quite conclusively that interactive-engagement through context-rich problem solving in groups is markedly beneficial for students in introductory physics courses. These kinds of activities aid in the comprehension of counter-intuitive physical principles, strengthen physical intuition and problem-solving skills, and help to repair damage caused by pre-conceived misconceptions. The bulk of the research presented here has, however, been geared towards mechanics-based introductory courses, and there appears to be a lack in focus on introductory

electricity and magnetism courses. Presumably this is due to the broader requirements of mechanics knowledge across all engineering disciplines, but this experimenter can only speculate due to a lack of resources on the subject.

Electricity and magnetism courses, as opposed to mechanics courses, present a unique challenge due to the abstract nature of the concepts being studied. From an extremely young age, most students will have experiential knowledge of mechanical principles (a ball colliding with a baseball bat, for example, or an innate sense of the nature of gravity). However, this experience can cause deeply rooted misconceptions to take hold, misconceptions which have been proven repairable through interactive-engagement.

In contrast, instructors teaching introductory electricity and magnetism courses are able to work with a somewhat cleaner slate. It is quite unlikely, for example, that students will have interacted in any memorable way with an electron that will help develop an intuition about their physical principles. Students may have interacted with magnets and perhaps formed simple circuits, but this work can only be done at a very peripheral level students are exposed to higher level physics and mathematics concepts. This is both a boon and a hindrance, as students may have difficulty grasping brand new concepts with which they have no prior real-world experience.

It can be challenging to impart fundamental principles of electromagnetism to students that may not have a strong calculus background, and it is difficult for students to learn these concepts experientially as many principles require laboratory conditions to exemplify (for example, the effects of an electric field on an accelerating proton). To repair this disconnect, and determine the best way to convey these concepts to students, action-research PER is required. The goal of this project is to use PER to determine whether context-rich problem solving in groups is beneficial to students in an introductory electricity and magnetism course.

Methodology

Hypothesis and experiment overview

A context-rich problem is specially crafted to encourage critical thinking. The purpose of this project is to determine whether or not the act of solving these problems will help students to understand the fundamental concepts of electricity and magnetism in an introductory physics course. My hypothesis is that students who are exposed to context-rich problem solving in groups will show strong improvement on exams than students who are not.

This hypothesis was tested in Professor Qi Wen's PH 1121 (Principles of Physics – Electricity and Magnetism, B term 2013) course, a calculus-based introductory physics course that covers the fundamentals of electricity and magnetism. Over the course of the term, students were given problems to solve on three separate occasions (in between the occurrences of the second and third exams) in their conference. Following this, students were given a survey to collect demographic data. After the conclusion of the term, grades were compared from the second and third exams (the third exam happened to be the course's final exam) to analyze gains due to context-rich problem solving in groups.

Part 1: Crafting the problems

How to craft a context-rich problem

Context-rich problems are particularly powerful in that they encourage critical thinking. These kinds of problems force a student to consider the nature of the

problem contextually, and then apply physical principles, ultimately determining a mathematical framework in which to work and solve the problem. A well-crafted context-rich problem will have a somewhat obscured objective that will reveal itself upon consideration by the problem solver.

To create a context-rich problem, first let us consider a standard problem statement:

The nucleus of a tin atom in a vacuum has a charge of $+50e$. (1) What is the potential energy on a proton $1 \cdot 10^{-9}$ meters away from the tin nucleus? (2) What is the potential energy if the proton is now 0.3048 meters away from the tin nucleus? (3) If the proton was released from rest at an initial distance of $1 \cdot 10^{-9}$ meters from the tin nucleus, what is the velocity of the proton when it is 0.3048 meters away from the tin nucleus? (4) What is the proton's velocity relative to the speed of light?

Fig. 1 - An example control problem statement. This is what one might find in a textbook.

This is very similar to a problem one might find in a college textbook. It is presented in a step-by-step fashion, guiding the student through its solution. Also, physical jargon is used (“potential energy”, “velocity”, etc.) extensively. Finally, the problem is presented completely without context, making it entirely separate from the student's experience. This decreases motivation to solve the problem.

A student reading this would have a very simple time reading a problem such as the problem presented in Fig. 1 and quickly determining the correct equation to use from the phrasing of the problem, without every really having to consider the physics involved. While practice in solving simple problems and becoming comfortable with the equations is certainly useful, to truly impart a concept a student must become accustomed to first applying physical principles, and then deciding which is the proper set of equations to work with based on the problem at hand. Thus, the standard problem statement above must be converted into a context-rich problem.

The problem statement must be structured in such a way that encourages a deeper level of thought than the standard statement. Context is crucial. Students will be encouraged to solve a problem if they knew why this particular problem *matters*, and how the physics involved relates to their own lives. This is difficult, and not always obvious, particularly in the realm of electricity in magnetism. However, every effort must be made to add context to a problem. Presented now is a context-rich version of the problem from Fig. 1:

After lunch, Jack noticed that his can of breath mints was made of tin. He believes that this tin can could get a proton up to near light speed by the time it was a foot away from the tin. If Jack held a proton, completely stationary, one nanometer away from a tin nucleus in a vacuum (no electrons are present), and then released the proton, is it possible for the proton to achieve light speed?

Hint: A tin nucleus has a charge of $+50e$. The tin nucleus is completely stationary.

Fig. 2 - A context-rich problem encourages critical thinking.

The version of the problem in Fig. 2 puts the physics in the context of answering a question; Jack is curious to know if a certain thing is possible. Simply rephrasing the question in this manner provides motivation. Is it possible for the proton to achieve light speed? This adds a kind of encouraged curiosity that helps to prod students and guides them to the correct physical principles. Instead of being asked for various physical properties of the system as in Fig.1, the student is being asked a question about the system. This forces the student to consider how they might determine this answer, and they begin to work backwards until reaching the starting point (part 1 of the problem in Fig. 1).

Also of note, there are a few other minor alterations to encourage thought. First, the tin nucleus is given context (it is a tin can of breath mints that sparks the idea). This

again provides motivation and helps the student relate to the problem. Instead of solving a work-energy theorem problem about abstract, tiny particles, the student begins by thinking about a simple can of breath mints. Second, the problem avoids the use of verbose physical jargon, instead employing terms such as “a foot away” and “achieve light speed” instead of “0.3048 meters” and “velocity relative to the speed of light”. These simple phrasing differences alters the perception of a problem, and can cause the student to find it more approachable. Finally, while making clear the final goal of the problem, the steps to achieve this goal are not immediately apparent in the problem statement.

This slight obfuscation is quite possibly the single most important piece of crafting a context-rich problem. It prevents the dreaded “plug-and-chug” method of problem solving, which requires no real thought, and forces students to consider the physics at work. Ultimately, the steps laid out in Fig. 1’s problem are what needs to be solved to learn the answer to the final question of Fig. 2’s problem, but Fig. 2’s problem statement presents a much more powerful learning opportunity by forcing students to deduce these steps themselves. Through direct application of the concepts and physics learned in lectures, students gain a much deeper understanding of the problems they are solving, and gain real-world experience by applying their knowledge of physics to an actual question.

Topic selection

Having outlined specifically how to craft a context rich problem (based on data experiential, anecdotal, and experimental), the task was now to select the actual topics that would be covered in the problem sets distributed to students. This portion of developing the experiment requires significant cooperation between the experimenter and the course instructor. Professor Qi Wen was kind of enough to volunteer his course’s lecture notes thus making, after selecting which dates to administer the problems, topic selection relatively straightforward.

It is crucial to note, however, that lecture notes (and lesson plans, for that matter) are merely guidelines and cannot be taken as a perfect model of reality. It is entirely possible that an instructor will plan to cover something during class and, for reasons unforeseen at the time of their writing, portions of the lecture notes may not be covered on the specific days that were intended for them. This is critical for an experimenter attempting to construct problems that test topics from the lecture notes, as one cannot be perfectly certain that these topics are covered unless one was physically present at lecture, or the lecture was digested through some other means (digitally, perusing a student's notes, speaking with the professor after class has ended). The most logical of these is for the experimenter to attend lectures personally, which is quite possible in a group but unfortunately somewhat difficult during a project conducted by a single person.

Upon selection of the days of the conference sessions, the problem topics were selected by determining what had been covered during the lecture immediately prior to the conference session, through analysis of the lecture notes and one day before that. According to this methodology the topics were as follows:

- *Conference session 1* – resistivity and its relation to area, current, resistance
- *Conference session 2* – magnetic force on a long wire due to another wire
- *Conference session 3* – direction of magnetic field, force per unit length of magnetic field

The problems themselves are listed in Appendix A.

Part 2: Group structure and session flow

The group structure is based on research conducted by Heller, Keith, and Anderson. According to their findings, groups of three are preferential. Barring this possibility,

groups of two will statistically fare better than groups of four. This, they determined, was largely due to social structure and differing personalities among students. With too large a group, students can become distracted, and not all students may be actively contributing. With too small a group, the chances are increased significantly that a dominant person will pair with a submissive person, which is detrimental to both students as the dominant one will persist unchecked in their problem solving, while the submissive will typically not be engaged. Thus, the students were broken into groups of three when possible, and groups of two when necessary. This decision was made because it is, in the experimenter's opinion, more dangerous to have a group of four than a group of two.

The four conferences involved in this project were divided accordingly:

- Experimental sections:
 - Section B09
 - Section B11
- Control sections:
 - Section B10
 - Section B12

As stated, on three separate conference days in between the occurrences of the second and final exams of the course, the students were broken into their assigned groups, and given a problem to solve.

Control sections were given a standard problem, similar to their homework problems, taken from their textbook. Experimental sections were given the same problem, except that it had been restructured into a context-rich problem through the means outlined above. Students in both sections were given ten minutes to solve the problems (15 minutes were allowed, if necessary), at which point each individual student in the class will turn in a solution sheet, with their group

members noted at the top of the sheet. Requiring that each student complete their own sheet is a measure taken to ensure that all students are fully engaged, and not simply passive observers of the derivation of solutions.

Following the session, the solution sheets were collected from the conference instructors via the physics departmental mailboxes, graded, and photocopied. After grading the solutions, the students' grades were entered a spreadsheet for tracking purposes. At this point, the originals were returned to the students and the photocopies were retained for further review if necessary (to be destroyed at the completion of the project). Then, the process will restart with the next session. New problems will be given, and they will be completed and recorded in the same manner.

Part 3: Analysis and demographic data

With time, all sections are expected to improve as students gain more experience with problem solving. Analysis of the data from the sessions will reveal whether or not the experimental sections improved significantly more than the control sections. However, the results of the sessions are not enough to show that context-rich group problem solving is more beneficial than standard group problem solving. To prove this hypothesis, it is necessary to show that *individual* problem solving is augmented by the context-rich group problem-solving sessions. Thus, I will use the results of the final exam compared against the results of the second exam to calculate an overall average individual gain to either confirm or deny the hypothesis.

Additionally, following the culmination of the three sessions, demographic data is to be collected from the students via a survey (expected year of graduation, declared majors, gender, conference section). This will provide insight as to whether or not solving context-rich problems can be beneficial for various subgroups of students,

or even if there are specific subgroups for which it is not quite as effective as for others. This insight may assist future instructors in offering targeted assistance for students that might be struggling with the course.

To protect the privacy of the students, a simple database of both the students' names, and randomly generated numbers (separated by conference sections) was created. The course instructor then entered students' exam grades into this spreadsheet, and removed the names. Thus, analysis of the raw data obtained during the course of the experiment will yield absolutely no insight into the identity of individual students beyond the sections in which they were enrolled, and their anonymity will be preserved.

Implementation

The initial plan was to have three conference sessions in which students in the experimental sections were given context-rich problems to solve in groups, and students in control sections were given standard problems to solve in groups. Unfortunately, due to disconnect between the lesson plan and the actual topics covered in class, the second session had to be cancelled. Students were given a problem covering topics they had not yet learned in earnest from lecture, and thus were assisted greatly by the conference instructors. As a result, data collection was cancelled for that session, as the purpose of the experiment is to help students, not hinder them. The problems for Conference Session 2 (referenced earlier in this section) are not present in Appendix A.

The two conference sessions that comprised the actual project occurred on Thursday December 5th, and Thursday December 12th. Students in sections B09 and B11 were given context-rich problems, and students in sections B10 and B12 were given standard problems. The problems were distributed to the students via the conference instructor, and collected by the instructor once they had been completed. This ensured no actual contact between the students and the

experimenter that might taint the results. Also, students were unaware of whether they were receiving a control or experimental problem, to preserve the integrity of the experiment.

Upon completion, the problems were returned to the experimenter (again via the conference instructor), graded, and recorded. Following the two sessions, the students took their final exams. These grades were compared against the grades from their previous exam (prior to any problem-solving conference sessions) to determine if context-rich problem solving in groups is beneficial to students enrolled in an introductory electricity and magnetism course.

Results

Overview of results

This project studied 52 students (23 control, 29 experimental) across four sections (B09, B10, B11, B12) in Professor Qi Wen's PH 1121 course during B term of 2013. Sections B10 and B12 (N=46) were control groups given standard textbook problems, while B09 and B11 (N=58) were given context-rich problems. Due to the relatively small N overall, the results yield little statistical significance (there are several notable exceptions, which will be discussed in detail in this section). Despite this, however, there is much to be said with regard to these results in a qualitative respect.

The results indicate that the experiment was successful. On average, students exposed to context-rich problem solving in groups scored earned a gain of five points more (when the scores of their third exams were compared against their second) than students that were not exposed to these problems. This result was smaller than one standard deviation, however, so it cannot be stated conclusively. Also, the results from the subgroup analysis suggest that overall that engineering majors with, what this experimenter has deemed, a "physics-heavy" curriculum (many course requirements in the physics program and fields that directly apply physical principles) such as Aerospace Engineering (AE) and Electrical and Computer Engineering (ECE) benefitted greatly from context-rich problem solving. Conversely, engineering majors with less focus on physical rigor and more on general engineering principles, such as Biomedical Engineering (BME) and Chemical Engineering (CHE), suffered when exposed to context-rich problem solving. Additionally, males had a much higher average gain than females. Finally, upperclassmen had a higher average gain than freshmen.

In this section, the process by which these results were calculated from the raw data will be discussed, and then the results themselves shall be examined. The results will be shown in context of various subgroups based off survey data collected from the students (gender, class year, and declared major), as well as across all of the sections without regard to subgroups. Throughout this section, numerical results will be expressed as such: result \pm standard deviation. The qualitative significance of individual results will be discussed in the Analysis section directly following this.

Explanation of data analysis

The success or failure of context-rich problem solving methods hinges upon, as do many measures of academic success, exam results. For this particular project, the results of the second and final exams of those students in the aforementioned sections were recorded and analyzed. The raw grades from these exams were entered into a spreadsheet, divided by control and experimental sections. This, along with the demographic survey data collection from the students, comprises the raw data collected by the experiment.

The raw data was then sorted into three subgroups according to demographic data (gender, class year, and declared major) as well as a general group to analyze the full control and experimental data sets. Once these groups were defined and the data had been sorted, the gain between the second and final exams was calculated by simple subtraction:

$$E_3 - E_2 = G,$$

where E_3 and E_2 are the results from the final and second exams, respectively, and G is equal to the *gain*, or the difference between the two. This number is representative of the student's improvement (or loss) from the second to the third

exam. Standard deviation is calculated from the individual gains across the group. Next, the individual gains were averaged to determine the *average individual gain* for that particular group:

$$\frac{\sum_{t=1}^N G_t}{N} = \langle G \rangle.$$

This number is representative of the average gain from the second to the third exam of individual students within a specific group (control or experimental). With the average gain from both control and experimental groups calculated, the *complete average gain* can be calculated:

$$\langle G_E \rangle - \langle G_C \rangle = \overline{\langle G \rangle},$$

where $\langle G_E \rangle$ and $\langle G_C \rangle$ are the experimental gain and control gain, respectively, for that subgroup, and $\overline{\langle G \rangle}$ is taken to be the complete average gain for that group. A positive value of $\overline{\langle G \rangle}$ indicates that the experimental group, on average, had a higher gain across the second and final exams than the control group, indicating that context-rich problem solving was ultimately, beneficial. Standard deviation for $\overline{\langle G \rangle}$ is calculated using all of the individual gains from both the control and experimental sections.

Overall results

The results of the overall group (N=52) are displayed in Fig. 3. Note that this chart expresses the average individual gain $\langle G \rangle$ as opposed to the complete average gain $\overline{\langle G \rangle}$. The experimental group had a small average individual gain ($\langle G \rangle = 0.172 \pm 14.7$), but it was still substantially larger than the control group's gain of $\langle G \rangle =$

-4.09 ± 14.3 . This indicates that, on average, overall, students benefit from context-rich problem solving in groups.

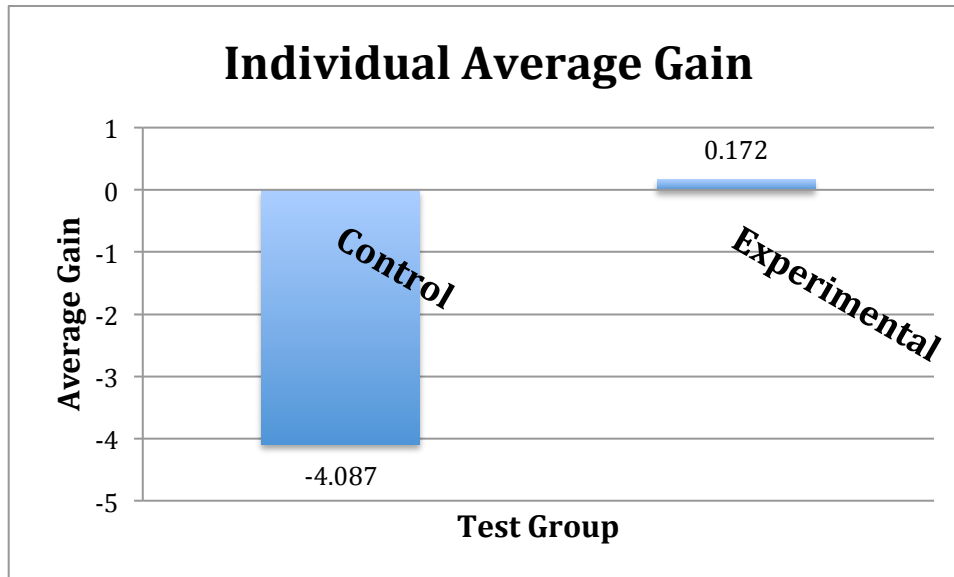


Figure 3 - A bar graph of the individual average gain for all students.

The complete average gain for this experiment was $\langle G \rangle = 4.26 \pm 14.5$. A positive value for $\langle G \rangle$ indicates success, but the standard deviation is much too large to prove anything conclusively for the entire group of students. With a larger N, a conclusive statement may be made, but this particular result lacks statistical significance.

When broken into smaller subgroups, however, statistically significant results emerge. Despite the large standard deviation, however, it is evident from Fig. 3 that context-rich problem solving in groups did have a beneficial effect for the students that participated in this study. A student earned, on average, approximately a five-point gain across their second and third exams when practicing context-rich problem solving than a student who did not.

Gender subgroup

The gender subgroup is composed of students from all sections, divided by gender (male or female). The N for each portion of the gender subgroup is as follows:

- Control:
 - Male N = 16
 - Female N = 2
- Experimental
 - Male N = 14
 - Female N = 7

As shown in Fig. 4, the complete average gain for the male portion of the gender subgroup was 13.1 ± 12.9 points, and the female complete average gain was -10.4 ± 16.0 points. Due to the relatively small number of females in the group, the standard deviation was quite large.

Despite the high standard deviation, it is evident from the results that females did benefit from context-rich problem solving as much as male did, but further study is needed to state this result conclusively. It is worth noting, however, that the control group of females performed quite well, with an average individual gain of 12 ± 28.3 points. The experimental female group had an average individual gain of -1.571 ± 12.6 points, more than 12 points below the control group.

The complete average gain for males was more than one full standard deviation above zero, indicating that males benefitted (on average) from context-rich problem solving. It should be stated, though, that while males had a higher $\overline{\langle G \rangle}$, the control group of males actually performed significantly worse than the control group of females ($\langle G \rangle = -7.63 \pm 11.7$ for males, more than 19 points worse than the

females). Conversely, the male experimental group performed approximately 6 points better than the female experimental group.

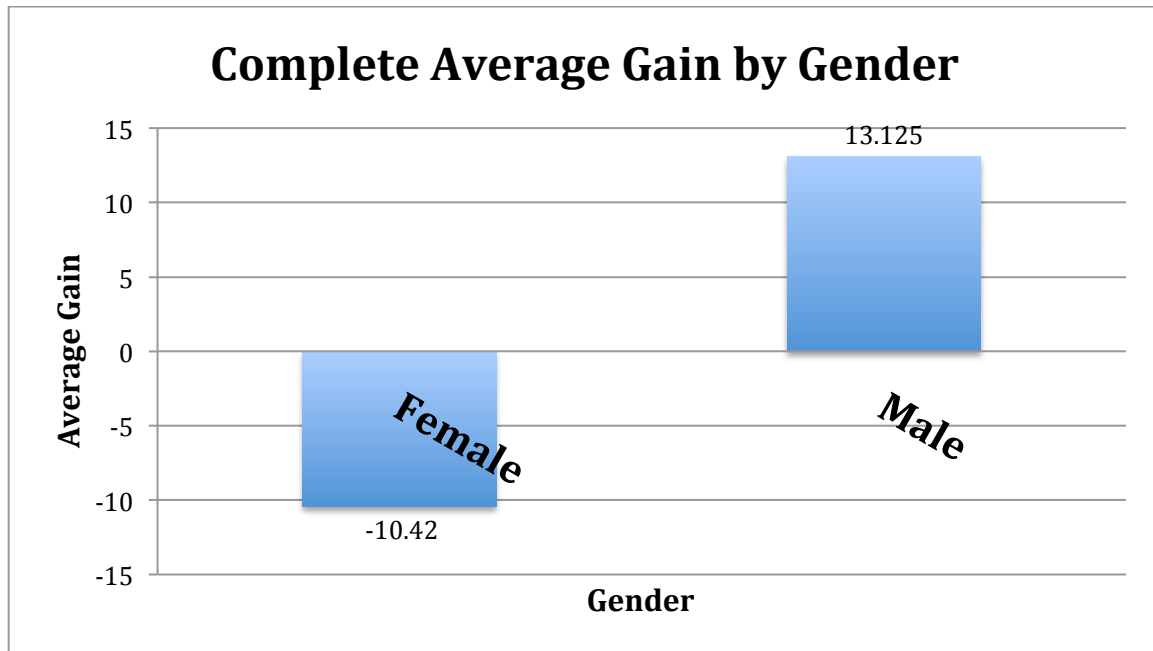


Figure 4 - A bar graph displaying the complete average gain for the gender subgroup.

Class year subgroup

The class year subgroup is composed of students from all sections, divided into a freshmen subgroup and upperclassmen subgroup. It was the original intent of the project to analyze data from each individual class year, but due to a small N in the non-freshmen class years, all upperclassmen were grouped together. The N for each group is as follows:

- Control:
 - Freshmen N = 17
 - Upperclassmen N = 1

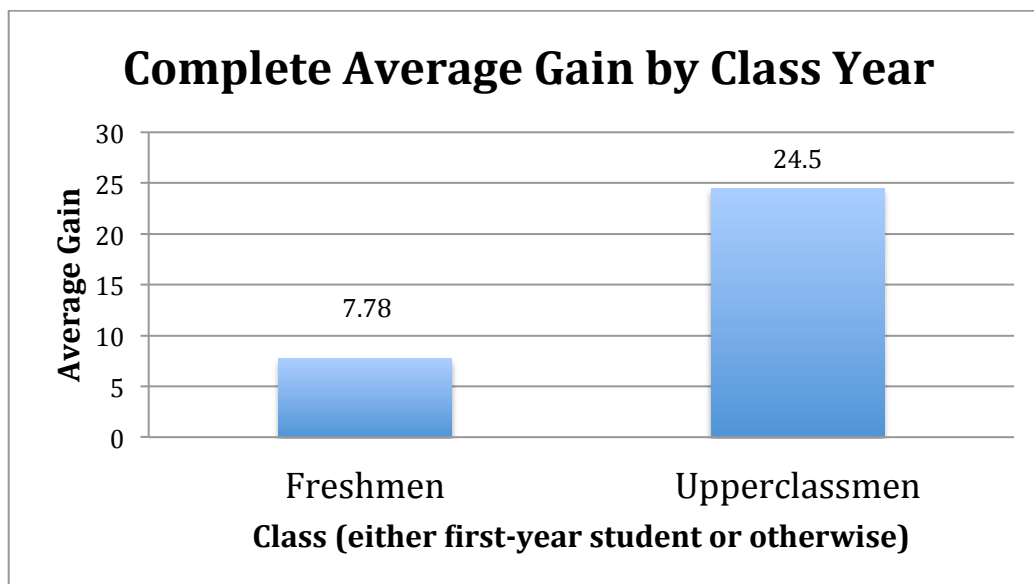


Figure 5 - A bar graph of the complete average gain by class year.

- Experimental
 - Freshmen N = 20
 - Upperclassmen N = 2

Fig. 5 displays the complete average gain for the two class year groups, both of which were positive, indicating that both groups benefitted from context-rich problem solving. The freshmen subgroup earned a complete average gain of 7.78 ± 13.3 points, while the upperclassmen had a gain of 24.5 ± 16.5 points.

The upperclassmen portion of the class year subgroup had a large positive gain, but it must be mentioned that the control group (N=1) had a very large, negative average individual gain of $\langle G \rangle = -21.0$. This average individual gain was tied for the lowest $\langle G \rangle$ for a control group in this experiment. The freshmen control group fared better, with $\langle G \rangle = -4.52 \pm 14.3$. Both freshmen and upperclassmen had relatively close average individual gains in their respective experimental sections ($\langle G \rangle = 3.25 \pm 11.6$ for freshmen versus $\langle G \rangle = 3.5 \pm 12.0$ for upperclassmen).

Declared major subgroup

The declared major subgroup is composed of students from all sections, divided by their declared majors. In the case of a student with multiple declared majors, their exam results were counted once for each major declared. Majors that are offered by WPI but not represented were not present in both the control and test groups. With further testing, and a larger N, each major will be able to be represented. The N for each represented major is as follows:

- Control:
 - AE (Aerospace Engineering) N = 1
 - BME (Biomedical Engineering) N = 1
 - CHE (Chemical Engineering) N = 1
 - CS (Computer Science) N = 2
 - ECE (Electrical & Computer Engineering) N = 2
 - ME (Mechanical Engineering) N = 3
 - PH (Physics) N = 3
 - RBE (Robotics Engineering) N = 6
- Experimental
 - AE N = 3
 - BME N = 2
 - CHE N = 1
 - CS N = 5
 - ECE N = 3
 - ME N = 5
 - PH N = 2
 - RBE N = 2

Fig. 6 displays the complete average gain across the declared major subgroup. This is organized in ascending order from the largest negative gain up through the

largest positive gain. As shown in Fig. 6, BME majors fared the worst when exposed to context-rich problem solving, with a complete average gain of -15.0 ± 8.66 . This was the lowest complete average gain of any subgroup. CHE and CS also performed negatively, with gains of $\langle \overline{G} \rangle = -3.00 \pm 2.12$ and $\langle \overline{G} \rangle = -0.90 \pm 16.5$, respectively.

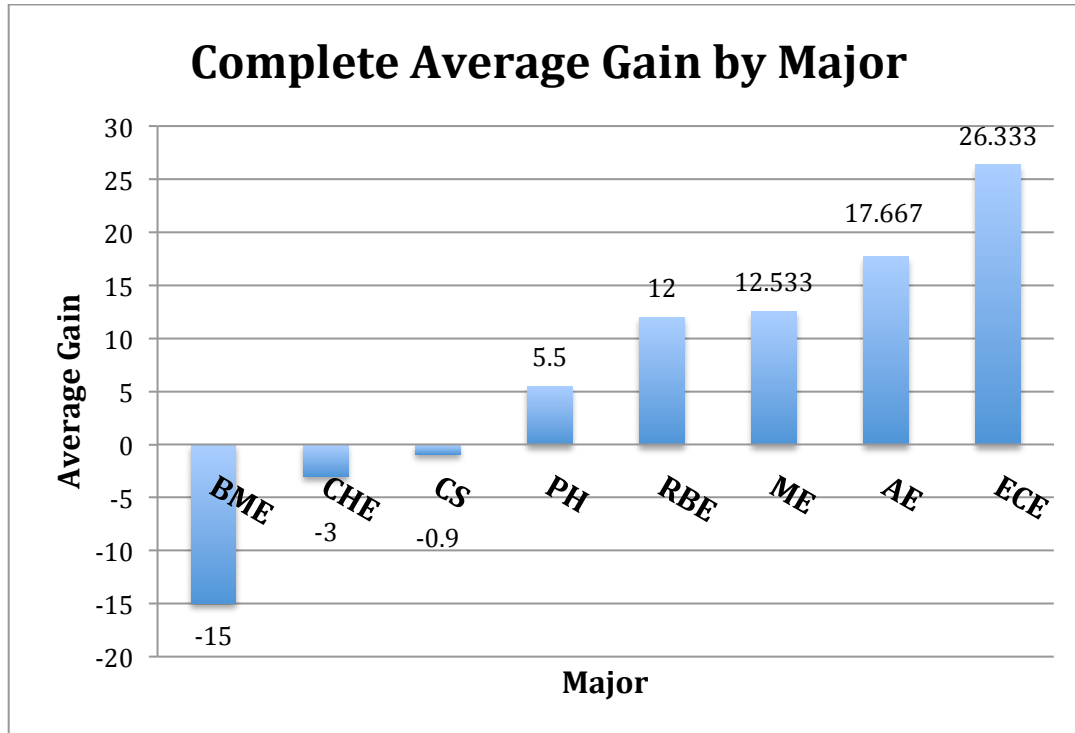


Figure 6 - A bar graph of the complete average gain by declared major.

Majors with a heavy physics background had positive gains, and some performed exceptionally after being exposed to context-rich problem solving. Of particular note are ECE ($\langle \overline{G} \rangle = 26.3 \pm 15.1$) and AE ($\langle \overline{G} \rangle = 17.6 \pm 12.1$). Interestingly, physics majors, while having a positive gain, did not perform the best. Physics majors had a gain of only 5.50 ± 12.7 , more than 20 points below the highest scoring major (ECE), and 6.5 points lower than the next highest scoring major (RBE). Physics majors had an N comparable to the other majors, so it is not certain whether a larger sample size would confirm or deny these results. RBE and ME

majors performed similarly, with RBE majors having $\overline{G} = 12.0 \pm 17.9$ and ME majors having $\overline{G} = 12.5 \pm 15.2$. The similar scores (RBE and ME differ by less than a standard deviation, only 0.5 points) might be due to the similar curriculum covered by the two majors, but more study is necessary to determine anything conclusively. All declared majors with a positive gain have curricula with a heavy focus on the direct application of physical principles, as well as a strong calculus-based mathematical focus.

Statistical conclusions

As mentioned, the relatively small sample size makes quantitative analysis somewhat difficult. It can be said (albeit, inconclusively) that context-rich problem solving in groups is beneficial to students taking PH 1121. On average, students in this course that were exposed to context-rich problem solving gained five more points on their final exam than students that were not exposed to these problems. As far as statistically significant results, certain subgroups did have gains that were greater than one standard deviation: male students, upperclassmen, AE majors, and ECE majors. Ergo, it can be said that these subgroups benefit from context-rich problem solving. Conversely, BME majors had a loss that was larger than one standard deviation, and seemed to suffer from context-rich problem solving. Further study is needed to substantiate these claims.

Analysis

The results of this project suggest that context-rich problem solving is beneficial to students in introductory electricity and magnetism courses. On average, students exposed to these problems gained approximately five points more on their final exams (as compared against their second exams) than students that were given normal, textbook style problems. These results are promising, but due to the relatively small sample size, are not statistically significant. Sources of error, and speculation as to the statistical insignificance of these results will be discussed.

Despite the statistical insignificance of the results, there is insight to be gained from a qualitative analysis of the data. There are definite trends in the data that can be readily observed. These trends will also be examined and discussed here.

Sources of error

Analysis of data must, first and foremost, discuss potentials for error (false negative, false positive, false correlation, etc.) that are present within the experiment. The most pertinent source of error is the fact that there were only two data points taken (the second and third exams) and that the second data point was, in fact, a final exam. A final exam will not give accurate results because final exams are separate from normal exams in that they foster a high-stress environment. If one is to include final exams, then one must account for the fact that students may be more stressed during the exam, that the exam may be more difficult than exams throughout the term, and that some students in danger of earning a C or failing grade may, due to the grading structure at WPI, intentionally fail the final exam to earn an NR for the course. This study was not comprehensive enough to account for these potential sources of error, and they cannot be ignored. This error could be

avoided in the future by performing a study that takes place across the entire term (due to extant circumstances, this study began approximately halfway through the term and could only examine the second and final exams) from beginning to end, encompassing every exam.

The relatively small sample size was a determining factor in the overall statistical significance of the results. Due to the fairly small N overall and across each of the subgroups, standard deviations were quite high, thus clouding the results. The only way to quell this error is to obtain a larger sample size. With support from the department, both PH 1121 and PH 1120 courses might be studied, and this could be ostensibly extended in a larger, multi-year study that would certainly mitigate any sample size concerns. Additionally, varying the particular PH 1121 course that is analyzed (for example, studying both the B term and D term courses) could increase the spread of data when considering subgroups. For example, the D term PH 1121 course may include more upperclassmen, whereas one might imagine the B term PH 1121 course would attract more freshmen.

An unexpected roadblock during this project was the creation of a context-rich problem that tested material that the students had not learned. As a result, this conference session had to be removed from the study. In the future, the experimenter (or experimenters) needs to maintain a very close relationship with the lecture instructor, and would be best to attend lectures themselves, to be acutely aware of the material that is being covered during a particular lecture. Curricula and lesson plans are, at best, an estimate or assumption of what will be covered and may not reflect reality in every case.

Overall results

The experiment was, overall, a success. While the overall complete average gain was less than one standard deviation above zero ($\overline{\langle G \rangle} = 4.26 \pm 14.5$), there was, in fact, a positive gain. This is indicative of success, but does not confirm it. With further research, this experimenter believes that this result will be corroborated and it will be shown conclusively that context-rich problem solving in groups is, overall, beneficial.

Fig. 7 is a histogram displaying the average individual gain of the control group versus that of the experimental group throughout the course of the experiment. The experimental histogram is very slightly shifted towards the right side of the graph, exemplifying the positive gains of that group. Unfortunately, these positive results came with an exceptionally large standard deviation. This standard deviation is largely due to the relatively small N (N=52 for the entire experiment). Also, as mentioned in the previous section, the second exam analyzed was, in fact, a final exam. This, in the experimenter's opinion, caused a dramatic increase in standard deviation simply due to the fact that the exam was a final as opposed to an ordinary exam. This conclusion is evidenced by the fact that the lowest exam grade recorded across both the control and experimental groups during the second exam was a 43 while the lowest exam grade on the final exam was a 0. This undoubtedly contributed quite heavily to the standard deviation, and suggests that future studies would do best to either exclude the final exam from the experiment, or include as many exams as possible (with as large a sample size as possible) throughout the term to lessen the effects of the final exam.

It cannot be ignored, however, that despite all of this, students in the experimental sections gained, on average, approximately five points more on their final exams than students in the control sections. If one ignores the aforementioned 0, the

complete average gain for the experiment becomes $\overline{G} = 6.16 \pm 12.7$, an increase of almost two points with a smaller standard deviation. Thus, overall, the experiment was a success, and students do benefit from exposure to context-rich problem solving in groups. Moreover, accounting for sources of error, particularly accounting for the final exam, it may be determined that these results show context-rich problem solving in groups is even more beneficial than currently believed.

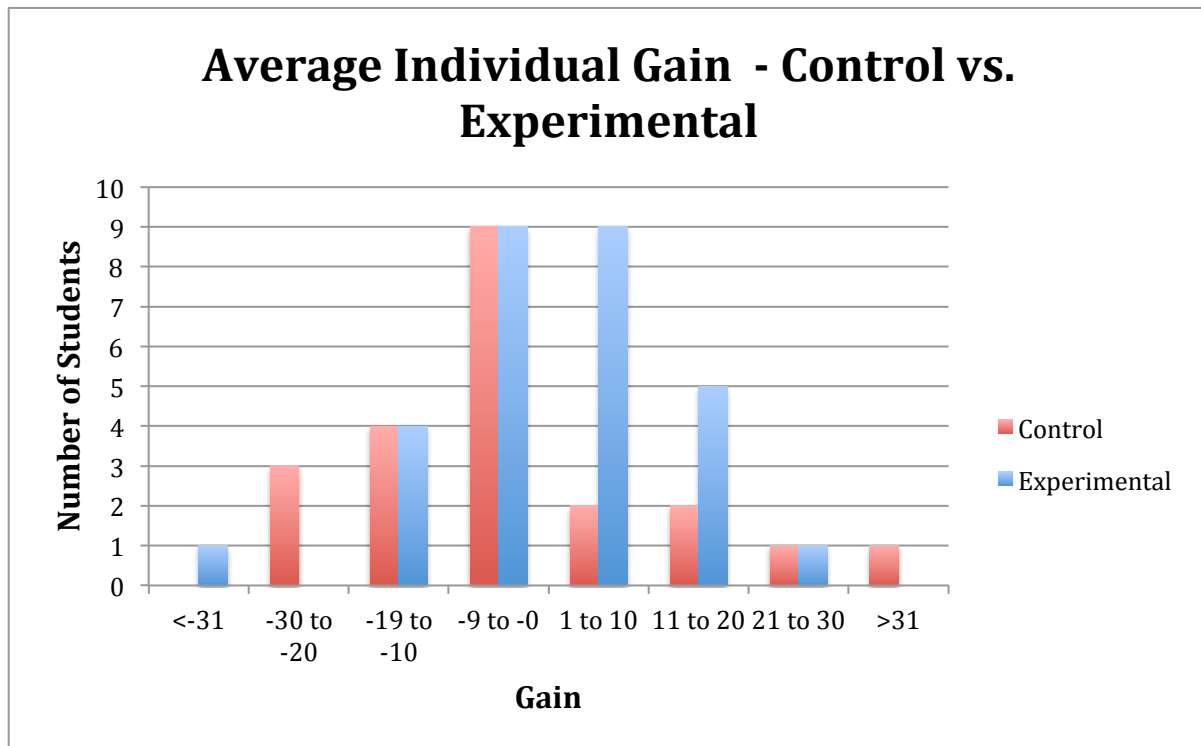


Figure 7 - Histogram displaying the average individual gain of the control group versus the individual average gain of the experimental group.

Gender subgroup

As shown in the previous section, males performed, on average, approximately 23 points higher than females throughout the course of this experiment. However, this result does not consider the difference between the experimental and control groups for each subgroup within the gender subgroup.

The male experimental average individual gain was six points higher than the female experimental average individual gain, but the female control average individual gain was almost 20 points higher than the male average individual gain ($\langle G \rangle = 12.0 \pm 28.3$ for females versus $\langle G \rangle = -7.63 \pm 11.7$ for males). This result implies that females benefit from standard textbook problems while males benefit from context-rich problem solving. However, it cannot be said conclusively that this is the case.

There are many other variables that could possibly influence the context-rich problem solving sessions, which will, in turn, affect exam performance. For example, group structure could be a factor. It may be that males are more dominant in a group than females (particularly given the fact that there were twice as many males as females in the experimental group). Thus males may be reaping a larger benefit from these group-themed exercises, and ergo would perform better on exams and earn higher gains. This is, of course, all speculation at this early phase of research, but underscores the fact that further research focused on group structure is needed to determine anything conclusive about the benefits context-rich problem solving for female students.

Speculation aside, it cannot be ignored that males had a significant gain in this experiment. This result shows conclusively that male students benefit from context-rich based problem solving. The gain was, however, only very slightly larger than standard deviation ($\langle G \rangle = 13.1 \pm 13.0$). This slight gain becomes more apparent when a histogram (Fig. 8, a histogram displaying the male experimental subgroup's average individual gains versus the male control subgroups average individual gains) is referenced.

Fig. 8 shows that the experimental average individual gain for males was slightly higher than the control average individual gain. Notably, no male students in the

experimental subgroup earned a gain less than -20, and four earned a gain of between 11 and 20 points.

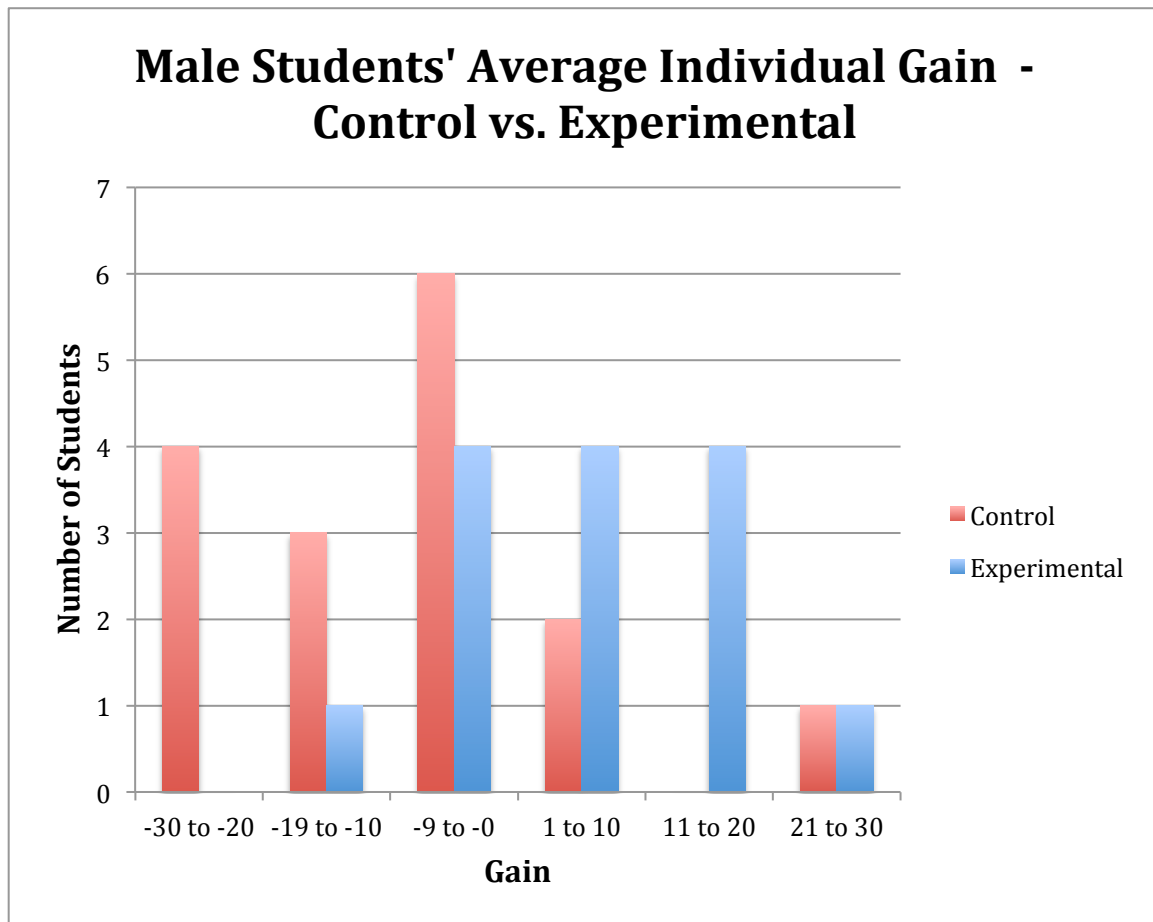


Figure 8 - A histogram comparing the male control average individual gain against the experimental male average individual gain.

It is the opinion of the experimenter that, these results are quite sound. The control sample size and experimental sample size for males were very nearly equal (N=16 for control, N=14 for experimental). Also, the lowest Exam 2 grade for the male control subgroup was higher than the lowest Exam 2 grade for the experimental male subgroup (57 versus 49). In light of this revelation, the gains of the experimental subgroup become even more significant. The control subgroup had an Exam 2 average grade of 75.1 while the experimental subgroup had an Exam 2

average of 70.1. Ergo, despite a lower Exam 2 average, the experimental subgroup gained, on average, 8 points more on their final exams, and thus earned a higher average individual gain. This implies, quite strongly, that male students in particular benefit from context-rich problem solving.

Class year subgroup

The upperclassmen earned a higher average individual gain than the freshmen. The freshmen did, in fact, earn a positive average individual gain. The freshmen subgroup gain was $\langle G \rangle = 7.78 \pm 13.3$, which is less than one standard deviation, and thus nothing conclusive can be stated about that particular subgroup. Despite the fact that the upperclassmen earned a gain more than one standard deviation in magnitude ($\langle G \rangle = 24.5 \pm 16.5$), the sample size of the upperclassmen subgroup cannot be ignored when considering the possibility of false positives. There were only three upperclassmen that participated in the project: two in the experimental subgroup and one in the control group. Additionally, it is quite possible that the upperclassmen had been exposed to these kinds of questions before, and it is likely that they have had more problem-solving experience than the freshmen group. The results must then be called into question, and it is recommended that further study be conducted to either confirm or deny the validity of this result. It is the hope of this experimenter that this result will be solidified with further study, but realistically this is very probably a false positive. The results collected from the freshmen subgroups are also suspect.

The freshmen subgroup sample size ($N=17$ for control, $N=20$ for experimental) was comparable to others that generated significant results (notably the male subgroup sample size), but the standard deviation cannot be ignored. The standard deviation of the freshmen subgroup was almost twice the measured average individual gain of that group, which is troubling, and emphasizes the diversity of the freshmen

subgroup. For example, the difference between the lowest and highest exam grades for the control and experimental groups was more than 30 points on Exam 2. This spread in the data directly contributes to the large standard deviation, and overpowers the benefits gained from the relatively large sample size of the group. The freshmen subgroup represents 10 different majors, while the upperclassmen have only three distinct majors among them, and thus the diversity of the freshmen cannot be ignored as a contributing factor in the large standard deviation seen. Their low average individual gain during this experiment can just as well be attributed to this vast diversity as it can be attributed to a failure of context-rich problem solving.

Declared major subgroup

Fig. 6 displays the complete average gains of students by major. The students with heavy, physics-based engineering backgrounds performed quite well, with ECE earning a complete average gain of $\overline{\langle G \rangle} = 26.3 \pm 15.1$ and AE earning a complete average gain of $\overline{\langle G \rangle} = 17.7 \pm 12.1$. Of particular note here, though, are the majors that did not perform so admirably. These results were discussed in the previous section, but are restated here for simplicity and emphasis. BME majors in particular earned the largest negative complete average gain of any subgroup studied ($\overline{\langle G \rangle} = -15.0 \pm 8.66$). Along with BME, another chemically themed major, CHE, performed poorly ($\overline{\langle G \rangle} = -3.00 \pm 2.12$). Finally, CS majors performed slightly better, but still showed a negative gain with $\overline{\langle G \rangle} = -0.90 \pm 16.5$. When one analyzes the majors with a much stronger background in physics (RBE, ME, and of course the PH major itself), a more positive picture emerges. This trend must be considered when determining the success or failure of context-rich problem solving techniques among particular majors.

It appears that students that may not have had much prior exposure (or indeed, much interest) in physical principles and their applications, gains are lower. Upon review of the curricula and degree requirements for these majors, it becomes clear that PH 1121 may very well be one of if not the only physics course these students take. It is difficult to say, then, whether this result is due to a failure of the educational apparatus presented during the course of the experiment, a lack of exposure to these kinds of problems, or a failure on the students' part due to a lack of interest and/or motivation because of the perceived lack of relation of the exercises to their declared major. These conclusions (most notably the last), of course, raise a much larger question about education in general: is a students' failure the fault of the student, or the fault of the instructional tools that are presented to them (educator, curriculum, etc.)?

It is the opinion of this experimenter that the onus is upon the educator to engage students on an intellectual level and augment their knowledge, problem solving abilities, and overall educational vivacity. This can only be done through active engagement, and it is again the responsibility of the instructor to encourage and foster this engagement. It may be that this particular method of context-rich problem solving is not an effective means of engagement for students across all walks of life, and that a more targeted approach to education may be beneficial. For example, it might be discovered that students of a particular major respond well to a particular sort of educational apparatus while students of other majors respond just as well to different techniques. The intrinsic problem with this sort of separation is of course the fact that these results (and indeed any study of this kind) will only inform as to the behavior and response of the *average* student. What is simultaneously the greatest and worst aspect of human beings is that they are, at their core, individuals. One can plan for, and teach to, the *average* student, but one can never hope to accommodate the needs of each individual student. The best that one can work towards, then, is the continuation of this sort of study to gain the most accurate perception possible of the behavior of the average student, and work to

actively engage this ideal student, regardless of their academic background, but be ready, willing, and abled to accommodate the needs of struggling students.

Further study is needed to confirm or deny these speculations.

Conclusions

Results of the experiment

As stated, the experiment was successful. It has been shown (although, with a minimal degree of certainty) that students, on average, benefit from exposure to context-rich problem solving in groups. The complete average gain, a measure of the average increase or decrease in exam scores of the experimental subgroup versus the control subgroup, for the experiment was $\langle \overline{G} \rangle = 4.26 \pm 14.5$. Due to the fairly small N in this experiment (N= 52), and due to the large spread of exam grades, there was a very large standard deviation. Consequentially, these results cannot be considered statistically significant.

A secondary, unexpected, insight was gained from this experiment as well. It is evident from the data that context-rich problem solving in groups, while beneficial overall, may not be beneficial and in fact may be a detriment for particular subgroups. Notably, BME students and female students appeared to fare poorly when subjected to this technique (with complete average gains of $\langle \overline{G} \rangle = -15.0 \pm 8.66$ and $\langle \overline{G} \rangle = -10.4 \pm 16.0$ respectively). These groups, however, were not very large and did not yield statistically significant results. Further research must be conducted to confirm whether or not this is a failure of context-rich problem solving.

Future research goals

Corroboration

The primary goal of further research would be, of course, to corroborate the results found here. These results are suggestive at best, and require further study to be proven conclusively. Future iterations of this experiment would do well to cover the entire term, as opposed to the second half of the term, and to expand across all sections of all PH 1121 courses. Ideally, this would be a multi-year study involving, at the very least, approximately 500 to 1000 students. This would provide a sufficiently large N for rigorous statistical analysis, and provide much deeper insight into how context-rich problem solving affects different subgroups.

Other variables to explore

The only variable explored during this experiment was context-rich problem solving in groups, and whether or not it affected test grades. It was found that, on average, test grades increase when students were exposed to this technique, but there are much deeper insights that can be attained through further study. There are likely many outside variables, as there often are in social studies, which influenced the results found here. Briefly touched upon in the previous section, these variables certainly merit further study.

First and foremost is, of course, to confirm the results found here. But, once it has been firmly established that students do benefit from context-rich problem solving in groups, there is quite a bit of study that needs to occur. This experimenter humbly suggests the following areas of study as potential points of interest: time of day of conference sessions, rigorous testing of subgroups, variance of the group structure, variance of problem presentation, variance of the conference instructor, analyzing certain aspects of course content, analyzing other problem types, and finally extending this research beyond the confines of PH 1121 to other physics courses.

Time of day

Varying the time of day of conference sessions and analyzing its effect on students seems obvious. One would assume that lethargic students in an 8:00am conference would be less apt to absorb information than students in an 11:00am conference, but then the 11:00am students might be more focused on an impending lunch hour. And what if the 8:00am students were provided coffee? There is certainly an abundance of testing and study that could occur if one focused solely on the time of day of conferences, and the mental conditions of students as a result of the time day, and their correlation with context-rich problem solving.

Subgroups

It seems nonsensical to ignore the very shallow dive made into subgroups in this experiment. This notion is at the heart of academic study: how certain educational techniques apply to different groups of students. It goes without saying that this particular area of research contains unplumbed depths of insight, and needs to be investigated further. There can be an entire study conducted, for example, on the effects of context-rich problem solving in groups for males versus females. Even, if one were so bold, a study on the effects of context-rich problem solving for *groups* of males versus *groups* of females. That is, breaking the students into groups of males and groups of females, as opposed to mixed gender groups, and studying the results.

Group structure

If one is interested in a male versus female project, it naturally follows then, that one might be interested in learning if group structure affects a students' ability to respond positively to context-rich problem solving. One could envision a study wherein students take a personality pretest, and then are broken into groups that vary dominant and submissive personalities. Or, perhaps, one might generate

groups based on academic performance on the first exam so that each group has a strong student, an average student, and a weak student. Varying the structure of groups is a very active area of current academic research, and there is no reason to ignore its application to context-rich problem solving here.

Problem presentation

This project analyzed the effects of context-rich problems versus standard problems, but not the actual presentation of the context-rich problems outside of the paragraph format described above in the Methodology section. It may be worthwhile, for example, to examine the effects of problems presented paragraph-style versus problems presented in a multiple-choice style format. There is insight to be gained with regards to problem solving pathways and how they relate to the presentation of the problem itself.

Instructors

There has been an interest, academically, in whether the attitude and idiosyncrasies of the instructors themselves has an impact on learning. This would be a logical point of study in any large research project that studies education, but should certainly be a point of focus in this body of research. Definitive answers as to whether or not students respond positively to a particular instructor would be a major benefit to administrators, as it would allow them to target particular areas within this department for professional development and growth. There certainly are ethical concerns to be raised by this form of research, as care may wish to be taken to protect the privacy of the instructors.

Particular topics in the curriculum

Further study needs to be conducted with respect to particular topics within the introductory electricity and magnetism curriculum. This particular project analyzed

only the overall result of context-rich problem solving, without experimental concern for the topics being covered. In fact, aside from the writing of the conference session problem sets, the topics covered were ignored altogether. It should be determined with further study whether or not there are particular topics which lend themselves to context-rich problem solving, as this could help instructors adapt their techniques for particular concepts.

Problem type variance

This project focused solely on investigating the potential benefits of context-rich problem solving, but it would be folly to limit the scope of this investigation into learning only to context-rich problems. This research showed, albeit somewhat inconclusively, that not all students respond well to context-rich problem solving in groups. How, then, do we encourage a stronger, more positive response? What kind of problem solving, or teaching style, will help these students become stronger problem solvers?

It has been shown that the traditional lecture method is, by itself, a relatively ineffective teaching method. While the focus of this research was context-rich problem solving in groups, there is no compelling reason to believe that other types of problem solving experiences offer no measurable benefit. What other learning experiences exist?

Other physics courses

If one were to analyze particular topics within the PH 1121, it follows logically that one would be interested in whether or not context-rich problem solving could benefit instructors in other courses. Immediately one might study the application of these techniques to higher-level electricity and magnetism courses, but research could be expanded to other introductory courses within the physics department, and indeed other higher-level courses within the department. There is no reason to

restrict context-rich problem solving to PH 1121. It must be noted, however, that as one studies higher-level courses, the sample size of the group will inevitably decline, making statistical analysis more difficult.

Final comments

The purpose of this experiment was to determine whether or not students' comprehension of introductory electricity and magnetism topics benefits from context-rich problem solving in groups, as opposed to being given standard textbook style problems. The results have been positive, and certainly intriguing, but further study is needed before this topic can be approached with anything but cautious optimism. Single projects will not be enough to fully explore this nascent area of study. There cannot be enough attention given to education research, particularly to research in the STEM fields. With the current problems that humanity is facing and will soon face (global warming, decline of natural resources, progress of modern industry, medical advancement, etc.), STEM education is more important than ever before. More efficient education in these areas will be incredibly beneficial to the human condition as a whole, and the importance of an intense research initiative cannot be emphasized enough.

This project was but a scratch upon the surface of physics education research at WPI and only the very first inklings of the depth of knowledge that can be gained from this kind of research. To put it quite plainly, the best investment one can make in one's future is a strong education, and as such there is no higher cause than that quest to better educate, enrich, and encourage students. The future rests upon their shoulders, and all they will create, all they will achieve, all they will be, is firmly based upon their education.

Works Cited

"Action Research." *Wikipedia, the free encyclopedia*. Wikipedia, 24 Jan 2014. Web. 26 Jan 2014. <http://en.wikipedia.org/wiki/Action_research>.

Antonenko, P. D., Jackman, J., Kumsaikaew, P., Marathe, R. R., Niederhauser, D. S., Ogilvie, C. A., & Ryan, S. M. (2011). Understanding student pathways in context-rich problems. *Education and Information Technologies*, 16(4), 323-342. Retrieved from <http://link.springer.com>

"Charles Hubbard Judd." *Wikipedia, the free encyclopedia*. Wikipedia, 20 Dec 2013. Web. 26 Jan 2014. <http://en.wikipedia.org/wiki/Charles_Hubbard_Judd>.

Comp, David. "A Brief History of Research on Education." *International Higher Education Consulting Blog a project by David Comp*. N.p., 13 Jul 2009. Web. 26 Jan. 2014. <<http://ihec-djc.blogspot.com/2009/07/brief-history-of-research-on-education.html>>.

Greenwood, D., and M. Levin. "Reconstructing the Relationships Between Universities and Society Through Action Research." *Trans. Array Handbook of Qualitative Research*. . 2nd ed. Thousand Oaks, CA: SAGE Publications, 2000. 143-144. Print.

Halloun, I., and D. Hestenes. "The initial knowledge state of college physics students." *American Journal of Physics*. 53. (1985): 1043-1055. Print.pl

Harke, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64-74. doi: 10.1119/1.8809

Heller, P., Keith, R., & Scott, A. (1992). Teaching problem solving through cooperative grouping. part 1: Group versus individual problem solving. *American Journal of Physics*, 60(7), 627-636. doi: 10.1119/1.17117

Heller, P., & Hollabaugh, M. (1992). Teaching problem solving through cooperative grouping. part 2: Designing problems and structuring groups. *American Journal of Physics*, 60(7), 637-644. doi: 10.1119/1.17118

Marušić, M., Erceg, N., & Sliško, J. (2011). Partially specified physics problems: university students' attitudes and performance. *European Journal of Physics*, 32(3), 711-722. doi: 10.1088/0143

"Physics education research." *Wikipedia, the free encyclopedia*. Wikipedia, 7 Nov 2013. Web. 26 Jan 2014. <http://en.wikipedia.org/wiki/Physics_education>.

Pring, Richard. *Philosophy of Educational Research*. 2nd ed. New York: Continuum, 2000. 90. Print.

Simpson, P. (2010). Personalized inquiry: help your students classify, generate, and answer questions based on their own interests or common materials. *Science and Children*, 48(4), 36. Retrieved from <http://www.nsta.org>

Travers, R.M.W. *How Research Has Changed American Schools: A History from 1840 to the Present*. Kalamazoo, MI: Mythos Press, 1983. 7. Print.

Appendix A – Conference problem sets

Included here are the two problem sets (control and experimental) for each of the conference two conference sessions.

Conference Session 1 – Experimental problem

Directions: Please include your section, your name, and the names of all group members at the top of this sheet. Each group member should submit their own sheet (please keep the sheets together). Include all work (you may use the opposite side of the paper).

Exercise: Jack is using his brand new 1 kW toaster. The heating element in his toaster is a two-meter length of wire made of nichrome that is only 0.5 mm thick! If Jack decides to make some toast, what is the current (in amperes) through the nichrome wire?

Hint: The resistivity of nichrome wire is $\rho = 110 \cdot 10^{-8} \Omega \cdot \text{m}$.

Conference Session 1 – Control problem

Directions: Please include your section, your name, and the names of all group members at the top of this sheet. Each group member should submit their own sheet (please keep the sheets together). Include all work (you may use the opposite side of the paper).

Exercise: A 1 kW toaster has a nichrome wire heating element which is two meters in length, and 0.5 mm in diameter. Nichrome wire has a resistivity of $110 \cdot 10^{-8} \Omega \cdot \text{m}$. 1) What is the cross-sectional area of the nichrome wire? 2) What is the resistance of the wire? 3) When the toaster is in use, what is the current (in amperes) through the nichrome wire?

Conference Session 2 – Experimental problem

Directions: Please include your section, your name, and the names of all group members at the top of this sheet. Each group member should submit their own sheet (please keep the sheets together). Include all work (you may use the opposite side of the paper).

Exercise: Jack noticed that power lines on telephone poles always seem to hang down and droop. Magnetic fields can affect current carrying wires, and he wondered what current it would take to keep the wires perfectly straight. The linear mass density of the wire is 20 g/m, Jack has generated a magnetic field of 5 T in the $-z$ direction, and Earth's gravity acts on the wire in the $-y$ direction. Considering the force per unit length due to gravity and the force per unit length due to the magnetic field, what current must the wire carry (magnitude and direction) in order to be suspended perfectly straight?

Hint: What magnetic force could balance the gravitational force?

Conference Session 2 – Experimental problem

Directions: Please include your section, your name, and the names of all group members at the top of this sheet. Each group member should submit their own sheet (please keep the sheets together). Include all work (you may use the opposite side of the paper).

Exercise: A long straight wire of linear mass density 20 g/m and carrying some current is immersed in a magnetic field of 5 T in the $-z$ direction, with Earth's gravity acting on the wire in the $-y$ direction. 1) What is the force per unit length due to gravity? 2) What is the force per unit length due to the magnetic field? 3) What current must the wire carry (magnitude and direction) in order to be suspended perfectly straight?

Appendix B – Survey

This survey was distributed to students that participated in the project during their last conference meeting (following the conclusion of the problem sessions):

This is a survey to capture demographic data about this course. Please answer each question as well as you can.

Name:

Section:

Gender:

Intended year of graduation:

Major (please list your intended major if you have not yet declared):

Appendix C – Table of results

A table of all complete average gains from this experiment, with statistically significant results in bold, and negative results in red. Results are organized by overall experimental category (gender, class year, and declared major) and then by subgroup within that experimental category.

All Results

<u>Category</u>	<u>Subcategory</u>	<u>\overline{G}</u>	<u>Std. Dev.</u>
Gender	Male	13.125	12.950
Gender	Female	-10.429	15.953
Class Year	Freshmen	7.779	13.342
Class Year	Upperclassmen	24.500	16.503
Major	AE	17.667	12.148
Major	BME	-15.000	8.660
Major	CHE	-3.000	2.121
Major	CS	-0.900	16.527
Major	ECE	26.333	15.123
Major	ME	12.533	15.156
Major	PH	5.500	12.696
Major	RBE	12.000	17.881