Exploratory learning activities in the physics classroom: contrasting cases versus a rich dataset.

Campbell Rightmyer Bego

University of Louisville

Follow this and additional works at: https://ir.library.louisville.edu/etd

Part of the Cognitive Psychology Commons, Educational Methods Commons, Engineering Education Commons, Higher Education Commons, and the Other Psychology Commons

Recommended Citation


https://doi.org/10.18297/etd/3321

This Doctoral Dissertation is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact thinkir@louisville.edu.
EXPLORATORY LEARNING ACTIVITIES IN THE PHYSICS CLASSROOM: CONTRASTING CASES VERSUS A RICH DATASET

By

Campbell Rightmyer Bego
B.S., Columbia University, 2008
M.S., University of Louisville, 2016

A Dissertation Submitted to Faculty of the College of Arts and Sciences of the University of Louisville in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Experimental Psychology

Department of Psychological and Brain Sciences University of Louisville Louisville, Kentucky

December 2019
EXPLORATORY LEARNING ACTIVITIES IN THE PHYSICS CLASSROOM:
CONTRASTING CASES VERSUS A RICH DATASET

By

Campbell Rightmyer Bego
B.S., Columbia University, 2008
M.S., University of Louisville, 2016

A Dissertation Approved on

October 30, 2019

By the following Dissertation Committee

Dissertation Director
Marci S. DeCaro

Cara H. Cashon

Nicholas C. Hindy

Keith B. Lyle

Patricia S. Ralston
ACKNOWLEDGEMENTS

This work, and my doctoral degree, would not have been possible without the mindful guidance and thorough feedback of my mentor Dr. Marci DeCaro. I am grateful for the thoughtful reading assignments, the gradual increase in responsibility and control of experimental design, and the direct push to think deeply about learning mechanisms. In addition to these personal experiences, I am thankful for Dr. DeCaro’s constant example of scientific rigor, objectivism, persistence, and enthusiasm.

I also owe an incredible debt to Dr. Raymond Chastain for many of the ideas as well as the materials in this document. Dr. Chastain’s desire to improve his already powerful classroom instruction is contagious and inspiring, and his willingness to set aside time for discussion and development seems to be unlimited. He not only offered his classroom for collaborative projects, but also increased my understanding of teaching and learning by several orders of magnitude.

I am especially thankful for Dr. Patricia Ralston at the J. B. Speed School of Engineering, and the members of the interdisciplinary research group GEARS (the Guild for engineering Education, Achievement, Retention, and Success). This collaboration has introduced me to an incredible variety of ideas and given me copious research opportunities that helped me develop my research identity as well as my future career path.
So many other people have been personal supports throughout this graduate program including my experimental psychology cohort (especially Dr. Samantha Patton and Dr. Joanna Weaver), graduate and undergraduate labmates, engineering teaching collaborators (Dr. Jeffrey Hieb and Dr. Angela Thompson), and my officemate, point of reference, education guru, and lifeline, the future-Dr. Terresa Tinnell.

I am eternally grateful for my parents, Mrs. Dr. Rightmyer and Mr. Dr. Rightmyer, for helping me develop a growth mindset, a passion for learning, and respect for all people. Their generosity with love, food, discussion, and editing has helped me with every writing project for my entire life.

Lastly, I thank my husband Cem Bego for the daily doses of coffee and laughter, for inviting me to have bigger dreams than I thought possible, and for pushing alongside me to make them happen. I could not have done this without you.
In exploratory learning, students engage in an exploration activity on a new topic prior to instruction. This inversion of the traditional tell-then-practice order has been shown to benefit learning outcomes, especially conceptual knowledge and preparation for future learning, but not always. In three studies, the current work examines whether the type of exploration activity impacts learning mechanisms and outcomes, on the topic of gravitational field in undergraduate physics classrooms. Activities using either contrasting cases (CC) or a rich dataset (RD) are compared in two instructional orders, explore-first (EF) and instruct-first (IF). Learning outcomes measured procedural knowledge, conceptual knowledge, and performance using a dynamic preparation for future learning (PFL) assessment that included a learning resource within the posttest. In addition, the current studies investigated process level measurements of interest and enjoyment, knowledge gap awareness, and cognitive load. Study 1 revealed that students in the EF-CC condition had better conceptual knowledge and PFL scores than students in the IF-CC condition. Study 2 investigated learning outcomes following an RD activity in EF and IF orders in two separate physics classrooms (algebra-based physics for pre-
medicine majors, and calculus-based for engineering majors). Procedural and conceptual knowledge overall was improved by the EF order, but only in the calculus-based course; PFL showed no differences. Study 3 compared learning outcomes from the two activities directly in a 2 (order: EF, IF) × 2 (activity: CC, RD) study design. Overall, the CC activity resulted in better student learning than the RD activity. However, an effect of instructional order was not found, potentially due to a methodological error. Across the three studies, interest and enjoyment did not differ by condition and did not predict learning outcomes. Knowledge gap awareness was higher for students in the EF order than the IF order for both activities, and negatively predicted learning outcomes. Cognitive load was negatively correlated with conceptual and PFL learning outcomes, but only in Study 3. Overall, results indicate that the type of activity could moderate learning outcomes, with any instructional order. More work is needed to investigate the boundary conditions impacting the benefit of exploratory learning, including activity type, amount of guidance during the activity, and the type of instruction.

Keywords: exploratory learning, productive failure, STEM education, undergraduate education, higher education.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .................................................................................................................. iii

ABSTRACT ....................................................................................................................................... v

LIST OF FIGURES .......................................................................................................................... ix

LIST OF TABLES ............................................................................................................................. x

CHAPTER I: INTRODUCTION ........................................................................................................... 1
  Exploration Activities .................................................................................................................. 5
  Learning Outcomes .................................................................................................................... 9
  Learning Mechanisms ............................................................................................................... 11
  Current Studies .......................................................................................................................... 15
  Theoretical and Practical Implications ..................................................................................... 20

CHAPTER II: STUDY 1 ................................................................................................................... 22
  Method ....................................................................................................................................... 22
  Results ....................................................................................................................................... 29
  Discussion .................................................................................................................................. 33

CHAPTER III: STUDY 2 ................................................................................................................ 37
  Method ....................................................................................................................................... 38
  Results ....................................................................................................................................... 42
  Discussion .................................................................................................................................. 48

CHAPTER IV: STUDY 3 ................................................................................................................ 53
  Method ....................................................................................................................................... 55
  Results ....................................................................................................................................... 58
  Discussion .................................................................................................................................. 63

CHAPTER V: GENERAL DISCUSSION ......................................................................................... 70
  Theoretical Implications ........................................................................................................... 74
  Limitations & Future Work ....................................................................................................... 78
LIST OF FIGURES

Figure 1: Contrasting cases activity example (Schwartz et al., 2011). ........................................ 6
Figure 2: Rich dataset activity example (Kapur, 2014). ............................................................. 8
Figure 3: IF-CC Activity. ........................................................................................................ 25
Figure 4: IF-CC and EF-CC Activity Instructions ................................................................. 26
Figure 5: Study 1 posttest performance by order. Error bars = ± SE ................................... 30
Figure 6: Study 1 PFL assessment performance by order. Error bars = ± SE. ................... 31
Figure 7: EF-RD Activity. ..................................................................................................... 40
Figure 8: IF-RD Activity Instructions. Note: The rest of the activity was the same as in
the EF-RD condition. ........................................................................................................... 40
Figure 9: Study 2 posttest performance by order and course. Error bars = ± SE .............. 43
Figure 10: Study 2 PFL assessment performance by subscale and order. Error bars = ± SE. ............................................................ 44
Figure 11: Study 2 distribution of the number of RD activity strategies. .......................... 46
Figure 12: Study 3 posttest performance by order and activity. Error bars = ± SE .......... 59
Figure 13: Study 3 posttest performance by activity, collapsed across orders. Error bars = ± SE. .................................................................................................................... 60
Figure 14: Study 3 student work in the IF-CC condition .................................................... 68
Figure 15: Studies 1 & 3 CC conditions, posttest performance by order and study .......... 72
Figure 16: Studies 2 & 3 RD conditions, posttest performance by order and study .......... 73
LIST OF TABLES

Table 1: Study 1 Number of Participants by Order .......................................................... 23
Table 2: Instructional Order in Explore-first and Instruct-first Conditions ..................... 23
Table 3: Study 1 Regression Results, CC Activity Success and Learning Outcomes ...... 32
Table 4: Study 1 Regression Results, Interest and Enjoyment Scale and Learning Outcomes ................................................................. 32
Table 5: Study 1 Regression Results, Cognitive Load and Learning Outcomes ............. 33
Table 6: Study 2 Number of Participants by Order and Course ....................................... 39
Table 7: Study 2 Regression Results, RD Number of Strategies and Learning Outcomes. ............................................................................................................................... 47
Table 8: Study 2 Regression Results, Interest and Enjoyment and Learning Outcomes. 47
Table 9: Study 2 Regression Results, Knowledge Gap Awareness and Learning Outcomes. ......................................................................................................................... 48
Table 10: Study 2 Regression Results, Cognitive Load and Learning Outcomes ........... 48
Table 11: Study 3 Number of Participants by Order and Activity. .............................. 56
Table 12: Study 3 CC Activity Success and Learning Outcomes. ................................. 61
Table 13: Study 3 RD Number of Strategies and Learning Outcomes. ....................... 62
Table 14: Study 3 Interest and Enjoyment and Learning Outcomes. ............................ 62
Table 15: Study 3 Knowledge Gap Awareness and Learning Outcomes. ..................... 63
Table 16: Study 3 Cognitive Load and Learning Outcomes ............................................ 63
CHAPTER I: INTRODUCTION

Undergraduate courses in Science, Technology, Engineering and Math (STEM) are traditionally taught using direct instruction methods, in which lectures are followed by individual practice at home. Research has revealed, however, that reversing the order can sometimes contribute to learning (e.g., DeCaro & Rittle-Johnson, 2012; Kapur, 2008; Schwartz & Bransford, 1998; Weaver, Chastain, DeCaro, & DeCaro, 2018; Westermann & Rummel, 2012). In exploratory learning, students explore a novel problem first and are given instruction afterwards. This order capitalizes on the positive elements of constructivist learning techniques (e.g., active learning and engagement; Schwartz, Lindgren, & Lewis, 2009) while avoiding the pitfalls of these methods (e.g., incorrect learning and confusion; Kirschner, Sweller, & Clark, 2006). Research has frequently demonstrated learning gains for students in exploratory learning conditions versus direct instruction conditions, including greater conceptual knowledge and transfer (Kapur, 2014; Schwartz & Bransford, 1998) and preparation for future learning (Schwartz, Chase, Oppezzo, & Chin, 2011; Schwartz et al., 2009; Schwartz & Martin, 2004).

The activities in exploratory learning research have taken many forms. For example, one commonly-used activity type is a novel problem presented with contrasting cases. Contrasting cases (CC) are sets of examples that, by design, vary along specific, identifiable problem features (e.g., Schwartz et al., 2011). Another activity type is a novel
problem with a *rich dataset* (RD), a list of data that includes both critical and extraneous information. In RD activities, the solution to the given problem, and the path to the solution, is not obvious. A third type of activity used in exploratory learning research is a series of novel problems with immediate feedback on student solutions (e.g., DeCaro & Rittle-Johnson, 2012).

No studies at this point have compared activity types within an exploratory learning context. The process of exploration could differ depending on the activity (e.g., discovery, strategy generation, amount of cognitive load), and therefore activity type could moderate the benefits of exploration. Comparing results between studies with different activities is difficult, because the specific learning outcomes also differ between studies. Two similar outcome measures are generally used across exploratory learning studies: procedural and conceptual knowledge. In a review of this research, Loibl, Roll, and Rummel (2016) found that, overall, exploratory learning conditions result in higher conceptual knowledge but equal procedural knowledge compared to more traditional tell-then-practice conditions. However, beyond those two measures, the process level and other learning outcome measures diverge between studies that use CC or RD activities. Studies using CC activities sometimes analyze deep structure acquisition (Glogger-Frey, Fleischer, Grüny, Kappich, & Renkl, 2015; Glogger-Frey, Gaus, & Renkl, 2017), whereas studies using RD activities tend to measure the number of representations and solution methods. There are no consistent measures of cognitive load, knowledge gap awareness, or performance on the activity (i.e., number of strategies or success).

In addition, different learning mechanisms have been proposed based on the different activities. Kapur (2008) observed that, although students in an exploratory
learning condition with an RD activity failed to generate the solution, they learned more than students in a direct instruction condition. Kapur named this effect *productive failure*. During an RD activity, students struggle and fail to reach the canonical solution and become aware of their knowledge gaps. Researchers using RD activities and novel problem sequences typically focus on failure, knowledge gap awareness, and knowledge revision as learning mechanisms (DeCaro & Rittle-Johnson, 2012; Kapur, 2012; Loibl & Rummel, 2014a). This awareness could result in a “need to know” feeling following the activity, which could make students more engaged during the instruction (e.g. DeCaro & Rittle-Johnson, 2012).

In contrast, researchers who have used CC activities tend to emphasize deep structure learning and new knowledge construction as potential learning mechanisms (Chin et al., 2016; Loehr et al., 2014; Schwartz et al., 2011). During a CC activity, students are guided to recognize the important dimensions in the new conceptual domain due to the problem features designed into the activity (Schwartz et al., 2011). In addition, students are often provided enough information to construct the solution. Students who successfully derive a canonical solution may experience a feeling of being in control of their learning, and may enjoy the success of invention.

Lastly, within the variety of exploratory learning studies, researchers have applied different degrees of experimental control. In particular, many productive failure experiments using RD activities have lacked control of experimental materials. For example, Kapur (2010) compared a lecture and practice condition to a productive failure condition. The total amount of time spent in class on the study was the same between conditions at seven, 55-minute class periods. In the lecture and practice condition, the
teacher introduced content guided by the course workbook, worked through example problems, and students then worked individually on well-defined problems, in class and as homework. In the productive failure condition, students were placed in collaborative groups of two or three, and first worked together on a complex problem. In the next period, the students were asked to solve two extension problems individually. These extension problems were designed to force students to consider the impact of specific parameters in the group problem. The lecture and practice condition did not see either of these types of problems. No homework was assigned in the productive failure condition. Students in the productive failure condition performed slightly better on well-defined problems and much better on application problems than students in the lecture and practice condition. Kapur concluded that there may be some hidden productivity from exploration, however, the effect of instructional order was confounded by using different materials. Other studies with RD activities had similar confounds (Kapur, 2008, 2012, 2014; Kapur & Bielaczyc, 2012).

The objective of the current research was to assess the impact of different activities on learning outcomes in an exploratory learning context, using a tightly controlled in vivo experimental design. Three studies compared CC and RD activities with multiple learning outcomes and process level measures to investigate learning mechanisms due to exploratory learning. These studies further the understanding of why exploration can be effective for learning and help to identify best practices for designing exploration activities, both for future research and educators’ use in the classroom.
Exploration Activities

The two most common activities in the exploratory learning literature use contrasting cases (CC) or rich datasets (RD). The activity types vary primarily in the amount and type of information given to students and may differ in the learning mechanisms they invoke.

CC Activities

Contrasting cases are carefully designed sets of examples that vary along the critical problem features or dimensions of a concept (Schwartz et al., 2009). The activity instructions typically include an invention prompt that encourages students to integrate the cases, observe the designed contrasts, and recognize dimensions of the new conceptual domain. It is possible, but not required, for students to generate a canonical formula or solution to the problem. Even without inventing the canonical solution, students can make learning gains by recognizing the factors that are important in the solution.

For example, Schwartz, Chase, Oppezzo and Chin (2011) used contrasting cases of “clown crowdedness” on busses to help teach eighth-grade students about density. Six contrasting cases were provided (see Figure 1). The cases in this example varied by number of clowns and size of bus, which analogically refer to mass and volume, the two components used to calculate density. Students were asked to “create an index of clown crowdedness.”
The design of cases within a CC activity is critical. Within a set of cases, the dimensions must be clear, not confounded, and discoverable (Schwartz et al., 2009). Schwartz, Lindgren, and Lewis (2009) suggest that adults can handle contrasting cases that target three to four conceptually central concepts at a time. The benefits of contrasting cases would likely depend on how the cases differ (Lee & Anderson, 2013a).
RD Activities

RD activities differ from CC activities in the amount and type of information presented to students. In RD activities, a richly contextualized problem is presented with a cover story, many data points, and a novel goal (Loibl et al., 2016; Loibl & Rummel, 2014a). Unlike in CC activities, where the problem features are discoverable, the large amount of data in RD activities hides the underlying concepts. Therefore, students who work on RD activities do not commonly develop the canonical solution. The large amount of data in RD activities also presents more opportunities for students to try out methods from prior knowledge to accomplish the goal. The number of student-generated representations and solution methods is often reported to be a significant predictor of performance outcomes (Kapur, 2012, 2014; Toh & Kapur, 2017).

For example, Kapur (2014) used a rich dataset to introduce ninth-grade students to standard deviation (see Figure 2). The goal of the activity was to determine the most “consistent” basketball player based on two players’ scores during many different games. At this point in their mathematical training, students were familiar with calculating averages and performing basic operations but had not yet learned standard deviation; calculating “consistency” was a novel task. When given this activity in an exploratory learning context, students generated many strategies but failed to derive the equation for standard deviation. Students in this condition demonstrated better conceptual understanding and transfer than students in a direct instruction condition. The number of student-generated solutions during the activity was predictive of learning outcome performance in the exploratory learning group.
Activity Instructions

A few studies have investigated the importance of the activity instructions within exploratory learning. One study using a CC activity found that a prompt of “invention” results in better knowledge transfer than “compare and contrast” (Chin et al., 2016). Another study using a CC activity found that including “use math” as opposed to “use words” in the instructions results in better qualitative and conceptual understanding (Schwartz, Martin, & Pfaffman, 2005). Although no studies have investigated the instructions in RD activities, instructions have varied between studies, from invention (“generate a quantitative index”; Kapur, 2012) to strategy generation (“generate as many solutions as possible”; Kapur, 2014), with similar conceptual learning benefits observed.
Learning Outcomes

In their review, Loibl et al. (2016) found that exploratory learning does not have a systematic effect on procedural knowledge, whereas most studies show an improvement in conceptual knowledge over a direct instruction condition. Procedural knowledge is defined as sequential actions that can be used to solve problems (Jonassen, 2009; Loibl et al., 2016; Rittle-Johnson, Siegler, & Alibali, 2001). This type of knowledge can be represented as production rules (e.g. if(a), do function f(b, c); Anderson, 1996). Procedural knowledge can be assessed directly by giving a problem that can be solved with the same step-by-step process taught in the instruction. In contrast, conceptual knowledge is abstract and relational, and represents principles in a domain as well as the connection between associated concepts (Jonassen, 2009; Loibl et al., 2016; Rittle-Johnson et al., 2001). Conceptual knowledge assessments typically require learners to describe principles using relational words. Conceptual and procedural knowledge are not separate and isolated; Rittle-Johnson, Siegler, and Alibali (2001) demonstrated that procedural knowledge and conceptual knowledge are linked and learned iteratively, and each is able to change based on the development or level of the other.

Additionally, researchers have acknowledged that there are goals for educational methods beyond knowledge acquisition, such as generalized domain understanding (Kalyuga & Singh, 2016). Differences in knowledge generalization, depth and flexibility can appear in transfer measures. Transfer requires the learner to adapt existing conceptual and procedural knowledge to a new situation or a different type of problem (Loibl et al., 2016). Because of the high cognitive demands, transfer tasks are sensitive to differences in understanding (Michael, Klee, Bransford, & Warren, 1993; Schwartz &
Generalized domain understanding can also be measured by “preparation for future learning assessments” as developed by Schwartz and Bransford (1998). In preparation for future learning (PFL) assessments, a learning resource is provided within the assessment, and students are tested on whether they learn from the new resource and apply the new knowledge again within the assessment.

Schwartz and Martin (2004) used a PFL assessment to study exploratory learning and direct instruction methods in a ninth-grade algebra class. The lesson was about variance, an early concept in statistics. On the learning assessments, half of the students in each condition received a problem that required them to invent a technique for normalization, a concept related to variance. A separate transfer question (at least two questions later in the assessment) required the students to recognize that normalization was required in another scenario. In this “double-transfer paradigm,” students in the exploratory learning condition with the embedded learning resource on their exam were twice as likely to be able to solve this problem than the other conditions (exploratory learning without a resource, direction instruction with a resource, direct instruction without a resource).

Schwartz and colleagues (2009) argue that there are benefits of constructivist methods that can only be observed with PFL assessments. PFL transfer questions are thought to detect the extent to which a student was prepared to learn from the resource (Schwartz et al., 2009). Students that have a more generalized domain understanding (e.g., due to an exploratory learning condition) should be more prepared to learn and apply their new knowledge (Schwartz & Martin, 2004). Research in exploratory learning
supports this point (Schwartz & Bransford, 1998; Schwartz et al., 2009; Schwartz & Martin, 2004; Sears, 2006).

**Learning Mechanisms**

The benefits of exploratory learning can be explained in several ways, some of which may be activity-specific. Activation of prior knowledge and learner interest and enjoyment are referenced across the literature (Capon & Kuhn, 2004; Kapur, 2014; Loehr et al., 2014; Schwartz & Bransford, 1998), whereas failure and attention to deep structure are emphasized in RD and CC activity research respectively.

**Activation of prior knowledge**

As learners work towards a solution to a novel problem, they activate, probe, and differentiate their existing knowledge (Kapur, 2011, 2012; Schwartz & Bransford, 1998). For example, when determining the most “consistent” basketball player from two lists of data (Figure 2), students often use strategies like calculating the average, finding the most common score (the mode), using subtraction and addition, and several other mathematical operations that they already know. Whether they are using trial-and-error methods or building new solutions, when learners attempt something new, they draw upon known information. This activation and reprocessing of prior knowledge may result in stronger connections between old and new information, perhaps providing a better overall organizational schema (Weaver et al., 2018). Activation of prior knowledge likely occurs for both RD and CC exploratory learning activities.

**Interest and enjoyment**

It is possible that having an opportunity to explore new content could increase learners’ interest and enjoyment both during the activity and during the instruction. Most
studies that have examined interest and enjoyment, however, have shown no difference between instructional orders (Glogger-Frey et al., 2015; Kapur, 2014; Newman & DeCaro, 2019). Weaver and colleagues (2018) found increased interest and enjoyment in one of their two studies. Despite predominantly null interest and enjoyment results, the idea is discussed throughout the exploratory learning literature. This is because a lack of an effect is still important – it shows that even when students are presented with a difficult assignment, general motivational factors are not disrupted.

Exploratory learning conditions have been shown to increase learners’ cognitive load (Kapur, 2014; Newman & DeCaro, 2019; Toh & Kapur, 2017). Cognitive load is the mental effort needed to perform a task (Paas, 1992). According to cognitive load theory (Sweller, Ayres, & Kalyuga, 2011), the best instructional techniques for domain knowledge acquisition reduce the extraneous cognitive load of the learners. However, Kalyuga and Singh (2016) discuss the importance of defining a learning goal when determining the best instructional design. They explicitly point to exploratory learning, stating that if the goal of an exploration activity is to activate and differentiate prior knowledge in preparation for instruction (as opposed to the acquisition of knowledge), minimal cognitive load may not be optimal.

RD and CC activities likely vary in their impact on cognitive load due to their design. RD activities are designed to be complex, in order to provide opportunities for students to try multiple (incorrect) solution methods. This means that RD activities include extraneous information that hides the canonical solution and may also make the underlying problem features more difficult to find. CC activities, on the other hand, are designed primarily to highlight key problem features.
Knowledge Gap Awareness

Kapur (2008) coined the term “productive failure” to describe the learning improvements from exploratory learning despite the failure of students to derive the canonical solution during the activity. By failing to solve a novel problem, students become aware of their knowledge gaps, and desire to learn the material that they do not know (see Kapur, 2016; Loibl & Rummel, 2014; Ohlsson, 1996; Schwartz et al., 2009). The perception of knowledge gaps has been assessed in some exploratory learning studies. Perception of knowledge gaps is typically measured after the activity in a survey (e.g., “Compared to most other people, I know less about calculating [consistency]”; Glogger-Frey et al., 2015; Newman & DeCaro, 2019). Loibl and Rummel (2014) found that awareness of specific knowledge gaps (elicited during instruction) was of critical importance to learning outcomes.

However, Glogger-Frey et al. (2015) found mixed results. In two studies, they found that perceived knowledge gaps correlated with learning outcomes in opposite directions. Student teachers’ learning outcomes increased with their perceived knowledge gaps, whereas eighth-grade students’ learning outcomes were negatively correlated with knowledge gap awareness. Loibl & Rummel (2014b) also found differences between instructional order in perceived knowledge gaps, however they concluded that knowledge gap awareness alone was not sufficient to facilitate learning. They proposed that two mechanisms would support learning: (1) prior knowledge activation and perception of knowledge gaps, followed by (2) a specific type of instruction that helps students differentiate between the canonical solution and typical erroneous solutions.
The theory of error correction (Ohlsson, 1996) supports the idea that failure leads to learning. Ohlsson suggests that learners revise their knowledge structures only after activating their existing ideas and recognizing that they are faulty or incomplete. Chi (2000) also suggested that learners first must detect flaws in their own imperfect mental models before they can repair the models. A recent meta-analysis of the productive failure literature focuses on learning from failure as the explanation for the overall improvement of conceptual knowledge in these studies (Darabi, Arrington, & Sayilir, 2018). Although authors found only twelve controlled experimental studies on productive failure, they concluded that there was a positive effect of learning from failure.

**Discernment of problem features**

Instead of reporting error correction as the primary mechanism for exploratory learning benefits, researchers using CC activities discuss how exploration focuses learners’ attention on the deep structure of the new conceptual environment (Chin et al., 2016; Loehr et al., 2014; Schwartz et al., 2011). Exploration includes processes of active observation, interpretation, and representation. These processes applied to a novel conceptual domain can allow students to learn not only the single canonical solution that they are deriving, but also the importance of underlying problem features. Schwartz et al. (2011) as well as Glogger-Frey et al. (2015) found that students who attended to the deep structure of the contrasting cases had greater learning outcomes. Introducing the canonical solution after highlighting the deep features (by providing contrasting cases during problem solving or by building instruction on student solutions) enables students to organize the target knowledge by its deep features (Loibl et al., 2016).
In CC activities, the individual cases are designed to clearly delineate the problem features; the deep structure of the concept is used to design the activity. In the CC example with clown cars (see Figure 1), the cases differ by number of clowns and number of cars such that students see that clown quantity as well as volume are two different dimensions of the solution (Schwartz et al., 2011). Students must use convergent thinking to put the available dimensions together to find the canonical solution.

In RD activities, the critical features are not obvious within the large amount of data. As students generate strategies, they may or may not discover the underlying concept features. However, Kapur and Bielaczyc (2012) also suggest that productive failure RD activities be designed with problem features in mind. They state that the identification of problem features is part of the learning process, alongside differentiation of prior knowledge, development of multiple incorrect strategies, strategy review, and consolidation into the canonical solution method. They found that the more strategies students generated, the better they learned. Therefore, the mechanism of problem feature discernment could be the primary mechanism of exploratory learning with RD activities as well as with CC activities.

**Current Studies**

The current studies compare student learning outcomes and process level measures between randomized, controlled experimental conditions that vary by activity type (CC or RD) and instructional order (exploratory learning or direct instruction). The goal is to investigate learning mechanisms and determine whether activity type moderates the learning benefits of exploratory learning. Study 1 investigated a CC activity in a
novel physics domain. Study 2 investigated an RD activity in the same domain. Study 3 then compared the activities in a 2 (instructional order: explore-first, instruct-first) × 2 (activity type: CC, RD) design.

Study 1 tested a CC activity on the topic of gravitational field in an undergraduate physics classroom. This study used newly designed physics materials (activity, instruction, and assessments) combined with established process level survey measures. In the Explore-First (EF)-CC condition, students were given the CC activity followed by instruction. In the Instruct-First (IF)-CC condition, students received instruction and then worked on the activity. Students in both conditions took a survey after completing the activity and completed a posttest at the end of class. Process level survey measures included success on the activity and survey items on interest and enjoyment and cognitive load. Learning outcome measures included procedural and conceptual knowledge and PFL procedural knowledge and transfer.

The primary learning outcome hypotheses were that students in both conditions would perform the same on the procedural assessment, and that students in the EF-CC condition would perform better on conceptual knowledge and PFL assessments than students in the IF-CC condition. These findings would support prior research using CC activities (e.g. Chin et al., 2016; Schwartz et al., 2011; Schwartz & Martin, 2004). Because the CC activity has not been used before, analyses of success on the activity would be exploratory. However, it was expected that the students in the IF-CC condition would have higher success than students in the EF-CC condition. Activity success could result in either higher learning outcomes or no difference in learning outcomes. A positive relationship was expected, because success on the activity would mean that
problem features were correctly identified. Cognitive load was expected to be greater for students in the EF-CC condition, because invention is more difficult than applying learned knowledge. Interest and enjoyment could vary in several ways. Compared to students in the IF-CC condition, students in the EF-CC condition could experience (1) higher interest and enjoyment due to successful discovery, (2) equal interest and enjoyment, or (3) lower interest and enjoyment because the activity was more difficult. Because the CC activity design was expected to enable discovery of many problem features, it was expected that students in the EF-CC condition would experience equal or greater interest and enjoyment than students in the IF-CC condition.

Study 2 tested examined the use of an RD activity on the same topic as Study 1, in EF and IF conditions in two undergraduate physics classrooms. The RD activity instructions were designed to match the CC instructions using the word “invent.” Process level and outcome measures were modified from Study 1. Hypotheses were loosely based on productive failure research with RD activities (e.g. Kapur, 2012, 2014), with the knowledge that these studies lack tightly controlled methods. One possibility is that students in the EF-RD condition could demonstrate greater conceptual knowledge than students in the IF-RD condition, and all students could have equivalent procedural knowledge. This finding would support the idea that reversing the order of instruction and an RD activity results in conceptual learning gains. An alternative hypothesis is that the procedural and conceptual learning outcomes would be the same across conditions. Null results might indicate that prior research findings were due to the differences in materials and other experimental confounds as opposed to the instructional order manipulation. There was no prior research on the effect of an EF-RD condition on PFL. It
is possible that, on the PFL assessment, students in the EF-RD condition would (a) perform the same as, or (b) perform better than students in the IF-RD condition. Similar performance between instructional orders in Study 2 (RD) might indicate that the learning mechanisms engaged by an RD activity are different than those engaged by a CC activity, and that this difference affects performance on PFL. A learning benefit in the EF-RD condition over the IF-RD condition might indicate that similar learning mechanisms are activated in both activities, and that the instructional order causes an increase in PFL.

Hypotheses regarding the process level measures were also exploratory in Study 2. The number of strategies was expected to be higher in the EF-RD condition than the IF-RD condition, because a specific strategy had not yet been specified. Then, based on previous analyses of number of strategies (e.g., Kapur, 2014), a greater number of strategies was expected to positively predict learning outcomes. This possibility would support the mechanism of prior knowledge activation at work for students the EF-RD condition. Unlike the CC activity, the RD activity was designed to be difficult and result in failure. It was therefore anticipated that students in the EF-RD condition would report higher perceived knowledge gaps and cognitive load following the activity than students in the IF-RD condition. Perceived knowledge gaps were expected to positively predict learning outcomes based on the theory of error correction. Higher experienced cognitive load, however, could result in either (a) lower performance, (b) higher performance, if the load resulted in additional motivation for error correction, or (c) no relationship to performance. It was expected that the load would not be detrimental, and that higher cognitive load would also activate the error correction mechanisms. Finally, based on the
expectation of failure on the RD activity, students in the EF-RD condition were expected
to experience equal or lower interest and enjoyment levels than in the IF-RD condition.

Study 3 was designed to replicate the results of Studies 1 and 2, as well as directly
compare the CC and RD activity types, using a 2 (activity type: CC, RD) × 2
(instructional order: EF, IF) design in two physics courses. Learning outcomes and
process level measures were the same as in Study 2.

Students were expected to obtain the same levels of procedural knowledge in all
conditions. On conceptual knowledge, students in the EF-CC condition were expected to
have higher conceptual knowledge than students in the IF-CC condition. It was possible
that students in the EF-RD condition would have higher conceptual knowledge than
students in the IF-RD condition equal to the difference between students in the EF-CC
and IF-CC conditions. This finding would indicate that there is a general mechanism
related to instructional order, unrelated to activity type. However, it was also possible that
the EF-RD and IF-RD conditions would result in equivalent performance, indicating that
the results observed in the productive failure literature were based on other variations
outside of instructional order.

In addition to expected differences between instructional orders, there were many
possible effects of activity type. For example, it was possible that students in the EF-CC
condition would outperform students in the EF-RD condition. Because the CC activity
was designed for students to identify the problem features, and the EF-RD condition was
designed to create a sense of knowledge gap awareness and failure, a difference in
performance between these groups would indicate that one mechanism may be more
important than another, and that different activities activate different mechanisms. The
primary hypothesis was that an EF order would significantly improve conceptual learning, and students in the EF-CC would perform the same as or higher than students in the EF-RD condition.

Because there was no prior research on PFL assessments and RD activities, primary hypotheses were as follows: either the EF conditions would result in significantly higher PFL performance, or only the EF-CC condition would result in higher PFL performance. A significant difference for only one condition would indicate that activity type is a moderator of the PFL benefits of exploratory learning.

The examination of effects of activity type on cognitive load and interest and enjoyment is also empirically novel. Because the RD activity is likely more difficult, students in the EF-RD condition were expected to experience the highest cognitive load, followed by the EF-CC condition, and lastly the IF conditions. In addition, because of the experience of failure, students in the EF-RD condition were also expected to have the highest perceived knowledge gaps, above students in the EF-CC condition and then followed by students in the IF conditions. Number of strategies and success on the activity were expected to be positively correlated with conceptual understanding and PFL performance across conditions.

**Theoretical and Practical Implications**

Only a few studies have demonstrated the effectiveness of exploratory learning in an undergraduate classroom (Newman & DeCaro, 2019; Weaver et al., 2018; Westermann & Rummel, 2012). The current work extends the literature by testing exploratory learning on a new undergraduate STEM topic, and is the first to use a PFL assessment in an undergraduate classroom. Also, the current studies are the first to
examine exploratory learning mechanisms by comparing different exploration activities. The studies were designed using controlled experimental methods that vary only the type of activity and the order of activity and instruction. Different types of exploration activities have not yet been compared in the same study with the same outcome measures.

Prior research has shown that exploratory learning improves conceptual knowledge, and preparation for future learning, and these benefits are thought to be due to activation of prior knowledge, knowledge gap awareness, and/or focus on problem features (DeCaro & Rittle-Johnson, 2012; Kapur, 2008, 2010, 2011, 2012, 2014; Loibl et al., 2016; Loibl & Rummel, 2014a; Newman & DeCaro, 2019; Schwartz & Bransford, 1998; Schwartz et al., 2011, 2009; Weaver et al., 2018). It is possible that some of these learning mechanisms are activated in both RD and CC activities, whereas others are activated by just one type of activity. This set of studies will provide valuable information to develop the theory behind exploration in the classroom.
CHAPTER II: STUDY 1

Study 1 tested learning outcomes following a CC activity on the topic of gravitational field in an undergraduate physics classroom. This study extended the recent application of exploratory learning in the undergraduate physics classroom (Weaver et al., 2018) with a different topic, and was the first to utilize a PFL assessment in an undergraduate classroom exploratory learning context. Process level measures included interest and enjoyment, cognitive load, and success on the activity. Based on prior studies using CC activities, it was expected that students in the EF-CC condition would score higher on conceptual knowledge and the PFL assessment but also experience higher cognitive load and lower success on the activity than students in the IF-CC condition. Performance on the procedural knowledge assessment was expected to be equal between groups. Interest and enjoyment responses were expected to be equal between groups or higher in the EF-CC condition.

Method

Participants

Participants (N = 129) were undergraduate students enrolled in one section of Introductory Mechanics, Heat and Sound, a calculus-based first-semester physics course for engineering and physics majors. All students who attended class on the day of the study were participants, but four were removed from the dataset for not attempting the
PFL assessment. Two students did not turn in their activity sheets, six students did not turn in surveys, and one student did not answer the cognitive load question, so the sample size was reduced for process level analyses that depended on those data. The participants with missing process-level data were not excluded from the learning outcome analyses, in order to maximize statistical power; analyses without these participants showed similar results. The number of participants by condition are shown in Table 1. One student did not respond to the knowledge gap awareness question.

Table 1: Study 1 Number of Participants by Order

<table>
<thead>
<tr>
<th></th>
<th>EF-CC</th>
<th>IF-CC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>82</td>
<td>129</td>
<td></td>
</tr>
</tbody>
</table>

Materials

Study 1 included five phases (instruction, CC activity, activity review, survey, and assessment), and the order of the phases varied by condition (see Table 2). Materials used in each phase are described below.

Table 2: Instructional Order in Explore-first and Instruct-first Conditions.

<table>
<thead>
<tr>
<th>Phase</th>
<th>EF-CC</th>
<th>IF-CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>EF-CC Activity</td>
<td>Instruction</td>
</tr>
<tr>
<td>2</td>
<td>Survey</td>
<td>IF-CC Activity</td>
</tr>
<tr>
<td>3</td>
<td>Instruction</td>
<td>Survey</td>
</tr>
<tr>
<td>4</td>
<td>Activity Review</td>
<td>Activity Review</td>
</tr>
<tr>
<td>5</td>
<td>Assessment</td>
<td>Assessment</td>
</tr>
</tbody>
</table>

Instruction. The course instructor gave a 12-minute presentation on gravitational field (four primary learning objectives for this topic are delineated in Appendix A). In the instruction, gravitational field was first described with respect to gravitational force, which students had learned in the previous class: a source mass creates a gravitational field, and if a test mass is added, the gravitational force is the test mass times the field \( F = mg \). Instruction then led students through the calculation of gravitational field
magnitude \( (g = Gm/r^2) \) and direction (towards the source mass). Next, the instructor discussed how vector addition can be used to calculate the gravitational field for points in space due to multiple source masses. Finally, an example problem was given in which magnitude and direction of the gravitational field was calculated.

The instruction was given using Microsoft PowerPoint slides, as well as use of the whiteboard. Though not scripted, the instruction was based on the detailed slides. The instructor had also given this same lecture numerous times in the past. Thus, though the two lectures may have varied slightly, they were instructionally equivalent.

**CC Activity.** The IF-CC activity is illustrated in Figure 3. The CC activity was designed to highlight the first three learning objectives of gravitational field: magnitude, direction, and vector nature (see Appendix A for a more detailed description of the learning objectives; a fourth objective relates the new concept to a previous concept).

Three sets of axes showed scenarios with different arrangements of masses \((m_1\) to \(m_5)\) and points \((P_A\) to \(P_J)\). A table was also given with a list of gravitational field vectors \((g_1\) to \(g_{10})\), and their magnitudes and directions. The sets of cases were designed to allow students to discover relations between them. For example, points \(P_A\) and \(P_B\) were located below a single source mass \(m_1\). From prior experience, students were likely to hypothesize that the field at those two locations has one of two vertical directions (either up or down). Only two vectors \(g_7\) and \(g_9\) in the table have a vertical direction (Up). While considering these two points, students can discover that the direction of a gravitational field from a single source mass points towards the mass.
Gravitational Field Activity

The figures below on the left show a series of particles, labeled \( m_1 \) through \( m_5 \), and points in space, labeled \( P_A \) through \( P_f \). Values for masses and distances are given in each figure.

<table>
<thead>
<tr>
<th>Vector</th>
<th>Magnitude</th>
<th>Direction</th>
<th>Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \vec{g}_1 )</td>
<td>1.75( G )</td>
<td>60° below (+x) axis</td>
<td>( P_F )</td>
</tr>
<tr>
<td>( \vec{g}_2 )</td>
<td>10( G )</td>
<td>To the left</td>
<td></td>
</tr>
<tr>
<td>( \vec{g}_3 )</td>
<td>( \frac{10}{3} G )</td>
<td>To the right</td>
<td></td>
</tr>
<tr>
<td>( \vec{g}_4 )</td>
<td>2.5( G )</td>
<td>To the left</td>
<td></td>
</tr>
<tr>
<td>( \vec{g}_5 )</td>
<td>0</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>( \vec{g}_6 )</td>
<td>( \frac{130}{4} G )</td>
<td>To the right</td>
<td></td>
</tr>
<tr>
<td>( \vec{g}_7 )</td>
<td>10( G )</td>
<td>Up</td>
<td></td>
</tr>
<tr>
<td>( \vec{g}_8 )</td>
<td>2.96( G )</td>
<td>4.3° below (+x) axis</td>
<td></td>
</tr>
<tr>
<td>( \vec{g}_9 )</td>
<td>2.5( G )</td>
<td>Up</td>
<td></td>
</tr>
<tr>
<td>( \vec{g}_{10} )</td>
<td>3.54( G )</td>
<td>45° below (+x) axis</td>
<td></td>
</tr>
</tbody>
</table>

Each row in the table above gives a gravitational field vector \( \vec{g} \) with its magnitude and direction. Use the information given in the table to practice the calculation of gravitational field \( \vec{g} \). The gravitational field \( \vec{g} \) is a vector quantity that is related to the gravitational force.

- Use the mathematical formula you have just learned to calculate the magnitude and direction of the gravitational field \( \vec{g} \) for every point in the figures above.
- In the far right column of the table, write the point \( P_A \) through \( P_f \) that corresponds to each vector. \( P_F \) has already been given, as we calculated it as part of the instruction.

Figure 3: IF-CC Activity.
The EF-CC and IF-CC activities differed only in the instructions (Figure 4). IF-CC activity instructions asked students to “use the mathematical formula you have just learned to calculate the magnitude and direction of the gravitational field,” whereas the EF-CC activity instructions asked students to “invent a mathematical formula to describe the magnitude of the gravitational field, and a rule to describe the direction that works for every point in the figures.” The words in the instructions were otherwise chosen to be as parallel as possible.

**CC Activity Review.** A completed version of the activity, with answers, was supposed to be briefly reviewed by instructors immediately prior to the assessment. In this implementation, students in the IF-CC condition were shown the answers very briefly on the screen but, due to an error, students in the EF-CC condition were not shown the answers to the activity. The answers were presented so briefly in the IF condition that any differences between conditions were likely minimal.

**Survey.** The post-activity questionnaire included interest and enjoyment and cognitive load questions, as well as other survey items for another study that will not be discussed further. The interest and enjoyment scale (Cronbach’s $\alpha = .82$) was adapted
from Ryan (1982; e.g., “I enjoyed this learning activity”). Participants responded to four items on a Likert scale from 1 (strongly disagree) to 5 (strongly agree): Today’s activity has been interesting; I have enjoyed today’s activity; Today’s activity really captured my attention; Today’s activity kept me engaged.

Cognitive load was measured with the Mental Effort Rating Scale (Paas, 1992). The question read: “In completing the learning activity today I invested:”, and the response was on a Likert scale ranging from 1 (very, very low mental effort) to 9 (very, very high mental effort).

Assessments. The assessments measured procedural knowledge, conceptual knowledge, and PFL. All assessments were multiple choice (see Appendix F). The procedural knowledge items (9 items; $\alpha = .62$) asked students to calculate magnitude and direction of the gravitational field at different points in space with different mass scenarios. These procedural calculations were embedded in the activity and taught explicitly in the instruction. The conceptual knowledge items (10 items; $\alpha = .76$) were True/False questions that queried relational and verbal understanding of the learning objectives. Embedded in these questions were several common misconceptions.

The PFL assessment included a learning resource, new procedural questions (9 items; $\alpha = .82$), and a final transfer question (1 item; Appendix C). The learning resource was on electric field, a topic that had not yet been taught to these students, including both a description and a worked-example problem. The procedure for calculating electric field magnitude and direction is very similar to gravitational field and direction, and the underlying concepts are similar as well. The transfer question asked students to predict
the equation for electric force, which has the same relationship to electric field as gravitational force and field.

**Procedure**

The study took place over one 75-minute class period. Students were randomly assigned to condition and split into two different classrooms. Students in the IF-CC condition completed the study in their regular classroom, which was a large lecture hall (188 fixed seats). Students in the EF-CC condition completed the study in a 60-seat active learning classroom (with design attributes like rolling tables for 4, and white noise). No classrooms of similar layout and size were available on the date and time of the study. Students were instructed to work in groups of 4 for the activity, and were allowed to self-select their groups. During the activity, activity sheets and blank worksheets were passed out to every other student.

Three instructors switched between the two classrooms for different phases of the study. The course instructor gave the instruction, a secondary instructor (graduate student) led the activity, and a third instructor (psychology professor) gave the survey. In addition, the course TA helped with the activity. In the IF-CC condition, students were given the instruction (20 min), followed by the activity (20 min) and survey (5 min). Then, the secondary instructor (graduate student) briefly reviewed the activity (1 min) and administered the assessments (10 min for the procedural and conceptual items, and 15 min for the PFL). In the EF condition, students completed the activity (20 min) and survey (5 min), followed by the instruction (20 min). Then, the course instructor gave the assessment (10 min for the procedural and conceptual items, and 15 min for the PFL). A research assistant in each classroom took notes and kept track of time.
Results

Students in the EF-CC condition were expected to outperform students in the IF-CC condition on the conceptual knowledge and PFL assessments, and performance on the procedural assessment was expected to be the same in both conditions (i.e., Loibl et al., 2016; Schwartz et al., 2009). In addition, it was hypothesized that interest and enjoyment and cognitive load would be equal or higher in the EF group (i.e., Toh & Kapur, 2017; Weaver et al., 2018). It was also hypothesized that all of the process measures (interest and enjoyment, cognitive load, and success on the activity) would predict learning outcomes.

Learning outcomes by condition

Performance on the procedural and conceptual assessments was examined using a 2 (type of knowledge: procedural, conceptual) × 2 (order: EF, IF) mixed-factorial analysis of variance (ANOVA), with order as a between-subjects factor and type of knowledge as a within-subjects factor. The main effect of order was not significant, $F < 1, p = .586$. There was a significant main effect of type of knowledge, $F(1, 127) = 25.75, p < .001, \eta^2_p = .17$, with procedural questions ($M = 72.76\%, SE = 1.67$) answered more accurately than conceptual questions ($M = 61.09\%, SE = 2.18$).

The main effect was qualified by a significant order × type of knowledge interaction, $F(1, 127) = 7.57, p = .007, \eta^2_p = .06$. Simple main effects (Figure 5) were evaluated using confidence intervals, and estimates of Cohen’s $d$ effect size are reported for significantly different results. Procedural knowledge performance was not significantly different between students in the EF order ($M = 70.45\%, SE = 2.66, 95\% CI [65.19, 75.71]$) and students in the IF order ($M = 75.07\%, SE = 2.01, 95\% CI [71.09,
However, conceptual knowledge performance was significantly higher for students in the EF order ($M = 65.11\%, SE = 3.48$, 95% CI $[58.23, 71.98]$) than students in the IF order ($M = 57.07\%, SE = 2.63$, 95% CI $[51.87, 62.28]$, $d = .36$).

![Figure 5: Study 1 posttest performance by order. Error bars = ±SE.](image)

Performance on the PFL assessment was also examined using a 2 ($order$) × 2 ($type of knowledge$: PFL procedural, PFL transfer) mixed-factorial ANOVA, with $order$ as a between-subjects factor and $type of knowledge$ as a within-subjects factor. There was a main effect of $order$, $F(1, 127) = 5.20$, $p = .024$, with students in the EF-CC condition ($M = 53.78\%, SE = 3.90$, 95% CI $[46.06, 61.51]$) performing better than students in the IF-CC order ($M = 42.62\%, SE = 2.96$, 95% CI $[36.77, 48.46]$). There was also a significant main effect of $type of knowledge$, $F(1, 127) = 115.44$, $p < .001$, $\eta^2_p = .48$, with procedural questions ($M = 70.39\%, SE = 2.48$, 95% CI $[65.48, 75.30]$) answered more accurately than the transfer question ($M = 26.01\%, SE = 3.79$, 95% CI $[18.51, 33.51]$). These effects were qualified by a significant $order \times type of knowledge$ interaction, $F(1, 127) = 4.91$, $p = .029$, $\eta^2_p = .04$. 

30
As shown in Figure 6, students in the EF-CC condition ($M = 36.17\%$, $SE = 6.04$, 95% CI [24.21, 48.13]) scored significantly higher on the transfer question than students in the IF-CC condition ($M = 15.85\%$, $SE = 4.57$, 95% CI [6.80, 24.91], $d = .33$). PFL procedural performance did not significantly differ between students in the EF-CC condition ($M = 71.40\%$, $SE = 3.96$, 95% CI [63.57, 79.22]) and students in the IF-CC condition ($M = 69.38\%$, $SE = 3.00$, 95% CI [63.45, 75.30], $d = .05$).

**Process level measures**

Activity success, interest and enjoyment, and cognitive load were (1) compared between instructional order conditions using 1-way ANOVAs, and (2) used as predictor variables in regression models for each learning outcome (procedural knowledge, conceptual knowledge, PFL procedural knowledge, and PFL transfer).

**CC Activity Success.** CC activity success was scored by counting the number of points correctly matched to the vectors and dividing by the total number of points and vectors (10) to calculate a percent correct score. As expected, the IF group ($M = 94.38\%$, $SE = 2.28$, 95% CI [89.86, 98.89]) had significantly greater success on the activity than
the EF group ($M = 84.26\%, \ SE = 2.98, \ 95\% \ CI \ [78.36, \ 90.15]$), $F(1, \ 125) = 7.27, \ p = .008, \ \eta^2_p = .06$. Regression results in Table 3 show that activity success did not significantly predict any of the learning outcomes.

Table 3: Study 1 Regression Results, CC Activity Success and Learning Outcomes.

<table>
<thead>
<tr>
<th>Knowledge Type</th>
<th>$R^2$</th>
<th>$F$</th>
<th>$\beta$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedural</td>
<td>&lt; .01</td>
<td>.04</td>
<td>-.02</td>
<td>.842</td>
</tr>
<tr>
<td>Conceptual</td>
<td>&lt; .01</td>
<td>1.11</td>
<td>-.09</td>
<td>.295</td>
</tr>
<tr>
<td>PFL Procedural</td>
<td>.01</td>
<td>.84</td>
<td>.08</td>
<td>.362</td>
</tr>
<tr>
<td>PFL Transfer</td>
<td>.03</td>
<td>3.63</td>
<td>-.17</td>
<td>.059</td>
</tr>
</tbody>
</table>

In addition to matching vectors to points in the activity, students in the EF-CC condition were asked to derive the equation for gravitational field. Of the 47 students in the EF condition that were included in the final dataset, 13 derived the correct equation (27.7%).

**Interest and enjoyment.** Mean responses to the interest and enjoyment scale did not differ significantly between students in the IF-CC ($M = 4.92, \ SE = .18, \ 95\% \ CI \ [4.56, \ 5.28]$) and EF-CC ($M = 4.61, \ SE = .14, \ 95\% \ CI \ [4.33, \ 4.89]$) conditions, $F(1, \ 121) = 1.79, \ p = .184$. No outcome measures were significantly predicted by the interest and enjoyment scale ($p > .173$; Table 4).

Table 4: Study 1 Regression Results, Interest and Enjoyment Scale and Learning Outcomes.

<table>
<thead>
<tr>
<th>Knowledge Type</th>
<th>$R^2$</th>
<th>$F$</th>
<th>$\beta$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedural</td>
<td>&lt; .01</td>
<td>.10</td>
<td>-.03</td>
<td>.749</td>
</tr>
<tr>
<td>Conceptual</td>
<td>&lt; .01</td>
<td>.01</td>
<td>-.01</td>
<td>.931</td>
</tr>
<tr>
<td>PFL Procedural</td>
<td>&lt; .01</td>
<td>.49</td>
<td>-.06</td>
<td>.485</td>
</tr>
<tr>
<td>PFL Transfer</td>
<td>.02</td>
<td>1.88</td>
<td>-.12</td>
<td>.173</td>
</tr>
</tbody>
</table>

**Cognitive Load.** Cognitive load was not significantly different between students in the IF-CC ($M = 5.88, \ SE = .22, \ 95\% \ CI \ [5.45, \ 6.31]$) and EF-CC ($M = 5.53, \ SE = .17,$
95% CI [5.20, 5.86]) conditions, \( F(1, 120) = 1.69, p = .196 \). No outcome measures were significantly predicted by cognitive load \( (p > .213; \text{Table } 4) \).

**Table 5: Study 1 Regression Results, Cognitive Load and Learning Outcomes.**

<table>
<thead>
<tr>
<th>Knowledge Type</th>
<th>( R^2 )</th>
<th>( F )</th>
<th>( \beta )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedural</td>
<td>&lt; .01</td>
<td>1.13</td>
<td>.03</td>
<td>.722</td>
</tr>
<tr>
<td>Conceptual</td>
<td>.01</td>
<td>1.24</td>
<td>-.10</td>
<td>.268</td>
</tr>
<tr>
<td>PFL Procedural</td>
<td>.01</td>
<td>1.57</td>
<td>-.11</td>
<td>.213</td>
</tr>
<tr>
<td>PFL Transfer</td>
<td>.01</td>
<td>.82</td>
<td>-.08</td>
<td>.368</td>
</tr>
</tbody>
</table>

**Discussion**

The learning outcome results supported hypotheses. Students in both conditions learned procedural knowledge equally well, and students in the EF-CC condition outperformed those in the IF-CC condition on the measure of conceptual knowledge. On the PFL assessment, students in both conditions performed equally well on the PFL procedural knowledge items, and students in the EF-CC condition performed better than students in the IF-CC condition on the PFL transfer item.

These findings support and extend early research on exploratory learning in an undergraduate STEM classroom (Newman & DeCaro, 2019; Weaver et al., 2018). This experiment used new materials, including a new CC activity, on a new physics topic. One goal in STEM courses is to have students adapt and apply procedures to novel questions. This study demonstrates how a simple procedure – switching the order of instruction and an activity – can make a difference in conceptual understanding. In addition, this study was the first to assess undergraduate students with a PFL task, thus expanding the exploratory learning PFL research beyond high school populations (Schwartz et al., 2009; Schwartz & Martin, 2004). The PFL transfer question asked about electric force on a particle, and students had to take information from the lesson (how gravitational force...
relates to gravitational field) as well as information from the PFL resource (electric field) and derive a function for the electric force. This study shows that exploratory learning can benefit students’ future learning. In STEM disciplines, it is especially important to connect earlier and later concepts throughout multiple semesters. Instructional methods that enable future learning may be especially useful, and this study provides evidence that exploratory learning equips students to make this transfer.

As expected, students had more success on the activity in the IF-CC condition than the EF-CC condition. However, responses to the interest and enjoyment scale and the cognitive load question did not differ between the two instructional orders. The interest and enjoyment scale result replicates previous findings (Glogger-Frey et al., 2015; Kapur, 2014). However, the cognitive load results do not – several researchers have found increased cognitive load when giving activities before versus after instruction (Glogger-Frey et al., 2015; Kapur, 2014; Newman & DeCaro, 2019). In addition, the process level measures did not significantly predict learning outcomes. These null results are still noteworthy; an EF condition that does not increase cognitive load or decrease interest and enjoyment is a positive result, because cognitive load theory predicts that exploration is innately more difficult (Sweller et al., 2011). The null results also indicate that students’ perceived/reported experiences during the activity (as measured by the survey) did not correspond to their learning.

Limitations

There were a few limitations in the implementation of Study 1. First, two different types of classrooms were used for the experiment. The EF condition took place in a specially-designed active learning classroom. It is possible that the classroom type caused
the differences in learning outcomes, rather than order of instruction. For example, four group members faced each other across a table and were able to work together more easily than in the lecture hall, potentially facilitating the learning process. However, it seems unlikely that the physical environment fully accounts for the superior learning in the EF condition, given that these results replicate many other findings from studies examining exploratory learning. It is also possible that the active learning classroom increased interest and enjoyment and decreased cognitive load for students in the EF-CC condition, perhaps eliminating a difference between orders that could have been observed if the classrooms were the same.

The second limitation is that both the procedural and conceptual knowledge subscales were included on the same timed section of the assessment, with the conceptual knowledge items last in the section. Some students did not answer any of the conceptual knowledge assessment \((n = 10)\), and the majority of these students were in the IF-CC condition \((n = 9)\). Unanswered questions were treated as incorrect. Scores from blank conceptual assessments were included in the final sample, because it is unknown whether the students did not complete the items due to a lack of time or lack of knowledge. If these students did only run out of time, it could mean that conceptual knowledge is underestimated for these students.

Finally, the IF-CC group was the only group to receive an activity review. Loibl et al. (2016) indicate that activity reviews are necessary, but especially for RD activities. However, if anything, this procedural mistake would work against the EF condition.
Conclusions

Study 1 provided support for the idea that completing a novel activity prior to instruction improves conceptual knowledge and transfer on a PFL assessment over the traditional instructional order. The findings support previous research, extending exploratory learning to additional content matter in an undergraduate physics classroom. In addition, the finding that students were better able to learn from a new, related topic is quite interesting in the context of undergraduate STEM learning. If exploratory learning helps students understand more deeply and transfer to new learning, then this type of intervention may have a lasting impact.

Studies 2 and 3 were designed to replicate and extend these findings, with attention to modifying procedural details that were limitations in Study 1. Specifically, the assessment was modified to separate the conceptual and procedural assessments, to ensure that students had time to attempt both subscales. In addition, Studies 2 and 3 were conducted in classrooms that were more equivalent and with an activity review for both conditions.
CHAPTER III: STUDY 2

Study 2 investigated an RD activity in both IF and EF instructional orders in two introductory physics classrooms. Productive failure studies using an RD activity have found a conceptual learning benefit of the EF instructional order (e.g., Kapur, 2011; Loibl et al., 2016). However, many productive failure studies did not control the instructional activities given between conditions. It was therefore an empirical question whether condition would significantly impact conceptual knowledge when using a tightly-controlled, experimental design. One possibility is that students in the EF-RD condition will gain higher levels of conceptual knowledge than students in the IF-RD condition. An alternative possibility is that students given an RD activity would perform the same across the two instructional orders. All were expected to achieve the same procedural knowledge.

No previous productive failure studies have used a PFL assessment. However, exploratory learning studies using CC activities showed PFL benefits in an EF order. Therefore, it was hypothesized that students in the EF-RD condition would either outperform or perform the same on the PFL assessment as students in the IF-RD condition. If students in the EF-RD had higher performance than students in the IF-RD condition on any part of the PFL assessment, then the results would indicate that exploratory learning improves learning transfer regardless of activity type. Alternatively,
no difference between the EF-RD and IF-RD groups on PFL questions could be the first to show that productive failure exploration activities do not improve PFL outcomes. In combination with Study 1 results, a no-difference finding would support the idea that isolating or discovering problem features, those that are embedded into CC activities, is required for PFL.

Process level hypotheses were based on previous exploratory learning research with RD activities (i.e., Glogger-Frey et al., 2015; Toh & Kapur, 2017; Weaver et al., 2018): interest and enjoyment, knowledge gap awareness, and cognitive load were expected to be equal or higher in the EF-RD condition than the IF-RD condition. In addition, interest and enjoyment, knowledge gap awareness, cognitive load, and number of strategies generated during the activity for students in the EF-RD condition were expected to predict learning outcomes.

**Method**

**Participants**

Participants ($N = 92$) included all students from two introductory physics courses: *Fundamentals of Physics I*, an algebra-based, first-semester physics course for liberal arts majors (predominantly pre-medicine), and *Introductory Mechanics, Heat and Sound*, a calculus-based, first-semester physics course for engineering students and physics majors (the course as in Study 1, in a different semester). Both courses had the same instructor of record.

The physics courses met three times a week for 50 minutes, and therefore Study 2 required participation over two consecutive class days. Participants who were in the EF-RD and IF-RD conditions and attended both days were included in analyses. An
additional condition was examined as a pilot for another study that will not be discussed or analyzed here. The number of participants by course and condition are shown in Table 6. One student did not respond to the knowledge gap awareness question.

Table 6: Study 2 Number of Participants by Order and Course

<table>
<thead>
<tr>
<th>Course</th>
<th>Condition</th>
<th>EF-RD</th>
<th>IF-RD</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algebra-based</td>
<td>EF-RD</td>
<td>30</td>
<td>23</td>
<td>53</td>
</tr>
<tr>
<td>Calculus-based</td>
<td>IF-RD</td>
<td>22</td>
<td>17</td>
<td>39</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>52</td>
<td>40</td>
<td>92</td>
</tr>
</tbody>
</table>

Materials

Activity. The RD activity is illustrated in Figure 7. The activity included a coordinate plane with five masses ($M_1$ to $M_5$) and two points ($P_A$ and $P_B$), a table with a list of values for mass and coordinates, and instructions to determine which point was influenced more by the masses. The distribution of masses and points did not include symmetrical or mirrored properties such that a calculation would be simple. The word “influence” was intended to activate student intuition, similar to the word “consistency” used in exploratory learning research about standard deviation (Kapur, 2014). The EF-RD activity also instructed students to invent a formula to determine the influence of the masses on the points, whereas the IF-RD activity instructed students to practice calculating the field at the different points. The activity instructions were otherwise similar across the EF-RD and IF-RD conditions and matched the instructions in the EF-CC and IF-CC activities used in Study 1 (Figure 8).
Figure 7: EF-RD Activity.

Gravitational Field Activity

The figure below on the left shows a series of particles in space, labeled \( m_1 \) through \( m_5 \), and points in the same space, labeled \( P_A \) and \( P_B \). Values for masses and locations for the masses and points are given in the table to the right. The masses are given in kilograms and the \( x \) and \( y \) components for the locations of both the masses and the points are given in meters.

Each row in the table above gives information about the masses or points in space in the diagram. Use the information given in the table to learn about gravitational field \( \vec{g} \). The gravitational field \( \vec{g} \) is a vector quantity that is related to the gravitational force and represents the influence that masses have on a point in space.

- Determine which point (\( P_A \) or \( P_B \)) is influenced more by the masses (\( M_1 \) to \( M_5 \)) in the figure above.
- Invent a formula or list the steps you might take to determine the influence of the masses (\( M_1 \) to \( M_5 \)) at points \( P_A \) and \( P_B \).

<table>
<thead>
<tr>
<th>Mass or Point</th>
<th>Mass (kg)</th>
<th>( x ) coord. (m)</th>
<th>( y ) coord. (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m_1 )</td>
<td>10</td>
<td>-2</td>
<td>0</td>
</tr>
<tr>
<td>( m_2 )</td>
<td>20</td>
<td>-3</td>
<td>5</td>
</tr>
<tr>
<td>( m_3 )</td>
<td>30</td>
<td>-4</td>
<td>-4</td>
</tr>
<tr>
<td>( m_4 )</td>
<td>40</td>
<td>2</td>
<td>-4</td>
</tr>
<tr>
<td>( m_5 )</td>
<td>50</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>( P_A )</td>
<td>–</td>
<td>-4</td>
<td>0</td>
</tr>
<tr>
<td>( P_B )</td>
<td>–</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 8: IF-RD Activity Instructions. Note: The rest of the activity was the same as in the EF-RD condition.

**Survey.** In addition to the interest and enjoyment scale and cognitive load question included in the Study 1 survey, a perceived knowledge gap awareness scale was added. The scale included four items (Cronbach’s \( \alpha = .87 \)) adapted from Flynn and
Goldsmith (1999), with responses on a 5-point Likert scale (1 = strongly disagree to 5 = strongly agree): I do not feel very knowledgeable about calculating gravitational field; I know pretty much about calculating gravitational field (Reverse coded); Compared to most other people, I know less about calculating gravitational field; When it comes to calculating gravitational field, I really don’t know a lot. This scale was used by Newman and DeCaro (2019) in an exploratory learning study and found differences between students who learned with invention, worked examples, or a completion problem. The knowledge gap awareness scale was interleaved with the interest and enjoyment scale. The final survey is shown in Appendix B.

**Assessments.** Modifications to the Study 1 assessments were as follows: a section break was added before the conceptual assessment (10 items, $\alpha = .33$); conceptual questions about electric field were added to the PFL assessment (4 items, $\alpha = .25$); and the PFL procedural scale (4 items; $\alpha = .53$) was modified by removing items with low variability to make time for the conceptual questions. The new conceptual PFL questions were True/False, and phrased similarly to the gravitational field conceptual questions. Study 2 assessments are shown in Appendix D.

**Procedure**

In each course, students were randomly assigned to the experimental conditions, EF-RD and IF-RD. On the first day of the study, students participated in the activity, survey, instruction, and activity review in two separate, adjacent classrooms based on condition. Students were instructed to sit in groups of 3 around folders that contained 1 activity sheet, 2 blank worksheets, 1 activity review, and 3 surveys. Students could self-select their groups within the randomized conditions.
Similar to the procedure in Study 1, the physics instructor led the instruction in both conditions, and the secondary instructor (graduate student) led the activity in both conditions. All course TAs helped with the activity in both courses (two TAs per classroom). Because the classrooms were adjacent to each other, a third instructor was not needed for the administration of the survey, as in Study 1. The physics instructor administered the survey in the EF-RD room, the secondary instructor administered the survey in the IF-RD room. In both cases, the survey was administered following the activity. On the second day of the study, students worked on assessment packets individually within the same, original classroom, and this phase was led by the primary instructor.

Results

Learning outcomes by condition

A 2 (type of knowledge: procedural, conceptual) × 2 (order: EF, IF) × 2 (course: algebra-based, calculus-based) mixed-factorial analysis of variance (ANOVA) was used to examine performance on the assessments, with order and course as between-subjects factors and type of knowledge as a within-subjects factor. The main effects were not significant: type of knowledge, $F < 1, p = .356$, order, $F(1, 88) = 2.72, p = .102$, or course, $F(1, 88) = 3.61, p = .061$. The three-way interaction was also not significant, $F < 1, p = .848$.

However, there was a significant order × course interaction, $F(1, 88) = 4.30, p = .041, \eta^2_p = .05$. Simple main effects are illustrated in Figure 9. In the calculus-based course, students in the EF condition ($M = 74.72\%, SE = 3.84, 95\% CI [66.95, 82.50]$) performed significantly better than students in the IF condition ($M = 60.98\%, SE = 4.37,$
95% CI [52.14, 69.83], \( d = .75 \). In the algebra-based course, the difference was not significant between students in the EF condition (\( M = 60.06\% \), \( SE = 3.08 \), 95% CI [53.88, 66.23]) and students in the IF condition (\( M = 61.62\% \), \( SE = 3.51 \), 95% CI [54.57, 68.67]). Note that the lack of an interaction with type of knowledge means that this effect occurred for the posttest as a whole, with both procedural and conceptual knowledge subscales combined.

![Figure 9: Study 2 posttest performance by order and course. Error bars = ± SE.](image)

A 3 (type of knowledge: PFL procedural, PFL conceptual, PFL transfer) × 2 (order) × 2 (course) mixed-factorial ANOVA was used to examine performance on the PFL assessment, with order and course as between-subjects factors and PFL type of knowledge as a within-subjects factor. Mauchly’s test indicated that the assumption of sphericity had been violated, \( \chi^2(2) = 47.69, p < .001 \). The degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (\( \varepsilon = 0.70 \)). There was a significant main effect of PFL type of knowledge, \( F(1.41, 123.77) = 52.30, p < .001, \eta^2_p = .37 \). A comparison of confidence intervals revealed significantly higher conceptual knowledge (\( M = 80.66\% \), \( SE = 1.77 \), 95% CI [77.15, 84.17]) than procedural knowledge.
(M = 43.13%, SE = 2.38, 95% CI [38.40, 47.86], d = 1.91) or transfer (M = 34.61%, SE = 5.16, 95% CI [24.35, 44.87], d = 1.28). The other main effects were not significant: order, F < 1, p = .418, course, F < 1, p = .755. The interactions were also not significant, Fs < 1. Means (collapsed across course) are illustrated in Figure 10. The results were not sensitive to the removal of course as a variable.

Figure 10: Study 2 PFL assessment performance by subscale and order. Error bars = ± SE.

Process level measures

Number of strategies was coded for the EF-RD condition and added as a predictor in regression models to predict learning outcomes. Interest and enjoyment, knowledge gap awareness, and cognitive load were first compared between instructional order conditions using ANOVAs, and then used to predict each learning outcome in separate regression models. Course was included as a variable in exploratory analyses, however was not a significant predictor of any of the process level measures. Therefore, course was not included in the analyses reported here.
Number of strategies and Success on the RD Activity. Coding of the RD activity was exploratory in this study. Overall, the groups in the IF-RD condition were not able to solve for the gravitational field at either point A or B. Only one out of 19 groups completed the calculations for a single point. Also, only one group made any written reference to the comparison question (whether $P_A$ or $P_B$ was more influenced), and in this case, one student merely underlined the comparison in the instructions. All groups followed the procedure that was demonstrated in the instruction, so there was no variability in the strategies used. These parameters were not analyzed further for the IF order.

The EF-RD groups’ work had different characteristics; groups tried different sets of strategies that they had learned for gravitational force and other vector problems. The following strategies were observed:

- Free body diagrams
- Distance calculations
- Tables
- Vector addition (a verbal reference)
- Gravitational force equation: \( F = GM_1M_2/r^2 \)
- Setting an arbitrary value for \( M_2 \) (e.g. “assume \( M_2 = 1 \)”)
- Ignoring \( M_2 \)
- Center of Mass
- Invented equations, e.g.:
  - \( g = GM/r^2 \) (correct)
  - \( G = Fg*d \)
The number of strategies was tallied for each group. Number of strategies ranged from 1 to 5 ($M = 2.5, SE = .15$), and the distribution is shown in Figure 11.

Figure 11: Study 2 distribution of the number of RD activity strategies.

Regression was used to determine whether number of strategies significantly predicted performance on the assessment for any of the different types of knowledge (procedural, conceptual, PFL procedural, PFL conceptual, and PFL transfer). The analysis was restricted to the EF-RD condition, because the number of strategies did not vary in the IF-RD condition. Number of strategies did not predict performance on any sections of the assessment ($p > .216$; Table 7).
Table 7: Study 2 Regression Results, RD Number of Strategies and Learning Outcomes.

<table>
<thead>
<tr>
<th>Knowledge Type</th>
<th>$R^2$</th>
<th>$F$</th>
<th>$\beta$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedural</td>
<td>&lt; .01</td>
<td>.20</td>
<td>- .06</td>
<td>.655</td>
</tr>
<tr>
<td>Conceptual</td>
<td>&lt; .01</td>
<td>.01</td>
<td>- .01</td>
<td>.940</td>
</tr>
<tr>
<td>PFL Procedural</td>
<td>&lt; .01</td>
<td>.20</td>
<td>- .06</td>
<td>.659</td>
</tr>
<tr>
<td>PFL Conceptual</td>
<td>.03</td>
<td>1.57</td>
<td>- .17</td>
<td>.216</td>
</tr>
<tr>
<td>PFL Transfer</td>
<td>&lt; .01</td>
<td>.16</td>
<td>- .06</td>
<td>.687</td>
</tr>
</tbody>
</table>

Similar to the IF-RD groups, the EF-RD groups rarely completed a series of calculations. However, unlike the IF-RD groups, 8 groups out of 20 addressed the second bullet point in the activity instructions and performed a comparison between $P_A$ and $P_B$; one group correctly identified that $P_A$ was influenced more, 4 groups incorrectly identified $P_B$, and 3 groups wrote down the word “compare” in a sentence about their incomplete work.

**Interest and enjoyment.** Interest and enjoyment were not significantly different between students in the IF-RD ($M = 3.70$, $SE = .11$, 95% CI [3.49, 3.91]) and EF-RD ($M = 3.89$, $SE = .09$, 95% CI [3.70, 4.07]) conditions, $F(1, 90) = 1.75$, $p = .189$. No outcome measures were significantly predicted by interest and enjoyment ($p > .182$; Table 8).

Table 8: Study 2 Regression Results, Interest and Enjoyment and Learning Outcomes.

<table>
<thead>
<tr>
<th>Knowledge Type</th>
<th>$R^2$</th>
<th>$F$</th>
<th>$\beta$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedural</td>
<td>.01</td>
<td>.68</td>
<td>.09</td>
<td>.412</td>
</tr>
<tr>
<td>Conceptual</td>
<td>.02</td>
<td>1.34</td>
<td>.12</td>
<td>.250</td>
</tr>
<tr>
<td>PFL Procedural</td>
<td>.02</td>
<td>1.81</td>
<td>.14</td>
<td>.182</td>
</tr>
<tr>
<td>PFL Conceptual</td>
<td>.01</td>
<td>.91</td>
<td>.10</td>
<td>.343</td>
</tr>
<tr>
<td>PFL Transfer</td>
<td>.01</td>
<td>.55</td>
<td>.08</td>
<td>.460</td>
</tr>
</tbody>
</table>

**Knowledge gap awareness.** Knowledge gap awareness was significantly higher for students in the EF-RD condition ($M = 3.30$, $SE = .112$, 95% CI [3.08, 3.52]) than
students in the IF-RD condition ($M = 2.64, SE = .13, 95\% CI [2.38, 2.89]), $F(1, 89) = 15.16, p < .001, \eta^2_p = .15$. Knowledge gap awareness significantly predicted procedural, conceptual, and PFL conceptual knowledge (Table 8). The beta values indicated an inverse relationship; students with lower perceived knowledge gaps had higher performance.

Table 9: Study 2 Regression Results, Knowledge Gap Awareness and Learning Outcomes.

<table>
<thead>
<tr>
<th>Knowledge Type</th>
<th>$R^2$</th>
<th>$F$</th>
<th>$\beta$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedural</td>
<td>.14</td>
<td>14.08</td>
<td>-.37</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Conceptual</td>
<td>.05</td>
<td>4.48</td>
<td>-.22</td>
<td>.037</td>
</tr>
<tr>
<td>PFL Procedural</td>
<td>&lt; .01</td>
<td>.13</td>
<td>-.04</td>
<td>.718</td>
</tr>
<tr>
<td>PFL Conceptual</td>
<td>.10</td>
<td>9.35</td>
<td>-.31</td>
<td>.003</td>
</tr>
<tr>
<td>PFL Transfer</td>
<td>.02</td>
<td>2.11</td>
<td>-.15</td>
<td>.150</td>
</tr>
</tbody>
</table>

**Cognitive load.** Cognitive load was not significantly different between students in the IF-RD ($M = 6.35, SE = .19, 95\% CI [5.98, 6.72]) and EF-RD ($M = 6.39, SE = .16, 95\% CI [6.06, 6.71]) conditions, $F(1, 90) = .02, p = .889$. Cognitive load did not directly predict any of the outcome variables ($p > .145$; Table 10).

Table 10: Study 2 Regression Results, Cognitive Load and Learning Outcomes.

<table>
<thead>
<tr>
<th>Knowledge Type</th>
<th>$R^2$</th>
<th>$F$</th>
<th>$\beta$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedural</td>
<td>&lt; .01</td>
<td>.22</td>
<td>.05</td>
<td>.642</td>
</tr>
<tr>
<td>Conceptual</td>
<td>&lt; .01</td>
<td>.17</td>
<td>.04</td>
<td>.677</td>
</tr>
<tr>
<td>PFL Procedural</td>
<td>.01</td>
<td>.99</td>
<td>.10</td>
<td>.324</td>
</tr>
<tr>
<td>PFL Conceptual</td>
<td>&lt; .01</td>
<td>.145</td>
<td>.04</td>
<td>.704</td>
</tr>
<tr>
<td>PFL Transfer</td>
<td>&lt; .01</td>
<td>.02</td>
<td>-.02</td>
<td>.879</td>
</tr>
</tbody>
</table>

**Discussion**

Study 2 investigated an RD activity in two instructional orders, in two introductory physics courses. Analysis of procedural and conceptual knowledge revealed significantly higher overall performance in the EF-RD condition than the IF-RD
condition, but only in the calculus-based course. There were no differences between conditions on the PFL assessment. In conjunction with the observed PFL benefits in Study 1, the null PFL result could indicate that activity type is a moderator of exploratory learning; namely, that CC activities result in PFL learning improvements, but RD activities do not. If this is the case, discernment of problem features could be a specific mechanism that prepares students for future learning. However, performance on learning outcomes in Studies 1 and 2 cannot be directly compared to each other because of the methodological differences between them. Study 1 occurred in a single 75-minute session, whereas Study 2 occurred over two consecutive 50-minute class periods. In addition, Study 2 participants came from two courses, and the two samples of students did not have the same learning outcomes.

In this study, number of activity strategies, interest and enjoyment, and cognitive load were neither different between conditions nor predictive of learning outcomes. However, perceived knowledge gap awareness was negatively correlated with performance. Knowledge gap awareness is one of the most frequently proposed mechanisms in the productive failure literature, with most researchers hypothesizing that failure on the activity leads to knowledge gap awareness, which then leads to knowledge updating (i.e., learning; Darabi et al., 2018). This argument continues despite research finding both positive and negative relationships between perceived knowledge gaps and performance. Glogger-Frey and colleagues (2015) reported two different results in the same article: in one study, graduate students’ knowledge gap awareness positively predicted their learning, whereas in a second study with eighth graders, greater knowledge gap awareness corresponded with lower performance. The negative
relationship observed in the current study could indicate that the undergraduate students were more discouraged by their knowledge gap awareness rather than made more curious. Or, alternatively, it could be argued that the activity was too hard; students in both instructional orders had trouble completing the activity. Either way, greater knowledge gap awareness did not result in better learning. Although students in the EF-RD condition reported more perceived knowledge gaps than students in the IF-RD condition, and there was an observed learning benefit of the EF-RD group in the calculus-based course, knowledge gap awareness did not cause the learning benefit.

**Course effects**

The two courses in this study were very similar; the instructor was the same, and the concepts and procedures were the same throughout the semester. However, the courses did differ in the mathematics used to describe the concepts as well as the student populations. The students in the algebra-based physics course were primarily “pre-med” majors, whereas the students in the calculus-based course were primarily engineering students. Differences between the samples may have altered the effectiveness of the EF condition. For example, the engineering students could have had more prior knowledge that enabled them to benefit more from the EF condition. The ENGR students likely had more experience in and focus on math, which could have helped them with this activity. It is also possible that the engineering students had different motivations to learn the material than pre-med majors. One study found that intrinsic and extrinsic motivation profiles for learning physics differed between students in different majors (Maurer, Allen, Gatch, & Shankar, 2013), although this study did not look specifically at engineering and pre-med students. However, not much can be concluded from course differences because
the sample size is small when broken down by course. Sample populations may be a potential moderator and should be researched in the future, but the current study has limited ability to comment on differences in samples.

Limitations

A significant limitation of this study was the low reliability of the conceptual assessment (10 items, $\alpha = .33$). In Study 1, the reliability was much higher ($\alpha = .76$). Studies 1 and 2 had the same questions on this assessment, but there were some differences between studies that could have affected the scale’s reliability. For one, participants in Study 2 included students from two different samples, one that had been studied previously (engineering students in the calculus-based course), and the other that had not (pre-med students in the algebra-based course). The reliability of the scale was analyzed for each course separately, and in fact, reliability was higher in the calculus-based course ($\alpha = .55$) than the algebra-based course ($\alpha = .28$). Also, the assessment methodology was different in Study 2: a section break separated the conceptual assessment from the procedural assessment. Because of this break, and the fact that time was allotted for the conceptual assessment, more students answered all of the questions. In Study 1, ten participants left all conceptual questions blank, fourteen more left 1-9 conceptual questions blank, and these blank responses were scored as zeros. It is possible that the number of zeros falsely increased the scale reliability in Study 1.

It is possible that the conceptual scale reliability score is low because the questions query different sub-concepts within the whole concept of gravitational field. Individual questions targeted the magnitude, direction, and vector nature of gravitational field. These sub-concepts are distinct in that students can understand one while not
understanding another. Therefore, if different students correctly understood different sub-concepts, and answered different sets of questions correctly, the overall scale might appear to be unreliable. It is also possible that the T/F question format, with a 50% chance of students getting a question correct, is not reliable. The reliability issues appeared on the conceptual questions on the PFL assessment as well (4 items, $\alpha = .25$), and these were also True/False.

**Conclusions**

Study 2 provided a controlled experiment investigating exploratory learning using an RD activity in the concept of gravitational field. This study was limited due to a mixed sample of two different populations, and low reliability of conceptual scales. Study 2 results cannot be compared to Study 1 results directly because of methodological differences (e.g., in Study 2, the posttest was given on a following class day). Preliminary evidence from this study, however, indicates that the EF instructional order benefitted procedural and conceptual learning in the calculus-based course. Mean PFL performance did not differ between students in EF and IF instructional orders.

Study 3 was designed to replicate the findings of Studies 1 and 2 and directly compare the two activity types within the same, calculus-based course. The same conceptual assessment was used, with the assumption that a greater number of participants from the same calculus-based course sample would increase the reliability of the scale.
CHAPTER IV: STUDY 3

Study 3 compared the impact of two exploratory learning activities in two instructional orders on learning outcomes and process level measures in a 2 (order: EF, IF) × 2 (activity: CC, RD) between-subjects design. This study is the first to compare two activity types within an exploratory learning context in a single experiment.

Research has indicated that an exploratory learning (EF) instructional order benefits conceptual knowledge (see Loibl et al., 2016), although many studies, especially those using an RD activity, have lacked experimental control. In Study 1, students in the EF-CC condition performed better on the conceptual assessment than students in the IF-CC condition. In Study 2, students in the EF-RD condition performed better than students in the IF-RD condition on the procedural and conceptual assessment combined (in the calculus-based course), however, the assessment was given after a delay of two days (one class period). The first of two hypotheses was therefore that students in the EF order would outperform those in the IF order on conceptual knowledge but not procedural knowledge, resulting in a significant \( \text{knowledge type} \times \text{order} \) interaction. The second hypothesis was that a more complex, three-way, \( \text{knowledge type} \times \text{order} \times \text{activity} \) interaction would be observed, in which students in the EF-CC condition would benefit more than the other three conditions on conceptual knowledge.
In Study 1, students in the EF-CC condition performed better on the PFL transfer question than students in the IF-CC condition, similar to results reported in the prior PFL literature (e.g., Schwartz & Martin, 2004). However, using RD activities in Study 2, order was not a significant factor in PFL performance. Therefore, a significant \textit{knowledge type} \times \textit{order} \times \textit{activity} interaction was expected in Study 3, with students in the EF-CC condition performing significantly higher than the other three conditions on the conceptual and transfer portions of the PFL assessment. This result would indicate that an EF-CC condition is unique in some way, perhaps in the focus on problem features. This finding could indicate that discernment of problem features is a key mechanism for preparation for future learning.

Activity success for the CC activity was shown to positively correlate with performance on the conceptual assessment and PFL assessment in Study 1, and was therefore expected to correlate with performance in this study in the CC conditions. Number of strategies (RD) did not correlate with performance on Study 2, however, the study was run over two class periods. Therefore, a positive correlation was possible but not expected.

Students in the EF instructional order were expected to report more perceived knowledge gaps than students in the IF conditions. In addition, because the CC activity was designed for students to discover problem features and experience success during the activity, and the RD activity was designed for productive failure, it was expected that students in the EF-RD condition would report higher perceived knowledge gap awareness than students in the EF-CC condition, resulting in a significant \textit{order} \times \textit{activity} interaction.
Cognitive load was not significantly different between conditions in either Studies 1 or 2. However, it is possible that the activities lead to different levels of cognitive load in both instructional orders. Hypotheses were therefore twofold: cognitive load was expected to either be the same across all conditions or higher in the RD conditions than the CC conditions. Lastly, responses to the interest and enjoyment scale were expected to be similar for students in all conditions, based on the null results from Studies 1 and 2.

Method

Participants

Participants (N = 256) included all physics students from Introductory Mechanics, Heat and Sound (calculus-based) from one semester. The primary instructor from Studies 1 and 2 was the instructor of record for four of the eight sections. Another physics professor was the instructor of record for the other four sections. However, the same primary instructor and secondary instructor (graduate student) from the previous studies led Study 3 in all sections. Similar to the inclusion criteria in Studies 1 and 2, participant data were excluded if the students were absent (N = 21), missed part of the class period (N = 5), had a blank PFL section or blank full page of the assessment indicating lack of participation (N = 4), or were observed to not participate in the activity (N = 1). In addition, in this study, students were excluded if English was their second language (N = 18); this item was added to the survey to remove the possibility of additional reading difficulties, particularly on the PFL assessment. Students were also excluded if they had participated in Studies 1 or 2 in a previous semester (N = 1). Additionally, outliers were removed (N = 5). The number of participants by order and activity is shown in Table 11.
Table 11: Study 3 Number of Participants by Order and Activity.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Order</th>
<th>EF</th>
<th>IF</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td></td>
<td>68</td>
<td>64</td>
<td>132</td>
</tr>
<tr>
<td>RD</td>
<td></td>
<td>64</td>
<td>60</td>
<td>124</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>132</td>
<td>124</td>
<td>256</td>
</tr>
</tbody>
</table>

Materials

Materials for this study included the four activities and instruction described in Studies 1 and 2 (EF-CC, IF-CC, EF-RD, and IF-RD), the updated assessment packet used in Study 2 (including procedural and conceptual knowledge, and PFL procedural knowledge, conceptual knowledge, and transfer), two new activity review packets (CC and RD, described below), and the same survey used in Study 2 (perceived knowledge gaps, interest and enjoyment, cognitive load) with an additional question asking students whether English was their second language (ESL). The ESL question was added to remove any potential effect of reading ability from analyses, especially on the PFL assessment.

Activity review packets. Activity review packets were necessary for this study because two different activities were given to different student groups in the same room,

1 Students in the EF-CC condition in Study 3 received activity packets with one incorrect value. This mistake did not likely impact the results, because students in the EF-CC condition were exploring as opposed to calculating, and this particular vector corresponded to the most difficult scenario. It is therefore not likely that any students looked at this value directly in their work on the activity. Students in the EF-CC and IF-CC received the activity review packets with the same incorrect value.
and therefore the instructor could not lead students through an activity review verbally. Instead, review packets were distributed in envelopes and students were instructed to open them at the appropriate time. Review packets consisted of hand-written solutions to the activities (Appendices E and F).

**Procedure**

Study 3 was run on two separate days, one day for each instructor of record’s sections. The dates fit into the instructors’ curricula when each would typically teach gravitational field. The primary and secondary instructor were responsible for leading the classes on these dates. In addition, the primary instructor also led an in-class lecture in the secondary instructor’s sections in the class period prior to the study, on the topic of gravitational force. This guest lecture was intended to ensure that all students received the same content immediately prior to the study, as well as increase familiarity with the primary instructor prior to the study. Both instructors gave homework credit to students who attended class on the date of the study as well as the day prior.

The primary instructor’s sections regularly met in an active learning classroom, and a comparable, adjacent classroom was also reserved for the study. The other sections regularly met in a traditional lecture hall. To control for environment as much as possible, two adjacent active learning classrooms were reserved for the sections with the different instructor of record. The remaining procedures were the same for both courses.

Students were randomly assigned to one of four conditions labeled with nondescript letters A (EF-CC), B (EF-RD), X (IF-CC), and Y (IF-RD). The students were instructed to go to a specific classroom and sit on the side of the room that matched their assigned letter, in groups of 3 around the folders at the tables as described in Study 2.
TAs at the entrance to each room helped guide students to the correct location. Groups were self-selected within each condition. In the IF room, the primary instructor gave the instruction, and then the secondary instructor came in to lead the activity, survey, and assessments. In the EF room, the secondary instructor led the activity, and the primary instructor came in to give the survey, instruction, and assessments. One TA also helped in each classroom, and the two TAs were assigned to different conditions for the two days of the study to counterbalance any effect of TA.

Results

Learning outcomes by condition

A 2 (type of knowledge: procedural, conceptual) × 2 (order: EF, IF) × 2 (activity: CC, RD) mixed-factorial analysis of variance (ANOVA) was used to examine performance on the assessments. Order and activity were between-subjects factors, and type of knowledge was within-subjects. The main effect of activity was significant, $F(1, 252) = 6.29, p = .013, \eta^2_p = .02$, with student performance in the CC conditions ($M = 72.78\%, SE = 1.30$) significantly higher than student performance in the RD conditions ($M = 68.09\%, SE = 1.34$). The main effects of order, $F(1, 252) = 1.74, p = .139$, and type of knowledge, $F(1, 252) = 2.56, p = .111$, were not significant. The three-way interaction was not significant, $F < 1, p = .339$. Estimated marginal means are illustrated in Figure 12.
These results were qualified by a significant \textit{type of knowledge} × \textit{activity} interaction, $F(1, 252) = 9.33, p = .002, \eta_p^2 = .04$. A comparison of means (Figure 13) and confidence intervals revealed that students’ procedural knowledge was significantly higher in the CC condition ($M = 75.61\%, SE = 1.73, 95\% \text{ CI}[72.20, 79.01]$) than in the RD condition ($M = 67.21\%, SE = 1.79, 95\% \text{ CI}[63.69, 70.72], d = .32$). Conceptual knowledge was not significantly different as a function of activity type ($M_{CC} = 69.95\%, SE = 1.35, 95\% \text{ CI}[67.29, 72.61]; M_{RD} = 68.97\%, SE_{RD} = 1.40, 95\% \text{ CI}[66.23, 71.72]$).
A 3 (type of knowledge: PFL procedural, PFL conceptual, PFL transfer) × 2 (order) × 2 (activity) mixed-factorial ANOVA was used to examine performance on the PFL assessment, with order and activity as between-subjects factors and PFL type of knowledge as a within-subjects factor. Mauchly’s test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 43.94, p < .001$. The degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = 0.88$). There was a significant main effect of PFL type of knowledge, $F(1.72, 434.26) = 319.52, p < .001$, $\eta^2_p = .56$. Students had significantly higher PFL conceptual knowledge ($M = 82.60\%, SE = 1.06, 95\% CI [80.52, 84.68]$) than procedural knowledge ($M = 66.61\%, SE = 1.88, 95\% CI [62.90, 70.32], d = .66$) and transfer ($M = 21.99\%, SE = 2.60, 95\% CI [16.88, 27.10], d = 1.68$). The main effects of order and activity were not significant, $Fs < 1, ps = .518$ and .347, respectively. The two-way interactions were not statistically significant, $Fs < 1, ps > .348$, in addition to the three-way interaction, $F(1.72, 434.26) = 1.72, p = .184$. 

Figure 13: Study 3 posttest performance by activity, collapsed across orders. Error bars = ±SE.
Process level measures

Students’ work on the activity was coded and analyzed by condition; both EF-CC and IF-CC activities were scored for success, and the EF-RD activities were coded for number of strategies. Interest and enjoyment, knowledge gap awareness, and cognitive load were compared across order and activity. All process level measures were then used in regression to predict learning outcomes.

Success on the CC Activity. Activity success was not significantly affected by order (EF: M = 73.24%, SE = 3.60, 95% CI [66.11, 80.34]; IF: M = 80.00%, SE = 3.74, 95% CI [72.60, 87.4]), F(1, 129) = 1.70, p = .195. Activity success significantly predicted procedural, conceptual, and PFL procedural knowledge; students with higher success had higher performance. Regression results are listed in Table 12.

Table 12: Study 3 CC Activity Success and Learning Outcomes.

<table>
<thead>
<tr>
<th>Knowledge Type</th>
<th>R²</th>
<th>F</th>
<th>β</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedural</td>
<td>.10</td>
<td>14.54</td>
<td>.32</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Conceptual</td>
<td>.05</td>
<td>6.23</td>
<td>.22</td>
<td>.014</td>
</tr>
<tr>
<td>PFL Procedural</td>
<td>.05</td>
<td>7.10</td>
<td>.23</td>
<td>.009</td>
</tr>
<tr>
<td>PFL Conceptual</td>
<td>&lt; .01</td>
<td>.03</td>
<td>-.01</td>
<td>.875</td>
</tr>
<tr>
<td>PFL Transfer</td>
<td>.02</td>
<td>2.20</td>
<td>.13</td>
<td>.140</td>
</tr>
</tbody>
</table>

Number of Strategies in the RD Activity. Number of strategies predicted performance on the PFL procedural assessment. Regression results are listed in Table 13.
Table 13: Study 3 RD Number of Strategies and Learning Outcomes.

<table>
<thead>
<tr>
<th>Knowledge Type</th>
<th>$R^2$</th>
<th>$F$</th>
<th>$\beta$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedural</td>
<td>.04</td>
<td>2.76</td>
<td>.21</td>
<td>.102</td>
</tr>
<tr>
<td>Conceptual</td>
<td>.01</td>
<td>.64</td>
<td>.10</td>
<td>.425</td>
</tr>
<tr>
<td>PFL Procedural</td>
<td>.17</td>
<td>13.01</td>
<td>.42</td>
<td>.001</td>
</tr>
<tr>
<td>PFL Conceptual</td>
<td>.03</td>
<td>2.18</td>
<td>.18</td>
<td>.145</td>
</tr>
<tr>
<td>PFL Transfer</td>
<td>.03</td>
<td>1.96</td>
<td>.18</td>
<td>.167</td>
</tr>
</tbody>
</table>

**Interest and enjoyment.** Order, $F(1, 250) = 1.28, p = .258$, and activity, $F(1, 250) = 1.25, p = .264$, were not significant predictors of responses to the interest and enjoyment scale, and the interest and enjoyment scale did not predict any of the learning outcomes ($p > .293$; Table 13).

Table 14: Study 3 Interest and Enjoyment and Learning Outcomes.

<table>
<thead>
<tr>
<th>Knowledge Type</th>
<th>$R^2$</th>
<th>$F$</th>
<th>$\beta$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedural</td>
<td>&lt; .01</td>
<td>1.50</td>
<td>.08</td>
<td>.222</td>
</tr>
<tr>
<td>Conceptual</td>
<td>&lt; .01</td>
<td>.36</td>
<td>.04</td>
<td>.551</td>
</tr>
<tr>
<td>PFL Procedural</td>
<td>&lt; .01</td>
<td>1.11</td>
<td>.07</td>
<td>.293</td>
</tr>
<tr>
<td>PFL Conceptual</td>
<td>&lt; .01</td>
<td>&lt; .01</td>
<td>-.01</td>
<td>.967</td>
</tr>
<tr>
<td>PFL Transfer</td>
<td>&lt; .01</td>
<td>.47</td>
<td>-.04</td>
<td>.493</td>
</tr>
</tbody>
</table>

**Knowledge gap awareness.** There was a significant effect of order (EF: $M = 3.24, SE = .08, 95\% CI [3.08, 3.40]$; IF: $M = 2.36, SE = .09, 95\% CI [2.19, 2.53]$) on perceived knowledge gaps, $F(1, 247) = 54.69, p < .001, \eta_p^2 = .18$. The effect of activity (CC: $M = 2.81, SE = .08, 95\% CI [2.64, 2.98]$; RD: $M = 2.79, SE = .09, 95\% CI [2.62, 2.96]$) was not significant, $F < 1, p = .888$, and the interaction between order and activity was not significant, $F < 1, p = .353$. Knowledge gap awareness significantly predicted all knowledge types (Table 15). The negative beta values indicated an inverse relationship; students with lower perceived knowledge gaps had higher performance.
Cognitive load. Cognitive load was not significantly predicted by order (EF: \(M = 5.95, SE = .13, 95\% CI [5.69, 6.21]\); IF: \(M = 6.07, SE = .14, 95\% CI [5.80, 6.33]\)) or activity (CC: \(M = 5.96, SE = .13, 95\% CI [5.70, 6.22]\); RD: \(M = 6.06, SE = .13, 95\% CI [5.79, 6.32]\)), and there was no interaction between the two factors, \(Fs < 1, ps > .530\).

Cognitive load inversely predicted performance on the conceptual knowledge assessment as well as the PFL transfer assessment: the higher the cognitive load, the lower the performance. Regression results are listed in Table 16.

Table 16: Study 3 Cognitive Load and Learning Outcomes.

<table>
<thead>
<tr>
<th>Knowledge Type</th>
<th>(R^2)</th>
<th>(F)</th>
<th>(\beta)</th>
<th>(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedural</td>
<td>.01</td>
<td>1.89</td>
<td>-.09</td>
<td>.170</td>
</tr>
<tr>
<td>Conceptual</td>
<td>.04</td>
<td>10.44</td>
<td>-.20</td>
<td>.001</td>
</tr>
<tr>
<td>PFL Procedural</td>
<td>.01</td>
<td>2.90</td>
<td>-.11</td>
<td>.090</td>
</tr>
<tr>
<td>PFL Conceptual</td>
<td>.04</td>
<td>9.14</td>
<td>-.19</td>
<td>.003</td>
</tr>
<tr>
<td>PFL Transfer</td>
<td>.03</td>
<td>7.77</td>
<td>-.17</td>
<td>.006</td>
</tr>
</tbody>
</table>

Discussion

Study 3 was the first study to compare different activity types in an exploratory learning context. There was a significant main effect of activity and a significant type of knowledge \(\times\) activity interaction, with students in the EF-CC and IF-CC conditions performing significantly better on the procedural assessment than those in the EF-RD and IF-RD conditions. Neither activity nor order significantly affected PFL outcomes. The interest and enjoyment scale response and number of strategies did not vary by condition.
or predict learning outcomes. Activity success was significantly affected by order (IF-CC > EF-CC) and positively predicted procedural, conceptual, and PFL procedural learning outcomes. Cognitive load negatively predicted conceptual and PFL transfer learning outcomes. Knowledge gap awareness was significantly affected by order (EF > IF) and negatively predicted all learning outcomes.

One possible explanation for the main effect of activity is that the discernment of problem features is a stronger learning mechanism than failure and error correction. In RD activities, the canonical solution is hidden such that students generate many incorrect strategies (used later for reconciliation). Students therefore may or may not uncover problem features during work on an RD activity, depending on the number and type of strategies that they attempt. Contrastingly, the examples used in CC activities are created to vary along the important problem features. Enough information is presented to the students to allow discovery of the problem features, and possibly the canonical solution. The clarity of problem features in the CC activity could cause learners in both EF and IF instructional orders to discern them. In addition, asking students to invent a broad solution encourages them to process the dimensions simultaneously to account for all cases (Schwartz et al., 2009).

The benefit of learning from problem features in a tell-then-practice order is supported by the literature on comparative instruction—instruction that is based on a presentation of multiple cases side by side. Comparative instruction has been shown to benefit learning above instruction using a single example or a series of examples presented one at a time (e.g., Gentner, Loewenstein, & Thompson, 2003; Rittle-Johnson & Star, 2007). Rittle-Johnson and Star (2009) specify that the best sets of cases to be used
in comparative instruction are ones that differ on relevant problem features, just like the ones embedded into the CC activity. The current study differs from comparative instruction studies because the contrasting cases are given as an unguided exploration activity and as an activity following instruction. The activity instructions that get students to synthesize across cases are key; Chin et al. (2016) found that instructions to merely “compare and contrast” did not increase learning as much as instructions to “invent a general solution that satisfies all cases.” In comparative instruction, teachers are responsible for getting students to understand problem features by guiding them through important similarities and differences between cases. The fact that IF-CC conditions led to better learning than IF-RD conditions speak to the idea that the CC activity is better for getting students to perceive problem features, even in an instruct-first order.

The process level results also indicate that discernment of problem features could be a more important mechanism of exploratory learning than failure and error correction. In this study, success on the CC activity was positively related to performance on the assessment. Because problem features were embedded into the CC activity one by one, it is likely that the students who had more success on the activity correctly identified more of the problem features. Perceived knowledge gaps were negatively correlated with performance. In this study, higher knowledge gap awareness resulted in lower performance on every learning outcome. If failure led to knowledge gap awareness, then knowledge gap awareness was not necessarily followed by error correction and deeper learning.

It is important to remember that the observed benefit in the CC conditions over RD conditions was in procedural knowledge. Discernment of problem features has been
thought to aid in conceptual knowledge more than procedural knowledge. It may be the case that discernment of problem features helped students to break the problem down, aiding in procedural knowledge development, but it is also possible that there is another difference between the activities is more likely to explain these results.

A second explanation is that the RD activity was too difficult. Consistent with Study 2, students in EF-RD and IF-RD conditions in Study 3 rarely completed the activity. This “impossibility” of completion could have made the intended failure during the activity more discouraging than useful in an error correction paradigm, and difficulty may be an activity moderator of exploratory learning benefits. However, there were no significant differences between activity conditions on interest and enjoyment, knowledge gap awareness, or cognitive load. Activity difficulty then may not affect performance by discouragement or demotivation, but rather by blocking students from practicing a complete problem solution. This argument is reasonable for students in the IF order; following instruction, students who completely solved through a gravitational field problem would do better on the procedural assessment than students who only completed part of the solution process. Students in the EF conditions on the other hand did not necessarily calculate all the way through with either activity. Therefore, there may be different reasons for the superiority of an EF-CC condition over an EF-RD condition and an IF-CC condition over an IF-RD condition.

Ultimately, the current results have limited value in investigating mechanisms of exploratory learning, because an effect of instructional order was not observed. Because an order effect was expected based on prior literature as well as Study 1 and 2 results, all available data was thoroughly reviewed for methodological errors. The audio recordings
and activity worksheets did reveal one possible methodological issue: the teaching assistants (TAs) in Study 3 instructed students to continue working even when they appeared to be finished with the activity. This instruction particularly affected students in the IF-CC condition, who were able to quickly complete the matching exercise in the CC activity. The TAs instructed these students to explain their reasoning on the activity worksheet with sentences. This instruction caused student work in the IF-CC condition to be significantly different from student work in Study 1. In Study 3, most of the groupwork in the IF-CC condition included complete, well-constructed sentences about the problem features of gravitational field. In comparison, students in the IF-CC condition in Study 1 rarely wrote any notations outside of the matching table. Typical student work in the Study 3 IF-CC condition is illustrated in Figure 14.

Having students write sentences to explain their reasoning is a teaching method called self-explanation. Many studies demonstrate learning benefits from self-explanation (e.g., Chi, De Leeuw, Chiu, & Lavancher, 1994). In fact, Sidney, Hattikudur, and Alibali (2015) investigated the benefits of self-explanation with contrasting cases in a study of undergraduate psychology students learning fraction division. Using a 2 (self-explanation, no self-explanation) × 2 (contrasting cases, sequential cases) between-subjects design, they found that self-explanation prompts significantly affected learning whereas the contrasting cases manipulation did not. Sidney and colleagues concluded that the learning gains observed in other contrasting case studies were due to the self-explanation that often naturally occurs with contrasting case materials. In this case, the
IF-CC condition in Study 3 could have benefitted as much from the self-explanation prompt as the EF-CC condition benefitted from exploration, thereby eliminating the possible effect of order.

These findings connect back into the mechanism of problem feature discernment. Students in the EF-CC condition, who got the CC activity prior to instruction, had to identify the problem features to get the correct answers on the activity. Students in the IF-CC condition were encouraged to explain their solutions with words, and were encouraged to articulate these problem features. The students in the EF-RD and IF-RD conditions did not have the opportunity to discover or articulate problem features.
Limitations

A significant limitation of Study 3 was the low reliability of the assessments: procedural knowledge, 9 items, $\alpha = .65$; conceptual knowledge, 10 items, $\alpha = .36$; PFL procedural knowledge, 4 items, $\alpha = .66$; and PFL conceptual knowledge, 4 items, $\alpha = .22$. The low Cronbach’s alpha scores indicate that the scale items were not closely related to each other. As discussed in Study 2, each posttest question targeted a sub-concept of gravitational field (either the magnitude, direction, or vector nature), so it is possible that despite the low reliability, the scales are acceptable in measuring the different types of knowledge of gravitational field overall. However, it is also possible that the items were of low quality, which would impede the ability to draw many conclusions from this data.

Conclusions

In the first comparative study of exploratory learning activity types, the CC activity appeared to be better for learning than the RD activity. However, there was no effect of instructional order, and the observed benefit was in procedural knowledge, so these results have limited implications for exploratory learning research. However, this study gives an initial impression that differences in exploration activities could moderate learning, whether the differences are in activity difficulty, activity materials, or instruction and guidance during the activity.
CHAPTER V: GENERAL DISCUSSION

Studies 1, 2, and 3 investigated learning outcomes and process level measures following two different activity types, given both before and after instruction in an introductory physics classroom. In Study 1 (EF-CC, IF-CC), there was a significant effect of type of knowledge (procedural > conceptual), and a significant type of knowledge × order interaction; students in both conditions performed equally well on the procedural assessment, and students in the EF-CC condition scored higher on the conceptual assessment than students in the IF-CC condition. There were significant effects of PFL type of knowledge (PFL procedural knowledge > PFL transfer) and order (EF > IF), as well as a significant type of knowledge × order interaction; students in the EF-CC condition performed significantly better on the PFL transfer question than students in the IF-CC condition. These results supported hypotheses based on prior CC exploratory learning research. Success on the activity differed between orders (IF > EF) but did not predict learning outcomes. Neither the interest and enjoyment scale nor cognitive load question showed differences between conditions or predicted learning outcomes.

In Study 2 (EF-RD, IF-RD), which included samples from two different introductory physics courses, there was a significant order × course interaction; students in the EF-RD condition performed better overall on the procedural and conceptual assessment than students in the IF-RD condition in the calculus-based course, and
students in both conditions performed equally in the algebra-based course. Study 2 PFL results included a significant effect of type of knowledge (conceptual knowledge > procedural knowledge and transfer), however order and the PFL type of knowledge × order interaction were not significant. Knowledge gap awareness was higher for students in the EF-RD condition than in the IF-RD condition, and negatively predicted learning outcomes. Number of strategies, the interest and enjoyment scale responses, and cognitive load were not different between conditions and did not predict learning outcomes.

Study 3 (EF-CC, IF-CC, EF-RD, IF-RD) provided a direct comparison of the CC and RD activity types in the EF and IF instructional orders. There was a significant main effect of activity (CC > RD) and a significant type of knowledge × activity interaction; students in the CC conditions performed significantly better on the procedural assessment than those in the RD conditions. Neither activity nor order were significant predictors of PFL outcomes. Activity success positively predicted procedural, conceptual, and PFL procedural learning outcomes for students in the CC conditions. Knowledge gap awareness negatively predicted all learning outcomes, and cognitive load negatively predicted conceptual knowledge and PFL transfer. Only knowledge gap awareness was different between conditions, with more perceived knowledge gaps in the EF order than in the IF order.

Study 3 also attempted to replicate both Studies 1 and 2. Posttest procedural and conceptual knowledge means are illustrated in Figure 15 for Study 1 and the CC conditions of Study 3. Although all of the means vary slightly between studies 1 and 3, the greatest difference appears to be in conceptual knowledge of students in the IF-CC
conditions. The type of knowledge × order interaction found in Study 1 (EF-CC > IF-CC in conceptual knowledge) did not occur in Study 3 because the IF-CC condition performed the same as the EF-CC group.

Figure 15: Studies 1 & 3 CC conditions, posttest performance by order and study.

Means for Studies 2 and 3 (RD conditions) are illustrated in Figure 16 below. Again, performance was slightly different between studies in all conditions, but the greatest difference between studies was in the procedural knowledge of students in the IF-RD condition. If anything, the relationship of procedural knowledge performance between students in the EF-RD and IF-RD conditions appears to have flipped from Study 2 (EF > IF) to Study 3 (IF > EF). The change occurred only in the IF-RD condition; performance of students in the EF-RD condition was roughly equivalent in the two studies.
As described in the discussion section of Study 3, the students in the IF condition in Study 3 received additional instructions from the TAs to continue working and to explain their work. In Study 3, students in the IF-CC condition also appear to have been more likely to participate in the well-established learning technique of self-explanation (e.g., Chi, De Leeuw, Chiu, & Lavancher, 1994). These students gained as much conceptual knowledge as the students in the EF-CC condition gained from exploration. The RD activity was more difficult than the CC activity, and students spent all of their activity time on procedural calculations. Students in the IF-RD condition therefore did not end up writing sentences of justification; however, it is possible that these students benefitted in procedural knowledge because the encouragement to continue working may have motivated them continue to practice the calculations.

There were also other methodological differences between the studies that could have changed the results. For example, the order effects observed in Study 1 could have been due to classroom differences instead of instructional order. Studies 2 and 3
controlled for classroom type with all conditions in similar active learning classrooms, but Study 1 had students in the EF-CC group in an active learning room and students in the IF-CC group were in the normal lecture hall. Similarly, the addition of a section break between the procedural and conceptual knowledge assessment could have changed the results between studies. The most reliable comparison is between conditions within Study 3, because all parameters were held constant across conditions.

**Theoretical Implications**

Prior exploratory learning research has revealed that reversing the tell-then-practice order can be beneficial for learning, especially for conceptual knowledge (Loibl et al., 2016). Researchers have proposed that exploration helps students (a) activate their prior knowledge, (b) perceive the gaps in their knowledge, and (c) identify problem features in the new concept. Researchers have developed different types of activities that correspond to different proposed learning mechanisms; researchers who use CC activities typically emphasize problem feature discernment in their discussions (see Schwartz et al., 2011) whereas researchers who use RD activities focus on error correction (e.g., Kapur, 2014). Researchers have also chosen to assess different learning outcomes and process level measures beyond procedural and conceptual knowledge in their studies. Finally, not all exploratory learning research has been done with thorough experimental control.

The current set of studies is the first to compare two different exploration activities with the same learning outcomes and a controlled experimental design. It is therefore the first to test whether activity type is a moderator of exploratory learning. The results of Study 1 indicate that CC activities in an EF order activate mechanisms that improve conceptual learning and PFL (e.g., Schwartz & Martin, 2004). The new CC
activity was designed (1) to activate prior knowledge and (2) to enable students to discover the problem features within gravitational field. The sub-concepts of gravitational field (the magnitude, direction, and vector nature of fields) were built into the contrasting cases. Although student success on the activity was lower in the EF condition (84%) than the IF condition (94%), means were generally high, indicating that students in the EF-CC condition were successful in problem feature discovery. Study 1 results therefore indicate that problem feature discovery could be an important learning mechanism in exploratory learning.

One possible reason for the generally null effects of instructional order in Study 2 is that, in a tightly controlled experiment, an RD activity in an exploratory learning condition does not result in learning benefits. If an EF-CC condition results in student learning gains and an EF-RD condition does not, then the learning benefits of exploratory learning could be due to the unique mechanisms evoked by CC activities; problem feature discernment would be more likely to be the driving mechanism than knowledge gap awareness and error correction (a mechanism unique to RD activities), or activation of prior knowledge (a mechanism that is similar in both activities). Indeed, in Study 2, knowledge gap awareness inversely predicted posttest performance, also indicating that error correction may not be the primary mechanism evoked in exploratory learning.

Another explanation for the Study 2 results, however, is that the RD activity used in these experiments was too difficult. The RD activity was designed (1) to activate prior knowledge, (2) to allow students to explore multiple strategies, and (3) to cause failure and knowledge gap awareness. In the end, students were unable to complete the calculations, even in the IF condition. The difficulty of the RD activity may have
prohibited students’ procedural practice in the IF condition as well as conceptual learning in the EF condition. Study 2 therefore may have revealed that activity difficulty, as opposed to type, could affect learning outcomes. Furthermore, activity difficulty could interact with knowledge gap awareness. If the activity is beyond a certain difficulty threshold, knowledge gap awareness could decrease learning, whereas if the activity is not too difficult, knowledge gap awareness could be beneficial to learning. This type of interaction could explain the mixed results found in (Glogger-Frey et al., 2015). Although a certain amount of difficulty is required to result in failure, there may be limits to how much failure students should experience. More studies would be needed to confirm that the boundary condition of activity difficulty exists, and to investigate the possible interaction between knowledge gap awareness and activity difficulty.

In Study 3, the null effects of instructional order may have been due to a benefit of self-explanation for students in the IF-CC condition. In this case, students in both EF-CC and IF-CC conditions may have been helped to focus on problem features. The consolidation of problem features in an IF order could be as helpful as the discernment of problem features in an EF order. Both consolidation and discernment of problem features could cause improvements in conceptual knowledge above a typical tell-then-practice instructional method. If discernment of problem features is the leading learning mechanism behind exploratory learning, and if students in the IF-CC condition also focused on problem features, then the Study 3 null conceptual knowledge results between instructional orders makes sense. Helping students appreciate the importance of underlying sub-concepts within a complex concept with interrelating parts is likely valuable for conceptual knowledge, which is fundamentally relational. Other branches of
research outside of exploratory learning have made similar propositions about the benefits of identifying problem features. Learning gains have been observed due to comparative instruction (e.g., Rittle-Johnson & Star, 2009) and analogy (e.g., Catrambone & Holyoak, 1989), which both require that learners synthesize critical features of new material. Exploratory learning may be one of several ways to help students to discern, identify, or focus on problem features, and the discovery of problem features could be the driving mechanism behind the learning benefits of an EF instructional order.

Lastly, it is important to mention that the instruction was not optimized for the RD activity. Loibl and colleagues (2016) found that, across many exploratory learning studies (mostly using RD activities), results were sensitive to the instructional method; positive results were found when the instruction was based on incorrect student solutions. Controlling the post-activity instruction in the current studies was necessary to adequately compare the learning mechanisms occurring during exploration. If the driving learning mechanisms for RD activities occur during the instruction rather than during exploration, they would be different from the mechanisms at work in exploratory learning with CC activities.

Overall, results from these three studies indicate that there may be different learning mechanisms at work when students participate in different types of exploratory activities, and that discernment of problem features could be one of the important mechanisms.
Limitations & Future Work

The studies reported here have several limitations, and therefore the results must be interpreted carefully. The limitations of each study are described in detail in the respective chapters and are summarized here. First, Study 3 did not replicate Studies 1 and 2. Although there were a few methodological changes between studies that could justify the different results, it remains that none of the results have been replicated. It is possible to use the differences in learning outcomes to argue that the benefits of exploratory learning are sensitive to implementation methods outside of the raw materials; learning benefits may only be observed if materials are implemented in a specific way. Differences in learning outcomes can also be associated with the second limitation of the current work, which is the low reliability of the assessment scales. Low reliability, as discussed in detail in Study 2 and Study 3, limits the ability to judge student learning from these assessments.

Much more work is needed to determine whether activity type is a moderator of exploratory learning. Lab studies may be a good starting place for manipulating experimental materials only and would be a good place to study CC versus RD activities. In addition to activity type, other moderators of exploratory learning are worth investigating (e.g., self-explanation instruction or other guidance during the activity, active learning classrooms, and student populations). The observed differences in the results between studies indicates that there are more moderators to explore. After several studies investigate individual moderators of exploratory learning, and when more information is known about the best implementation practices, it would be interesting to compare some best-practice exploratory learning combinations against each other. For
example, an experiment could compare the learning outcomes of students in conditions with (1) a CC activity followed by problem-feature-focused instruction to (2) an RD activity followed by student-error-focused instruction. These could both be compared to a tell-then-practice condition with an activity that was optimized to challenge students as well as help them engage in comparison in and around the dominant problem features.

Conclusions

There is currently a strong drive to apply active learning methods in STEM education, and exploratory learning seems like a promising method due to the observed benefits of conceptual understanding, transfer, and preparation for future learning. The future learning benefits are especially useful in STEM domains where information is cumulative, and expertise depends on strong memory and organization of related concepts. However, there is much more work to be done to investigate the boundary conditions and moderators impacting exploratory learning. The current findings indicate that activity type may make a difference in quality of exploratory learning in an undergraduate physics classroom. In a direct comparison of CC and RD activities in IF and EF instructional orders, students in the CC conditions performed better on a procedural assessment than students in the RD conditions. Although these studies are limited and findings should not be widely generalized, the results indicate that the design of in-class activities is important, and that when designing an exploration activity, a good goal would be to get students to identify problem features. Learning mechanisms are likely dependent on choices made in the activity design. Many more studies are required to pinpoint the learning mechanisms and best practices for exploratory learning.
REFERENCES


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


The four learning objectives for gravitational field are as follows:

A) A mass in space creates a gravitational field at all points in space around it with a magnitude equal to the mass divided by the distance from the mass to the point squared.

\[ | \vec{g} | = \frac{m_{source}}{r^2} \]

B) Gravitational field is a vector quantity with a magnitude and direction. The gravitational field produced by a single source mass points towards the mass.

C) Multiple masses in the same space influence each point in that space. The magnitude and direction of the gravitational field can be calculated with vector addition.

D) If another test mass \( m_{test} \) is added to a point space with a gravitational field of \( \vec{g} \), the resulting force on that mass is the vector quantity of \( m \) multiplied by \( \vec{g} \).

\[ \vec{F} = m_{test} \vec{g} \]
**APPENDIX B: Survey (Studies 2 and 3)**

**Instructions:** We would like to ask you some questions about your perceptions of the learning activity. We are interested in your honest opinion. There are no right or wrong answers.

**Please indicate how much you agree or disagree with the statements below:**

<table>
<thead>
<tr>
<th>Question</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. Today’s activity has been interesting.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2. I do not feel very knowledgeable about calculating gravitational field.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3. I have enjoyed today’s activity.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4. I know pretty much about calculating gravitational field.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5. Today’s activity really captured my attention.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6. Compared to most other people, I know less about calculating gravitational field.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>7. Today’s activity kept me engaged.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>8. When it comes to calculating gravitational field, I really don’t know a lot.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

**Please indicate how much mental effort you invested when completing the learning activity.**

In completing the learning activity today I invested

1. Very, very low mental effort
2. Very low mental effort
3. Low mental effort
4. Rather low mental effort
5. Neither low nor high mental effort
6. Rather high mental effort
7. High mental effort
8. Very high mental effort
9. Very, very high mental effort
Procedural Assessment

Directions: For each question below, write the letter for the correct answer in the space provided. Where calculations are necessary, use \( G = 6.67 \times 10^{-11} \text{ N} \cdot \text{m}^2/\text{kg}^2 \).

Mass \( m_1 = 25 \text{ kg} \) is a distance \( d = 3 \text{ m} \) away from point \( P_A \). Point \( P_B \) is an additional distance of \( d = 3 \text{ m} \) away from \( P_A \).

1. What is the magnitude of the gravitational field at point \( P_A \)?
   A. \( 5.56 \times 10^{-10} \text{ N/kg} \)
   B. \( 4.63 \times 10^{-9} \text{ N/kg} \)
   C. \( 1.85 \times 10^{-10} \text{ N/kg} \)
   D. \( 1.39 \times 10^{-8} \text{ N/kg} \)
   E. 0 N/kg

2. What is the direction of the gravitational field at point \( P_A \)?
   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. What is the magnitude of the gravitational field at point \( P_B \)?
   A. \( 1.16 \times 10^{-11} \text{ N/kg} \)
   B. \( 6.95 \times 10^{-9} \text{ N/kg} \)
   C. \( 2.78 \times 10^{-10} \text{ N/kg} \)
   D. \( 4.63 \times 10^{-11} \text{ N/kg} \)
   E. 0 N/kg

4. What is the direction of the gravitational field at point \( P_B \)?
   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
Two masses, both with a mass of 10 kg, are separated by a distance $8d$, where $d = 1$ m. Point $P_C$ lies at the midpoint along the line connecting the two masses. Point $P_D$ lies a distance $3d$ above point $P_C$.

5. What is the magnitude of the gravitational field at point $P_C$?

A. $3.34 \times 10^{-10}$ N/kg  
B. $3.34 \times 10^{-9}$ N/kg  
C. $8.34 \times 10^{-10}$ N/kg  
D. $8.34 \times 10^{-11}$ N/kg  
E. $1.04 \times 10^{-11}$ N/kg  
F. 0 N/kg

6. What is the direction of the gravitational field at point $P_C$?

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

7. What is the direction of the gravitational field at point $P_D$?

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
A particle with mass \( m_1 \) sits a distance \( d \) away from the point \( P \). Initially you place a second particle mass \( m_2 \) at point \( P \).

\[
\begin{array}{c}
\bigcirc \\
\hspace{1cm} m_1 \\
\hspace{1cm} d \\
\bullet P
\end{array}
\]

8. If you replace \( m_2 \) with a third particle that is twice as massive as \( m_2 \), what happens to the magnitude of the gravitational field at point \( P \) due to \( m_1 \)?

A. **The magnitude of the gravitational field remains the same.**
B. The magnitude of the gravitational field doubles.
C. The magnitude of the gravitational field is four times greater.
D. The magnitude of the gravitational field is reduced by half.
E. The magnitude of the gravitational field is reduced by a factor of four.

9. When you replace \( m_2 \) with a third particle that is twice as massive as \( m_2 \), how does the magnitude of the gravitational force on \( m_2 \) compare to the magnitude of the gravitational force on \( m_3 \)?

A. The magnitude of the gravitational force is the same on \( m_2 \) and \( m_3 \).
B. The magnitude of the gravitational force on \( m_2 \) is twice as large.
C. The magnitude of the gravitational field on \( m_2 \) is four times as large.
D. **The magnitude of the gravitational force on \( m_3 \) is twice as large.**
E. The magnitude of the gravitational field on \( m_3 \) is four times as large.

**Conceptual Assessment**

**Directions:** For each statement below, indicate whether the statement is true (T) or false (F) in the space provided.

10. The field around a mass points away from the mass. **F**

11. The magnitude of the gravitational field is dependent upon the mass of two objects and the distance between them. **F**

12. When multiple masses contribute to the gravitational field at a point in space, the direction of the gravitational field at the point can be calculated using vector addition because the gravitational field is a vector quantity. **T**

13. The gravitational force on a mass placed in a gravitational field points in the opposite direction as the gravitational field. **F**
F 14. The gravitational field is the force felt between two masses in space.
F 15. The direction of the gravitational field at a point in space always points toward the nearest mass.
T 16. The gravitational force on an object placed at point $P$ depends only on the gravitational field at point $P$ and the object’s mass.
T 17. A mass in space creates a gravitational field around it.
T 18. The direction of the gravitational field is dependent on all masses in the system.
F 19. A mass $M$ is located a distance $D$ away from point $A$. Assuming there are no other masses in the vicinity of point $A$, if the mass is increased to $2M$, the gravitational field at point $A$ is increased by a factor of 4.

Please stop at this point. We will let you know when to begin part two.
The Electric Field

The electric field generated by charged particles is very similar to the gravitational field generated by masses. Like the gravitational field, the electric field is a vector quantity that tells us how a particular charge influences the space around it.

The magnitude of an electric field $E$ at a point in space due to a single charge $Q$ is:

$$E = k \frac{|Q|}{r^2}$$

where $Q$ is given in units of Coulomb’s (C), $k$ is the Coulomb’s constant with a value of $8.99 \times 10^9 \text{ N} \cdot \text{m}^2/\text{C}^2$ and $r$ is the distance from the point in space to the charge. Charges are different from masses in that they can be positive or negative. Notice that in our equation for the magnitude of the electric field, we need the absolute value of $Q$ to insure that we calculate a positive magnitude for the field vector.

The direction of the electric field is based on the sign of the charge: the electric field vector points toward negative charges, and away from positive charges. In the figure below, the panel on the left shows the electric field at point $P$ due to a single charge $+Q$. The panel on the right show the electric field at the same point $P$, but the charge $+Q$ has been replaced by a second charge $-Q$. Notice that the electric field points away from the positive charge on the left and towards the negative charge on the right.

When there are multiple charged particles, the magnitude and direction of the electrical field at a point can be calculated with vector addition. For example: two particles with charges $+Q$ and $-Q$ are located along the $x$-axis, a distance $d$ away from each other, as shown in the figure below. Point $P$ is located along the $y$-axis, a distance $d$ away from the positive charge.
What is the magnitude and direction of the electric field at point $P$? First we need to determine the contribution to the electric field at point $P$ from each charge. The $+Q$ charge contributes a field vector that points upwards, while the contribution from the $-Q$ charge points down and to the right, as shown in the figure below.

![Diagram showing electric field contributions](image)

Notice that the negative charge is farther away from point $P$ than the positive charge by a factor of $\sqrt{2}$. Because both particles have the same amount of charge, one positive, the other negative, the magnitude of the electric field vector from the negative charge will be \( \frac{1}{(\sqrt{2})^2} = \frac{1}{2} \) the magnitude of the electric field vector from the positive charge.

The net electric field at point $P$ will just be the vector sum of the electric field vectors from each individual charge. Using the usual process for vector addition, we can add together $\vec{E}_+$ and $\vec{E}_-$ to get the net electric field vector, $\Sigma \vec{E}$, as shown in the figure below. Therefore, the electric field at point $P$ points up and to the right.
PFL Assessment

Procedural

Directions: For each question below, write the letter for the correct answer in the space provided. Where calculations are necessary, use \( k = 8.99 \times 10^9 \text{ N}\cdot\text{m}^2/\text{C}^2 \).

A particle with a charge of \( +Q = +3 \times 10^{-9} \text{ C} \) is a distance \( d = 2 \text{ m} \) away from point \( P_1 \).

\[ +Q \quad P_1 \quad +Q = +3 \text{ C} \]
\[ d = 2 \text{ m} \]

_____ 1. What is the magnitude of the gravitational field at point \( P_1 \)?

A. \( 6.74 \times 10^9 \text{ N/C} \)   D. \( 4.05 \times 10^{10} \text{ N/C} \)
B. \( 1.35 \times 10^{10} \text{ N/C} \)   E. 0 N/kg
C. \( 2.02 \times 10^{10} \text{ N/C} \)

_____ 2. What is the direction of the gravitational field at point \( P_1 \)?

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

A particle with a charge of \( -Q = -4 \times 10^{-9} \text{ C} \) is a distance \( d = 5 \text{ m} \) away from point \( P_2 \).

\[ -Q \quad P_2 \quad -Q = -4 \text{ C} \]
\[ d = 5 \text{ m} \]

_____ 3. What is the magnitude of the gravitational field at point \( P_1 \)?

A. \( 2.88 \times 10^{10} \text{ N/C} \)   D. \( 1.44 \times 10^9 \text{ N/C} \)
B. \( 7.19 \times 10^9 \text{ N/C} \)   E. 0 N/kg
C. \( 5.75 \times 10^9 \text{ N/C} \)
4. What is the direction of the gravitational field at point \( P_1 \)?

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

**Directions:** For each question below, particles with either a positive charge \( +Q \) or negative charge \( -Q \) (with an equal absolute value) are arranged in space as shown in each figure. Estimate the direction of the electrical field at the specified points.

Two particles, both with charge \( +Q \), are separated by a distance of 4 m. Points \( P_3 \) and \( P_4 \) lie along the \( y \) axis, with \( P_3 \) at the origin and \( P_4 \) 1 m away from the origin below the \( x \) axis. Point \( P_5 \) lies along the \( x \) axis 1 m away from the charge to the right of the origin.

5. What is the direction of the electric field at point \( P_3 \)?

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

6. What is the direction of the electric field at point \( P_4 \)?

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  

---

95
B. Up and to the right
C. Up
D. Up and to the left
E. Horizontally to the left

7. What is the direction of the electric field at point $P_5$?

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

Two particles, one with charge $+Q$ and the other with charge $-Q$, are separated by a distance of 2 m. Points $P_6$ and $P_7$ lie along the $y$ axis, with $P_6$ at the origin and $P_7$ 1 m above the origin.

8. What is the direction of the electric field at point $P_6$?

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. What is the direction of the electric field at point $P_7$?

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

Transfer

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

_____ 10. Which of the following equations describes the electric force exerted on a particle with a charge $q$ that is placed in the electric field $E$ created by another charge $Q$?

A. $\vec{F}_e = q\vec{E}$  
B. $\vec{F}_e = Q\vec{E}$  
C. $\vec{F}_e = qQ\vec{E}$  
D. $\vec{F}_e = q^2\vec{E}$  
E. $\vec{F}_e = Q^2\vec{E}$
APPENDIX D: Studies 2 & 3 Assessments

Procedural and Conceptual Assessments

The Procedural and Conceptual Assessments were the same as in Study 1 (pages 88-98). A section break was added between the two assessments, as follows:

Please stop at this point. We will let you know when to begin part two.

PFL Learning Resource

The PFL Learning Resource was the same as in Study 1 (see page 92).
PFL Assessment

The Procedural, Conceptual and Transfer questions were mixed.

**Procedural:**

**Directions:** For each question below, write the letter for the correct answer in the space provided. Where calculations are necessary, use $k = 8.99 \times 10^9 \text{N} \cdot \text{m}^2/\text{C}^2$.

A particle with a charge of $-Q = -4 \text{ C}$ is a distance $d = 5 \text{ m}$ away from point $P_1$.

_____ 1. What is the magnitude of the electric field at point $P_1$?
   A. $2.88 \times 10^{10} \text{ N/C}$
   B. $7.19 \times 10^{9} \text{ N/C}$
   C. $5.75 \times 10^{9} \text{ N/C}$
   D. $1.44 \times 10^9 \text{ N/C}$
   E. $0 \text{ N/kg}$

_____ 2. What is the direction of the electric field at point $P_1$?
   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

**Transfer:**

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

_____ 3. Which of the following equations describes the electric force exerted on a particle with a charge $q$ that is placed in the electric field $E$ created by another charge $Q$?

   A. $\vec{F}_e = q\vec{E}$
   B. $\vec{F}_e = Q\vec{E}$
   C. $\vec{F}_e = qQ\vec{E}$
   D. $\vec{F}_e = q^2\vec{E}$
   E. $\vec{F}_e = Q^2\vec{E}$
**Conceptual:**

**Directions:** For each statement below, write a “T” or an “F” in the blank to the left. Write a “T” for statements you think are true and an “F” for statements you think are false.

___F___ 4. The electric field points towards a positive charge.

___T___ 5. The electric field is a vector quantity.

**Procedural:**

**Directions:** For each question below, particles with either a positive charge $+Q$ or negative charge $-Q$ (with an equal absolute value) are arranged in space as shown in each figure. Estimate the direction of the electrical field at the specified points.

Two particles, one with charge $+Q = +2 \text{C}$ and the other with charge $-Q = -2 \text{C}$, are separated by a distance of 2 m. Points $P_2$ and $P_3$ lie along the $y$ axis, with $P_2$ at the origin and $P_3$ 1 m above the origin.

___ 6. What is the direction of the electric field at point $P_2$?
   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

___ 7. What is the direction of the electric field at point $P_3$?
   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
Conceptual:

Directions: For each statement below, write a “T” or an “F” in the blank to the left. Write a “T” for statements you think are true and an “F” for statements you think are false.

8. To calculate the net electric field at a point in space, you just need to add together the magnitudes of the individual contribution of each charge to the electric field as numbers.

9. The magnitude of the electric field generated by a particle at a point in space is based on a constant, $k$, the absolute value of the charge of the particle, $|Q|$, and the distance from the particle to the point.
Incorrect:

103

Corrected:
\[\begin{align*}
\text{At } \gamma: \\
\frac{\partial}{\partial x} & \left( \frac{\partial \gamma}{\partial x} \right) + \frac{\partial}{\partial z} \left( \frac{\partial \gamma}{\partial z} \right) = \gamma + \beta
\end{align*}\]
\[ \begin{align*}
\eta = \text{component of } \eta = \\
2 \cos \theta_2 - \eta_2 - 2 \cos \theta_1 + \eta_1 - \eta_3 = \Sigma \eta_y \\
(\theta \cos \theta_1 - \theta \sin \theta_1) - 2 \sin \theta_1 - 2 \sin \theta_1 + \Sigma \eta_y = \Sigma \eta_y \\
\Sigma \eta_y = (\cos \theta_1 - 2 \sin \theta_1 - 2 \sin \theta_1 + \Sigma \eta_y) = -1.27 \text{ m}
\end{align*} \]

6. Determine magnitude of net \( \eta \)-field

\[ \Sigma \eta_y = \sqrt{\Sigma \eta_y} \]

\[ = 1.27 \text{ m} \quad \Rightarrow \quad \Sigma \eta_y = 1.27 \text{ m} \]

For distance of each mass to \( P_0 \):

\[ \begin{array}{c|c|c|c|c|c|c|c|c|c|c}
\text{Distance of } P_n & \text{Distance of } (x_n, y_n) & \Delta x & \Delta y & x_n & y_n \\
\hline
m_1 & 2 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\
m_2 & 2 & 0 & -3 & 5 & 0 & 0 & 0 & 0 & 0 \\
m_3 & 2 & 0 & -4 & -1 & 0 & 0 & 0 & 0 & 0 \\
m_4 & 2 & 0 & 2 & -4 & 0 & 0 & 0 & 0 & 0 \\
m_5 & 2 & 0 & 4 & 4 & 0 & 0 & 0 & 0 & 0 \\
\end{array} \]

\[ \begin{align*}
\Sigma x &= \sqrt{\Sigma x^2} + \sqrt{\Sigma y^2} \\
\Sigma x &= \sqrt{(-2)^2 + (5)^2 + (-4)^2 + (-4)^2 + (-1)^2} \\
\Sigma y &= \sqrt{(-2)^2 + (5)^2 + (-4)^2 + (-4)^2 + (-1)^2} \\
\Sigma x &= 3.13 \text{ m} \\
\Sigma y &= 7.21 \text{ m} \\
\end{align*} \]

\[\begin{align*}
\text{At point } P_0:
\eta_1 &= \frac{\eta_1}{\Sigma \eta_y} \\
\eta_2 &= \frac{\eta_2}{\Sigma \eta_y} \\
\eta_3 &= \frac{\eta_3}{\Sigma \eta_y} \\
\eta_4 &= \frac{\eta_4}{\Sigma \eta_y} \\
\eta_5 &= \frac{\eta_5}{\Sigma \eta_y} \\
\Sigma \eta_y &= \eta_1 - \eta_2 + \eta_3 - \eta_4 + \eta_5 \\
\Sigma \eta_y &= \Sigma \eta_y \\
\end{align*} \]

\[\begin{align*}
\eta_1 &= \frac{\eta_1}{\Sigma \eta_y} \\
\eta_2 &= \frac{\eta_2}{\Sigma \eta_y} \\
\eta_3 &= \frac{\eta_3}{\Sigma \eta_y} \\
\eta_4 &= \frac{\eta_4}{\Sigma \eta_y} \\
\eta_5 &= \frac{\eta_5}{\Sigma \eta_y} \\
\Sigma \eta_y &= \eta_1 - \eta_2 + \eta_3 - \eta_4 + \eta_5 \\
\Sigma \eta_y &= \Sigma \eta_y \\
\end{align*} \]

\[\begin{align*}
\eta_1 &= \frac{\eta_1}{\Sigma \eta_y} \\
\eta_2 &= \frac{\eta_2}{\Sigma \eta_y} \\
\eta_3 &= \frac{\eta_3}{\Sigma \eta_y} \\
\eta_4 &= \frac{\eta_4}{\Sigma \eta_y} \\
\eta_5 &= \frac{\eta_5}{\Sigma \eta_y} \\
\Sigma \eta_y &= \eta_1 - \eta_2 + \eta_3 - \eta_4 + \eta_5 \\
\Sigma \eta_y &= \Sigma \eta_y \\
\end{align*} \]

\[\begin{align*}
\eta_1 &= \frac{\eta_1}{\Sigma \eta_y} \\
\eta_2 &= \frac{\eta_2}{\Sigma \eta_y} \\
\eta_3 &= \frac{\eta_3}{\Sigma \eta_y} \\
\eta_4 &= \frac{\eta_4}{\Sigma \eta_y} \\
\eta_5 &= \frac{\eta_5}{\Sigma \eta_y} \\
\Sigma \eta_y &= \eta_1 - \eta_2 + \eta_3 - \eta_4 + \eta_5 \\
\Sigma \eta_y &= \Sigma \eta_y \\
\end{align*} \]
Plugging into component equations:

\[-q_y = -q_1 \cos \theta_1 - q_2 \cos \theta_2 + q_3 \cos \theta_3 = \Sigma q_y\]

\[-\frac{5}{6} G \cos \theta_1 - \frac{3}{5} G \cos \theta_1 = \frac{5}{6} G \cos \theta_2 + \frac{5}{6} G \cos \theta_3 = \Sigma q_y\]

\[\Sigma q_y = (-\frac{5}{6} - \frac{3}{5} \cos \theta_1 - \frac{5}{6} \cos \theta_2 + \frac{5}{6} \cos \theta_3) G = -0.270 G\]

\[q_2 \sin \theta_2 - q_3 \sin \theta_3 - q_4 \sin \theta_4 = \Sigma q_y\]

\[\frac{2}{5} G \sin \theta_1 - \frac{4}{5} G \sin \theta_1 = \frac{2}{5} G \sin \theta_2 + \frac{2}{5} G \sin \theta_3 = \Sigma q_y\]

\[\Sigma q_y = \left(\frac{2}{5} G \sin \theta_1 - \frac{4}{5} G \sin \theta_1 - \frac{2}{5} G \sin \theta_2 + \frac{2}{5} G \sin \theta_3\right) G = -0.802 G\]

For magnitude of net \(q\)-field:

\[|\Sigma q| = \sqrt{\Sigma q_x^2 + \Sigma q_y^2} = \sqrt{(-0.270 \text{ G})^2 + (-0.802 \text{ G})^2} = 0.85 G\]

At point \(P_a\), net \(q\)-field has a magnitude of 4.05 G.
At point \(P_b\), net \(q\)-field has a magnitude of 0.90 G.

\(|\Sigma q|\) at \(P_a > |\Sigma q|\) at \(P_b\) and masses have a greater influence at \(P_a\) than at \(P_b\).

\[q - R - 0.5\]
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>Contrasting Cases</td>
</tr>
<tr>
<td>EF</td>
<td>Explore-first</td>
</tr>
<tr>
<td>IF</td>
<td>Instruct-first</td>
</tr>
<tr>
<td>PFL</td>
<td>Preparation for Future Learning</td>
</tr>
<tr>
<td>RD</td>
<td>Rich Dataset</td>
</tr>
</tbody>
</table>
CURRICULUM VITAE
Campbell Rightmyer Bego

EDUCATION

*University of Louisville: Ph.D., Experimental Psychology*  
Dec. 2019

M.S., Experimental Psychology, GPA: 4.0  
*May 2016*

*Columbia University: B.S., Mechanical Engineering, GPA: 3.65*  
*May 2008*

PUBLICATIONS


**CONFERENCE POSTERS AND PRESENTATIONS**


Rightmyer, C. A. & Pani, J. R. “No Gender Effects in Real-World Spatial Ability.”

Poster presented at Graduate Research Symposium, University of Louisville, April, 2016

Rightmyer, C. A. “No Gender Effects in Real-World Spatial Ability.” Three-minute Thesis Competition Presentation at Graduate Research Symposium, University of Louisville, April 2016

GRANT AWARDS

Postdoctoral Researcher in NSF IUSE EHR Award #1912253: Implementing Spaced Retrieval Practice Across Multiple STEM Domains to Enhance Student Learning in Barrier Courses. $598,000. Oct. 2019 - Oct. 2022

GRANT APPLICATIONS

NSF Graduate Fellowship Application May 2016

BUSINESS & CONSULTING EXPERIENCES


Tunnel Ventilation Engineering Consultant


Mechanical Engineering Consultant
HONORS & DISTINCTIONS


*Pi Tau Sigma*, Columbia University.


*Professional Engineer*, New York State, License #091749