Learning a new physics concept by exploring analogous problems: an instructional intervention

Joanna Perry Weaver
University at Albany, State University of New York, jprweaver1@gmail.com

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Part of the Cognitive Psychology Commons, Educational Psychology Commons, and the Science and Mathematics Education Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive. Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
LEARNING A NEW PHYSICS CONCEPT BY EXPLORING ANALOGOUS PROBLEMS:
AN INSTRUCTIONAL INTERVENTION

by

Joanna P. Weaver

A Dissertation
Submitted to the University at Albany, State University of New York
in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Philosophy

School of Education
Department of Educational and Counseling Psychology
2019
ABSTRACT

This study tested the hypothesis that exploratory learning, with and without analogous problems, would improve students’ ability to make connections between conceptually-related topics. In this randomized experiment, undergraduates in introductory physics (N = 171) studied a new topic under three different instructional conditions. Order and type of instruction varied: Two experimental groups explored the concept before hearing a lecture; a control group followed the typical sequence of hearing a lecture before working with the concept. Within the experimental condition, students in the analogy-first group simultaneously explored analogous problems; students in the explore-first group explored only the new problem with a prompt to refer to prior knowledge of the analogous concept. Students in the instruct-first group heard a lecture on the new material before attempting to solve the problem. Ability to conceptualize and solve the new problem type differed between all three groups: On the learning activity, instruct-first students outperformed analogy-first students who outperformed explore-first students. However, posttests measuring procedural and conceptual knowledge at the end of the lesson, once all students had experienced both lecture and problem-solving phases, revealed no main effect of condition. Although all groups’ scores dropped on a retention test two weeks later, the instruct-first group’s scores declined more steeply than the others’. These results suggest that benefits of direct instruction may be short-lived. Students’ self-reported process-level measures indicated that exploring the analogous problems was no more cognitively loading than problem solving after instruction and was less so than exploring without the analogy. Equal levels of motivation between the three conditions show that exploring did not dampen enthusiasm for learning. Exploring analogous problems was an effective way of learning a physics concept in this study, but did not exceed the advantages of a traditional method in the short time frame measured.
DEDICATION

To Tim, Benjamin and Noah, for filling up my heart anew each day.

And to my parents, for always believing in and supporting me, whatever the endeavor.
ACKNOWLEDGMENTS

It takes a village to raise a scholar. I would like to express my gratitude toward the members of my village for their contributions to my professional development.

My reinvention as an academic in my fourth decade of life is due to the example and support of my doctoral mentor. Dr. Marci DeCaro gave me my first chance to conduct psychological research in the Learning and Performance Lab at the University of Louisville. As her lab coordinator and then graduate student, she taught me the fundamentals of experimental design, scientific research and scholarly writing. Under her guidance, I received the best training I could have hoped for. The time and energy she invested in me ensured I was prepared to meet every challenge I encountered over the course of my degree program.

I am grateful to Dr. David Dai for co-chairing my dissertation committee and engaging me in thoughtful discussions of the theoretical foundations of and the important debates in educational psychology. In this field, I found my intellectual home. Eternal thanks goes to Dr. Kim Colvin for teaching me key principles of psychological and educational measurement and coaching me in the statistical analyses I needed to accomplish my research goals. Under her advisement, I gained competence and confidence in the technical skills of my new profession.

Experimental research in field settings is made possible through partnerships between researchers and practitioners. I have been extraordinarily lucky to collaborate with an inquisitive and dedicated researcher who is also an innovative and passionate teacher. Dr. Ray Chastain was invariably willing to explain physics content, rework assessments, think through alternative explanations, and bring enthusiasm to every brainstorming session we had over the course of five years.

I am thankful for the numerous educational, research and teaching opportunities provided me by the Division of Educational Psychology and Methodology faculty at the University at Albany. My appreciation also goes to Dr. Patricia Ralston of the J.B. Speed School of Engineering for inviting me into the world of interdisciplinary research and for posing important educational problems to solve. I could not have reached this point without my lab mates and
cohort at UofL, and I will forever appreciate my friends at UA for their generosity in sharing resources, knowledge and especially companionship at department events, conferences and in B10!

This research would not have been possible without the financial support of the Karen R. Hitchcock New Frontiers Fund Award and the Dissertation Research Fellowship Award from the Office of Graduate Education at the University at Albany. I am extremely grateful for their vote of confidence in this research.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>CHAPTER I: INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Theoretical Framework</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>2</td>
</tr>
<tr>
<td>Statement of Purpose</td>
<td>7</td>
</tr>
<tr>
<td>Current Study</td>
<td>8</td>
</tr>
<tr>
<td>Exploratory Learning</td>
<td>16</td>
</tr>
<tr>
<td>Analogical Reasoning during Problem Solving</td>
<td>21</td>
</tr>
<tr>
<td>Expansive Framing</td>
<td>23</td>
</tr>
<tr>
<td>Motivation for Learning and Perceptions of Difficulty</td>
<td>24</td>
</tr>
<tr>
<td>CHAPTER III: METHOD</td>
<td>30</td>
</tr>
<tr>
<td>Participants</td>
<td>30</td>
</tr>
<tr>
<td>Materials</td>
<td>32</td>
</tr>
<tr>
<td>Procedure</td>
<td>45</td>
</tr>
<tr>
<td>Data Analysis Plan</td>
<td>48</td>
</tr>
<tr>
<td>CHAPTER IV: RESULTS</td>
<td>50</td>
</tr>
<tr>
<td>Preliminary Analyses</td>
<td>51</td>
</tr>
<tr>
<td>Descriptive Statistics</td>
<td>52</td>
</tr>
<tr>
<td>Statistical Analyses</td>
<td>55</td>
</tr>
<tr>
<td>CHAPTER V: DISCUSSION</td>
<td>66</td>
</tr>
<tr>
<td>Summary of Major Findings</td>
<td>68</td>
</tr>
<tr>
<td>Detailed Summary of Findings</td>
<td>68</td>
</tr>
<tr>
<td>Implications</td>
<td>75</td>
</tr>
<tr>
<td>Inconsistencies with Previous Work</td>
<td>76</td>
</tr>
<tr>
<td>Limitations</td>
<td>76</td>
</tr>
<tr>
<td>Future Directions</td>
<td>77</td>
</tr>
<tr>
<td>Conclusion</td>
<td>78</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>80</td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>97</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. Participant flow through the intervention
Figure 2. Electric force problem
Figure 3. Magnetic force problem
Figure 4. Mislabeled diagram
Figure 5. Mean score on the learning activity as a function of instructional condition
Figure 6. Mean percentage scores, on posttests, by condition
Figure 7. Knowledge subscales at posttest and retention test, by condition
Figure 8. Mean cognitive load ratings by condition
## LIST OF TABLES

Table 1. Reviewed Research of Use of Exploratory Learning and Analogies to Introduce New Material

Table 2. A Priori Coding Scheme for Learning Activity

Table 3. Comparison of Analogs in Current Study

Table 4. Instructional Sequence

Table 5. Prior Knowledge of Physics by Condition

Table 6. Descriptive Statistics for Learning Activity by Condition

Table 7. Descriptive Statistics for Knowledge Subscales by Condition at Posttest and Retention Test

Table 8. Summary of Correlations, Means, and Standard Deviations for All Survey Scales

Table 9. Descriptive Statistics for Survey Scales by Condition
CHAPTER I: INTRODUCTION

When psychology professors explain the workings of human memory to their students, they often do so by using an analogy. The memory system is likened to a computer with the brain described as an information processor composed of different types of storage. Though this model is no longer favored by cognitive psychologists, the analogy dies hard because of its intuitive appeal. It makes a complex concept comprehensible by building on something with which students are already familiar. Analogies have served this purpose for great communicators throughout time: Carl Sagan described the life of the universe in terms of a calendar year (Sagan, Druyan, & Soter, 1980); Winston Churchill used the symbol of an iron curtain to represent the divide between Soviet-dominated Eastern Europe and democratized Western Europe (Churchill, 1946). As powerful tools for simplifying complicated ideas, analogies are widely used in natural science classrooms to facilitate learning (Podolefsky & Finkelstein, 2007). Students studying the atom for the first time, for example, are taught to envision its structure like the solar system’s where electrons revolve around a nucleus the way that planets revolve around the sun (Rutherford, 1911).

How can analogies be used in the college classroom to help students connect familiar ideas with new concepts? Analogical reasoning has been the subject of much theoretical work and extensive laboratory research. It has been undertaken in classroom experiments as well. The current study extends this literature by uniquely situating analogies in the context of a particular instructional approach: student-centered exploratory learning of physics. In my dissertation, I found that using analogs during problem solving scaffolded students’ exploration of a new concept – but not more than being introduced to analogous concepts through direct instruction.

Theoretical Framework
From an information-processing perspective, the best way to learn something new is to integrate it with what is already known. When novel information is connected to previously learned material, a web of interrelated concepts grows. These associations increase the links between neural pathways connecting memory areas in the brain. The more elaborated the web of concepts, the more likely relevant connections will be activated (Anderson, 1983). Organization of newly learned material is, therefore, critical to its successful retrieval. Students want to access what they know about a topic and their ability to do so depends, in large part, on how well they have structured their knowledge in long-term memory. Simply put, how well information goes in determines how well it will come out. Principles of constructivism similarly stress the importance of engaging a learner in active knowledge construction by building on previous experience (Phillips, 1995). According to constructivists (e.g., Dewey, 1916), knowledge must be created by a learner, not passed down from a teacher. It follows that learning activities must be (a) designed to build knowledge on learners’ existing foundations, and (b) involve students in active knowledge construction.

**Statement of the Problem**

Despite the compelling rationale from information-processing and constructivist perspectives, college-level students often fail to see the connection between the topics they are taught. In physics, for example, only a few general principles govern the physical world and are thus able to explain all of the topics taught in introductory physics courses (Lin & Singh, 2011). Yet, because there are infinite applications of these principles, physics students often struggle to identify principles and relate new concepts back to them (Chi, Feltovich, & Glaser, 1981; Chin, Chi, & Schwartz, 2016; Hestenes, Wells, & Swackhamer, 1992). Part of the reason for students’ knowledge compartmentalization is the typical piecemeal instructional approach by which
professors fail to capitalize on the knowledge students accumulate throughout their coursework. When each new topic is introduced as if from the beginning of students’ knowledge development, students tend to see the topics as discrete (Richland, Stigler, & Holyoak, 2012).

This failure to draw relational comparisons may explain why students often turn to procedural and mathematical knowledge to solve problems, instead of using a conceptual explanation (Larkin, 1981; Richland et al., 2012). But these techniques result in superficial learning (Chi & VanLehn, 2012). Furthermore, Hestenes and colleagues (1992) identified that because physics students fare much better when solutions can be reached quantitatively, instructors may tend toward emphasizing problems that can be solved with minimal conceptual understanding. When students perform a disciplinary task easily, like using a formula to solve a problem, it can appear to their professors that they have understood the task well. The fluency with which something is learned or executed, however, often belies superficial understanding (Bjork, 1994). Encouraging students to wrestle more with difficult conceptual explanations is, therefore, likely to be more effective in supporting students in developing deep structural knowledge. An additional benefit to the more difficult conceptual approach is that greater mental effort expended during the learning process can lead to more durable memory (Bjork, 1994; Bjork & Bjork, 2011).

Moreover, when students rely on procedural knowledge, they may become efficient problem solvers but fail to become innovative problem solvers (Schwartz, Bransford, & Sears, 2005). This means that while they can quickly and accurately apply procedures to familiar problems, they may fail to recognize opportunities to transfer their skills between problem types or to adapt procedures to new requirements. Hestenes and colleagues (1992) argue that “certain concepts and modes of reasoning must be developed before problem-solving instruction can be
effective” (p. 15). The distinction between efficiency and innovation in expert learners has similarly been characterized as that of routine and adaptive experts (Hatano, 1988). Students who are trained to become routine experts will become proficient in a well-structured domain but are expected to flounder when encountering an ill-structured domain. While these students are prepared to solve routine problems, they have not acquired the rich conceptual knowledge that will enable them to understand how and why a given procedure works (Hatano, 1988). Students who are trained to become adaptive experts, on the other hand, will learn to process information differently, taking an integrative view of their domain by attending to relationships between ideas, problems, or systems, rather than focusing on individual facts or phenomena (Richland & Begolli, 2016).

Why does this matter? Educating youth for innovation and expertise is a national priority: Former President Obama’s Strategy for American Innovation (2009) made these objectives a primary policy and education goal for the 21st century. Students who enroll in introductory physics courses, for example, are often preparing for careers in the applied sciences, such as engineering and the health professions. Physics knowledge is, in fact, required for performing well on the MCAT, the entrance exam for medical school. Efficient problem-solving and routine expertise will not serve these professionals well in the event that they encounter circumstances beyond their experience. Furthermore, when physics students pursue careers in physics, astronomy, or other natural sciences, they are poised to push beyond the frontiers of our knowledge – an accomplishment that can only be reached with the vision that grows from adaptive expertise and innovation.

In undergraduate Science, Technology, Engineering, and Mathematics (STEM) education, innovative methods are increasingly recommended to improve performance and
engagement, especially in large introductory classes from which many students fail or withdraw (Dykstra, 2005; Felder, Woods, Stice, & Rugarcia, 2000; Prince, 2004). Which innovative methods should educators use to help their students develop expertise in their fields of study? Instructional approaches that emphasize development of flexible knowledge and conceptual understanding can improve students’ ability to go beyond rote use of procedures to more sophisticated understanding of the principles that lie beneath them (DeCaro, 2016; Richland & Begolli, 2016; Rittle-Johnson, Star, & Durkin, 2017). The literatures on transfer and conceptual knowledge development offer insight into how to design instruction that will address this problem. Learning activities must capitalize on students’ prior knowledge and allow them opportunities to “transfer in” what they already know (Schwartz et al., 2005).

Additionally, learning assessments must be designed to be sensitive to the state of students’ early knowledge by measuring knowledge in development, not only knowledge that has reached an independent level of mastery (Schwartz, Lindgren, & Lewis, 2009; Schwartz, Sears, & Chang, 2007). In this way, learners can be assessed in their *zone of proximal development*, when they have yet to master the learning objectives, but are on their way to doing so (Vygotsky, 1987). Instruction can then be aimed at this zone of optimal learning. That means using formative assessment throughout the learning process, not just as a summative measure at the end. Low-stakes formative assessment can detect misconceptions that form early or suboptimal strategies that can be corrected in subsequent instruction (Andrade & Heritage, 2018).

Exploratory learning is one method that addresses both the need to activate prior knowledge and assess students’ learning before they are expected to have mastered course content (DeCaro, DeCaro, & Rittle-Johnson, 2015; DeCaro & Rittle-Johnson, 2012; Weaver,
Exploratory learning encourages students to grapple with new material before receiving direct instruction. One well-studied version of this approach is *productive failure*, in which students try (but fail) to solve problems before being equipped with correct solution strategies. This line of research has found benefits for conceptual knowledge development, primarily in middle and high school STEM classes (Kapur, 2010, 2011, 2012). Another version of this approach researched extensively in middle and high-school STEM, *invention*, presents students with contrasting cases and asks them to invent problem-solving procedures before direct instruction (Chin et al., 2016; Schwartz, Chase, Oppezzo, & Chin, 2011; Schwartz & Martin, 2004). What these approaches share in common is combining constructivist learning activities with teacher-led instruction in an untraditional sequence. This combination is decidedly better than using one or the other in isolation (Alfieri, Brooks, Aldrich, & Tenenbaum, 2011; Kapur, 2016; Mayer, 2004). A review of studies involving problem solving prior to instruction from 2006-2016 found benefits for conceptual knowledge and transfer when one of two conditions were met: (a) contrasting cases were used during problem solving, or (b) student responses were addressed before the canonical solution was presented (Loibl, Roll, & Rummel, 2016).

A meta-analytic review of learning through case comparison confirmed the efficacy of using contrasting cases during problem solving (Alfieri, Nokes-Malach, & Schunn, 2013). To better understand what processes, contexts, and measurement intervals resulted in successful learning from case comparison, the authors identified reliable moderators of the effect. Of particular note are the process moderators, which included the objective of using the cases and presentation of a principle. When the objective was to identify similarities between cases, as opposed to differences, students gained almost one standard deviation in learning. The reviewers
also found that providing learners with a principle *after* they compared cases led to greater learning outcomes than providing the principle *before* case comparison, or not at all.

The findings above, regarding the importance of conceptual understanding, formative assessment of learning, the effectiveness of exploratory learning, and the strength of case comparison informed the research design of this dissertation study. Its purpose and predictions follow.

**Statement of Purpose**

The purpose of the current study was to investigate whether and how exploratory learning could be structured around analogous problems to help students transfer their understanding of a previously-learned physics concept to a new, related concept. The goal was to create learning conditions that prevented students from falling prey to compartmentalization of knowledge and overreliance on procedural execution at the expense of broader conceptual understanding. I chose exploratory learning as the instructional approach to test because of its success at promoting conceptual knowledge and transfer in formal domains, such as math, statistics, and physics, as indicated by the cognitive and educational psychology literature. I employed exploratory learning in two formats, in an ecologically-valid classroom context through collaboration with the course instructor so that its true viability in an educational setting could be evaluated. The first exploration condition was structured around solving analogous problems, timed to occur before the professor presented the lesson’s target principle. Students were presented with a familiar course concept and a new, related concept in a pair of analogous problems to solve. The second exploration condition received only a hint to refer to the prior course concept when solving the new problem. A control condition learned about the analogy during lecture prior to solving the new problem but did not see the analogous problem when they
solved the new problem. This conservative control was intended to hold constant the effect of exposure to the analogy across conditions but allow for isolation of the effect of exploring analogies, specifically. The contribution of this design is that it allows comparison of two conditions that use analogy, one in the course of exploratory learning and one in the typical tell-then-practice instructional sequence. Furthermore, students exploring without analogy serve as a comparison condition to isolate the effect of analogy on exploratory learning.

I selected the specific content for the experiment because of its natural applicability to the aim of the research. Magnetic force was the third topic in a series of topics on force; thus, I believed students would possess an adequate level of prior knowledge to build upon in the lesson – particularly when primed by the use of analogous problems. Additionally, learning to solve vector addition force problems has been shown to be particularly difficult for physics students (Larkin, 1981). The vector addition of forces states that the total force on a system is equal to the sum of the forces on it due to all individual objects with which it interacts. Though this is a fundamental physics principle, identifying the forces in a system changes with problem type and interpreting the symbols becomes a challenge. Finding a method that supports students in meeting these challenges could have real instructional implications because these types of problems are included in nearly all physics curricula and require conceptual understanding to master.

**Current Study**

This study was designed to test two versions of an exploratory learning activity, in an undergraduate physics course, and to compare them with the more traditional non-exploratory approach of lecture followed by practice. Students acquired the requisite prior knowledge needed for the lesson of interest during classes that occurred earlier in the course sequence. In the first
few weeks of the course, the professor of record taught the topic of electric force. A few weeks later, the collaborating instructor introduced the new topic of magnetic force by using one of two exploratory activities or a tell-then-practice approach.

In the primary experimental condition (analogy-first), an exploration activity primed students to use a “force schema” by providing two force problems for them to explore simultaneously. This format created the conditions necessary for the students to engage in analogical reasoning, a mechanism theorized to be critical for successful problem solving (Holyoak & Thagard, 1989). Two cases were presented on the same page as a scaffold for learners to help them induce the general principle that underlies both cases: All forces are interactions between two things. Students were directed to apply their prior knowledge of one problem to solving the other. The first problem concerned a previously-taught topic (electric force), closely related to the new concept (magnetic force). The second problem concerned magnetic force.

By providing both problems to the students, I aimed to eliminate a cognitively taxing stage of analogical reasoning: the search of long-term memory for an appropriate analog to the current problem-to-be-solved. Students would not have to retrieve and evaluate candidate analogs for possible relevance to the problem at hand, accepting or rejecting them as they attempted to find a match. Having the source problem presented alongside the target problem also reduced the cognitive load inherent in exploration by constraining the solution space to only solutions that could be generalized from the first problem to the second.

The structural similarity of the two force problems was meant to aid students in mapping relations between them. Both were rectangular diagrams with equal angles. In the first case, the diagram depicted point charges in an electrical field; in the second, current-carrying wires. Both
problem solutions are obtained through vector addition. To prime students to approach the magnetic force problem using what they had learned about electric force, the instructor presented it as an exploration activity, before providing direct instruction on the new concept. Both of the exploration conditions received a framing instruction to use the two problems in conjunction to assist in solving them.

In a comparison experimental condition (explore-first), students received only the new topic problem to see if they could spontaneously transfer in previously-learned material with just a written direction to do so. Students in this condition also explored, but only using the magnetic force problem, with the direction to “use what you have learned earlier in the course about electric force to help you solve this problem.” This expansive framing technique is a typical approach for encouraging students to induce relationships and rests on the notion that students are more likely to transfer relevant prior knowledge to a new topic when they are primed to do so (Engle, Lam, Meyer, & Nix, 2012). Multiple studies have used the exploration before instruction paradigm in the past (see Loibl et al., 2016 for a review of the literature). Finally, students in both experimental conditions heard a lecture on magnetic force.

In a control condition (instruct-first), students heard a lecture on the new topic of magnetic force, and then practiced applying their new knowledge to the same magnetic force problem as used in the two exploratory learning conditions. These students did not engage in exploratory learning.

After the learning phases of the lesson, I measured all students’ subjective ratings of cognitive load during the learning activity. By doing so, I collected evidence that could help explain a null result. If my findings demonstrated that the two exploration conditions did not exceed the instruct-first condition in developing conceptual knowledge, one explanation to
consider would be that the exploration was too cognitively loading to benefit schema acquisition. Additional process-level measures, such as interest, utility-value, self-efficacy, and awareness of knowledge gaps were also administered to provide insight into when and why exploratory learning is beneficial. The dissertation aimed to answer four research questions.

**Research Question 1 (R1).** Do students learn a new concept best when exploring the topic before receiving instruction on it (two exploratory learning conditions), compared to receiving direct instruction on the topic before solving problems related to it (instruct-first condition)?

**Hypothesis 1 (H1).** I expected that completing the activity as exploration would be more difficult than completing it as practice. Therefore, I hypothesized that the instruct-first students would become proficient problem-solvers and exhibit the most accuracy during the problem-solving activity. Conceptually, however, I did not expect them to have the deep grasp of the underlying general force principle that was at the heart of the lesson. By the end of the lesson, once all students had heard the lecture and received correct answers to the problem-solving activity, I predicted that all three group averages would be similar for the posttest problems assessing procedural knowledge (Loibl et al., 2016). However, I predicted a main effect of question type, such that overall scores on the procedural knowledge questions would be higher than overall scores on the conceptual knowledge questions. Furthermore, I hypothesized that the exploration groups would exceed the instruct-first group in conceptual understanding.

**R2.** Does solving analogous problems during exploration improve conceptual understanding, as compared with a simple framing instruction to connect two topics?

**H2.** I hypothesized that having only one problem to explore before receiving instruction on the topic would create a less-constrained problem space than having two analogous problems.
This lack of constraint could make it more difficult to navigate. Without the benefit of an analogous and familiar problem, novice students might become frustrated or experience high cognitive load while exploring. However, some students might rise to the challenge and benefit from the process of exploring the problem space. They would experience the exploration as a “desirable difficulty,” activating prior knowledge and applying it appropriately to the novel problem (Bjork & Bjork, 2011). While I hypothesized that students in the analogy-first condition would experience productive failure, I predicted that students in the explore-first condition might experience failure that was comparatively unproductive. I hypothesized that by allowing students to search the problem space for what they could bring to bear from prior knowledge, they would recognize the analogous representation of the two problem types. This recognition was hypothesized to help the students to identify the deep underlying structure common to both types of problems.

I hypothesize that, due to the lack of the analogy scaffold, the explore-first students will struggle more than the analogy-first students to accurately solve the problem during exploration. I anticipate that actually working through analogous problems will better support students because the focus on comparing two isomorphic problems will constrain the problem space and provide students with an instructional scaffold. But perhaps the most important contribution of this study will be to examine differences in performance of the two exploration groups on conceptual knowledge development. This is a new question, unanswered by the literature to date. Although I hypothesize that the analogy-first group will have more success on the activity than the explore-first group, I can only tentatively predict that this will be enough to produce a superior conceptual knowledge base at posttest.
R3. Does any positive effect of exploring the new concept before instruction develop or persist over time?

H3. There may be a “sleeper effect” in which the benefits of exploratory learning do not emerge until after the effect has time to take hold on the learner (Klahr, 2009). This would suggest that any differences between groups may not be evident until the retention test two weeks after the instructional intervention. It may also be the case that learning from exploration is more durable than learning through a tell-than-practice approach, in which case, scores would be expected to remain consistent over time.

R4. What effect does exploratory learning, with and without analogous problems, have on interest, utility value, self-efficacy for problem solving, awareness of knowledge gaps, and cognitive load, compared to an instruct-first condition?

H4. The distinct hypotheses for each process measure follow.

Interest. Previous research has shown that situational motivational beliefs (e.g., interest, enjoyment) are associated with student academic achievement in engineering undergraduate education. Specifically, increases in interest in engineering are associated with higher GPA (Liu, Snyder, & Ralston, 2015). However, my question is whether a particular instructional approach – exploratory learning – can elicit interest in learning a new concept. Based on previous research (Weaver et al., 2018), I hypothesize that exploring the concept of magnetic force will produce higher interest ratings in the learning activity than the direct instruction format will produce. I do not have specific hypotheses about whether ratings will differ between the exploration groups, but I tentatively predict that, because it will be more focused on comparison, the analogy-first group will endorse greater interest than the explore-first group. Those results would be important
findings in their own right, because of the relationship between intrinsic motivation and academic achievement.

**Utility value.** Generally, subjective task value (e.g., utility, attainment) is associated with performance in a subject and choice about whether to pursue a subject (Eccles & Wigfield, 1995; Wigfield & Eccles, 2000). Liu, Snyder, and Ralston (2015) found that attainment value for engineering was associated with higher GPAs in undergraduate engineering. Less is known about what precipitates feelings about task value. I will explore how different types of exploration versus the typical instructional sequence facilitate the development of differences in utility value for the lesson, the course, and the subject.

**Self-efficacy.** I anticipate that self-efficacy for learning about magnetic force will be impacted by instructional condition. This analysis is exploratory, as I have not seen literature to guide my predictions about the effect of group membership. It may be that exploring the concept and then listening to a lecture about it results in greater feelings of self-efficacy for learning about the topic than hearing the lecture first and then practicing the problems; however, I can also imagine the opposite result, whereby being fully prepared to solve the problems before trying to leads to greater self-efficacy at the end of the lesson.

**Knowledge gaps.** In a study comparing an invention condition to a worked example condition (Glogger-Frey, Fleischer, Grüny, Kappich, & Renkl, 2015), perceptions of knowledge gaps were greater in the invention condition ($d = .70$, a medium effect); however, that is not a direct parallel to the current study’s design, in which no worked examples will be provided in any condition. It is instructive, though, because it indicates that conditions in which students are left to discover principles for themselves experience greater knowledge gaps than conditions in which learners are provided with all they need to know. Therefore, I predict that because the
instruct-first group in the current study is most similar to the worked example group in Glogger-Frey and colleagues’ research, this group will be least aware of knowledge gaps. Certainly, they are expected to have the most fluency during learning because they will be provided with all the information they need during lecture, before practicing the problem. The analogy-first condition has the next highest level of scaffolding of knowledge, due to being provided a familiar analogous problem; therefore, this may induce them to recognize what it is they need to know better than the explore-first group, resulting in greater perceived knowledge gaps. The explore-first group, however, has no guidance other than a hint to use prior knowledge, so they may simply flounder without identifying their gaps, or they may have the greatest awareness of them. This question is exploratory.

**Cognitive load.** Having only one problem to explore before receiving instruction on the topic is predicted to be more cognitively loading than exploring analogous problems, due to lack of constraint that may make problem solving more difficult to navigate. In Newman and DeCaro’s (2018) recent study, order of instruction (explore-first or instruct-first) did not result in differences in cognitive load ratings by students in undergraduate psychology. However, students who received a worked example, either while exploring or practicing problems after instruction, reported less cognitive load than students who invented in either instructional sequence. Based on those findings, I would tentatively predict that the analogy-first condition will experience less mental effort than the explore-first condition. I consider it an open question whether the analogy-first group will experience more or less cognitive load than the instruct-first group.
CHAPTER II: LITERATURE REVIEW

In this chapter, I will review the relevant bodies of literature on which the current study rests: exploratory learning, analogous reasoning, and attendant benefits of these instructional approaches to conceptual knowledge and motivational factors. In addition, I will discuss the primary criticism of exploratory approaches to learning and describe potential disconfirming evidence of my hypotheses.

In the previous chapter I proposed that one way to address the problem of students’ impoverished conceptual understanding is through improved instruction. Teachers’ and students’ goals are the same: for students to learn new content in a way that “makes it stick” so that it can be drawn on again in the future. Applying knowledge, reasoning, or skill in novel contexts is defined as transfer (Loibl et al., 2016) and has long been considered the primary objective of formal education (Barnett & Ceci, 2002; Detterman, 1993). However, designing instruction that maximizes the likelihood of transfer can elude instructors. One reason this may be so is because instructors possess a level of expertise that makes obvious to them the underlying connections between concepts in their field (Goodstein, 1999; National Research Council, 2000). Novices, on the other hand, do not automatically see the big picture (Chi et al., 1981). Introductory physics students are not necessarily novices in the domain, but they are not yet experts (Chi et al., 1981). They have been characterized by some as “competent beginners” (Mason & Singh, 2013; National Research Council, 2000). Introductory physics courses should capitalize on and extend students’ previous content experience by utilizing teaching methods that train their students to think like experts in the domain.

**Exploratory Learning**
A method that has been shown to increase both conceptual knowledge (Weaver et al., 2018) and transfer in physics (Schwartz et al., 2011) is *exploratory learning*. Exploratory learning is one of many instructional approaches that creates an environment in which students are cognitively active during the learning process. In this dissertation, exploratory learning will refer to methods that involve students in grappling with a problem before receiving explicit instruction on the topic (Weaver et al., 2018), including invention with contrasting cases (Belenky & Nokes-Malach, 2012; Chin et al., 2016; Glogger-Frey et al., 2015; Schwartz et al., 2011; Schwartz & Martin, 2004), inversions of the typical instructional sequence that have students solve problems before being instructed on how to do so (DeCaro & Rittle-Johnson, 2012; Loehr, Fyfe, & Rittle-Johnson, 2014; Loibl et al., 2016), and productive failure (Kapur, 2010, 2011, 2012; Kapur & Bielaczyc, 2012; Loibl & Rummel, 2014). Importantly, all these methods include both an exploratory phase and an expository instructional phase. In the exploratory phase, a degree of difficulty is embedded in the learning task that is intended to make learning durable in the long term. In the instructional phase, learners’ misconceptions and questions are addressed and resolved. Exploratory learning is constructivist in nature because it relies on learners constructing their own knowledge through interaction with instructional materials and often, their peers. However, it diverges from early constructivist methods, such as pure discovery learning (Mayer, 2004), by emphasizing the critical role of direct instruction alongside exploration (Klahr & Nigam, 2004).

An essential mechanism for successful exploratory learning is activation of relevant prior knowledge. Activating prior knowledge has been hypothesized to be critical to helping students learn a new topic (Schwartz et al., 2007). In exploratory learning students are presented with new concepts, problems, or information which require them to ply long-term memory to address the
current object of study. In the process of accessing an existing schema, students connect new information to prior knowledge, and schemas may be updated or expanded. Students are better prepared to learn from subsequent instruction when they are already thinking about the material and are ready to make connections to what they already know (Schwartz & Bransford, 1998).

In Weaver and colleagues’ (2018) study conducted with over 200 undergraduates studying introductory physics, students were randomly assigned to learn about electric potential in one of two different instructional sequences. In the *explore-first* condition, students were given an unfamiliar problem to solve in small groups and told that they were not expected to arrive at the correct answer but were to apply prior knowledge to attack the problem. After fifteen minutes, the professor delivered a lecture on the concept of electric potential and went over the correct answers to the activity. In the *instruct-first* condition, students received the lecture first and then solved the same problems as practice, in line with a more traditional approach. After all students had debriefed the activity with the professor, they took an instructor-created quiz, which included procedural and conceptual items.

We found that the students who explored the problem before receiving instruction on how to solve it struggled to solve it accurately but ultimately outperformed the direct-instruction comparison group on measures of conceptual knowledge on the quiz. In addition to scoring significantly higher on conceptual understanding than the instruct-first group, students in the explore-first condition had equivalent problem-solving accuracy on the quiz. This difference was found despite the explore-first group’s worse (Study 1) or equal (Study 2) performance on the learning activity. Furthermore, students in the explore-first condition reported equal (Study 1) or greater (Study 2) interest and enjoyment in the lesson. We interpreted our results to mean that, though exploring the problems before receiving instruction was experienced as more difficult
than practicing them after the instruction, the sequence did not damage students’ intrinsic motivation for learning the content, nor did it harm their learning outcomes. On the contrary — overall averages were higher for conceptual knowledge in the exploration group.

Schwartz and colleagues (2011) reached a similar finding with a younger set of physics learners. In their study, eighth-graders were taught the physics concept of density. In a *tell-then-practice* condition, the students learned the concepts and formulas before practicing on a set of contrasting cases. In the *invention* condition, students used the contrasting cases to invent formulas before they learned them and were only instructed on them later. Both groups performed equally well at solving word problems using the formulas, but the invention condition exceeded the tell-then-practice condition at transferring their understanding of the ratio structure of density to a different physical phenomenon, spring constant, that also had a ratio structure.

Delaying instruction until after an initial problem-solving phase has benefited learning in domains other than physics, including math, statistics, and psychology (e.g., DeCaro & Rittle-Johnson, 2012; Kapur, 2012, 2014; Schwartz & Bransford, 1998; Schwartz & Martin, 2004; Westermann & Rummel, 2012). DeCaro and Rittle-Johnson (2012) had 159 second- through fourth-grade students solve problems either before (*solve-instruct* condition) or after (*instruct-solve* condition) being taught the concept of mathematical equivalence. All students’ procedural knowledge improved after instruction and was equal across groups. However, students in the solve-instruct condition demonstrated better understanding of the meaning of the equal sign and the structure of equations on an immediate posttest and two weeks later on a retention test.

A recent meta-analysis of learning-from-failure studies (Darabi, Arrington, & Sayilir, 2018) identified twelve experimental studies that found superior learning outcomes for students who attempted to solve problems without instructional scaffolds compared to students who had
instructional support before solving the problems. Learning from failure has a long history, from its roots in the trial and error method of learning to Piaget’s theory of cognitive disequilibrium (Darabi et al., 2018). The most common incarnation of this approach in the recent literature is known as productive failure (Kapur, 2008). Productive failure results when learners do not succeed at the learning task they have been given, but after an instructional phase, they benefit more than their peers who experienced the instructional phase before attempting the learning task. Essentially, this approach flips the typical instructional sequence on its head: Students attempt complex, ill-structured problems before they have been given scaffolds to help them solve them. This approach has led to ultimately superior learning outcomes in the domains of math, statistics, and science, with learners from junior high to high school (Kapur, 2008, 2010, 2012, 2014). The prevailing explanation of productive failure’s benefits is that students search the problem space more broadly than they would if the space had been constrained for them. In doing so, they represent the problems they are attempting to solve in multiple ways and try a variety of solution strategies. Though they do not succeed in reaching the correct solutions before they receive instruction, after instruction provides them with the information they were lacking, they proceed to consolidate knowledge more effectively than their peers on related posttests (Kapur, 2014). Effect sizes for the findings are in the medium range ($g = 0.43$, 95% CI [0.19, 0.68]; Darabi et al., 2018).

These methods (invention, flipping the instructional sequence, and productive failure) have been found to benefit learners’ conceptual knowledge and transfer without sacrificing their procedural competency (Loibl et al., 2016). To use these methods in the classroom, teachers must let go of some control over the learning process and allow students to struggle with new material. It is through the struggle, and its resolution, that meaningful learning occurs. To feel
comfortable with this process, teachers will surely want to know what structures will work best for their students.

**Analogical Reasoning during Problem Solving**

Two recent literature reviews (Kalyuga & Singh, 2016; Loibl et al., 2016) concluded that one of the methods that works best during exploratory learning is to allow students to invent problem-solving procedures using contrasting cases. Comparison methods have also been found to aid in exploration (Rittle-Johnson, Star, & Durkin, 2009). Approaches that involve students in comparing and contrasting types of problems or problem solutions benefit learning in multiple ways. The first is that comparing solution methods can facilitate knowledge development (e.g., mathematics) and procedural flexibility (Rittle-Johnson & Star, 2007). A second reason that viewing more than one problem can aid in problem solving is that simultaneous comparison of problems can invoke analogical reasoning (Begolli & Richland, 2016; Catrambone & Holyoak, 1989).

Analogical reasoning occurs in the course of problem solving when two or more problems share a relational structure that allows a learner to transfer what they understand about the relations at work in the one problem to another. The process of using the solution to one problem to attempt to solve another involves three steps. Theorists of analogical transfer agree that perceptions of similarity between problems is the initial stage of analogical reasoning, and thus analogous problems must be designed to highlight similarity to make it salient to a learner (Vosniadou & Ortony, 1989). In this first step, learners must recognize that a familiar problem (the source) can be useful in solving a novel problem (the target). Theories differ with respect to what occurs next, with some favoring automatic explanations, and others strategic explanations, of how learners proceed to extrapolate an abstract solution principle (Reeves & Weisberg, 1994).
In both structure-mapping theory (Gentner, 1983) and pragmatic schema model (Holyoak and colleagues, 1985, 1989), a learner must identify structural features of the source problem and then evaluate the target problem for those same features. Structural features are specified as those integral to the relational structure of the problem and are distinguished from surface features, which are superficial elements that have no bearing on the solution strategies. Focus on structural features rather than surface features is one way that expert problem solvers are distinguished from novices in their domain (Chi et al., 1981; Mayer, 1992). In this second step, features relevant to solving the target problem are found and learners map these features from one problem to the other. The third step is to extract the commonalities of both problems to induce a new knowledge structure, or schema. Two dominant theories posit that strategic abstraction of a common schema (Gick & Holyoak, 1980, 1983), or deep structure (Gentner, 1983, 1989), is what mediates transfer between analogs. When two (or more) similar problems or cases are given (i.e., analogs), a learner is likely to induce this content-based generalization. When the schema has been abstracted to the point that it can be applied to a new problem, learners are more likely to apply what they have learned from one analog to another (Engle et al., 2012).

Using analogous problems has been shown to be most successful for transfer of knowledge, when hints are given to consider both problems (Gick & Holyoak, 1980), or when more than one analogy is given (Gick & Holyoak, 1983). Use of contrasting cases has been found to improve recognition of deep structure in physics, particularly when helping learners to induce a general rule (Kuo & Wieman, 2015). As mentioned previously, introducing an overarching principle after learners have compared multiple examples of an idea has been shown to be most effective for learning (Alfieri et al., 2013).
Expansive Framing

Analogical reasoning, therefore, aids learners in inducing a generalized schema that can transfer between problems. Engle, Lam, Meyer, and Nix (2012) offered another, simpler, mechanism for transfer that may potentially be equally powerful for accomplishing the aims of this study. These scholars proposed a theoretical framework to explain mechanisms for successful content transfer. One such mechanism is to give learners an explicit prompt or hint to think of their knowledge as useful across time and contexts, what they call “expansive framing” (Engle et al., 2012). *Expansive framing* occurs when a teacher communicates to learners that what they are learning will be useful again in the future. This type of framing is contrasted with *bounded framing*, which is when no effort is made to encourage students to see the applicability of what they are learning to multiple future contexts. Framing can reach forward in time, as just outlined, or backwards in time as in a direction to think of previous relevant cases when learning something new. Framing, therefore, is a social interaction that can set up learners to integrate new ideas with prior and future knowledge.

The goal of this study is to determine how best to help learners recognize a general principle that underlies several physical phenomena. Expansive framing could enable students to meet that objective by one of two pathways, proposed by Engle and colleagues (2012). The first possibility is that a verbal instruction to think of what has been learned in an earlier lesson in the course will help students learn something new in the present because students will connect the two settings, thus fostering transfer between them. In this case, students would view their prior knowledge as useful in the current setting. The second pathway would be that expansive framing could promote student authorship by helping students view prior knowledge as “both relevant for
current learning and desired socially” (Engle et al., 2012, p. 220). In this case, students would transfer in their prior knowledge during learning.

My research questions concerned how to best leverage instruction to enhance recognition and transfer of underlying general principles. The study was designed to determine whether analogies need to be invoked directly through exploration of analogous problems, or indirectly, with just an expansive framing direction to use an analogous problem type to solve a new problem.

**Motivation for Learning and Perceptions of Difficulty**

Analogical reasoning and transfer are difficult cognitive processes. More than thoughtful instructional design is necessary to support learners in successfully engaging in these processes. For one, teachers must motivate students to engage in effortful and meaningful learning by cultivating interest and value in the tasks of the discipline (Deci, Vallerand, Pelletier, & Ryan, 1991). This should involve boosting students’ self-efficacy in the area of study and convincing them of its utility in the future (Wigfield & Eccles, 2000). Without this investment in the work that is required to learn at a deeper level and attendant confidence in their abilities, students may not be prepared to succeed under the instructional conditions that will move them beyond procedural execution to conceptual understanding (Hatano, 1988).

Exploratory learning has been found to activate learners’ attention, prior knowledge, interest, and motivation. These benefits are thought to explain its role in fostering learning. However, learners also experience more difficulty during exploratory learning than in instructional methods that equip them with all relevant information upfront. Therefore, the process-level measures I asked students to report in the current study (i.e., interest, utility value, self-efficacy, awareness of knowledge gaps, and cognitive load) were intended to shed light on
the comparative learning experience of students engaged in these different instructional approaches.

**Interest.** Instructional approaches that encourage students to explore problems may give rise to greater motivation to learn the material (Weaver et al., 2018; Wise & O'Neill, 2009). For example, situational interest created by involving students in challenging problem solving may spark the development of an emerging *individual interest* – a sustaining interest in the material (Hidi & Renninger, 2006). The downstream effects of a well-developed individual interest may lead students to learn more in the long run and persist longer in the face of failure. As previously mentioned, Weaver and colleagues (2018) found that undergraduate physics students who explored a novel concept before hearing a lecture on that concept enjoyed their lesson more than students in the same class who heard the lecture before solving the same problems, *even when they struggled more to solve those problems*. What might account for the greater intrinsic motivation experienced by the students who explored? One explanation is that the exploration phase created *situational interest* in the students who explored. Situational interest refers to both a high quality of attention and attendant affective reaction that are triggered by the environment (Hidi & Renninger, 2006). Though solving the problem was more difficult for students in the exploration condition (as evidenced by their lower success rate), the instructional environment – or the challenge in particular – created interest in that moment.

Interest has been found to have positive effects on affect and cognition (Hidi & Renninger, 2006). Whether exploration can create situational interest in the content of a lesson is an empirical question. It is certainly plausible that interest may focus learners’ attention on their classwork to a degree that enables them tune in differently to instruction. Or, it may be that exploring creates disequilibrium in learners, such that they feel a “need to know” (Hatano, 1998;
and enter the expository phase of the lesson with more interest in the concept.

**Utility value.** Expectancy-value theory (Wigfield & Eccles, 2000) is a broad motivational theory that seeks to explain people’s choices of particular tasks, their effort and persistence, and performance on these tasks. One component of the theory’s construct of subjective task value is utility value. When students expect their learning to be relevant beyond their immediate situation, they develop a sense of its *utility value*. Lessons that integrate concepts from course material (Engle et al., 2012), or that deeply engage students in learning (Hulleman, Durik, Schweigert, & Harackiewicz, 2008), may help students appreciate the value of the lesson. Utility value predicts interest and performance (Hulleman et al., 2008).

**Self-efficacy.** Furthermore, when students come through a challenge, they may develop a greater sense of self-efficacy in the domain, which is important to learning (Bandura, 1982; Zimmerman, 2002). This study aims to help establish whether the difficulty encountered during exploratory learning is motivating or demotivating, as measured by students’ self-efficacy ratings for learning about the concept after the lesson.

**Knowledge gaps.** Preparedness to learn from future instruction may depend on a learner’s awareness of his or her level of understanding of the topic. For example, Glogger-Frey and colleagues (2015) found that the more undergraduate students perceived gaps in their knowledge, the more attentively they focused upon subsequent instruction. Unfortunately, the metacognition literature amply demonstrates that when learners experience fluency during learning they are often susceptible to overinflated judgments of learning, or “illusions of knowing” (Bjork, 1994; Dunlosky & Rawson, 2012; Koriat, 1998). There is some evidence that direct instruction can promote such fluency and overconfidence, contributing to decreased effort.
and poorer learning outcomes (DeCaro & Rittle-Johnson, 2012; Renkl, 1999). In contrast, because exploration is more error-prone, it may enhance monitoring of learning and reduce overconfidence, preparing students to pay closer attention to subsequent instruction (DeCaro & Rittle-Johnson, 2012; Koriat & Bjork, 2005). Newman and DeCaro (2018) found evidence that students who explored the concept of statistical variance before receiving instruction had greater awareness of knowledge gaps than students who were taught the concept before applying it. My dissertation study seeks to clarify the effects of exploratory learning activities on recognizing knowledge gaps.

**Cognitive load.** Perhaps the most consistent criticism of exploratory learning approaches, and sometimes constructivist approaches in general, has come from cognitive load theorists (e.g., Clark, 2009; Kirschner, 2009; Mayer, 2004; Mayer, 2009; Sweller, 2009; Sweller, Kirschner, & Clark, 2007). Cognitive load theory (CLT) dates back more than three decades and has generated extensive applications of research on cognitive processes into instructional design (Paas, Renkl, & Sweller, 2003; Sweller & Chandler, 1991). The premise at the heart of CLT is that human cognitive architecture works most effectively when learners maximize their cognitive resources for mental processing in the most efficient and direct way possible (Kirschner, Sweller, & Clark, 2006). Reducing cognitive load by improving directions, integrating information, or eliminating distracting features, has been shown to free up working memory capacity for learners to direct at apprehending the target schema (Mayer & Moreno, 2003). CTL’s contributions to education are numerous; for example, it has led to development of instructional materials that are more focused on essential processing and trimmed of extraneously-loading images, text, or preliminary steps (Chandler & Sweller, 1991; Mayer & Moreno, 2003).
Traditionally, cognitive load theorists have leveraged claims that instructional approaches such as productive failure, invention, discovery learning, or problem-based learning overtax learners by involving them in searching for information or procedural solutions (Hsu, Kalyuga, & Sweller, 2015). The theorists amassed evidence that worked-examples or direct instruction before problem-solving lead to superior learning outcomes (Cooper & Sweller, 1987; van Gog, Paas, & van Merriënboer, 2006). However, in a recent reversal, some prominent cognitive load theorists conceded that CTL assumed only one instructional goal for all learning activities: the acquisition of domain-specific content schemas (Kalyuga & Singh, 2016). When the goal is different (e.g., a pre-instructional goal or a higher-level goal) then it may be appropriate to use more search-based exploratory activities (Kalyuga & Singh, 2016). I would add to this that the narrow focus on academic outcomes – at the exclusion of motivational outcomes – eliminates examination of some of the benefits offered by exploratory approaches.

Hsu, Kalyuga, and Sweller (2015) conducted two experiments in secondary school physics classrooms in Taiwan, measuring students’ ratings of cognitive load after one of two instructional sequences: worked-example-problem or problem-worked-example. In Study 1, they found a main effect of the sequence, such that students in the worked-example-problem sequence reported less cognitive load than those in the problem-worked-example sequence. However, simple effects tests showed that this effect only held for students with less knowledge of the topic; for more knowledgeable students there was no difference in cognitive load. In Study 2, the authors failed to find any main effect of sequence on cognitive load. This study differed from the first in that students had no prior knowledge of the topic to be learned; however, they provided half of the students in both sequences with guidance on the physics principle. Students in the problem-worked-example condition reported less cognitive load if they received the guidance
during their initial problem-solving phase compared with students who solved the problem without the guidance. This interaction suggests that students benefit from more scaffolding during problem solving, such as the guidance analogous problems could potentially provide.

In Table 1 I have classified the research discussed above into four dimensions. My dissertation study synthesizes two lines of research (exploratory learning and analogous reasoning) and employs a factorial design to investigate whether the benefit of using exploration and analogy stacks, above and beyond using one or the other. Thus, the study’s three conditions (bolded) can be found in three of the four dimensions.

Table 1

<table>
<thead>
<tr>
<th>Exploration</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analogy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>analogy-first condition</td>
<td>instruct-first condition</td>
</tr>
<tr>
<td></td>
<td>Invention with contrasting cases literature (see Schwartz and colleagues).</td>
<td>Comparing and contrasting literature (see Rittle-Johnson and colleagues). Analogous reasoning literature (see Gentner and colleagues, Holyoak and colleagues).</td>
</tr>
<tr>
<td>No</td>
<td>explore-first condition</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inversion of instructional sequence (see DeCaro and colleagues). Productive Failure literature (see Kapur and colleagues, Roll and colleagues).</td>
<td></td>
</tr>
</tbody>
</table>

Reviewed Research of Use of Exploratory Learning and Analogies to Introduce New Material
CHAPTER III: METHOD

The task motivating this research was to test the efficacy of, and hone, an instructional strategy – exploratory learning – in a classroom experiment during which students attempted to use analogies to recognize principles or general rules underlying and connecting their course material.

Participants

Participants were 171 undergraduate students (27% female) at the University of Louisville enrolled in two sections of an Introductory Electricity, Magnetism, and Light (Intro II) course. The course is the second in a calculus-based physics sequence; students were engineering majors. Students were included if they attended class on the day of the intervention and turned in the learning activity \((N = 189, 83\% \text{ of 228 enrolled students})\), and took both the posttest \((N = 184)\) and the retention test \((N = 171)\). Within each course section students were randomly assigned to one of three conditions: *analogy-first* \((n = 62)\), *explore-first* \((n = 55)\), or *instruct-first* \((n = 54)\). No procedures for obtaining consent were required by the exempt status granted by the Institutional Review Board at the University of Louisville to the Principal Investigator to conduct research in regular educational settings.

Missing data. There were no systematic missing data between learning activity, posttest, and retention test. Five students who left class before the posttest commenced were distributed equally among the three conditions and over the range of learning activity scores. Fourteen students who did not attend class on the day of the retention test were also equally distributed between conditions, and only one had also left class early on the day of the posttest. A flowchart of participants through each stage of the experiment is shown in Figure 1.
Figure 1. Participant flow through the intervention. This flowchart is adapted from the Publication Manual of the APA, 6th edition.

**Power analysis.** Effect sizes in the small range were consistently found in previous studies (Alfieri et al., 2011; Darabi et al., 2018; Weaver et al., 2018), so 0.25 was estimated to be the upper limit of a reasonable effect size to expect in this study. Power analysis was conducted in G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009) to determine a priori the necessary
sample size to find an effect size of .24, specified by Cohen (1988). The lowest power ratio that was deemed to be acceptable was to make commission of a Type II error four times more likely than a Type I error, and thus, Type II error threshold was set at $\beta = .20$. Accordingly, power was calculated at $(1 - \beta) = 0.80$. Type I error probability was set at $\alpha = 0.05$ for the omnibus $F$-tests. The total sample size returned was 234, which was determined to be close enough to the 228 students enrolled in the two sections of the course.

After the intervention, power analysis was run again, post-hoc, to ensure that the actual sample size that turned in complete data for the study would suffice. A sensitivity analysis was performed to compute the required effect size that would have to be present in the data in order to have enough power to detect it in our sample of 171. Number of groups was entered as three, one for each condition; number of measurements was entered as four, one for each of the knowledge subscales at posttest and retention test. Correlations among the repeated measures (subscales) was estimated to be .44, calculated by averaging the zero-order correlations between the two knowledge subscales at posttest and retention test. In order to run a mixed-factorial ANOVA with both within-subjects and between-subjects factors, the returned effect size was calculated to be 0.28 (according to Cohen, 1988). This meant that the study could be underpowered if an effect of this magnitude was not found.

Materials

**Prior Instruction.** Students in Intro II had studied electrostatics in the first few weeks of the semester and in the previous semester. In those earlier lessons, students solved problems asking them to calculate net electric force among three point-charges. The first two course exams, which measured this concept, were collected. Scores were averaged to use as a pretest of students’ prior knowledge of force. One of the new concepts taught in Intro II was magnetic
force, a concept closely related to electric force. The lesson of interest in this study was the first in the unit on magnetism.

**Learning Activity.** Students in all three conditions completed the learning activity in self-selected small groups, as collaborative learning has been found to be beneficial for STEM learning (Pai, Sears, & Maeda, 2015; Springer, Stanne, & Donovan, 1999). Constructivists have long acknowledged the benefit of social interaction during the learning process; Piaget (1973) and Vygotsky (1987) both highlighted the role that others play in an individual’s cognitive development. Peer-to-peer collaborations are opportunities for modeling thinking, learning from more knowledgeable individuals, and receiving immediate feedback on ideas (cf. Mazziotti, Loibl, & Rummel, 2014). Students were given 20 minutes to complete the activity. The activity required students to determine the magnitude and direction of the net magnetic force on a wire that sits parallel to two other wires arranged on the corners of a rectangle. Student work was coded in four steps, shown in Table 2. I determined this scheme a priori with the physics professor. The coding scheme emphasized correctly setting up the problem, as opposed to solving it correctly.

Table 2

**A Priori Coding Scheme for Learning Activity**

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
<th>Credit awarded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1*</td>
<td>Set up a free-body diagram</td>
<td>0, 1</td>
</tr>
<tr>
<td>Step 2</td>
<td>Use diagram to set up component equations</td>
<td>0, .5, 1</td>
</tr>
<tr>
<td>Step 3*</td>
<td>Plug in to component equation and/or calculate magnitude</td>
<td>0, .5, 1</td>
</tr>
<tr>
<td>Step 4</td>
<td>Solve for magnitude and direction</td>
<td>0, 1</td>
</tr>
</tbody>
</table>

*Note. *The exploration conditions had to rely on making guesses to perform these steps. 0=no credit, .5=partial credit, 1=full credit.
The problem(s) to solve during the learning activity were presented to students in packets (see Appendix A). This format was chosen to ensure that the students in the same classroom who were assigned to different conditions were not aware of the differences between their instructional materials. The packets also ensured that the instructor was blind to student assignment to the two exploration conditions, which shared a classroom. All students received the same oral instructions from the professor and were told that the activity was for research purposes. When they opened their packets, they saw either two analogous problems (**analogy-first**) or one new problem (**explore-first** and **instruct-first**). Each group worked from a single packet and recorded its answers on one solution sheet. The professor emphasized to students that they should use the packet as a place to document and refine their thinking about the activity. Packets in all three conditions began with a cover sheet presenting “Important Tips for Working in Groups.” The tips emphasized that students should select roles for themselves to take on a particular function in the group (e.g., recorder, facilitator) and that students should contribute questions, ideas and guidance to their group.

**Analogy-first condition.** Students in both exploration conditions began by reading a section on tips for exploration. These tips emphasized that students should focus on reasoning through the problem(s) and not on arriving at a correct solution. They were encouraged to consider a range of strategies or solutions. Students in the analogy-first (AF) condition were given two diagrams and asked to solve a problem represented by each. The first problem
concerned electric force and should have been familiar to students from previous coursework.

Electric force problem:
Three point charges are arranged on the corners of a rectangle, as shown in the figure below. Charge 1 has a charge of +1 nC. Charge 2 has a charge of −2 nC. Charge 3 has a charge of +3 nC. What is the magnitude and direction (relative to the +x axis) of the net electric force exerted on charge 1 due to the other two charges?

![Electric force diagram](image)

*Figure 2. Electric force problem, seen only by students in the analogy-first condition.*

The second problem concerned magnetic force, a concept to which students had not yet been exposed. Each diagram contained three points arranged on a rectangle. In the electric force problem, these points represented electrical point charges (see Figure 2). In the magnetic force problem, these points represented current-carrying wires (see Figure 3).

Magnetic force problem:
Three parallel wires are arranged on the corners of a rectangle, as shown in the figure below. Wire 1 carries 1 A of current into the page. Wire 2 carries 2 A of current out of the page. Wire 3 carries 3 A of current into the page. What is the magnitude and direction (relative to the +x axis) of the net magnetic force per length exerted on wire 1 due to the other two wires?

![Magnetic force diagram](image)

*Figure 3. Magnetic force problem, seen by students in all three conditions.*
The placement of the three points (charged particles in the first problem and current-carrying wires in the second problem) and distances between them were the same in both diagrams. In both cases, students were asked to determine the force exerted on the point (charge or wire) in the top-left-hand corner of the rectangle. The activity was designed to meet several instructional goals: (a) to activate students’ prior understanding of electric force; (b) to connect this prior knowledge to the new concept of magnetic force; and, (c) to highlight that vector addition is used to solve for forces in both electric and magnetic scenarios. The structural alignment of the analogous problems was intended to focus students’ attention on relational information common to both problem types. A comparison of the structural features in the study problems is displayed in Table 3. Surface features were absent from the problems to reduce learners’ extraneous cognitive load (Mayer & Moreno, 2003). Students were not explicitly instructed to map from one problem to the other but read the direction to “Use what you have learned earlier in the course about electric force to help you solve the magnetic force problem.”

Previous research shows that simultaneous presentation of problems or cases (as opposed to sequential) is sufficient to stimulate comparison (Begolli & Richland, 2016; Christie & Gentner, 2010).

Table 3

Comparison of Analogs in Current Study

<table>
<thead>
<tr>
<th>Source problem</th>
<th>Target problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Force</td>
<td></td>
</tr>
<tr>
<td>Electrostatic</td>
<td>Magnetic</td>
</tr>
<tr>
<td>Structural features</td>
<td>interaction between point charges</td>
</tr>
<tr>
<td></td>
<td>interaction between currents</td>
</tr>
<tr>
<td></td>
<td>current into the page</td>
</tr>
<tr>
<td></td>
<td>current out of the page</td>
</tr>
<tr>
<td>Solution</td>
<td></td>
</tr>
<tr>
<td>Magnitude and direction of net electric force</td>
<td>magnitude and direction of net magnetic force per length</td>
</tr>
</tbody>
</table>
The intention of providing the electric force problem was so that AF students would be able to use it as a guiderail for solving the magnetic force problem; this intention rested on the assumption that students would be able to complete the four steps first for the electric force problem (which was not scored) before attempting them for the new problem type. Examining what students did in each step allowed me to analyze how students made use of analogical comparison. For instance, in Step 1, if students mapped the rule they knew for electric force to magnetic force, their diagrams would be drawn backwards because of the transverse relations in these two systems. Therefore, they were not expected to draw an accurate diagram. Similarly, in Step 2, using the analogy as a guide would cause students to use an erroneous assumption to set up their component equations. Solving for magnitude and direction would also depend upon erroneous assumptions. The success of students in the AF condition, therefore, relied in large part on how willing they were to make guesses in order to move forward with the problem steps.

Explore-first condition. In the explore-first (EF) condition, students also explored a problem before hearing a lecture on the topic but were only given the magnetic force problem described above and depicted in Figure 2. This condition is more typical of college-level instruction, in which students are assumed to have learned previously-taught material and are expected to apply it appropriately when new material is presented. It is also the most frequently used structure in exploration research (Loibl et al., 2016). In this condition, students were instructed to “Use what you have learned earlier in the course about electric force to help you solve the magnetic force problem.” Time across both exploration conditions was equal. Students exploring with one problem had more time to try to solve the magnetic force problem but without the benefit of a familiar analogous problem type to scaffold their thinking and constrain the solution space.
**Instruct-first condition.** Students in the instruct-first (IF) condition followed a business-as-usual instructional sequence. First, they were introduced to the new concept of magnetic force through PowerPoint lecture and note-taking. Instruction included teaching students how to solve magnetic force problems. After the lecture, students were given the magnetic force problem and told to practice using the procedures they were just taught to solve it in small groups. They had five minutes less to solve the problem than the exploration conditions had because they had to wait for the professor to give the activity instructions in the other classroom before he commenced the lecture in their classroom. After the lecture, the IF group moved back to their regular classroom to join the other students for the remainder of the lesson.

**Metacognitive report.** Once students completed the learning activity, all students were given a sheet on which they were asked to individually list things they thought they knew about magnetic forces between wires, as well as questions to which they still needed answers. Students wrote their names onto the group packet and the metacognitive reports and turned them in at the end of the class period in a manila folder, which was collected by the primary researcher.

**Lecture.** The lecture covered the concept of magnetic force. The lecture was presented using 15 animated PowerPoint slides. Though not formally scripted, each slide provided a detailed structure for the material presented during the lecture. In addition, in order to cover all of the material presented in the slides during the time allotted for the lecture, the instructor strictly adhered to the slides. Thus, though the two lectures were not administered verbatim from a script, they were instructionally equivalent. In addition, the instructor had given this lecture numerous times before, decreasing the likelihood of practice effects between the two lectures. The lecture was audio-recorded each time it was given to assess time equivalency between conditions.
The learning objective of the lecture was for students to understand that magnetic force is caused by an interaction between currents, just like electric force is caused by an interaction between charges, and gravitational force is caused by an interaction between masses. At this point in the course, students had been taught that a charge influences the space around it by creating an electric field and that when a second charge is placed into the electric field of the first, it responds to the field through an electric force. That force can be either attractive or repulsive, depending on the signs of the two charges. The goal for this lesson was to lay the groundwork for a similar understanding of how currents create magnetic fields and respond to those fields through magnetic forces. The analogy instruction focused on how two long, straight current-carrying wires either attract or repel each other based on the direction of the current flow in both wires. The lecture exposed students to this analogy and reminded students that all forces are interactions between two objects. The intention was to isolate the effect of exploring analogies on learning in the AF group by making sure that students in the EF and IF conditions were also exposed to the analogy through direct instruction. The professor also contrasted electric force and magnetic force between current-carrying wires by emphasizing that there is an attractive force between wires when currents are traveling in the same direction and a repulsive force when those currents are traveling in opposite directions. This relationship is the reverse of what is true about charges, a contrast to which he drew students’ attention. Another point the professor highlighted was the meaning of the symbols “into the page” and “out of the page,” which students encountered in the magnetic force diagram. He taught them to solve for magnitude and for direction. He also presented an example, through which he walked the class as he modeled drawing a free body diagram and plugging components into a vector-addition equation to solve for net magnetic force. The professor also mentioned easy mistakes to make
when using the equations. During the lecture, the professor paused a number of times to solicit questions or participation from students.

**Measures.** Following the entire lesson (lecture and activity, in either order), students individually completed a questionnaire in the same room where they did both lesson components. The students were told that completing the questionnaire was voluntary and were given five minutes to do so. Appendix B contains the full scales of each measure described below.

**Interest.** Three survey items measured interest and enjoyment (also known as situational intrinsic motivation). Those items were: “I found this learning activity interesting”; “I enjoyed this learning activity”; and, “This learning activity was boring” [reverse-coded] (Ryan, 1982). Students responded to these items on a five-point Likert scale of 1 (*strongly disagree*) to 5 (*strongly agree*). Responses were averaged to form a mean score. Previous use of this scale by Weaver and colleagues (2018) yielded good reliability ($\alpha = .82$); the scale was originally validated by Harackiewicz (1979). One further item assessed students’ desire to pursue learning about the concept after the lesson. This behavioral measure of motivation was developed by the author, based on previous items used by Harackiewicz (1979) and in response to calls for alternatives to scale measures of motivation (DiBenedetto, & Schunk, 2018).

**Utility value.** Two items from the subjective task value scale, adapted from Wigfield et al. (1997) and Eccles and Wigfield (1995) measured utility value. A sample item read: “In general, how useful is what you learned in this class today?” Students responded to these items on a seven-point Likert scale of 0 (*not at all useful*) to 6 (*very useful*). Responses were averaged to form a mean score. The full five-item scale has a reliability of $\alpha = .85$ and was validated by Conley (2007).
**Self-efficacy.** Three survey items from the Motivated Strategies for Learning Questionnaire (MSLQ) measured self-efficacy for learning and performance. Those items included: “I'm confident I can understand the basic concepts taught in this lesson”; “I'm certain I can understand the most difficult material presented in this lesson”; and, “I expect to do well on this activity” ($\alpha = 93$; validated by Pintrich, Smith, Garcia, & McKeachie, 1991). Students responded to these items on a seven-point Likert scale of 0 (*not at all true of me*) to 6 (*very true of me*). Responses were averaged to form a mean score.

**Knowledge gaps.** Four items measured participants’ experience of knowledge gaps. A sample item read: “When it comes to solving magnetic force problems, I really don’t know a lot” adapted from Glogger-Frey et al. (2015). Students responded to these items on a five-point Likert scale of 1 (*strongly disagree*) to 5 (*strongly agree*). After reverse-coding one item to match the direction of the other three items, responses were averaged to form a mean score. When experience of knowledge gaps was assessed with nine items by Glogger et al. (2013), the scale had high reliability ($\alpha = .89$); when assessed with five items by Glogger-Frey et al. (2015), reliability was also high ($\alpha = .83$). The scale was originally validated by Flynn and Goldsmith (1999).

**Cognitive load.** One item measured cognitive load. This item read: “In solving the previous problem, I invested…” (full scale $\alpha = .90$; validated by Paas, 1992). Respondents selected a level of mental effort to complete the phrase on a scale of 1 (*very very low mental effort*) to 9 (*very very high mental effort*).

**Posttests.** Finally, during the last 20 minutes of the class, students took a posttest that captured the study’s primary outcome measures. Fourteen multiple-choice questions and one open-ended question assessed how well students understood magnetic force as presented in the
learning activity and lecture (see Appendix C for the assessment). Three versions of the quiz were created with answer choices scrambled between them. Two weeks later, students received the same assessment as a retention test, testing them again on these concepts and procedures. In the intervening two weeks, the regular course instructor proceeded with the unit on magnetism.

As recently called for by scholars advocating a more ecologically-valid approach to productive failure research (Chowrira, Smith, Dubois, & Roll, 2019), the posttests were developed in close consultation with, and primarily by, the physics professor who facilitated the lesson. Attention was paid to aligning learning objectives with performance criteria that would provide evidence that the objectives had been met. Reliability was maximized by writing more than one question for each learning target. Care was taken to discriminate between learning objectives in each question, though some overlap was unavoidable.

Questions targeted both procedural knowledge, defined by Rittle-Johnson, Siegler, and Alibali (2001) as action sequences for solving problems, and conceptual knowledge, defined as “implicit or explicit understanding of the principles that govern a domain and of the interrelations between units of knowledge in a domain” (Rittle-Johnson et al., 2001, p. 347). There were two new procedures targeted by the instruction: (a) how to calculate the magnitude of the magnetic force between two long, straight current-carrying wires; and, (b) how to determine whether the force will be attractive or repulsive. Procedural questions included calculations and simple applications of the algorithm for when the force between wires is attractive or repulsive. Instruction was also designed to emphasize two critical conceptual ideas. The first is that “a force is a force is a force,” and therefore the procedures that have been used in previous force problems are equally useful in a new context. The second is that thinking about the magnetic force between long, straight current-carrying wires should encourage students to
think about how a long-straight current-carrying wire creates a magnetic field. Conceptual questions required students to create and interpret the correct vector representation of free-body diagram. To do so, students had to demonstrate understanding of the underlying concept of force well enough to translate from the verbal and pictorial representation given in the problem statement to the diagrammatic representation which shows the forces and then to correctly interpret the diagram.

In sum, the posttests were optimized to capture the full range of students’ knowledge of the target concept. Easy questions were expected to be a minimal threshold for understanding the lesson, for example, by asking students to apply a rule or learned procedure. Other questions represented a range of item difficulties so that discriminability was maximized and variation among the students detectable. The most difficult question on the posttests was aimed at identifying students who understood the concept that “a force is a force is a force,” i.e., that whether gravitational, electric, or magnetic, all forces share the same basic properties. A wide representation of performance was expected on the posttests and was predicted to vary with instructional condition.

**Reliability analyses.**

**Learning activity.** Four items composed the learning activity (see Appendix A), which was completed in small groups before instruction in the two exploration conditions and after instruction in the IF condition. The format varied between the AF, EF and IF conditions to include either one magnetic force problem (EF and IF) or two analogous problems on electric and magnetic force (AF). The activity served as the focus of the instructional intervention, a learning opportunity, and a formative assessment of what students understood about the new concept being taught. Sample reliability for the activity was high: Cronbach’s α = .806. Learning
activities were de-identified by removing names and condition labels. The problems were not visible on the sheets on which students wrote their solutions. The professor that delivered the lesson graded the solution worksheets, blind to condition. Each step of the learning activity was scored “no attempt (0 points),” “attempt (1 point),” or in steps two and three, “partial attempt (.5 point).”

Posttests. The posttest was designed to contain three subscales of knowledge (see Appendix C). Seven questions were written to measure students’ procedural knowledge of calculating force. Six questions were aimed at assessing conceptual understanding. The final two questions were designed to test students’ ability to transfer their understanding of magnetic force to the related concept of magnetic field, which had not yet been taught. To maximize power to detect an effect of the intervention, the number of comparisons was reduced by combining the two transfer items with the conceptual knowledge subscale. Certainly, accurate responses to the transfer questions required conceptual understanding of magnetic force at the very minimum. To verify that the knowledge subscales (procedural and conceptual) worked the way the physics professor and I intended them to when we designed the assessment, reliability analyses were run on these subscales in our sample. Alphas greater than or equal to .70 are considered to have good reliability for tests of ability (Kline, 1999). The reliability of both subscales at posttest were good: procedural Cronbach’s $\alpha = .726$; conceptual Cronbach’s $\alpha = .726$. At retention test the procedural knowledge scale remained satisfactory ($\alpha = .657$), but the conceptual knowledge scale had less consistency ($\alpha = .551$), most likely due to almost floor effects on Question 13.

Surveys. Of the 171 students who took both posttests, 160 completed the entire survey. One hundred and sixty-four completed only the first page of the survey, and 162 completed two of the three measures on the second page. Therefore, sample sizes differ depending on the scale.
Four items asked students to report knowledge gaps; sample Cronbach’s alpha was good, $\alpha = 0.834$ ($N = 164$). Three items asked students about their interest in the learning activity; again, sample Cronbach’s alpha was strong, $\alpha = 0.852$ ($N = 164$). Three items asked students to report their self-efficacy; sample $\alpha = 0.814$ ($N = 163$). Two items assessed students’ utility value; sample $\alpha = 0.675$ ($N = 162$). The cognitive load and behavioral motivation measures were each assessed with only one item, so no reliability analyses were performed.

**Procedure**

On the day of the study, the instructor of record for the course did not teach the lesson; rather, another professor in the department who also teaches the same course taught the lesson. The collaborating professor had experience with facilitating exploratory learning and lecturing from a script; his delivery of the lesson was intended to ensure fidelity to the procedure. Furthermore, the instructor’s regular lesson was modified to vary type and order of instruction. The order of the first two lesson components (activity and lecture) varied between the exploration and instruct-first conditions. Students were randomly assigned to either begin with exploration in their regular classroom (both AF and EF) or to begin with direct instruction in an adjacent classroom (IF). To manage these two different lesson orders, the professor moved between the two classrooms to deliver the instruction and to provide directions for the activity. The professor began in the exploration classroom to introduce the activity. Afterward, he went to the IF classroom, lectured, and introduced the activity. Then, the professor returned to the exploration classroom and delivered the same lecture to the two exploration groups simultaneously. The only difference in the delivery of the lecture was the order in which students received it, which varied between the IF and two exploration conditions. This sequence is shown in Table 4.
Experimenters were present in both rooms to assist with materials and directions, audio record the lessons, and administer the questionnaire after the instructional phases of the lesson. They did not answer questions about the content or offer help on solving the problems. However, in the large lecture hall where the analogy-first and explore-first conditions completed the learning activity, an experimenter did get involved with facilitating the lesson when she discovered an error on the packet. The learning packets for these two conditions presented two structurally aligned problems, as intended. However, the labels on the magnetic force problem did not match the symbols depicted on the diagram, as shown in Figure 4. Therefore, the physics graduate student present in the exploration classroom instructed students to ignore the diagram in their packets and instead to look on the board, where she redrew the diagram to match the erroneous labels. This solution made the two diagrams appear to be misaligned to the analogy-first students who saw both, although one could argue that by matching the labels, instead of the symbols, the true relations between diagrams was preserved. This error only occurred in the first
section of the experiment and thus did not affect all students; however, this was the larger course section and therefore, the majority of students in the AF condition were affected.

**Magnetic force problem:**

Three parallel wires are arranged on the corners of a rectangle, as shown in the figure below. Wire 1 carries 1 A of current into the page. Wire 2 carries 2 A of current into the page. Wire 3 carries 3 A of current out of the page. What is the magnitude and direction (relative to the +x axis) of the net magnetic force per length exerted on wire 1 due to the other two wires?

![Diagram of magnetic force problem](image)

*Figure 4.* Mislabeled diagram. Label for Wire 2 says "into" but should say "out of" to match the symbol on the diagram. Wire 3 says "out of" but should say "into" to match the symbol on the diagram.

Following the questionnaire, the students rejoined in their regular classroom. At this point, the professor explained the correct answers to the activity, as well as common incorrect answers. He visually and verbally modeled how to solve the magnetic force problem. This debrief took five minutes. The professor then administered the fifteen-question posttest to students in all three conditions during the final 20 minutes of the class period. The entire class period was 75 minutes in duration. In all three conditions, 20 minutes were devoted to the learning activity, 15 minutes were devoted to the lecture, five minutes were allotted for the questionnaire and five for the transition back to the classroom, five minutes were spent debriefing the activity, and 20 minutes were reserved for the posttest. All additional time was spent giving directions. Two weeks later, the retention test was administered by the regular course instructor.
Data Analysis Plan

Statistical analyses. For each of the three conditions in this experiment, there were 11 dependent variables of interest, distributed between two families: performance and process-level measures. The five performance variables included learning activity accuracy, procedural posttest score, conceptual posttest score, procedural retention test score, and conceptual retention test score. The six survey measures included experience of knowledge gaps, interest, self-efficacy, utility value, cognitive load, and behavioral motivation. The independent variable in all analyses was instructional condition. The dependent variables (DVs) are delineated below. The Holm-Bonferroni correction was applied to control for the Type I error rate that was inflated by running multiple tests. Holm-Bonferroni is a modification of the Bonferroni correction and is less conservative (Abdi, 2010).

Learning activity accuracy. Learning activity accuracy was indicated by scores on exploration problem(s) or practice problem, depending on condition. Analysis of variance (ANOVA) tested the effect of instructional condition on success on the activity, with condition as the independent variable, score as the dependent variable. Students’ grasp of electric force, which was explicitly involved in the AF condition and implicitly suggested in the EF and IF groups, was assumed to be related to students’ subsequent acquisition of magnetic force. For any main effects of condition found, planned comparisons were carried out to reveal which conditions accounted for the significant omnibus difference between group means.

Posttest and retention test performance. A 3 (condition: AF, EF, IF) x 2 (question type: procedural, conceptual) mixed-factorial ANOVA tested the effect of instructional condition on posttest outcomes. Condition was a between-subjects variable, and question type a within-subjects variable. Main effects and interactions were examined.
**Survey measures.** Because of my interest in the difference between the three groups on each of six DVs, the plan was made to run a multivariate analysis of variance (MANOVA), with group membership as the IV and moderately-correlated survey scales as the DVs. MANOVA uses the covariance of the DVs to interpret the relationships and offers greater power to reject the null hypothesis than a series of ANOVAs, when the DVs are related. For a significant result from the MANOVA, simple and special contrasts were specified to determine which linear combination of the DVs was responsible. Survey scales which did not moderately correlated with the other scales were run separately in ANOVA.

**Effect sizes.** It is important to create a standardized measure of the significance of any effect so it may be compared with the magnitude of other studies’ findings (APA, 2010). Effect sizes were given for the experimental effect calculated by ANOVA using Pearson’s correlation coefficient $r$. Small effects were considered to be in the range of $r > .10$; $r > .30$ were interpreted as a medium effect, and large effects were $r > .50$. Effect sizes for planned comparisons were given by Cohen’s $d$ because this measure of effect is simple to understand and useful to meta-analysts. Cohen’s $d$ is computed in standard deviations. Specifically, it calculates the difference between the experimental and control group’s scores on outcome measures and divides the difference by the pooled (or control group) standard deviation, as shown in the following equation:

$$d = \frac{\bar{X}_1 - \bar{X}_2}{s}$$

In line with Cohen’s recommendations (Cohen, 1988), small effects were considered to be $d > 0.2$ standard deviations, medium effects to be $d > 0.5$, and large effects to be $d > 0.8$. Effect sizes for mixed-factorial ANOVA and MANOVA were reported as partial eta-squared.
CHAPTER IV: RESULTS

The research questions motivating this study and the hypotheses tested in the research are as follows:

Research Question 1 (R1): Do students learn a new concept best when exploring problems before receiving instruction on how to solve them, compared to receiving direct instruction on solution methods before practicing the problems?

Hypothesis 1 (H1): Initial acquisition of the concept, as measured by the learning activity, is likely to be most rapid in the instruct-first condition, due to all necessary information being presented in the lecture prior to problem solving. The exploration conditions are likely to acquire the concept less well at first (during learning activity phase) but hypothesized to catch up by posttest, once they receive the relevant information through the lecture.

R2: Does using an analog during exploration improve conceptual understanding of magnetic force, measured at posttest and retention test, as compared with a simple framing instruction to connect the two topics?

H2: I sought to discover whether analogies needed to be invoked directly, through exploration of analogous problems, or indirectly, with just a framing direction to use an analogous problem type to solve the new problem. Of the two exploration conditions, the analogy-first condition was expected to perform better than the explore-first condition, due to the opportunity for production of a general representation of a force concept offered by the analogous problems.

R3: Does any effect of instructional condition develop or persist over time?
H3: I predicted a “sleeper effect,” such that any benefit of condition on learning outcomes may not emerge immediately (at posttest), but rather after students have had more time to reflect on the new concept (at retention test).

R4: What effect does exploratory learning have on interest, utility value, self-efficacy for problem solving, knowledge gaps, cognitive load, and motivation, compared to a direct instruction condition?

H4: I hypothesized that students in the exploration conditions would rate more highly their interest and enjoyment in the learning activity and report greater motivation to continue to learn about the topic than students in the instruct-first group. The questions of utility value and self-efficacy were exploratory, so no specific hypotheses were advanced about differences between groups on these measures. Perceptions of knowledge gaps were expected to be greater in the exploration conditions than in the instruct-first condition, due to the greater fluency in learning the latter condition would experience. Finally, I predicted that students in the analogy-first condition would report less cognitive load than the explore-first condition and considered it an open question whether students in the analogy-first condition would experience more or less cognitive load than the instruct-first group.

Preliminary Analyses

Prior knowledge between groups. I computed a prior knowledge variable from the average of students’ first two exams in the course, out of a total of 100 percent ($N = 171, M = 63.25, Mdn = 65.00, SD = 13.43, 95\% CI: [61.44, 65.33]$). Conversion of this variable to a standardized score demonstrated that all students were within 2.5 standard deviations of the mean score. Inspection of histograms, Q-Q plots, and box-plots confirmed that prior knowledge was normally distributed in each condition in the sample. Levene’s statistic was not significant,
indicating that the assumption of homogeneity of variance between groups was met. A one-way analysis of variance (ANOVA) by condition on prior knowledge was run to test for systematic differences in prior knowledge between groups before the intervention. No significant results were obtained, $F < 1$, indicating that random assignment to condition was successful. Means, standard deviations, and 95 percent confidence intervals are shown in Table 5.

Table 5

*Prior Knowledge of Physics by Condition*

<table>
<thead>
<tr>
<th>Condition</th>
<th>$n$</th>
<th>Mean</th>
<th>$SD$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instruct-first</td>
<td>54</td>
<td>65.00</td>
<td>12.73</td>
<td>[61.53, 68.47]</td>
</tr>
<tr>
<td>Explore-first</td>
<td>55</td>
<td>62.23</td>
<td>14.11</td>
<td>[58.41, 66.04]</td>
</tr>
<tr>
<td>Analogy-first</td>
<td>62</td>
<td>62.62</td>
<td>13.49</td>
<td>[59.20, 66.05]</td>
</tr>
</tbody>
</table>

*Note. N = 171. CI = confidence interval. Prior Knowledge is shown as a percentage score.*

**Descriptive Statistics**

**Learning activity.** The learning activity asked students to solve a magnetic force problem. Means, standard deviations, and 95 percent confidence intervals for the sum of all four steps of the learning activity, delineated by condition, are shown in Table 6.

Table 6

*Descriptive Statistics for Learning Activity by Condition*

<table>
<thead>
<tr>
<th>Condition</th>
<th>$n$</th>
<th>Mean</th>
<th>$SD$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instruct-first</td>
<td>54</td>
<td>3.74</td>
<td>0.72</td>
<td>[3.54, 3.94]</td>
</tr>
<tr>
<td>Explore-first</td>
<td>55</td>
<td>0.94</td>
<td>0.86</td>
<td>[0.70, 1.16]</td>
</tr>
<tr>
<td>Analogy-first</td>
<td>62</td>
<td>2.87</td>
<td>1.07</td>
<td>[2.60, 3.14]</td>
</tr>
</tbody>
</table>

*Note. N = 171. CI = bias-corrected and accelerated bootstrap confidence interval. The learning activity was out of 4 points.*
**Posttests.** First tests of normality were examined for each of the posttest knowledge subscales in each condition. Inspection of skewness statistics showed that none of the conditions’ scales were severely skewed (i.e., dividing the statistic by its standard error did not result in values more than 2 standard deviations from the mean), though the histogram for the IF group at posttest appeared negatively skewed, as would be expected. Levene’s test for homogeneity of variance was not significant in any condition for the conceptual knowledge subscale or for the procedural knowledge at retention test. At posttest, Levene’s test was significant at \( p = .043 \), but the test based on the median was not, \( p = 1.00 \). Normal Q-Q plots of each subscale in each separate condition did not indicate any cause for concern. Histograms at retention test displayed normally distributed scores for each subscale in each condition. Means, standard deviations, and 95 percent confidence intervals for the posttest knowledge subscales are shown in Table 7.

**Procedural knowledge outliers.** Inspection of the scores for outliers revealed that 25 students (15\% of the sample) scored perfectly on the procedural knowledge subscale at posttest: nine in the IF condition, six in the EF condition, and ten in the AF condition. Only five students (3\%) scored 0 points on this subscale at posttest, two in the EF condition and three in the AF condition. Twelve students scored perfectly on this subscale at retention test: two in the IF condition, four in the EF condition, and five in the AF condition. While no students in the IF condition had scored 0 on the procedural knowledge subscale at posttest, three students in this condition scored 0 at retention test. No students in the EF condition scored 0 at retention test, but four students in the AF condition did.

**Conceptual knowledge outliers.** Eight students (5\% of the sample) scored perfectly on the conceptual knowledge subscale at immediate posttest: four in the IF condition, two in the EF condition, and two in the AF condition. Nine students scored a 0 at posttest: three in the IF
condition, two in the EF condition, and four in the AF condition. No one scored perfectly on this subscale at retention test; six students scored 0: two in the EF condition and four in the AF condition. The conclusion is that there were no outliers in the dataset.

Table 7

*Descriptive Statistics for Knowledge Subscales by Condition at Posttest and Retention Test*

<table>
<thead>
<tr>
<th></th>
<th>Instruct-first&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Explore-first&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Analogy-first&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>95% CI</td>
<td>M (SD)</td>
</tr>
<tr>
<td><strong>Posttest</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedural</td>
<td>67.72 (25.20)</td>
<td>[60.05, 75.40]</td>
<td>58.57 (29.56)</td>
</tr>
<tr>
<td>Conceptual</td>
<td>55.09 (27.63)</td>
<td>[47.80, 62.39]</td>
<td>49.55 (27.32)</td>
</tr>
<tr>
<td><strong>Retention Test</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedural</td>
<td>49.34 (24.24)</td>
<td>[42.19, 56.48]</td>
<td>51.17 (25.34)</td>
</tr>
<tr>
<td>Conceptual</td>
<td>37.50 (21.17)</td>
<td>[31.46, 43.54]</td>
<td>39.09 (22.96)</td>
</tr>
</tbody>
</table>

*Note. N = 171. CI = confidence interval. Posttest scores are shown as percentages.*

<sup>a</sup>n = 54. <sup>b</sup>n = 55. <sup>c</sup>n = 62.

**Surveys.** The decision was made to report correlations with pairwise deletion for missing data to obtain the most comprehensive view from students who participated in the lesson.

Correlations between the variables are shown in Table 8. The greater the knowledge gap students perceived, the less their interest in the learning activity and the lower self-efficacy they reported. In turn, they reported less utility value for the class and the course overall. There was no correlation between perception of knowledge gaps and cognitive load or behavioral motivation to study the topic further. Interest in the lesson was significantly associated with self-efficacy for the lesson and utility value for the class and course. Interest was also positively associated with cognitive load, such that the greater interest in the lesson students reported, the more effort they
reported to have put forth. In addition, interest was directly related to behavioral motivation to pursue the topic further outside of class. Self-efficacy was significantly positively associated with utility value, but unrelated to cognitive load or behavioral motivation. Utility value was not associated with cognitive load or behavioral motivation. A significant positive association between cognitive load and behavioral motivation was found, such that the greater the mental effort exerted on the activity, the greater the behavioral motivation students reported.

Table 8

Summary of Correlations, Means, and Standard Deviations for All Survey Scales

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Knowledge Gaps</td>
<td>2.91 (0.88)</td>
<td>-.20*</td>
<td>-.59***</td>
<td>-.42***</td>
<td>-.009</td>
<td>.054</td>
</tr>
<tr>
<td>2. Interest</td>
<td>3.49 (0.84)</td>
<td>.21**</td>
<td>.44***</td>
<td>.34***</td>
<td>.29***</td>
<td></td>
</tr>
<tr>
<td>3. Self-efficacy</td>
<td>4.10 (1.27)</td>
<td>.29***</td>
<td>-0.13</td>
<td>-0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Utility Value</td>
<td>3.95 (1.10)</td>
<td>.15</td>
<td>.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Cognitive Load</td>
<td>5.63 (1.28)</td>
<td>.23**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Behavioral Motivation</td>
<td>.68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. N_{knowledge gaps, interest, self-efficacy} = 164; N_{utility value, cognitive load} = 162; N_{behavioral motivation} = 160. Means (standard deviations) are shown on the diagonal.
*p < .05 (2-tailed). **p < .01. ***p < .001.

Statistical Analyses

For each of the three experimental conditions, I have two planned comparisons: (a) IF versus AF, and (b) EF versus AF; that is, a comparison between the control and the primary experimental group and a comparison between the two experimental groups.
Learning activity accuracy. Learning activity accuracy was indicated by scores on the magnetic force problem, completed either as exploration or as practice, depending on condition. A one-way analysis of variance (ANOVA) was run to test the effect of instructional condition on the activity total score. All of the students who attended class submitted learning activities ($N = 189$). Students worked in groups of two ($n = 3$), three ($n = 57$), or four ($n = 3$) members to complete the learning activity, 19 in the IF condition, 21 in the EF condition, and 23 in the AF condition. However, students’ learning activities were only included in analyses if they had submitted all study measures ($N = 171$, 90% of available activity data).

Tests of assumptions. An assumption of ANOVA is that scores within groups are normally distributed. Inspection of standardized scores within each condition confirmed that there were no outliers. The skewness and kurtosis scores for the IF group revealed that this distribution was significantly negatively skewed, as did the significant values of Kolmogorov-Smirnov test, $D(54) = .536$, $p < .001$. To correct for this, bias-corrected and accelerated bootstrapping procedures were used to sample 1000 times from the data to estimate the sampling distribution from these bootstrapped samples. Field (2013) recommends using robust procedures such as bootstrapping over transforming the data, because the latter procedure can make it more difficult to interpret the results. ANOVA also assumes that variances are equal across groups. Levene’s test of homogeneity of variance was significant, $F(2, 168) = 9.67$, $p < .001$, indicating that this assumption was violated in my sample. Certainly, this was due to the ceiling effect that being in the IF condition had on accuracy. Therefore, a corrected version of the $F$-ratio was used, and Welch’s $F$ is reported below.

---

1 Because ANOVA assumes that each data point is independent from every other, data was analyzed by groups as well. The same pattern of results was found as when scores were analyzed for individuals.
2 The pattern of results holds if all 189 students’ data is analyzed.
**One-way ANOVA.** The effect of condition was found to have a significant impact on learning activity accuracy, resulting in a large effect, $F(2, 111.16) = 171.43, p < .001, r = .79$. Planned contrasts showed that when equal variances were not assumed, students in the IF condition performed significantly better than students in the AF condition, $t(107.19) = 5.18, p < .001, d = 1.21$ and that students in the AF condition outperformed students in the EF condition, $t(113.86) = 10.80, p < .001, d = 2.25$ (see Figure 5).

![Figure 5. Mean score on the learning activity as a function of instructional condition. Error bars represent standard error of the mean.](image)

In order to establish whether students in the AF condition were making use of the analogy to solve the magnetic force problem, each step of the activity was examined separately, using a series of cross-tabulations. In the first step, in which students were expected to draw a free-body diagram, 89% of students in the IF condition made a diagram and 5.6% were awarded a half point for a partial attempt. Only three students made no attempt. In the AF condition, about an equal number of students did and did not make an attempt (43.5% vs. 46.8%); 9.7% earned
partial credit. In the EF condition, a majority of students did not attempt a diagram (65.5%). These differences were significant, \( N = 171, \chi^2(4) = 48.88, p < .001 \). This finding shows that many more students in the AF condition were willing to try Step 1, even though it involved making guesses about information they did not yet have. EF students were less likely to make an attempt.

In Step 2, again, significant differences were found between the three groups’ ability to set up component vector addition equations, \( N = 171, \chi^2(4) = 76.54, p < .001 \). In the IF condition, 87% did so, while another 7.4% earned partial credit. In the AF condition, 58.1% set up an equation, while another 27.4% received partial credit. In stark contrast, a majority of EF students did not set up an equation (69.1%).

In Step 3, considered a “pressure point” for being able to move on in the problem, students had to plug in to the component equations to be able to calculate magnitude. All IF students did this (94.4% earned 1 point and 5.6% earned \( \frac{1}{2} \) point). Similarly, 90.3% of AF students did this, while another 4.8% received partial credit for doing so. Though many of them used a magnitude equation that was wrong (because they had not been taught what it was yet), that was coded as evidence of using a magnitude expression, so successfully completing Step 3. Conversely, the EF group was unable to make the guesses necessary in the absence of instruction to move on in the problem. Only 25.5% of this group earned a point for Step 3. These differences were statistically significant, \( N = 171, \chi^2(4) = 87.22, p < .001 \).

This pattern held in Step 4. Solving for magnitude and direction of the force was attained by 94.4% of IF students, 62.9% of AF students, and merely 5.5% of EF students, \( N = 171, \chi^2(4) = 126.97, p < .001 \). Students who explored with the analogy were able to extrapolate from the electric force problem the steps they should take to solve the problem, while the students who
explored without the analogy could not constrain the problem space to a reasonable set of steps to take.

**Posttest performance.** A three (condition: IF, EF, AF) x two (question type: procedural, conceptual) x two (time of test: posttest, retention test) mixed-factorial ANOVA tested the effect of instructional condition on posttest outcomes at two time points. Condition was a between-subjects variable, and question type and time of posttest were within-subjects variables.

Levene’s test was not significant for all combinations of levels of the repeated-measures variables, confirming that the assumption of homogeneity of variance was met. No main effect of condition was found, $F < 1$. As predicted, there was a main effect of question type, such that overall scores on the procedural knowledge questions ($M = 55.81\%, SD = 27.62\%, 95\% CI [52.06, 59.57]$) were higher than overall scores on the conceptual knowledge questions ($M = 44.52\%, SD = 24.78\%, 95\% CI [41.27, 47.77]$), $F(1, 168) = 54.48, p < .001, \eta^2_p = .245$. There was also a main effect of time, such that scores fell for all three groups from posttest ($M = 55.85\%, SD = 28\%, 95\% CI [52.04, 59.65]$) to retention test ($M = 44.48\%, SD = 24.4\%, 95\% CI [41.28, 47.69]$), $F(1, 168) = 53.71, p < .001, \eta^2_p = .242$.

Contrary to prediction, there was no knowledge type by condition interaction; however, this appears to be due to low power to predict an effect (observed power was .065). As shown in Figure 6, a significant time by condition interaction was found $F(2, 168) = 4.61, p = .011, \eta^2_p = .052$. The IF group performed significantly better at posttest ($M = 61.41\%, SD = 23.12\%, 95\% CI [54.65, 68.17]$) than at retention test ($M = 43.42\%, SD = 19.68\%, 95\% CI [37.72, 49.12]$), falling below the exploration groups at retention test; whereas, the EF condition’s performance did not significantly change between posttest ($M = 53.76\%, SD = 26.21\%, 95\% CI [47.53, 60.53]$) and retention test ($M = 45.13\%, SD = 19.68\%, 95\% CI [39.48, 50.78]$); nor did the AF
condition’s performance significantly change between posttest ($M = 52.07\%,\ SD = 25.80\%,\ 95\%\ CI\ [45.41,\ 58.35]) and retention test ($M = 44.90\%,\ SD = 22.89\%,\ 95\%\ CI\ [39.58,\ 50.22])

Figure 6. Mean percentage scores, on posttests, by condition. Error bars represent standard error of the mean.

There was not a knowledge type by time interaction, $F < 1$, nor was there a three-way interaction between condition, knowledge type, and time, $F < 1$ (observed power = .089), as shown in Figure 7.
To test for differences between the three conditions on each of the survey measures, I ran a multivariate analysis of variance (MANOVA). MANOVA uses the covariance of the DVs to interpret the relationships and offers greater power to reject the null hypothesis than a series of ANOVAs, when the DVs are related. However, MANOVA requires that the DVs be moderately correlated. I found that only the first four survey measures could be described this way (see Table 8). Therefore, I excluded the last two measures from the multivariate analysis: cognitive load and behavioral motivation. The MANOVA tested group differences on experience of interest, utility value, self-efficacy, and knowledge gaps.

**Tests of assumptions.** MANOVA tests whether mean differences between groups are likely to have occurred by chance. The first assumption of MANOVA is that there is a linear relationship between dependent variables, measured by correlations (see Table 8). A second assumption of MANOVA is homogeneity of the variance-covariance matrix. This assumption is tested with Box’s $M$, and a non-significant result indicates that the assumption was met. This
assumption was confirmed in my data, Box’s $M = 27.75$, $F(20, 78942.57) = 1.34$, $p = .143$. A third assumption is the absence of singularity; that is, dependent variables should be distinct from one another so that they are measuring different constructs. However, they should be moderately correlated. This was true of four of the survey measures, as explained above. A fourth assumption is that data should display multivariate normality, which inspections of line graphs confirmed. Finally, a rule of thumb for sample size to run the test is that there should be more cases than DVs in every cell, which was the case.

Interpreting the $F$-ratio in MANOVA relies on test statistics that differ when there are more than 2 levels of the independent variable. Most researchers prefer to use one of the pooled statistics (i.e., Wilk’s $\lambda$ or Pillai’s criterion). Pillai’s is said to be more robust when unequal sample sizes are an issue (Olson, 1974), and so I have elected to interpret that statistic, since my three groups are not equal in size. Based on the one-way MANOVA results with interest, utility value, self-efficacy, and knowledge gaps as dependent variables, there was a significant multivariate effect of condition, Pillai’s Trace = .143, $F(8, 314) = 3.02$, $p = .003$, $\eta^2_p = .071$. The effect size signifies that 7% of the variance in the multivariate DV is explained by the independent variable (i.e., instructional condition). Univariate tests showed that average responses on all four survey measures differed by condition; however, instead of interpreting the univariate results I will interpret planned contrasts. Group means for all survey variables are shown in Table 9.
Table 9

*Descriptive Statistics for Survey Scales by Condition*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Instruct-first&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Explore-first&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Analogy-first&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>95% CI</td>
<td>M (SD)</td>
</tr>
<tr>
<td>Interest</td>
<td>3.74 (0.70)</td>
<td>[3.54, 3.95]</td>
<td>3.38 (0.90)</td>
</tr>
<tr>
<td>Utility Value</td>
<td>4.30 (0.96)</td>
<td>[4.01, 4.59]</td>
<td>3.71 (1.13)</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td><strong>4.38 (1.19)</strong></td>
<td>[4.03, 4.73]</td>
<td><strong>3.74 (1.46)</strong></td>
</tr>
<tr>
<td>Knowledge Gaps</td>
<td>2.59 (0.84)</td>
<td>[2.34, 2.83]</td>
<td>3.02 (0.91)</td>
</tr>
<tr>
<td>Cognitive Load</td>
<td><strong>5.42 (1.32)</strong></td>
<td>[5.02, 5.82]</td>
<td><strong>6.07 (1.30)</strong></td>
</tr>
<tr>
<td>Behavioral Motivation</td>
<td>.60 (.50)</td>
<td>[.45, .75]</td>
<td>.74 (.44)</td>
</tr>
</tbody>
</table>

*Note.* CI = confidence interval.

<sup>a</sup>n = 47, or 45 for Utility Value, Cognitive Load, & Behavioral Motivation. <sup>b</sup>n = 55, or 54 for Behavioral Motivation. <sup>c</sup>n = 62, or 61 for Behavioral Motivation.

Bold indicates statistically significant differences between groups.

**Interest.** No significant differences were found between groups on this measure.

**Utility value.** No significant differences were found between groups on this measure.

**Self-efficacy.** The IF group reported three-quarters of a point greater self-efficacy than the EF group 95% CI [0.07, 1.46]. There were no significant differences between the IF group and AF group on this measure, nor were there significant differences between the two exploration conditions.

**Knowledge gaps.** There was a significant difference in perception of knowledge gaps, with IF students reporting a half point lower gap in perception than students in the AF group.
95% CI [0.07, 1.01] and than the EF group, 95% CI [0.02, 0.99]. No difference was found between the two exploration conditions.

**Cognitive Load.** To determine whether the EF condition found the learning activity most cognitively loading of the three groups, as predicted by H2, a one-way ANOVA was run with cognitive load ratings as the dependent variable and condition as the independent variable. Planned comparisons examined differences between the IF and exploration conditions and between the EF and AF conditions.

Histograms of group mean scores on the cognitive load item were found to display normal distributions. No group had a skewness statistic > |1.5|, confirming the presence of a normal distribution. Kurtosis for the IF group was 2.06, indicating a significantly pointy and heavy-tailed distribution; no other group had a kurtosis score > 1.96 (the threshold for significance). However, the K-S test was significant for all three conditions, IF: $D(45) = .197, p < .001$; EF: $D(55) = .223, p < .001$; AF: $D(62) = .189, p < .001$. To correct for this, bias-corrected and accelerated bootstrapped confidence intervals are reported. Levene’s statistic was not significant, confirming that the assumption of homogeneity was met. Condition was found to have a significant effect on self-reported cognitive load, resulting in a medium effect, $F(2, 159) = 5.26, p = .006, r = .25$. Planned comparisons showed that students in the EF condition reported significantly greater cognitive load on the learning activity than students in the IF condition, $t(159) = 2.60, p = .010$, and in the AF condition, $t(159) = 3.00, p = .003$ (see Figure 8).
Behavioral Motivation. A chi-square test measured the association between condition and students’ behavioral motivation to learn more about the problem-solving activity after the lesson. There was not a significant association between condition and self-reported behavioral motivation (request for additional information about how to solve the learning activity problem), $N = 160$, $\chi^2(2) = 2.26$, $p = .323$. 

*Figure 8. Mean cognitive load ratings by condition. $N = 162$. 

[Graph showing cognitive load ratings by condition]
CHAPTER V: DISCUSSION

With this study, I sought to identify conditions under which an exploratory learning approach might differ from a traditional tell-then-practice instructional sequence for learning a new concept. The central objective was to find a “sweet spot” for how and when exploratory learning could enhance students’ conceptual understanding. Informed by Vygotsky’s idea of a zone of proximal development, this zone was thought to lie in a realm where level of cognitive engagement was sufficiently challenging to be effortful and motivating but not overly cognitively loading. Also following from sociocultural theory (Vygotsky, 1987), learning was scaffolded in the primary experimental condition by introduction of an analogy. Learners in all three conditions worked collaboratively to problem solve.

The comparison between the two experimental conditions was designed to help discern which ingredients were necessary for successful problem solving during exploratory learning, as compared with a control group following a direct instruction format. To this end, I compared three instructional sequences, which varied in order of lecture and problem solving and in the degree to which exploratory learning was supported by use of analogy. Achievement outcomes were measured at three time points: during acquisition of a new concept, after the lesson, and two weeks later. Furthermore, I examined process-level measures that students reported on questionnaires regarding their learning experiences under the three instructional conditions.

The central hypothesis guiding the research was that exploring analogous problems would facilitate learning of a general rule that would help students solve future problems related to the rule. Structural similarity between analogs was predicted to encourage students to notice that two problems were of the same type (force) and to represent them in the same way. The hypothesized mechanism of exploring the problems simultaneously was that it would alert
students to the vector addition procedure common to solving both electric and magnetic force problems. Recognition that the same procedure could be applied to both problem types was expected to foster students’ conceptual understanding that force problems are solved the same way because all force problems are interactions between two objects. Analogical reasoning was, therefore, theorized to be a mechanism in the development of conceptual knowledge.

The analogs were hypothesized to be more useful in an exploratory format than when presented in the midst of lecture, because exploration required students to use the analogy to solve the problem (Schwartz et al., 2011). Scholars of analogical reasoning argue that problem solving provides a pragmatic goal which makes analogical features more salient than they might be in another context (e.g., in a lecture; Gentner, 1989; Holyoak & Thagard, 1989). Instruction that emphasizes problem solving has also been shown to lead to more flexible and retrievable knowledge (National Research Council, 2000). Although lecture could highlight the similarities between problem types, students would not depend upon the analogy to solve the problems, because they would be taught a procedure for doing so. Teaching about the common principle of force through lecture first was, therefore, hypothesized to result in inert knowledge (i.e., students would not spontaneously recognize the applicability of electric force to magnetic force).

The students in the EF condition were expected to find exploring the magnetic force problem cognitively taxing, compared to the AF condition and certainly to the IF condition, due to the unconstrained search for candidate analogs from long-term memory (Begolli & Richland, 2016). Though they had more time to inspect just one problem, without having to examine another, they had no hint about where to access relevant prior knowledge, except in the form of a direction to use what they knew about electric force. Finally, based on previous research (Weaver et al., 2018), exploration was expected to have motivational benefits.
Summary of Major Findings

I found that exploratory learning with analogous problems provided undergraduates a more optimal level of instructional scaffolding than having them explore only a new problem, as evidenced by greater success on the learning activity. Compared to direct instruction, however, exploratory learning did not lead to enhanced conceptual knowledge development by posttest or greater motivation, as I had hypothesized. Students’ ratings of cognitive load provided initial evidence that exploring with analogous problems alleviated cognitive load imposed by exploring a new concept before instruction.

Detailed Summary of Findings

H1.

Learning activity. The pattern of results obtained by the one-way ANOVA of learning activity scores by condition provided strong support for the first part of H1. Students in the IF group experienced the most efficient acquisition of learning to solve the novel problem, likely because they were provided with the information they needed to do so before attempting the problem. It is also plausible that the AF students were unable to work the problem as accurately as IF students because they had more to do in the same amount of time. Only the AF group had two problems to examine and solve during the learning activity. However, even with more to do, students in this group outperformed students in the EF condition, who had only one problem to process and solve.

The results also indicate that the AF group benefited from the use of the analogous problems to scaffold their problem-solving, whereas the EF group struggled to search the problem space for solutions they might apply to the new problem. The superior performance of students in the AF condition to those in the EF condition shows that students in the latter were
less successful at spontaneously applying analogical reasoning, even though they were instructed to do so. This finding demonstrates that, to some degree, the EF students’ knowledge of electric force was context-bound. Therefore, EF students may have failed to recruit the relevant information from memory that could have assisted them in recognizing structural similarities between the previously learned concept and the magnetic force problem. This finding is anticipated by the literature, which finds that search for candidate analogs can instead result in the inappropriate selection of problems that are the same on the basis of mere appearances, a situation in which two problems share similar attributes but not structural relations (Gentner, 1989; Gentner & Markman, 1997). Without the analogy visually presented to students in the EF condition, they were unable to benefit from comparing the two problems, which is key to successful analogical transfer (Gentner, Loewenstein, & Thompson, 2003).

**Posttest and retention test performance.** The second part of HI was also partly supported by the findings at posttest. I had predicted that all three group averages for procedural knowledge would be similar by posttest and this was borne out by the data, which did not show significant differences on procedural questions between the three groups at posttest or retention test. However, my expectation that exploration groups would exceed the IF group on posttest questions of conceptual knowledge was not met. So, though the deficit experienced initially by the exploration groups was eliminated by the lecture and resulted in equivalent problem-solving accuracy at posttest, it did not reverse itself by pushing the explore groups past the IF group in terms of conceptual understanding.

It would be appropriate to say that exploratory learning was sufficient, but not necessary, for development of adequate conceptual knowledge of force. It is probable that this finding is due, in part, to the optimal lecture condition implemented in the study. The exploratory learning
conditions were pitted against a lecture that contained explicit reference to and depiction of the analogy between electric and magnetic force. This conservative control was selected in order to isolate the effect of *exploring* the analogy and in order to provide best practices to every student at some point in the lesson. However, by reducing the temporal difference in analogy exposure between the IF and AF condition (students in both conditions were exposed to it immediately), the effect of exploring the analogy may have been reduced.

**H2.**

With H2, I had predicted that the AF group would exceed the EF group’s performance on the conceptual knowledge subscale, while procedural knowledge would be equivalent between exploration conditions. This pattern of results was not found with this study sample on either posttest or retention test. A plausible explanation for the lack of differences between exploration groups may be that the lesson was too difficult in general. There is reason to think that all students could have benefited from more scaffolding of the new concept before taking the posttests (Paas, Renkl, & Sweller, 2003).

One of the greatest sources of difficulty in learning the new concept would have been perceptual. Up to this point in the course, students had been studying electrostatics, which exists in a two dimensional plane. The diagrams that they were accustomed to seeing and drawing would have been 2D. In magnetic force, however, charges are moving through currents, not remaining static. The magnetic force diagram presented to the students was not 3D – it was a cross-section of a 3-dimensional space, which is standard practice in teaching this material. Nonetheless, students may have required more guidance in how to visualize the forces and interpret the diagrams.
Moreover, in electrostatics opposite charges attract, but in magnetics, opposite currents repel. Although experts perceive the two problems students saw as the same, novices would have potentially perceived them as different (Chi et al., 1981). Without having been told this piece of missing information, students in the exploration conditions would not have been able to guess that the diagrams would have to be reversed. The coding scheme used to score the learning activities accounted for students not yet knowing the rule for opposite currents by awarding them a point for applying correct procedures, even without achieving the correct answer. However, posttest scoring did not allow for incorrect responses (though the presumption was that the intervening lecture would have provided all students with correct problem-solving procedures). In this regard, the analogous problem-solving condition differed from previous research designs, which generally equip problem solvers with all the relevant information they need to map similarities or alignable differences from a source to a target problem (Gentner et al., 2003; Gentner & Markman, 1997).

It is also conceivable that students would have had more success with this lesson if it had come later in the sequence on magnetism. First-year students are generally not thought to possess expertise in the area of magnetism; this may set them apart from students in other studies who may have had more advanced understanding of the concepts they were exploring. For instance, in the study by Chowrira and colleagues (2019), teaching assistants supported students in a productive failure condition by providing hints when they got stuck; students also studied in advance of being tested on the study’s outcome measures. In Weaver et al. (2018), the sample was composed of a mix of under- and upper-classmen, and even post-baccalaureate students, who may have brought to bear more relevant prior knowledge of the study’s concepts.

H3.
My third hypothesis concerned the effects of instructional condition over time. I had predicted a sleeper effect for exploration groups, but results showed that these students’ posttest scores remained stable over the course of two weeks. There was, however, a significant drop in scores for students in the IF condition from posttest to retention test, as shown by the interaction between condition and time of test. IF students’ posttest scores were the highest, numerically, and retention test scores numerically lowest. Bearing in mind that scores did not differ significantly between conditions, this finding seems to indicate that scores for the IF group were initially inflated, perhaps due to the fluency with which students learned the new concept. It has been repeatedly shown that students learning from direct instruction rely on declarative knowledge taught explicitly to acquire basic knowledge, but fail to elaborate on this knowledge by connecting it to prior knowledge (Jacobson et al., 2017; Kapur, 2016; Schmidt & Bjork, 1992). Without activation of relevant prior knowledge, schema abstraction does not ensue. In the absence of schema abstraction, advantages to having heard a recent lecture would soon fade. My finding also suggests that a sense of fluency at the end of a lesson does not indicate better retention of the learned material.

Furthermore, power was likely too low to detect an effect of a knowledge type × time interaction (observed power = .078). This points to the possibility that my measures did not capture hypothesized differences between groups. I attempted to tease apart conceptual and procedural understanding by separately comparing conditions on knowledge subscales on the posttest. I expected to see differences between groups in terms of how conceptual knowledge developed over the course of two weeks. It may be that a longer delay would have produced this difference, particularly if the course instructor continued to make connections between types of force sharing similar relational attributes. Perhaps final exam questions at the end of the semester
would have been more sensitive to differences produced by the experimental conditions. On the other hand, it is possible that the hypothesized difference in conceptual knowledge did not materialize because the learning activity actually emphasized procedural skill. It was the vector addition procedure that was analogous between electric and magnetic force, which may be less of a conceptual point than anticipated.

**H4.**

Awareness of knowledge gaps was greater in both of the exploration conditions than in the IF condition, confirming previous research (Loibl & Rummel, 2014) and part of H4, which predicted this difference. There was not a significant difference in reported knowledge gaps between the two exploration conditions. Because the survey was given after both phases of the lesson, the knowledge gaps that students had in the exploration conditions prior to hearing the lecture should have been filled by the time they reported knowledge gaps. However, because students who explored the problems were in greater need of the missing knowledge during the learning activity, the impression of gaps in their knowledge is likely to have remained highlighted for them at the time of the survey. Focusing on knowledge gaps is likely to have reduced students’ illusions of knowing (Dunlosky & Rawson, 2012; Kruger & Dunning, 1999) in the exploration conditions.

The expectation was that greater awareness of what they needed to know would have led students to be more attentive during the lecture, as they would be motivated to fill gaps in their knowledge. Awareness of knowledge gaps has been shown to spark greater interest in learning (Rotgans & Schmidt, 2014; but see Glogger-Frey, 2015). Interestingly, in this study, the two were inversely related.
Yet, there were not significant differences between groups in terms of interest, disconfirming part of H4. I had expected students in the exploration conditions to exceed IF students in interest and enjoyment in the activity. Previous research has shown increased motivation from exploring (Glogger-Frey et al., 2015; Glogger et al., 2013; Weaver et al., 2018), a finding thought to be tied to increased autonomy and feelings of competence that come from relying on one’s own prior knowledge to attack a new problem and satisfying epistemic curiosity (Lammina & Chase, 2017; Niemiec & Ryan, 2009; Rotgans & Schmidt, 2011; Silvia, 2008). However, no difference between groups was found. Previous research with introductory physics students has found a similar pattern of results (Weaver et al., 2018, Study 1).

An important contribution of this study is the finding that relational reasoning did not overtax cognitive resources. In this study, there was no reported difference in cognitive load between the AF and IF conditions; only the EF students rated their cognitive load as significantly greater than the other two conditions. Therefore, while the EF condition was potentially too taxing, having analogous problems alleviated the cognitive load imposed by exploring a new concept before instruction. This finding addresses the exhortation by Alfieri, Nokes-Malach, and Schunn (2013) in their meta-analytic review that future research should investigate the subjective cognitive load that learners experience during case comparison. Perhaps most surprisingly, the positive association between interest and cognitive load precludes the interpretation that the exploration groups found the learning activity too difficult to be interesting and enjoyable. Although, for example, the level of cognitive load reported by the EF group was significantly higher than that of AF or IF, confirming H4, these students did not report lower motivation for the lesson.
Part of H4 was exploratory, regarding the questions of self-efficacy and utility value. The strongest association between process-level variables was that between self-efficacy and knowledge gaps. The greater the awareness of knowledge gaps, the significantly less efficacious the students reported feeling. Condition also impacted self-efficacy, with the EF group reporting significantly less self-efficacy than the IF group. Exploring before instruction, therefore, did not increase students’ self-efficacy for the topic. Being fully prepared to solve problems before practicing them boosted self-efficacy in this experiment. Utility value for the learning task was also moderately negatively associated with knowledge gaps, though no differences were found between groups in terms of utility value.

**Implications**

Based on the results from this study, course instructors should feel free to incorporate more exploratory learning activities in their classrooms without fear that learning or performance will suffer. Assessments given at the end of class and after a typical two-week delay found no difference between conditions on learning a new concept taught either through a lecture format or an exploratory learning format. In addition to the potential benefits to students that varying class activities may bring, this variety in the routine may also benefit instructors who have a desire to flip their classrooms.

In this study, additional cognitive load was imposed by having students explore without the benefit of a comparison problem. This finding suggests that exploring analogous problems is a way to alleviate extraneous cognitive load imposed by the search requirements of having only a hint to use prior knowledge to help solve a new problem. Another resolution might be to include analogous problems in a practice session after lecture, although it is possible that they will not be
attended to as carefully as during exploration, because students will already have a schema upon which to rely (Holmes, Day, Park, Bonn, & Roll, 2014).

**Inconsistencies with Previous Work**

The findings from this study only partially replicate the studies of Weaver and colleagues (2018), in which students in an explore-first condition exceeded the performance of students in an instruct-first condition on measures of conceptual knowledge and, in one study, interest and enjoyment in the topic. Though this inconsistency is surprising, there are some likely explanations for the discrepant results. First, in the earlier studies, the sample was a mix of engineering, physics, and health pre-professional majors. In the current study, all students in the sample were engineering majors. It is likely that these students were oriented to the material in the lesson differently depending on their major field. For example, physics students are generally thought to be interested in the basic science behind the concepts they are learning, whereas engineering students are thought to be studying physics “as a background for their primary interest in other fields” (Feynman, 1963). Furthermore, engineering students are typically very adept at using procedures to solve problems, but may be less inclined to conceptualize the problems they are solving. Repeating this study in a sample of physics majors might yield different results.

**Limitations**

In this study, students in the AF condition were expected to solve the magnetic force problem by the mechanism of analogical reasoning. However, the instructional scaffold designed to support analogical reasoning may have compromised the hypothesized advantage for students in the AF condition. This is due to the experimenter error that mislabeled the magnetic force diagram in the first course section participating in the study. Examining the diagram with
compromised structural alignment of features may have disrupted analogical transfer at both the problem solving and schema abstraction phases. A critical step in successful analogical transfer is checking candidate inferences from the source problem for validity in the target problem. Where the charges and currents did not align, students may have dismissed their inferences as invalid, thus inhibiting their use of the source problem to solve the target problem and limiting schema induction. This error did not affect students in the IF condition who were in a different classroom, nor would it have particularly impacted students in the EF condition, as they were only viewing one problem; thus, the structural alignment of the two problems was not an issue for these groups.

**Future Directions**

Future research should investigate whether modeling the process of structure-mapping for students would improve their schema induction. Modeling could take the form of cognitive apprenticeship or think-aloud, led by the professor in an earlier lesson so that students would gain proficiency with the process. In this same vein, directions could explicitly direct students to locate the key features of each diagram to help them attend to similarities and essential features of the problems. A preliminary step to problem solving could be for students to compare and contrast the problems by listing similarities and differences between them (Holmes et al., 2014).

Providing feedback to students during problem solving could also enhance the utility of the exploratory learning (Chowrira et al., 2018). Future designs could allow for TAs to circulate and check work as students proceeded through the activity. This could serve the purpose of encouraging the groups to continue moving forward when they got stuck. The professor could also address student-generated solutions in the subsequent lecture, contrasting students’ solutions with the canonical solution to highlight differences. Doing so might help to prevent the “IKEA
effect,” by which students become enamored of their own creations (Norton, Mochon, & Ariely, 2012), and thus prevent them from persisting in incorrect strategies.

It would be equally important to test the success of exploration with analogies in a class that regularly engaged in analogical reasoning as part of course instruction. There is reason to believe that embedding this method in a semester-long intervention would increase differences between conditions, as each implementation of analogue exploration might bolster the effects of previous sessions. Students using analogies would, in theory, adopt a habit of mind when they broached related concepts if they became accustomed to thinking about the deep structures of problems they encountered. This skill would likely develop over time, and consolidation of conceptual knowledge would occur over the course of the semester and into the future if topics were revisited. Therefore, assessments should take place at longer intervals, such as final course exams.

Finally, individual differences between learners contribute to their ability to learn from different instructional approaches (e.g., Belenky & Nokes-Malach, 2012; DeCaro et al., 2015). Research designs that measure individual differences, such as test anxiety, achievement motivation orientation, or levels of expertise, would help ascertain aptitude-treatment interactions at work in exploratory learning with analogy.

**Conclusion**

In sum, this study contributed to the problem-solving-prior-to-instruction literature a randomized-controlled study conducted within a classroom ecology. Confounds were avoided by using a factorial design, in which order of instruction varied between two classrooms and learning materials varied within one classroom. An instruction-before-problem-solving condition was compared to both a problem-solving-before-instruction condition and a scaffolded-problem-
solving-before-instruction condition. Similar learning outcomes were found for all three groups. However, a faster decay in retention was evidenced by an instruct-first group. Use of analogous problems during an exploratory learning activity condition reduced cognitive load imposed by exploring and increased willingness to attempt and success at problem-solving.
REFERENCES


Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology, 95*(2), 393-408. doi:10.1037/0022-0663.95.2.393


APPENDIX A

Learning Activity Packet (version for AF condition)

Name:________________________ Date:_________

Directions: This packet contains an activity to help you learn about the concept of magnetic force between current-carrying wires. The structure and method for this activity is group exploration. Tips for how to be successful at working in this format are below. Please read and follow them.

Important Tips for Effective Exploration:

You are NOT being thrown into the deep end.
- Use whatever you have learned in the past that may help you in the task.

The goal is NOT to have the right answer.
- Having the right answer, but not understanding why it is correct, is a problem. Focus the majority of your time trying to explain the concepts and reasoning that are the basis for your answer.
- Also make sure that you show your work and write out your explanations to the fullest extent possible.

Being wrong is NOT actually a problem.
- Trying to explain something shows you whether you know it and learning what doesn’t work is just as important as learning what does.
- Ask specific questions about what you don’t understand. Good questions often begin with What if? Why? and How? And remember “I don’t know” is not actually a question.

Consider a range of strategies or solutions that might work.

Important Tips for Working with Groups:

Assign these roles to complete the necessary tasks for the group.
- Discussion leader keeps group moving forward and solicits everyone’s contribution.
- Recorder turns in a neat copy of the group’s process and answers with everyone’s names on it.
- Checker goes over the work and answers to be sure they represent the group’s best efforts.
- Facilitator insures that everyone gets the group’s work copied into their notes.

The key is not what you get from the other students in your group but what you contribute to the group. So what can you “give” to the group?
- Leadership
- Specific questions
- Structure/focus
- Guidance
- Ideas
- Failure (Remember, learning what doesn’t work can often be as important as learning what does.)
Directions: Examine the two force problems below. Then, complete the following steps.

1. Solve each problem and show your work. Use what you have learned earlier in the course about electric force to help you solve the magnetic force problem. Write our your solution on the sheet titled “Work for Learning Activity” included in your folder. Make sure the name of each group member is written at the top of the page.

2. If you cannot completely solve the problem, write out as much of the solution as you can and identify any remaining information that you need.

Electric force problem:
Three point charges are arranged on the corners of a rectangle, as shown in the figure below. Charge 1 has a charge of $+1 \text{ nC}$. Charge 2 has a charge of $-2 \text{ nC}$. Charge 3 has a charge of $+3 \text{ nC}$. What is the magnitude and direction (relative to the $+x$ axis) of the net electric force exerted on charge 1 due to the other two charges?

Magnetic force problem:
Three parallel wires are arranged on the corners of a rectangle, as shown in the figure below. Wire 1 carries 1 A of current into the page. Wire 2 carries 2 A of current out of the page. Wire 3 carries 3 A of current into the page. What is the magnitude and direction (relative to the $+x$ axis) of the net magnetic force per length exerted on wire 1 due to the other two wires?
Names: _____________________  ____________________
                           _____________________  ____________________

Work for Learning Activity:
After you have completed the activity, please generate a list of things you think you know about magnetic forces between wires and questions to which you still need answers in order to fully understand it. You should continue to work with your groups, but each of you will turn in your own individual sheet. These will be turned in to and reviewed by the professor to see how much you learned from the activity and how much still needs to be taught.

Things I think I know:

Questions I still need answers to:
APPENDIX B

Questionnaire

Name:_______________________________ Date:________

Please reflect on today’s class activity. Respond to each statement – do not skip any. After you have completed this survey, place it in the envelope.
**Please indicate how much you agree or disagree with each of the following statements by circling a number.**

<table>
<thead>
<tr>
<th></th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neither Agree or Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>I do not feel very knowledgeable about solving magnetic force problems.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2.</td>
<td>I found this learning activity interesting.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3.</td>
<td>When it comes to solving magnetic force problems, I really don’t know a lot.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4.</td>
<td>Compared to most other people, I know less about solving magnetic force problems.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5.</td>
<td>I enjoyed this learning activity.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6.</td>
<td>I know pretty much how to solve magnetic force problems.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7.</td>
<td>This learning activity was boring.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

**If you think the statement is very true of you, circle 6; if a statement is not at all true of you, circle 0. If the statement is more or less true of you, determine the number between 0 and 6 that best describes you.**

<table>
<thead>
<tr>
<th></th>
<th>Not At All True of Me</th>
<th>Very True of Me</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.</td>
<td>I’m certain I can understand the most difficult material presented in this lesson.</td>
<td>0</td>
</tr>
<tr>
<td>9.</td>
<td>I’m confident I can understand the basic concepts taught in this lesson.</td>
<td>0</td>
</tr>
<tr>
<td>10.</td>
<td>I expect to do well on the learning activity.</td>
<td>0</td>
</tr>
</tbody>
</table>
Please circle the number that corresponds with your answer to the question about how useful or important your learning is. Check 6 if your answer is “very;” check 0 if your answer is “not at all.” If your response is somewhere in between, determine the number between 0 and 6 that best captures how useful or important you find it to be.

<table>
<thead>
<tr>
<th></th>
<th>Not At All</th>
<th>Somewhat</th>
<th>Very</th>
</tr>
</thead>
<tbody>
<tr>
<td>11. In general, how useful is what you learned in this class today?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>12. Compared to most of your other college courses, how useful is what you learn in this class?</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Please indicate how much mental effort you invested when solving the problem(s).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>13. In solving the previous problem(s) I invested:</td>
<td>Very, very low mental effort</td>
<td>Very low mental effort</td>
<td>Low mental effort</td>
<td>Rather low mental effort</td>
<td>Neither low nor high mental effort</td>
<td>Rather high mental effort</td>
<td>Higher mental effort</td>
<td>Very high mental effort</td>
<td>Very, very high mental effort</td>
</tr>
</tbody>
</table>

14. When you leave class today, you will have the opportunity to receive additional practice problems on the topic of study (by email). Would you be interested in receiving this assignment?

- [ ] Yes
- [ ] No

Thank you for completing this survey.
APPENDIX C

Directions: For each question below, write the letter for the correct answer in the space provided.

Throughout the document, we will use the notation \( F_{12} \) to indicate the magnitude of the magnetic force exerted on wire 1 by wire 2.

Two long, straight wires run parallel to each other and are separated by a distance of \( d = 3 \, \text{cm} \), as shown in the figure below. Wire 1 carries a current of \( I_1 = 4 \, \text{A} \) to the left. Wire 2 carries a current of \( I_2 = 6 \, \text{A} \) to the right.

\[
\begin{array}{c}
\downarrow \quad d \\
I_1 = 4 \, \text{A} \\
\downarrow \\
I_2 = 6 \, \text{A} \end{array}
\]

\[ d = 3 \, \text{cm} \]

1. What is the direction of the magnetic force that wire 1 exerts on wire 2?
   A. Horizontally to the right
   B. Up and to the right
   C. Up
   D. Up and to the left
   E. Horizontally to the left
   F. Down and to the left
   G. Down
   H. Down and to the right
   I. No direction because the magnitude is zero

2. What is the direction of the magnetic force that wire 2 exerts on wire 1?
   A. Horizontally to the right
   B. Up and to the right
   C. Up
   D. Up and to the left
   E. Horizontally to the left
   F. Down and to the left
   G. Down
   H. Down and to the right
   I. No direction because the magnitude is zero

3. If you move wire 2 up so that the distance between the wires is reduced to 1 cm, how will \( F_{21} \), the magnitude of the force per length of wire that wire 1 exerts on wire 2, have changed?
   A. \( F_{21} \) will have increased by a factor of 2
   B. \( F_{21} \) will have increased by a factor of 3
   C. \( F_{21} \) will have increased by a factor of 4
   D. \( F_{21} \) will have increased by a factor of 9
   E. \( F_{21} \) will not have changed
   F. \( F_{21} \) will have decreased by a factor of 2
   G. \( F_{21} \) will have decreased by a factor of 3
   H. \( F_{21} \) will have decreased by a factor of 4
   I. \( F_{21} \) will have decreased by a factor of 9
Four long, straight wires, labeled wires 3, 4, 5, and 6, run parallel to each other into and out of the page. The wires lie along a single line, with each sequential wire a distance $d$ further to the right, as shown in the figure below. Wire 3 carries a current of $I_3 = I$ into the page. Wire 4 carries a current of $I_4 = 2I$ out of the page. Wire 5 carries a current of $I_5 = 3I$ out of the page. Wire 6 carries a current of $I_6 = 6I$ into the page.

$$I_3 = I \quad I_5 = 3I \quad I_4 = 2I \quad I_6 = 6I$$

4. How does the magnitude of the force per length wire 3 exerts on wire 5, $F_{53}$, compare to the magnitude of the force per length wire 3 exerts on wire 6, $F_{63}$?

A. $F_{53}$ is greater than $F_{63}$
B. $F_{53}$ is the same as $F_{63}$
C. $F_{53}$ is less than $F_{63}$
D. The relative magnitudes of $F_{53}$ and $F_{63}$ cannot be determined from the information given.

5. How does the magnitude of the force per length wire 4 exerts on wire 5, $F_{54}$, compare to the magnitude of the force per length wire 4 exerts on wire 6, $F_{64}$?

A. $F_{54}$ is greater than $F_{64}$
B. $F_{54}$ is the same as $F_{64}$
C. $F_{54}$ is less than $F_{64}$
D. The relative magnitudes of $F_{54}$ and $F_{64}$ cannot be determined from the information given.

6. What is the direction of the net magnetic force exerted on wire 5 by the other three wires?

A. Horizontally to the right
B. Up and to the right
C. Up
D. Up and to the left
E. Horizontally to the left
F. Down and to the left
G. Down
H. Down and to the right
I. No direction because the magnitude is zero
Four long, straight wires carrying current into and out of the page lie along an equilateral triangle with sides of length $d$, as shown in the figure to the right. Wires 7 and 8 lie at the vertices along the base of the triangle and wire 9 lies at the remaining vertex. Wire 10 lies at the midpoint of the base of the triangle. All three wires carry a current $I$, with wires 7, 9, and 10 carrying it out of the page and wire 8 carrying it into the page.

7. What is the direction of the net magnetic force exerted on wire 9 by the other three wires?
   A. Horizontally to the right  
   B. Up and to the right  
   C. Up  
   D. Up and to the left  
   E. Horizontally to the left  
   F. Down and to the left  
   G. Down  
   H. Down and to the right  
   I. No direction because the magnitude is zero

8. What is the direction of the net magnetic force exerted on wire 10 by the other three wires?
   A. Horizontally to the right  
   B. Up and to the right  
   C. Up  
   D. Up and to the left  
   E. Horizontally to the left  
   F. Down and to the left  
   G. Down  
   H. Down and to the right  
   I. No direction because the magnitude is zero

9. Which ranking below correctly ranks the magnitude of the forces exerted on wire 7 by the other three wires?
   A. $F_{7,8} = F_{7,9} = F_{7,10}$  
   B. $F_{7,8} > F_{7,9} > F_{7,10}$  
   C. $F_{7,10} > F_{7,9} > F_{7,8}$  
   D. $F_{7,10} > F_{7,8} = F_{7,9}$  
   E. $F_{7,8} = F_{7,10} > F_{7,9}$  
   F. $F_{7,9} > F_{7,10} > F_{7,8}$
10. Suppose you remove wire 10 so that only the wires on the vertices of the triangle remain, as shown in the figure on the right. What is the direction of the net magnetic force exerted on wire 9 by the other two wires?

A. Horizontally to the right  
B. Up and to the right  
C. Up  
D. Up and to the left  
E. Horizontally to the left  
F. Down and to the left  
G. Down  
H. Down and to the right  
I. No direction because the magnitude is zero

Three long, straight wires lie parallel to each other in the plane of the page. Wire 11 carries a current $I$ up the page. Wire 12 carries a current $2I$ up the page. Wires 11 and 12 are a distance $d$ away from each other and wire 13 is an additional distance $d$ away from wire 12.

11. If the current in wire 13 runs up the page, how much current would wire 13 need to carry for the net magnetic force on wire 11 to be zero?

A. $I$  
B. $2I$  
C. $3I$  
D. $4I$  
E. $5I$  
F. $6I$  
G. $7I$  
H. $8I$  
I. The net magnetic force on wire 11 cannot be zero in this situation.

12. If the current in wire 13 runs down the page, how much current would wire 13 need to carry for the net magnetic force on wire 11 to be zero?

A. $I$  
B. $2I$  
C. $3I$  
D. $4I$  
E. $5I$  
F. $6I$  
G. $7I$  
H. $8I$  
I. The net magnetic force on wire 11 cannot be zero in this situation.
13. The magnetic field is represented by the letter $B$, with $B_1$ being the magnetic field caused by wire 1 and $B_2$ being the magnetic field caused by wire 2. Which of the following expressions seems the most reasonable expression for describing the force per length that wire 1 with a current $I_1$ exerts on wire 2 with a current $I_2$ in terms of the magnetic field?

A. $\frac{F_{21}}{L} = I_1 B_2$  
B. $\frac{F_{21}}{L} = I_1 I_2 B_1$  
C. $\frac{F_{21}}{L} = I_1 B_1 B_2$  
D. $\frac{F_{21}}{L} = \frac{I_1 I_2}{B_1}$  
E. $\frac{F_{21}}{L} = I_2 B_1$  
F. $\frac{F_{21}}{L} = I_1 I_2 B_2$  
G. $\frac{F_{21}}{L} = I_2 B_1 B_2$  
H. $\frac{F_{21}}{L} = \frac{I_1 I_2}{B_2}$

14. Which of the following expressions could be used to calculate the magnitude of the magnetic field at a distance $d$ away from a long, straight wire that carries a current $I$?

A. $B = \frac{\mu_0 I}{2\pi d}$  
B. $B = \frac{\mu_0 I^2}{2\pi d^2}$  
C. $B = \frac{\mu_0 I}{2\pi d^2}$  
D. $B = \frac{\mu_0 I^2}{2\pi d}$

Two wires carry the same current. Wire 14 carries the current out of the page and wire 15 carries the current into the page. Both wires are the same distance away from the origin, with wire 14 on the $+y$ axis and wire 15 on the $+x$ axis, as shown in the figure to the right.

15. Suppose the magnetic field points away from a current flowing out of the page and towards a current flowing into the page, which of the vectors below correctly indicates the direction of the net magnetic field at the origin? Clearly circle your answer.

![Diagram with vectors]