In press in *Contemporary Educational Psychology*. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. The final article will be available via its DOI at https://doi.org/10.1016/j.cedpsych.2017.12.003

Reverse the Routine: Problem Solving Before Instruction Improves Conceptual Knowledge in Undergraduate Physics

Joanna P. Weaver[a, b], Raymond J. Chastain[a], Daniel A. DeCaro[a], and Marci S. DeCaro[a]

University of Louisville

Author Note

[a]Joanna P. Weaver and Marci S. DeCaro, Department of Psychological and Brain Sciences, 317 Life Sciences Building, University of Louisville, Louisville, Kentucky 40292 USA; Phone: (502) 852-8273. Raymond J. Chastain, Department of Physics and Astronomy, University of Louisville; Daniel A. DeCaro, Departments of Urban and Public Affairs and Psychological and Brain Sciences, University of Louisville.

[b]Joanna Weaver is now in the Division of Educational Psychology & Methodology, University at Albany, State University of New York.

Correspondence concerning this article should be addressed to Joanna P. Weaver (phone: 518-512-3841; email: jweaver2@albany.edu) or Marci S. DeCaro (phone: 502-852-8273; email: marci.decaro@louisville.edu).

Abstract

STEM undergraduate classrooms are increasingly adopting instructional methods to enhance student engagement and improve learning outcomes. For example, in exploratory learning, students explore novel problems before they are taught the underlying concepts and procedures. The current studies examined the benefits of exploratory learning in undergraduate physics instruction. In Studies 1 and 2, students worked collaboratively in groups to complete a learning activity before lecture (*explore-first condition*) or after (*instruct-first condition*). The two studies were conducted in different semesters, with different physics courses and instructors of record. Students' conceptual understanding and procedural knowledge (problem-solving accuracy) were assessed using an instructor-created quiz. Performance on the learning activity indicated that students in the explore-first condition struggled as much as (Study 2) or more (Study 1) than students in the instruct-first condition. However, after the learning activity, students in the explore-first condition exhibited better conceptual understanding and equal procedural knowledge, compared to students in the instruct-first condition. In addition, self-reported interest and enjoyment was either equal (Study 1) or greater (Study 2) in the explore-first condition. Study 3 tested the effects of exploring alone versus in a collaborative group. Learning outcomes were equal across conditions, suggesting that there is no added benefit of exploring collaboratively compared to individually. However, interest and enjoyment were higher when students explored collaboratively, which may have long-term educational benefits. Exploratory learning, with or without collaboration, offers a useful method to improve student engagement and performance in essential undergraduate STEM courses.

Keywords

exploratory learning; problem solving; interest and enjoyment; STEM; undergraduate

Reverse the Routine: Problem Solving Before Instruction Improves Conceptual Knowledge in Undergraduate Physics

1. Introduction

In undergraduate Science, Technology, Engineering, and Math (STEM) education, there is an increasing emphasis on innovative, student-centered teaching methods to improve performance and engagement in traditional lecture-based courses (Prince, 2004). Constructivist-inspired learning techniques, such as discovery learning, problem-based learning, exploration, and inquiry learning, are theorized to improve conceptual understanding of important topics by better engaging students in the learning process, compared to traditonal lecture formats (Duffy & Jonassen, 1992; Loyens, Jones, Mikkers, & van Gog, 2015; Schwartz, Lindgren, & Lewis, 2009; Wise & O'Neill, 2009). Constructivist methods require students to take a more active role in creating their own knowledge (Piaget, 1926, 1973; Vygotsky, 1978). By increasing learning and engagement, constructivist-inspired approaches can improve both academic achievement and student retention in core STEM undergraduate courses (Felder, Woods, Stice, & Rugarcia, 2000; Prince, 2004). For example, Hake's (1998) classic study demonstrated that students' conceptual understanding of mechanics and problem-solving performance improved by more than two standard deviations in core physics courses that used a constructivist learning activity (e.g., collaborative peer instruction, modeling) to supplement lessons taught in the more traditional lecture format.

Instructional methods inspired by constructivism have been used and promoted by educational scholars for decades. However, empirical evidence for these methods is not as strong as may be assumed, and there are significant boundary conditions to consider when implementing them in classrooms (Alfieri, Brooks, Aldrich, & Tenenbaum, 2011; Capon & Kuhn, 2004; Kirschner, Sweller, & Clark, 2006; Mayer, 2004; Schwartz et al., 2009). Specifically, a large body of research indicates that constructivist methods can sometimes be too difficult for students (e.g., DeCaro, DeCaro, & Rittle-Johnson, 2015; Kirschner et al., 2006) and that more traditional direct instruction methods may yield better learning outcomes in some cases (cf. Kirschner et al., 2006; Klahr & Nigam, 2004; Mayer, 2004; Sweller, Kirschner, & Clark, 2007). Proponents of direct instruction argue that active knowledge construction increases the cognitive demands of the learning activity, which may not necessarily improve learning (Chandler & Sweller, 1991; Mayer & Moreno, 2003; Sweller, van Merrienboer, & Paas, 1998). Students may acquire incorrect, incomplete, or disorganized knowledge when constructing knowledge for themselves, and these shortcomings may be too difficult to overcome (Kirschner et al., 2006). Direct instruction guides students' attention to the most relevant information, reducing cognitive load, confusion, and difficulty (Kirschner et al., 2006; Sweller et al., 2007).

Other studies reveal the value of combining both direct instruction and constructivist-inspired methods (Alfieri et al., 2011; Kapur, 2016; Mayer, 2004). One approach, termed *productive failure*, has students solve problems before being taught the correct procedures, and then compare their solutions to the canonical one (e.g., Kapur, 2010, 2011, 2012; Kapur & Bielaczyc, 2012; Loibl & Rummel, 2014). Another approach, termed *invention*, has students invent problem-solving procedures based on their review of a set of contrasting cases (e.g., Belenky & Nokes-Malach, 2012, Chin, Chi, & Schwartz, 2016; Glogger-Frey, Fleischer, Grüny, Kappich, & Renkl, 2015; Schwartz, Chase, Oppezzo, & Chin, 2011; Schwartz & Martin, 2004). We characterize these and similar approaches (e.g., problem-solve-instruct conditions, Loibl, Roll, & Rummel, 2016; explore-instruct approach, Loehr, Fyfe, & Rittle-Johnson, 2014) under the broader label *exploratory learning* (e.g., DeCaro et al., 2015; DeCaro & Rittle-Johnson, 2012). With exploratory learning methods, students begin with a constructivist-inspired activity, such as solving novel problems or viewing contrasting cases, prior to receiving direct instruction on the relevant procedures and/or concepts. Thus, our use of the term “exploratory learning” includes both an exploration phase as well as instruction—combining both constructivist-inspired and direct instruction approaches.

Exploratory learning attempts to deepen conceptual understanding by giving learners an opportunity to explore aspects of novel topics on their own prior to receiving more traditional instruction (DeCaro & Rittle-Johnson, 2012; Schwartz & Bransford, 1998). This method has been shown to improve learning for students in mathematics (e.g., DeCaro & Rittle-Johnson, 2012; Kapur, 2014; Loehr et al., 2014; Loibl & Rummel, 2014; Schwartz & Martin, 2004) and physics (Schwartz et al., 2011, cf. Hsu, Kalyuga, Sweller, 2015). Beyond learning to use procedures, exploratory learning is intended to help learners create new knowledge beyond instruction (Schwartz et al., 2009). Consistent with this goal, the benefits of exploratory learning are typically limited to measures of conceptual knowledge or transfer, when compared to more traditional tell-then practice conditions. Students learn procedures equally well in either learning condition (Loibl et al., 2016; Schwartz et al., 2009).

A number of prior studies on exploratory learning, and in STEM education research more generally (Ruiz-Primo, Briggs, Iverson, Talbot, & Shepard, 2011), fail to use careful experimental controls to isolate the effects of learning condition from other factors, such as differences in the instructions or the problems used (Hsu et al., 2015; Loibl et al., 2016; Schwartz et al., 2011; Sweller, 2009). Thus, more rigorous experimental studies investigating the use of constructivist-inspired learning strategies in STEM are needed, to ascertain their effectiveness and to justify any additional demands they may place on instructors. Indeed, few experimental tests of exploratory learning in STEM college classrooms have been conducted (but see Schwartz & Bransford, 1998; Westermann & Rummel, 2012). The current studies empirically examine the impact of exploratory learning in undergraduate physics instruction, using careful experimental controls.

1.1 Mechanisms of Exploratory Learning

There are multiple mechanisms that may contribute to the benefits of exploratory learning. First, during exploratory learning, a learner must *activate prior knowledge* and attempt to apply it appropriately to new material (Kapur & Bielaczyc, 2012; Loibl et al., 2016; Schwartz & Bransford, 1998; Schwartz, Sears, & Chang, 2007). This process of active discovery may not always lead to correct answers initially. However, as learners search for solutions, they may develop and explore alternative representations of the problem or topic and differentiate their prior knowledge (Kapur, 2011). Thus, learners activate existing concepts and schemas during exploration, as they attempt to organize and integrate new information introduced during exploration into their already developed memory structures (DeCaro & Rittle-Johnson, 2012; Schwartz et al., 2009). This generative activity may be challenging and prone to mistakes, but it may also result in stronger connections between old and new information (e.g., Anderson, 1983; Carpenter, 2009). In this way, initial "failure" can sometimes be "productive," and early challenges can pose a *desirable difficulty*, because they may better prepare learners for future instruction (DeCaro & Rittle-Johnson, 2012; Kapur, 2012, 2014, 2016; Schmidt & Bjork, 1992).

Second, when something is easy to learn, learners are susceptible to believing that they know it well (Bjork, 1994). This *fluency* can lead to overconfidence, which may contribute to poor metacognitive awareness of one's own level of understanding, reducing necessary effort and practice time (Bjork, 1994; Dunlosky & Rawson, 2012). There is some evidence that direct instruction can promote such fluency and overconfidence, contributing to decreased effort and poorer learning outcomes (DeCaro & Rittle-Johnson, 2012; Renkl, 1999). In contrast, because exploration is typically more error prone, exploratory learning may enhance metacognitive awareness and reduce overconfidence, preparing students to pay closer attention to subsequent instruction and improving their learning outcomes (DeCaro & Rittle-Johnson, 2012; Koriat & Bjork, 2005). Exploratory learning may therefore also motivate learners by creating a desire to understand (Wise & O'Neill, 2009). The goal to learn more, and to address gaps in one's own understanding, could motivate learners to attend to key concepts and conceptual connections (Capon & Kuhn, 2004; Glogger-Frey et al., 2015; Loibl et al., 2016; Schwartz & Martin, 2004).

Some researchers have suggested that the discovery aspect of exploration and other constructivist-inspired methods has the potential to increase task interest and enjoyment, thereby encouraging deeper processing of the material and improving conceptual learning, persistence, and performance (Belenky & Nokes-Malach, 2012; Glogger-Frey et al., 2015). Specifically, interest and enjoyment are often piqued by tasks, and learning situations, that introduce some level of novelty, complexity, and personal direction (Rotgans & Schmidt, 2011; Silvia, 2008), supporting a sense of exploration and volition (Niemiec & Ryan, 2009). These factors are important in STEM education, and generally characterize exploratory learning (e.g., Mitchell, 1992). Many factors influence situational task interest and enjoyment (Renninger & Hidi, 2011). Under favorable conditions (e.g., sufficient prior knowledge), individuals may perceive an opportunity for exploration as feasible and rewarding, encouraging them to deeply engage in the learning activity (Silvia, 2008). Such engagement may have long-term benefits for retention and future educational attainment (Hidi & Renninger, 2006). It is therefore hypothesized that exploratory learning also influences perceptions of interest and enjoyment—or situational intrinsic motivation, as they are sometimes called (Hidi & Renninger, 2006; Niemiec & Ryan, 2009).

1.2 Collaboration During Exploratory Learning

One feature of many exploratory learning studies is collaboration among students, an instructional method that may have its own cognitive benefits (Johnson, Johnson, & Smith, 1991; Prince, 2004; Smith, Sheppard, Johnson, & Johnson, 2005; Webb, 1991). Student collaboration involves verbal communication and exchange of ideas, which may encourage students to restructure their own ideas, potentially improving their understanding. As students elaborate or explain their ideas to peers, their own thinking may be clarified and reinforced (Webb, 1991). Sharing multiple viewpoints can also correct misconceptions held by individual students, potentially leading to better group solutions. Students are more likely to benefit this way in areas such as STEM, where systematic principles or logic (e.g., formulas) can be used to derive accurate solutions and communicate ideas to group members (Laughlin, Zander, Knievel, & Tan, 2003). Collaborative learning can therefore be beneficial for STEM learning (Pai, Sears, & Maeda, 2015; Springer, Stanne, & Donovan, 1999). Given these benefits, any positive effect of exploratory learning done in groups could be partly due to the collaborative nature of the activity (Mazziotti, Loibl, & Rummel, 2014, 2015).

Whether exploratory learning is further enhanced by collaborative learning has received little attention. Both individual and collaborative methods have been used in previous exploratory learning studies, but the few studies comparing these methods have led to inconclusive results (Loibl et al., 2016; Mazziotti et al., 2014, 2015). For example, Sears (2006) found no added benefit of collaboration on exploratory learning for comprehension, but did find benefits for transfer. However, Sears also noted that important methodological problems may have limited the findings. Wirkala and Kuhn (2011) systematically examined the impact of collaboration using a similar instructional method, problem-based learning. Sixth grade students in a social studies course explored one problem in teams and then a different problem individually. Learning measures did not differ as a function of collaboration.

Indeed, use of collaborative learning as a more general instructional method does not always improve learning (e.g., Yetter et al., 2006). The benefits of group collaboration depend on many factors. Individuals in groups are more likely to improve their own conceptual understanding, and generate superior solutions as a group, when problems are well-defined, having a clear, objective answer that can be derived from systematic principles, like those commonly encountered in non-theoretical STEM courses (Laughlin et al., 2003; Slavin, 1996). Benefits are more likely when there is at least one group member with high ability, to guide others' thinking (e.g., Wiedmann, Leach, Rummel, & Wiley, 2012, 2015). Incentive structures that reward group members individually for their work reduce commitment to the group, whereas incentive structures that hold the group accountable for the success of its members result in improved learning outcomes (Slavin, 1996). Group performance is also affected by factors such as motivation (Wittenbaum Hollingshead, & Botero, 2004), leadership (Peterson, 1997), group norms (Postmes, Spears, & Cihangir, 2001), and social status of group members (Hollingshead, 1996). Potential pitfalls of collaborative learning include competitive learning environments (Slavin, 1996), social loafing (Latané, Williams, & Harkins, 1978), boredom of higher-skilled students (Aronson, 2000), and – in groups that do not function well – increased cognitive load imposed by having to organize and communicate within the group (Sweller et al., 2007).

Even if collaborating during exploratory learning does not directly increase learning outcomes, groups may still be a good structure in which to organize exploratory learning. For instance, cooperating with peers during learning at the undergraduate level tends to promote positive attitudes towards learing and the subject area, which could have lasting benefits to a student's overall academic progression (Johnson, Johnson, & Smith, 2014).

1.3 Current Studies

The benefits of exploratory learning prior to instruction have been demonstrated for STEM students ranging from second through ninth grade (e.g., DeCaro & Rittle-Johnson, 2012; Kapur, 2010, 2014; Loehr, Fyfe, & Rittle-Johnson, 2014; Schwartz & Martin, 2004; Schwartz et al., 2011). There is increasing interest in teaching professional skills to STEM students at the college level, and these involve many of the same skills that are often utilized in exploratory learning (e.g., critical thinking, problem solving, and collaboration; ABET, 2016). However, the benefits of exploration have not been as fully tested in STEM courses at the university level (but see Pease & Kuhn, 2011, for problem-based learning). There are particular challenges that might complicate exploratory learning in STEM college classes, such as large class sizes, physical limitations of lecture halls (e.g., fixed inclined seating oriented to lecturn), and students' and instructors' perceptions that lecture is the most expedient method for imparting instruction. It is important to rigorously test the potential advantages of exploratory learning at the college level, because exploratory learning may offer benefits for conceptual understanding, compared to more traditional lecture (Schwartz, Bransford, & Sears, 2005; Schwartz et al., 2009).

The current studies examined the benefits of exploratory learning for undergraduate students enrolled in introductory physics courses at a large urban public university. Study 1 examined whether exploratory learning prior to direct instruction improves conceptual understanding. Study 2 provided a replication of this study in a different semester, with two different physics courses and instructors of record. In both Studies 1 and 2, students completed a collaborative learning activity either before lecture (*explore-first condition*) or after (*instruct-first condition*). We hypothesized that students in the explore-first condition would struggle more initially during exploration but develop comparatively better conceptual knowledge. Conceptual knowledge was measured by the ability to consolidate and differentiate the components of a target principle on a quiz developed by the instructor (e.g., Rittle-Johnson, Siegler, & Alibali, 2001). Consistent with prior studies (e.g., DeCaro & Rittle-Johnson, 2012; Kapur, 2014; Schwartz et al., 2009, 2011), we did not expect experimental condition to impact the use of taught procedures to solve problems (i.e., procedural knowledge; Rittle-Johnson & Koedinger, 2009; Rittle-Johnson et al., 2001).

We also examined self-reported interest and enjoyment (situational intrinsic motivation) following the lesson (i.e., the lecture and activity). Even though students in the explore-first condition were expected to struggle more (e.g., Kapur, 2008), they should also experience a greater sense of discovery, personal direction, and complexity. Therefore, we anticipated that exploratory learning would be more engaging and intrinsically motivating (Glogger-Frey et al., 2015; Vansteenkiste, Niemiec, & Soenens, 2010). Thus, we expected self-reported interest and enjoyment in the explore-first condition to be either equal to or higher than the instruct-first condition.

Study 3 examined whether collaboration drives the benefits of exploratory learning. We hypothesized that conceptual knowledge would be comparable or better when exploring collaboratively compared to exploring individually, given the mixed results of group collaboration demonstrated in prior work (Mazziotti et al., 2014, 2015; Sears, 2006). We also examined whether interest and enjoyment would be enhanced by exploring collaboratively. Because retention in large gateway courses is an ongoing problem in STEM fields (Chen, 2013), motivational benefits may have real value above and beyond achievement.

2. Study 1

2.1 Method

2.1.1 Participants

Participants were undergraduate students (N = 209; M_{age} = 19.33, SD = 2.18; 20.1% female [18.2% unreported], 68.9% first- and second-year students), enrolled in one section of an *Introductory Electricity, Magnetism, and Light (Intro)* course, who completed both the activity and the quiz. This is the second course taken in a calculus-based physics sequence; students were primarily physics and engineering majors. Students were randomly assigned to one of two conditions: instruct-first (n = 119) or explore-first (n = 90). Sample size between conditions was unequal due to capacity limits in one classroom. Gender and class level did not differ by condition, ps > .777.

2.1.2 Materials

Lecture. The lecture covered the concept of *electric potential.* First, electric potential was defined as the work done per unit charge against the electric field to move a charged particle from point A to point B. Then, the instructor explained the relation between electric potential and electric potential energy. Next, students were presented with the equation for calculating the change in the electric potential in moving from point A to point B in the presence of a charged particle, and the contingencies for whether the potential is increasing or decreasing when moving from one point to another. Then, the relation between electric potential and the related concept of electric potential energy was revisited, to explain how different charges move when they are released within an electric field. The instructor then described procedures for calculating electric potential at a single point in the presence of multiple charged particles and discussed common errors made by students when performing that calculation.

The lecture was presented using animated PowerPoint slides. Though not formally scripted, each slide provided a detailed structure for the material presented during the lecture. In addition, in order to cover all of the material presented in the slides during the time allotted for the lecture, the instructor needed to strictly adhere to the slides. Thus, though the two lectures were not administered verbatim from a script, they were instructionally equivalent. In addition, the instructor had given this lecture numerous times before, decreasing the likelihood of practice effects between the two lectures.

Activity. Students in both conditions completed the learning activity together in randomly-assigned, two- to four-person groups. As part of the general structure of the course, students had received tips for working with groups from the instructor on the first day of class, and group work was a routine part of the class format. Expectations for group work were addressed specifically in the course syllabus and emphasized that group work should be seen as an opportunity to check understanding and ask questions about points of confusion. Students were given an additional set of "tips for exploration," which encouraged them to engage with the activity without focusing on whether or not they were getting the correct answer.

The activity required students to determine whether the electric potential at a point in space is positive, negative, or zero, and how the electric potential changes in moving from one point to another. Students were given two diagrams, each containing a pair of charged particles on a grid (Figure 1). The amount of charge on both particles, the placement of the charged particles on the grid, and the four points in space were the same in both diagrams. In one scenario, one of the particles was positively charged and the other was negatively charged. In the other scenario, both particles were positively charged. Students were asked to determine whether the *electrical potential difference* between three pairs of points was positive, negative, or zero. Additionally, students were asked if the electric potential at a given point was positive, negative, or zero. The activity builds upon several exercises presented by O'Kuma, Maloney, and Hieggelke (2000) and is a standard lesson in physics courses in the United States. The activity was designed to meet several instructional goals: (a) to activate students' prior understanding of electric potential energy and to differentiate it from electric potential; (b) to address a common student misconception that a point in space is a charge; and (c) to highlight the relationship between potential and magnitude of the field.

Each group was given a single worksheet to use. It was emphasized to students in both conditions that they should use the worksheet as a place to document and refine their thinking about the activity. Students wrote their names onto the group's worksheet, which was turned it in at the end of the class period.

Interest and Enjoyment Questionnaire. Following the entire lesson (lecture and activity, in either order), students individually completed a questionnaire. Three items, measuring interest and enjoyment (situational intrinsic motivation), were relevant to the current research ("I found this learning activity interesting; I enjoyed this learning activity; This learning activity was boring" [reverse-coded]; Ryan, 1982; Coefficient α = .82). Students responded to these items on a 5-point Likert scale (*1= strongly disagree to 5 = strongly agree*). Responses were averaged.

Quiz. The quiz assessed both procedural and conceptual knowledge. These items were developed by the course instructor.

Procedural Knowledge. Seven multiple-choice questions targeted *procedural knowledge* directly taught in the lecture (e.g., *Is the electric potential at point 3 positive, negative, or zero?; The electric potential difference in moving from point 2 to point 1 is* ___.; Cronbach's α = .63).[1]

Conceptual Knowledge. Four multiple choice questions targeted students' *conceptual knowledge*. These items required students to apply procedures used to calculate electric potential to demonstrate understanding of the concept of electric potential energy (e.g., *If you were to move a* ***positive charge*** *from point 5 to point 4 its electric potential energy would____; A* ***negative charge*** *at point 4 would have ____ electric potential energy;* Cronbach's α = .20. Students typically have difficulty in both differentiating between electric potential and electric potential energy and in describing how they relate to each other. Quizzes were scored for accuracy. A second rater independently coded 20% of the open-ended quiz items (kappa = 79%).

2.1.3 Procedure

The study was conducted during one seventy-five minute class period.[2] On the day prior to the study, the instructor lectured on the topic of *electric potential energy* due to a collection of point charges. Conceptually, electric potential energy is the work that must be done against an electric force to move a charge within an electric field. Students were taught how to calculate the electric potential energy stored in the interaction between charges

[1] Students were asked to explain their answers to the first two multiple choice questions, but these explanations were not factored into the score; the findings are the same if these explanations are used.

[2] Students in Study 1 had completed Study 3 four weeks prior. These studies were planned to test different hypotheses and are described out of chronological order for presentation purposes.

and how to determine the change in the electric potential energy of a charge as it is moved in the presence of other charges. The related concept of *electric potential* was not discussed.

On the day of the study, the instructor's regular lesson was modified to vary order of instruction. Students were randomly assigned to either begin with lecture in their regular classroom (*instruct-first* condition), or to begin with the activity in another classroom (*explore-first* condition). To manage these two different lesson orders, the instructor first went to the explore-first classroom and introduced the activity (shown in Figure 1). Afterward, the instructor went to the instruct-first classroom, lectured, and introduced the activity. Then, the instructor returned to the explore-first classroom and lectured. The activity and lecture each lasted approximately fifteen minutes and were the same between conditions. Experimenters were present in both rooms. Finally, both groups rejoined in their regular classroom and completed the interest and enjoyment questionnaire. Following the questionnaire, the instructor explained the correct answers to the activity, as well as common incorrect answers, and gave students an opportunity to ask questions. Students were then given fifteen minutes to complete the quiz, which served as the primary outcome measure.

2.2 Results and Discussion

Although we expected that students in the explore-first condition would perform at a lower level during the exploratory learning activity (e.g., Kapur, 2010, 2011, 2012; Kapur & Rummel, 2012), we expected their quiz scores to be better than in the instruct-first condition. We also expected equal or higher self-reported interest and enjoyment. The benefits of exploratory learning are most often found for measures of conceptual knowledge, rather than for problem-solving procedures (cf. Schwartz et al., 2009). Therefore, we specifically hypothesized that exploratory learning would benefit performance on the conceptual knowledge questions. We predicted that students in both conditions would perform similarly on the procedural knowledge questions.

2.2.1 Activity

Groups in the explore-first condition scored significantly lower on the learning activity (n = 23; M = 3.09 out of 8, SE = 0.42) than groups in the instruct-first condition (n = 32; M = 6.16, SE = 0.35), $F(1, 53) = 31.81$, $p < .001$, $\eta_p^2 = .38$. As expected, completing the activity as an exploratory learning process was more difficult than completing it as practice after instruction.

2.2.2 Quiz

There was no significant correlated error caused by individual membership in groups (ICCs < = 0.02, *p*s > .712); therefore, we analyzed quiz scores individually. Quiz scores were analyzed using a 2 (condition: explore-first, instruct-first) × 2 (question type: procedural, conceptual) mixed-factorial ANOVA, with condition between-subjects and question type within-subjects. No main effect of condition was found, $F(1, 207) = 1.68$, $p = .196$, $\eta_p^2 = .01$. A main effect of question type was found, with procedural questions (*M* = 89.72%, *SE* = 1.17) answered more accurately than conceptual questions (*M* = 72.58%, *SE* = 1.69), $F(1, 207) = 100.74$, $p < .001$, $\eta_p^2 = .33$. This effect was qualified by a significant condition × question type interaction, $F(1, 207) = 4.08$, $p = .045$, $\eta_p^2 = .02$. As shown in Figure 2 and Table 1, no difference was found for procedural questions between students in the explore-first (*M* = 89.52%, *SE* = 1.77, CI: 86.03-93.02) and instruct-first conditions (*M* = 89.92%, *SE* = 1.54, CI: 86.88-92.95), *d* = 0.02). However, students in the explore-first condition (*M* = 75.83%, *SE* = 2.55, CI: 70.81-80.86) scored higher on the conceptual questions than students in the instruct-first condition (*M* = 69.33%, *SE* = 2.22, CI: 64.96-73.70), *d* = 0.27).

2.2.3 Interest and Enjoyment Questionnaire

No difference was found between conditions on the measure of self-reported interest and enjoyment (instruct-first *M* = 3.31, *SE* = 0.07, CI: 3.18-3.44; explore-first *M* = 3.20, *SE* = 0.08, CI: 3.05-3.36), $F(1,196) = 1.03$, $p = .312$, $\eta_p^2 = .01$. This finding is consistent with the hypothesis that interest and enjoyment would be equal or better between conditions, even though the learning activity was more difficult for students who explored.

Thus, in Study 1, exploring a challenging physics topic prior to direct instruction led to worse performance on the learning activity, but ultimately benefitted conceptual understanding. However, these results are limited to one course in one semester. Study 2 was conducted to replicate and extend these findings to two additional courses in another semester, with different instructors of record and sample characteristics.

3. Study 2

3.1 Method

3.1.1 Participants

Participants (*N* = 153; 46% female) were undergraduate students enrolled in the second semester of first-year physics courses who completed both the activity and quiz used for this study. Students participated in two different physics courses. Seventy participants were enrolled in *Intro,* taught by a different professor and in a different style than in Study 1. As in Study 1, this class was predominantly male (24% female) and composed of first- and second-year physics and engineering majors (72% first- and second-year students). Eighty-three participants were enrolled in *Fundamentals of Physics II (Fundamentals),* which is the second course in an algebra-based physics sequence. This course was taught by the same instructor from Study 1. This class was predominantly female (63%) and composed of primarily third- and fourth-year and post baccalaureate students working towards degrees with a pre-professional health science focus (16% first- and second-year students).[3] Gender and class level did not differ by condition, $ps > .499$.

3.1.2 Procedure

Students participated during their regular instructional period (*Fundamentals*) or recitation section (*Intro*). As in Study 1, all students learned about electric potential energy prior to the study. Study activities and lecture were delivered by the Study 1 instructor and were the same as in Study 1, except that students self-selected into groups of four to five upon arrival to class. Hence, students were not randomly assigned to their collaborative groups. Students were randomly assigned to explore-first (*Intro* $n = 35$; *Fundamentals* $n = 43$) or instruct-first (*Intro* $n = 35$; *Fundamentals* $n = 40$) conditions. After the lecture and activity were completed, the class rejoined in the same classroom and were given a survey to complete individually, which included the interest and enjoyment questions. Then, the instructor explained correct answers and common incorrect answers on the activity, and students were given the opportunity to ask questions. The quiz was administered during the last fifteen minutes of the 75-minute *Fundamentals* class and at the beginning of the next day in the 50-minute *Intro* class. The quiz was identical to that of Study 1, except that two questions were added to the *conceptual knowledge* scale to increase scale reliability (conceptual knowledge: Cronbach's $\alpha = .47$; procedural knowledge: Cronbach's $\alpha = .61$). Two raters independently coded 20% of the quiz items requiring an explanation (kappa = 70%).

3.2 Results and Discussion

Predictions were the same as in Study 1. Preliminary analyses revealed no differences in outcome measures as a function of condition between the two courses (i.e., no condition × course interactions), so data from the two courses were combined for analyses.

3.2.1. Activity

[3] Gender was self-reported and was not provided by nine participants (6%). Class level was self-reported and was not provided by 20 participants (13%).

Forty-one groups completed the learning activity (instruct-first condition, $n = 21$, $M = 4.43$, $SE = 0.55$; explore-first condition, $n = 20$, $M = 3.25$, $SE = 0.56$). Group accuracy on the activity did not differ by condition, $F(1, 39) = 2.27$, $p = .140$, $\eta_p^2 = .06$.

3.2.2 Quiz

There was no significant correlated error caused by individual membership in groups (ICCs < = 0.04, *p*s > .309). Therefore, we analyzed quiz scores individually. A 2 (condition) × 2 (question type) mixed-factorial ANOVA revealed no main effect of condition, $F < 1$. A main effect of question type was found, $F(1, 151) = 320.35$, $p < .001$, $\eta_p^2 = .68$, with procedural questions ($M = 81.73\%$, $SE = 1.67$, CI: 78.42-85.04) answered more accurately than conceptual questions ($M = 42.74\%$, $SE = 1.90$, CI: 38.98-46.49). This effect was qualified by a significant condition × question type interaction, $F(1, 151) = 5.56$, $p = .020$, $\eta_p^2 = .04$ (Figure 3; Table 1). Again, no difference was found for procedural question accuracy between the instruct-first ($M = 83.24\%$, $SE = 2.39$, CI: 78.51-87.96) and explore-first ($M = 80.22\%$, $SE = 2.34$, CI: 75.59-84.85) conditions, $d = 0.15$. However, as in Study 1, students in the explore-first condition scored higher on conceptual questions ($M = 46.37\%$, $SE = 2.66$, CI: 41.11-51.63) than students in the instruct-first condition ($M = 39.11\%$, $SE = 2.71$, CI: 33.75-44.47), $d = 0.31$.

3.2.3 Interest and Enjoyment Questionnaire

Students in the explore-first condition ($M = 3.54$, $SE = 0.10$, CI: 3.54-3.75) reported significantly higher interest and enjoyment than students in the instruct-first condition ($M = 2.75$, $SE = 0.11$, CI: 2.53-2.97), $F(1,141) = 27.29$, $p < .001$, $\eta_p^2 = .16$.[4]

In summary, students who explored a physics concept before receiving instruction demonstrated better conceptual understanding and equal procedural knowledge compared to students who received direction instruction first. These findings replicate Study1, using two additional classes with different participant characteristics.

Performance on the learning activity did not differ based on condition. However, after the entire lesson (both activity and lecture), students who explored first reported greater interest and enjoyment for the learning activity, indicating a potential motivational benefit of using the activity as an exploration opportunity. These two findings differ from Study 1, where students in the explore-first condition performed worse on the learning activity, but reported equal levels of intrinsic motivation. Importantly, across both studies, regardless of patterns of performance on the activity or self-reported interest and enjoyment, conceptual learning remained higher in the

[4] Ten surveys were missing.

explore-first condition. Thus, the benefits of exploring may not depend exclusively on activity failure or success, or interest and enjoyment in the process (cf. Loibl et al., 2016).

In both Studies 1 and 2, students explored in groups. It is possible that exploring with others reduces the difficulty of exploration, by promoting greater elaboration and clarification of ideas and increasing the diversity of solution approaches (Johnson et al., 1991; Smith, Johnson, & Johnson, 1981; Webb, 1991). Exploring in a group context may also increase interest and enjoyment (Rotgans & Schmidt, 2011). In Study 3, we investigated whether the benefits of exploring in a collaborative setting exceed that of exploring individually—both in terms of learning and motivational outcomes.

4. Study 3

Both Studies 1 and 2 demonstrated that conceptual understanding increased in an explore-first relative to an instruct-first condition. Study 3 did not seek to replicate this effect further. Rather, Study 3 investigated a specific feature of the previous explore-first method more closely, comparing collaborative exploration to individual exploration. If exploring collaboratively versus individually leads to the same learning outcomes (e.g., Sears, 2006), then the benefits of exploratory learning are most likely due to exploration itself, not the group collaboration. However, if collaborative exploration improves learning and performance beyond individual exploration, then the results would suggest that instructors should consider implementing exploratory learning activities as group work in order to maximize the potential learning benefits.

4.1 Method

4.1.1 Participants

Participants were the same as in Study 1, and completed the study as part of their *Intro* course. However, because this study was conducted on a different day, the *N* differs slightly (N = 208; M_{age} = 19.00, SD = 2.18, 25% female). Students were randomly assigned to one of two conditions: individual exploration (n = 126) or collaborative exploration (n = 82). Sample size between conditions was unequal due to a capacity limit in one of the classrooms.

4.1.2 Materials

Activity. Students were asked to rank the magnitude of net gravitational force on four point masses presented in different arrangements on four grids representing empty space (see Figure 4). In each case, the force exerted on the lettered mass by the particle labeled as mass 1 is the same. The goal of the activity is for students to recognize the differences in the force exerted on the lettered mass by mass 2 in each case and understand how those differences result in a different net force when added to the force from mass 1. Although students had not yet received instruction on the Universal Law of Gravitation, they only needed to know that the gravitational force is attractive and how to add vectors to come up with a ranking, which they had learned in a prior, prerequisite course.

Students worked individually (*individual exploration condition*) or collaboratively (*collaborative exploration condition*) on the activity. In both conditions, students were given a worksheet to document their thinking about the activity. The worksheet contained four tasks for the students to complete: (a) rank the four scenarios from greatest to least in terms of the net gravitational force on the lettered mass; (b) explain the reasoning used to determine their rankings; (c) express level of confidence in their rankings on a scale of 1 (“basically guessed”) to 10 (“very sure”); and (d) list additional factors they might need to consider in determining their rankings. Students in both conditions received “tips for exploration.” Students were informed that they (a) should rely on their past knowledge to help them in the task, (b) did not need to have the right answer, but should instead try to explain the concepts, and (c) should ask specific questions about what they did not understand. As students had not been asked to work in groups prior to this activity, students in the collaborative condition also received a set of tips for how to make their collaborative efforts productive.

Lecture. The lecture covered the concepts underlying the Universal Law of Gravitation, which describes the gravitational force between two point masses. Students were given the equation used to calculate the magnitude of the gravitational force between point particles and a discussion of its direction. A significant portion of the instruction reviewed the process of vector summation, which is needed for calculating the net gravitational force exerted on a particle by a collection of point masses.

Interest and Enjoyment Questionnaire. After the activity, all students completed the same questionnaire as in Studies 1 and 2.

Quiz. On the day after the activity and lecture, all students took a quiz comprised of twelve questions assessing procedural understanding (13 points total; Cronbach’s $\alpha = .81$) and three questions assessing conceptual understanding (Cronbach’s $\alpha = .39$). Procedural knowledge items asked students to perform procedures that were explicitly taught during the lesson: (a) draw the free-body diagram for each of four masses, (b) explain the reason for drawing one of the free-body diagrams with the method used, (c) compare individual forces due to differences in distance between particles, and (d) compare net forces (vector addition). Conceptual knowledge items required students to consolidate what they learned by asking them to compare net forces due to differences in mass between particles.

4.1.3 Procedure

The instructor's regular lesson on gravitational force was modified to vary the structure of the class. All students began with an exploratory activity. Immediately prior to the activity, the instructor gave the three "tips for effective exploration." Students in the *individual exploration* condition worked alone in their regular classroom, whereas students in the *collaborative exploration* condition were randomly assigned to work in groups of three to four in another classroom. The instructor first introduced the activity to students in the collaborative exploration condition, then to students in the individual exploration condition. Experimenters were present in both rooms to answer clarification questions as students worked, and the instructor circulated between the two rooms.

After the activity, students completed the same interest and enjoyment questionnaire as in Study 1 and then rejoined in their regular classroom, where the instructor lectured on gravitational force. The activity and lecture each lasted approximately 15 minutes. After the lecture, the instructor gave the correct answers to the activity and discussed their explanations. At the next day's class, all students took an unannounced quiz on the topic.

4.2 Results and Discussion

We tested whether performance on the activity, quiz, and interest and enjoyment questionnaire differed between the collaborative exploration and individual exploration conditions. The nature of the performance differences between conditions (equal or better when working in groups) would indicate the extent to which collaboration enhances learning benefits derived from exploration. The effect of condition on interest and enjoyment was an empirical question, as we were unaware of prior research specifically testing this question.

4.2.1 Activity

Groups exploring collaboratively scored significantly higher on the learning activity (n = 25, M = 1.52 out of 2 points, SE = 0.13, CI: 1.26-1.78) compared to students exploring individually (n = 126, M = 0.91 out of 2 points, SE = 0.06, CI: 0.79-1.02), $F(1, 149) = 19.26$, $p < .001$, $\eta_p^2 = .11$.

4.2.2 Quiz

A 2 (condition) × 2 (question type) mixed-factorial ANOVA revealed a main effect of question type, $F(1, 206) = 34.50$, $p < .001$, $\eta_p^2 = .14$. Procedural questions (M = 75.13%, SE = 1.62%, CI: 71.94-78.32%) were answered more accurately than conceptual questions (M = 61.69%, SE = 2.27%, CI: 57.22-66.15%). No main effect of condition or condition × question type interaction was found, *F*s < 1 (see Figure 5; Table 2). Thus, performance was equal between conditions, suggesting that collaborative exploration provided no advantage compared to individual exploration on assessments of either procedural or conceptual knowledge.

4.2.3 Interest and Enjoyment Questionnaire

Self-reported interest and enjoyment for the learning activity was greater for students in the collaborative-exploration condition (M = 3.65, SE = 0.08, CI: 3.49-3.81) than for students in the individual-exploration condition (M = 3.44, SE = 0.07, CI: 3.31-3.57), $F(1, 206) = 4.08$, $p = .045$, $\eta_p^2 = 0.02$.

Though students exploring collaboratively, on average, scored at a higher level on the learning activity and expressed higher interest and enjoyment, they did not achieve higher quiz scores than students exploring individually. Thus, articulating the concepts aloud to others in an exploratory learning setting did not strengthen students' conceptual understanding more than exploring individually. Several explanations for these findings are possible. The first possibility is simply a measurement issue. Reliability of the conceptual knowledge measure was relatively low (Cronbach's $\alpha = .39$), which may have limited our ability to find group differences on this measure. Further, to the extent that collaborative learning might have been beneficial, these benefits may be reduced in an exploratory learning situation. For example, students with higher knowledge may have answered the question without sharing their explanation with their group members, so when assessed individually later, the other group members did not benefit fully from the collaboration and did not know the answer (Schwartz, 1995). Finally, although collaborative performance on the activity was higher than individual performance, collaboration may have inflated how much some group members thought they knew.

5. General Discussion

5.1 Exploring Prior to Lecture

In STEM classes, a primary objective is to help students identify general principles that govern the physical world. However, there is a concern that students may tend to mindlessly "plug and chug" formulas rather than understand the underlying concepts (Catrambone, 1998; Richland, Stigler, & Holyoak, 2012). In Studies 1 and 2, we investigated whether exploring a topic before receiving direct instruction (lecture) helped to promote such conceptual understanding. Groups who explored prior to instruction performed a problem-solving activity less accurately (Study 1), or as accurately (Study 2), as groups that attended a lecture and then practiced the problems. As expected, exploratory learning did not impact rote use of procedures. However, exploratory learning improved conceptual knowledge in both studies. Learners' self-reported interest and enjoyment (situational intrinsic motivation) towards the learning activity was equally strong in both conditions (Study 1), or stronger in the explore-first condition (Study 2). Thus, students in the explore-first condition demonstrated as least as much interest and enjoyment for the learning activity, even though they likely struggled more.

These findings demonstrate that exploring prior to lecture benefits conceptual understanding. The opportunity to activate prior knowledge, identify their own knowledge gaps, generate questions, and test alternative strategies, in a relatively self-directed manner, likely made exploratory learning a desirable difficulty and sparked a "need to know," which prepared students to learn at a deeper conceptual level (Bjork, 1994; DeCaro & Rittle-Johnson, 2012; Glogger-Frey et al., 2015; Wise & O'Neill, 2009). Indeed, students who explored novel problems before being taught the procedure to solve them were better at connecting the procedure (calculation of electric potential) to the overarching concept (the relation between electric potential and electric potential energy) on a later quiz. This conceptual integration may have occurred because, without access to the formulas typically provided in prior instruction, students explored the problem space more fully, relying on their previous knowledge of relevant concepts instead (Kapur & Bielaczyc, 2012; Loibl et al., 2016; Schwartz & Bransford, 1998; Schwartz & Martin, 2004). Specifically, students in the explore-first condition may have activated and applied prior knowledge of electric potential energy, which was learned in a previous lesson, to solve these novel problems about electric potential. By doing so, they may have been better prepared to recognize and differentiate novel concepts, improving their ability to encode and integrate these concepts with their prior knowledge, once encountered in lecture (Schwartz et al., 2007).

These findings run counter to more standard arguments based on cognitive load theory that state that direct instruction should lead to more optimal learning outcomes (e.g., Kirschner et al., 2006). However, recently, Kalyuga and Singh (2016) have offered a revised view of cognitive load theory that acknowledges the benefits of exploratory learning for certain goals. Specifically, Kalyuga and Singh state that direct instruction will offer the most efficient transmission of domain-specific knowledge, such as facts and procedures. But exploratory learning methods are useful for other goals, such as activating prior knowledge, increasing awareness of the problem situation and knowledge gaps, increasing attention to deeper structures, or motivating and engaging students (see also Kapur, 2016). The current studies align with this perspective, demonstrating similarities in procedural knowledge between conditions, but conceptual (and in Study 2, motivational) benefits to exploring first.

These studies contribute to the limited literature examining the impact of exploratory learning combined with direct instruction in STEM college classrooms. Exploratory learning improved conceptual understanding much like in prior work with younger students, despite potential differences in classroom environments (e.g., large class sizes, fixed seating in lecture halls, lecture as the standard instructional method). In addition, these studies add empirical strength to the overall literature on exploratory learning. As noted by Hsu and colleagues (2015), many related studies lack rigorous experimental control, because they change too many variables at once (e.g., both task type and activity order; see also Loibl et al., 2016; Sweller, 2009). To address this problem, we used the same learning activities in Studies 1 and 2 and only manipulated the order in which lecture and problem-solving were presented. With these tighter experimental controls, we found a causal benefit of exploring problems prior to lecture, replicating and extending prior research.

5.2 Exploring Collaboratively

In Studies 1 and 2, all students completed the activity collaboratively in groups, either as exploration or practice. In Study 3, we examined collaboration more closely, because it is often used in studies testing the effects of exploratory learning. We questioned whether the observed learning benefits are driven by potential benefits of collaboration, rather than due to exploration itself. Students in both conditions completed an exploratory learning activity before lecture, but some students explored collaboratively while others explored individually. No differences in learning outcomes were found between conditions. However, students who explored collaboratively had higher exploratory learning activity scores and rated their interest and enjoyment for the activity as higher than students who explored individually.

These findings suggest that collaborative learning was not an essential component of exploratory learning for learning outcomes in our instructional environment. Hence, use of either individual or group work during exploration appears to lead to comparable results (see also Mazziotti et al., 2014, 2015; Sears, 2006). The relative effects of exploring collaboratively likely depend on the situation. Factors associated with collaboration, such as task difficulty and type, motivation, and group/class dynamics and norms, are known to impact collaborative learning (Laughlin et al. 2003; Niemiec & Ryan, 2009).

It is worth noting the motivational benefit of exploring collaboratively observed in our research. Although exploring collaboratively did not have an added impact on learning in the short term, the improvement in interest and enjoyment posed by collaboration could impact longer-term outcomes. Exploratory learning can be demanding and require persistence in the face of challenge (DeCaro et al., 2015). Interest and enjoyment (situational intrinsic motivation) sustains students throughout their academic career, helping them overcome temporary setbacks and remain in a degree program longer, thereby achieving more (Black & Deci, 2000; Guay, Vallerand, & Blanchard, 2000; Hidi & Renninger, 2006). Thus, if an instructor utilizes collaborative exploration more often, any motivational benefits may become more apparent (Rotgans & Schmidt, 2011). More research is needed to investigate this idea.

5.3 Methodological Considerations

In order to overcome reluctance on the part of instructors who may resist additional burdens placed on their preparation time, instructional methods should be simple enough to implement easily. In addition, the benefit accrued to students must be worth the effort of adopting a new approach (Dees, 1991). In these studies, we performed a minimally intrusive intervention by working with an instructor's existing problem-solving materials and lectures and preferred format for group work. Our findings suggest that the benefits of exploring prior to instruction can be achieved in college STEM classrooms without extensive modifications to teachers' instructional methods, by simply switching the order of lecture and activities.

However, before implementing exploratory learning methods, instructors should be mindful of important design features thought to benefit learning from this method, as described by Kapur (2016; see also Loibl et al., 2016). All were present in our studies. First, at least in Study 1, the problem-solving activity was sufficiently difficult, but was also within students' capabilities. Students in the explore-first condition did significantly worse, but remained interested in and enjoyed the experience as much as their peers in the instruct-first condition. Second, students searched the problem space for solutions. Third, students needed to leverage prior knowledge to solve the problem. Importantly, students in all of the current studies had received instruction on similar topics prior to the exploratory learning activity, so they had some basis for exploration and further knowledge development. Finally, at the end of each session (i.e., following both the activity and lecture), the instructor explained the canonical solution and common student errors, although students were not asked to contrast these directly. The current studies were designed to investigate exploratory learning as a method to facilitate student achievement in STEM undergraduate classrooms, so these factors were not systematically tested. However, they warrant increased attention in future research.

Importantly, it can be assumed that for most students, the challenge posed by exploration was manageable, given their prior knowledge. Hence, the degree of challenge did not overwhelm task interest and enjoyment. These findings imply that instructors need to carefully consider students' prior knowledge and individual capacity when planning the difficulty of exploration (DeCaro et al., 2015; Kapur, 2016; Loibl et al., 2016). Otherwise, the constellation of factors that seem to have contributed to beneficial learning may become imbalanced. Tasks that are overwhelmingly difficult are rarely interesting, or sustaining. By the same token, tasks that are too easy, or routine, are hardly interesting or conducive to development (Silvia, 2008).

The conceptual benefits of exploratory learning may have been more evident in the current study because the quiz was designed to assess both procedural and conceptual knowledge. Previous research suggests that benefits of exploration are primarily for conceptual knowledge and transfer (Loibl et al., 2016; Schwartz et al., 2009). Furthermore, students in our studies learned correct problem-solving procedures regardless of educational format. Receiving instruction on the procedure, as students did in both conditions, was sufficient to correctly apply it. Thus, in order to carefully investigate the benefits of exploratory learning, assessments should ideally capture both procedural knowledge and higher-level conceptual understanding.

5.4 Limitations

Despite this promising evidence for the use of exploratory learning in undergraduate STEM courses, several limitations were present in the current studies. First, the reliability of our quiz items, particularly conceptual knowledge, was low. Low scale reliability increases the difficulty of detecting effects that exist, but does not invalidate the findings (Bacon, 2004; Nunnally, 1978). In addition, we used instructor-developed quizzes. Thus, our assessments had face- and ecological-validity, and were designed with disciplinary degree requirements in mind. But, they have not been validated using conventional test and measurement approaches.

Second, we did not measure long-term retention from the learning session. Our outcome measure (i.e., quiz score) was given either immediately, or at the beginning of the next class period. Therefore, we cannot be certain that the benefits of instructional order would be maintained over time. In addition, students in the explore-first condition could have had better performance simply because they heard the lecture closer in time to the quiz. It is worth noting that the two courses examined in Study 2 showed no significant differences in quiz measures, even though they took the quiz at different time points (i.e., immediately versus the following day). The explore-first condition still received the lecture last in both conditions, but any lecture information would not have been immediately available in working memory for the students who took the quiz on another day.

Finally, further research is needed to extend these findings to other topics, exploratory activities, and STEM disciplines. In addition, more research is needed to fully understand the cognitive mechanisms underlying exploratory learning's benefits. Related to this idea, the nature of the lecture might matter: differences between conditions may be weaker if lecture methods are used that instantiate similar cognitive mechanisms as exploratory learning (e.g., activate students' prior knowledge, increase metacognitive awareness of gaps in students' knowledge, and increase a desire to understand). The current work joins a growing body of literature demonstrating that exploratory learning accomplishes these tasks. Given the simplicity of this method, and its demonstrated ability to improve conceptual understanding from a single lesson, exploratory learning holds great promise for use in undergraduate STEM classes.

Table 1

Mean Quiz Scores (Percent Correct) as a Function of Question Type and Condition in Studies 1 and 2. Standard Deviations in Parentheses.

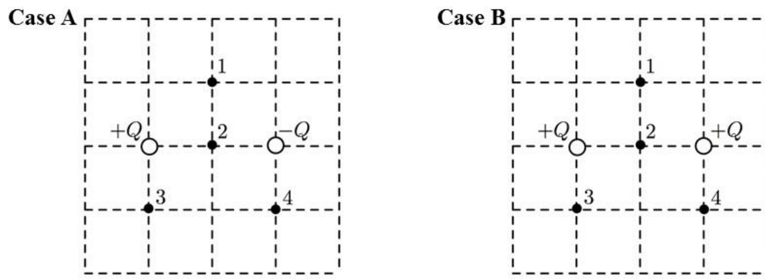
	Condition	
Study 1	**Explore-First**	**Instruct-First**
Procedural Questions	89.52 (16.42)	89.92 (17.08)
Conceptual Questions	75.83 (22.63)	69.33 (25.30)
Study 2		
Procedural Questions	80.22 (20.88)	83.24 (20.52)
Conceptual Questions	46.37 (24.57)	39.11 (22.34)

Table 2

Mean Quiz Scores (Percent Correct) as a Function of Question Type and Condition in Study 3. Standard Deviations in Parentheses.

	Condition	
	Collaborative Exploration	**Individual Exploration**
Procedural Questions	74.34 (23.55)	75.92 (22.30)
Conceptual Questions	59.35 (33.95)	64.02 (30.58)

Two fixed arrangements containing two point charges each are shown in the figure below. In each arrangement, both charges have the same amount of charge and have been placed on a grid at the points where the lines of the grid intersect. In all four cases, the lines of the grid are all separated by a distance d. In Case A, one charge is positive and the other is negative. In case B, both charges are positive. Assume that each arrangement has been placed in a region of empty space, so the only electrical interactions you need to consider are those between the two charges in that particular arrangement.



In each arrangement, determine whether the electrical potential difference between each pair of points below is positive, negative, or zero.

- From 1 to 2
- From 2 to 4
- From 3 to 4

In each arrangement, is the electric potential at point 2 positive, negative, or zero?

Figure 1. Learning activity in Studies 1 and 2.

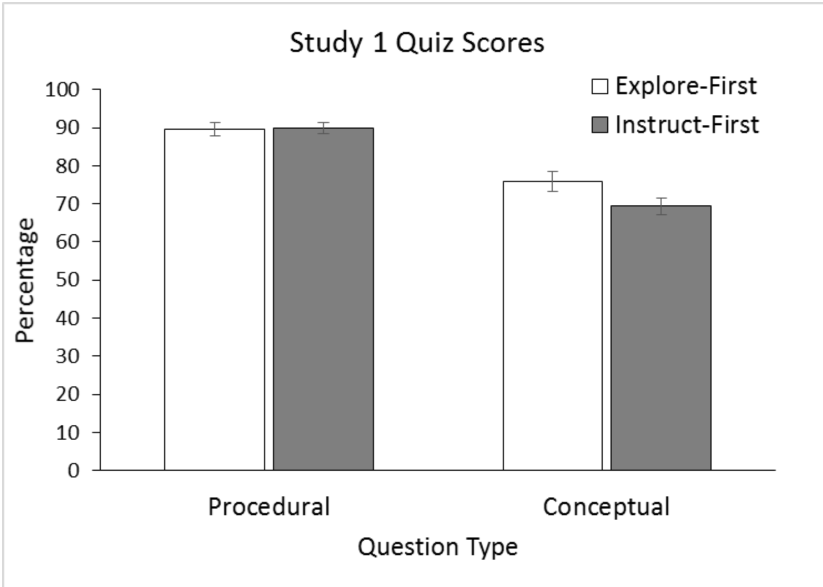


Figure 2. Study 1 mean quiz scores as a function of question type and condition. Error bars represent standard error.

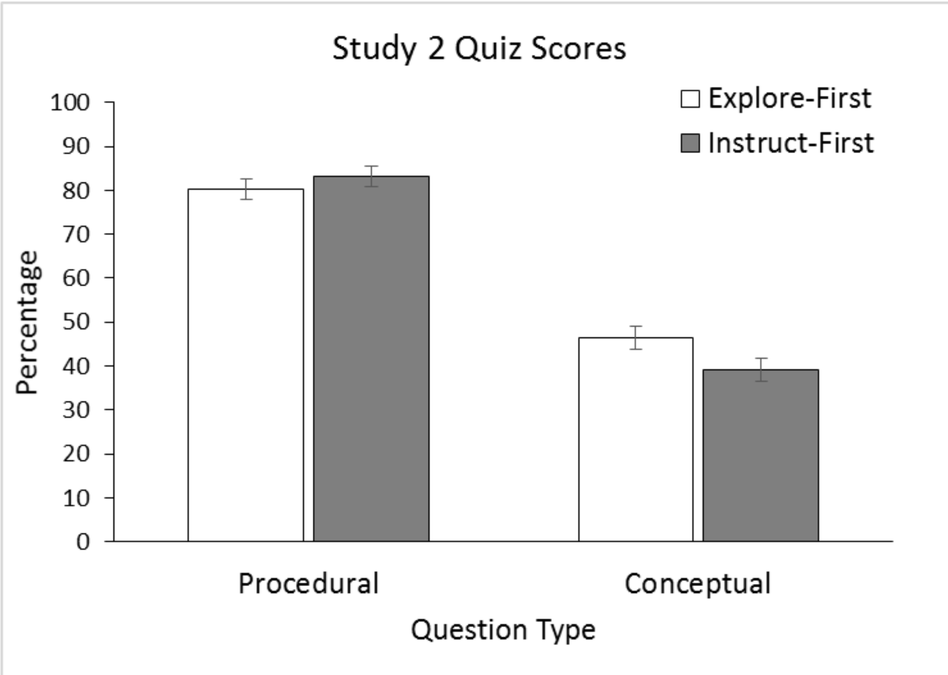
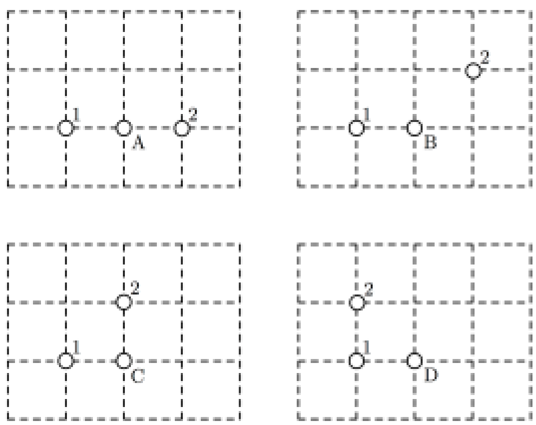


Figure 3. Study 2 mean quiz scores as a function of question type and condition. Error bars represent standard error.

Four arrangements of point masses are shown in the figure below. In each arrangement, the three masses all have a mass *m* and have been placed on a grid at the points where the lines of the grid intersect. In all four cases, the lines of the grid are all separated by a distance *d*. Assume that each grid has been placed in a region of empty space, so the only gravitational interactions you need to consider are those between the three particles in that particular arrangement.



In each arrangement, one of the masses has been labeled with a letter (A, B, C, or D).

Rank the magnitude of the net gravitational force on the lettered mass in each arrangement, from greatest to least. Explain the reasoning you used to determine your rankings. Be as explicit as you can.

Figure 4. Learning activity in Study 3.

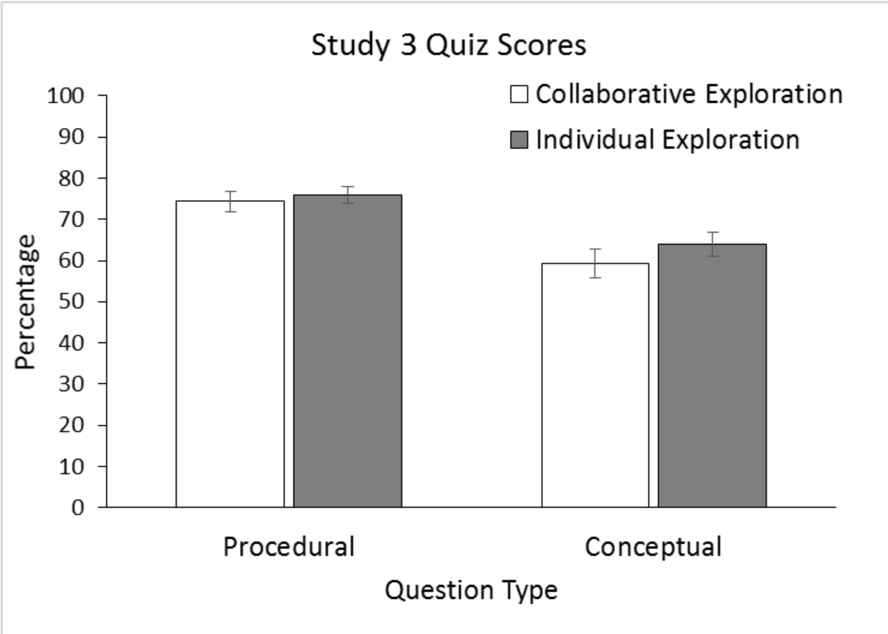


Figure 5. Study 3 mean quiz scores as a function of question type and condition. Error bars represent standard error.

References

ABET (2016). *ABET Accreditation Policy and Procedure Manual (APPM)*. Retrieved from http://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2016-2017/#general Accessed 4/12/17.

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*(1), 1-18, doi:10.1037/a0021017.supp.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Aronson, E. (2000). *The jigsaw classroom.* Retrieved from https://www.jigsaw.org/ Accessed 9/24/17.

Bacon, D. (2004). The contributions of reliability and pretests to effective assessment. *Practical Assessment Research & Evaluation, 9*(3) Retrieved from http://PAREonline.net/getvn.asp?v=9&n=3. Accessed 12/8/13.

Belenky, D. M., & Nokes-Malach, T. J. (2012). Motivation and transfer: The role of mastery-approach goals in preparation for future learning. *Journal of the Learning Sciences*, *21*(3), 399-432.

Black, A. E. & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Science Education*, *84*(6), 740-756.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe and A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. l85-205). Cambridge, MA: MIT Press.

Capon, N., & Kuhn, D. (2004). What's so good about problem-based learning? *Cognition and Instruction*, *22*(1), 61-79, doi:10.1207/s1532690Xci2201_3.

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(6), 1563-1569, doi:10.1037/a0017021

Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General*, *127*, 355-376.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8,* 293-332.

Chen, X. (2013). STEM attrition: College students' paths into and out of STEM fields (NCES 2014-001). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Chin, D. B., Chi, M., & Schwartz, D. L. (2016). A comparison of two methods of active learning in physics: Inventing a general solution versus compare and contrast. *Instructional Science, 44*(2), 177-195. doi: 10.1007/s11251-016-9374-0

DeCaro, D. A., DeCaro, M. S., & Rittle-Johnson, B. (2015). Achievement motivation and knowledge development during exploratory learning. *Learning and Individual Differences, 37*, 13-26. doi:10.1016/j.lindif.2014.10.015

DeCaro, M. S., & Rittle-Johnson, B. (2012). Exploring mathematics problems prepares children to learn from instruction. *Journal of Experimental Child Psychology, 113*(4), 552-568. doi:10.1016/j.jecp.2012.06.009

Dees, R. L. (1991). The role of cooperative learning in increasing problem-solving ability in a college remedial course. *Journal for Research in Mathematics Education*, *22*(5), 409-421. doi:10.2307/749188

Duffy, T. M., & Jonassen, D. H. (1992). Constructivism: New implications for instructional technology. In T. M. Duffy & H. H. Jonassen (Eds.), *Constructivism and the technology of instruction: A conversation* (pp. l-16). Hillsdale, NJ: Erlbaum.

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, *22*(4), 271-280, doi:http://dx.doi.org/10.1016/j.learninstruc.2011.08.003.

Felder, R. M., Woods, D. R., Stice, J. E., & Rugarcia, A. (2000). The future of engineering education II: Teaching methods that work. *Chemical Engineering Education*, *34*(1), 26-39.

Glogger-Frey, I., Fleischer, C., Grüny, L., Kappich, J., & Renkl, A. (2015). Inventing a solution and studying a worked solution prepare differently for learning from direct instruction. *Learning and Instruction, 39*, 72-87, doi:10.1016/j.learninstruc.2015.05.001.

Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The Situational Motivation Scale (SIMS). *Motivation and Emotion, 24(3)*, 175-213.

Hake, R. (1998). Interactive-engagement vs. traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64-90.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111-127.

Hollingshead, A. B. (1996). Information suppression and status persistence in group decision making: The effects of communication media. *Human Communication Research*, *23(2),* 193-219

Hsu, C.-Y., Kalyuga, S., & Sweller, J. (2015). When should guidance be presented in physics instruction? *Archives of Scientific Psychology*, 3(1), 37-53, doi:10.1037/arc0000012, 10.1037/arc0000012.supp (Supplemental), 10.3886/ICPSR35626.v1 (Supplemental)

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1991). Cooperative learning: Increasing college faculty instructional productivity. *ASHE-ERIC Higher Education Report No. 4*. Washington, D.C.: The George Washington University, School of Education and Human Development.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (2014). Cooperative learning: Improving university instruction by basing practice on validated theory. *Journal on Excellence in College Teaching, 25*(3-4), 85-118.

Kalyuga, S. & Singh, A. (2016). Rethinking the boundaries of cognitive load theory in complex learning. *Educational Psychology Review, 28*, 831-852. doi:10.1007/s10648-015-9352-0

Kapur, M. (2008). Productive failure. *Cognition and Instruction*, *26*(3), 379-424, doi:10.1080/07370000802212669.

Kapur, M. (2011). A further study of productive failure in mathematical problem solving: Unpacking the design components. *Instructional Science, 39*(4), 561-579, doi:10.1007/s11251-010-9144-3.

Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science, 40*(4), 651-672. doi:10.1007/s11251-012-9209-6

Kapur, M. (2014). Productive failure in learning math. *Cognitive Science, 38*(5), 1008-1022. doi:10.1111/cogs.12107

Kapur, M. (2016). Examining productive failure, productive success, unproductive failure, and unproductive success in learning. *Educational Psychologist*, *51*(2), 289-299, doi:10.1080/00461520.2016.1155457.

Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *Journal of the Learning Sciences, 21*(1), 45-83.

Kapur, M. & Rummel, N. (2012). Productive failure in learning from generation and invention activities. *Instructional Science*, *40*, 645-650. https://doi.org/10.1007/s11251-012-9235-4

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75-86. doi:10.1207/s15326985ep4102_1

Klahr, D. & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science, 15*(10), 661-667.

Koriat, A., & Bjork, R. A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(2), 187-194, doi:10.1037/0278-7393.31.2.187

Latané, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology, 37*(6), 822-832. http://dx.doi.org/10.1037/0022-3514.37.6.822

Laughlin, P. R., Zander, M. L., Knieval, E. M., & Tan, T. K. (2003). Groups perform better than the best individuals on letters-to-numbers problems: Informative equations and effective strategies. *Journal of Personality and Social Psychology, 85,* 684-694.

Loehr, A. M., Fyfe, E. R., & Rittle-Johnson, B. (2014). Wait for it . . . Delaying instruction improves mathematics problem solving: A classroom study. *The Journal of Problem Solving, 7*(1), doi:10.7771/1932-6246.1166

Loibl, K., & Rummel, N. (2014). Knowing what you don't know makes failure productive. *Learning and Instruction, 34*, 74-85.

Loibl, K., Roll, I., & Rummel, N. (2016). Towards a theory of when and how problem solving followed by instruction supports learning. *Educational Psychology Review.* doi: 10.1007/s10648-016-9379-x

Loyens, S. M., Jones, S. H., Mikkers, J., & van Gog, T. (2015). Problem-based learning as a facilitator of conceptual change. *Learning and Instruction, 38,* 34-42.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist, 59*(1), 14-19, doi:10.1037/0003-066X.59.1.14

Mayer, R. E. & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38(1),* 43-52.

Mazziotti, C., Loibl, K., & Rummel, N. (2014, June). Does collaboration affect learning in a productive failure setting? In J. L. Polman, E. A. Kyza, D. K. O'Neil, I. Tabak, W. R. Penuel, A. S. Ju-row, K. O'Connor, T.

Lee, T., & L. D'Amico, *Proceedings of the 11th international conference of the learning sciences (ICLS 2014). Vol. 3* (pp. 1184-1185). International Society of the Learning Sciences, Inc.

Mazziotti, C, Loibl, K, & Rummel, N. (2015). Collaborative or individual learning within productive failure: Does the social form of learning make a difference? In O. Lindwall, P. Häkkinen, T. Koschman, P. Tchounikine, & S. Ludvigsen (Eds.) *Exploring the material conditions of learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015* (Vol. 2, pp. 570-575). Gothenburg: ISLS.

Mitchell, M. (1992). Situational interest: Its multifaceted structure in the secondary mathematics classroom. *Journal of Educational Psychology, 85(3),* 424-436.

Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory and research to educational practice. *Theory and Research in Education*, *7(2),* 133-144.

Nunnally, J. C. (1978). Assessment of reliability. In J. C. Nunnally (Ed.), *Psychometric theory* (pp. 225–255). New York, NY: McGraw-Hill.

O'Kuma, T., Maloney, D., & Hieggelke, C. (2000). *Ranking task exercises in physics* (p. 140). Boston, MA: Addison-Wesley.

Pai, H.-H., Sears, D., & Maeda, Y. (2015). Effects of Small-Group Learning on Transfer: a Meta-Analysis. *Educational Psychology Review, 27*(1).

Pease, M. A. & Kuhn, D. (2011). Experimental analysis of the effective components of problem-based learning. *Science Education*, 95, 57–86. doi:10.1002/sce.20412

Peterson, R. S. (1997). A directive leadership style in group decision making can be both a virtue and a vice: Evidence from elite and experimental groups. *Journal of Personality & Social Psychology, 72(5),* 1107-1121.

Piaget, J. (1926). *The language and thought of the child.* New York, NY: Harcourt Brace.

Piaget, J. (1973). *The child and reality: Problems of genetic psychology.* (Trans. Arnold Rosin). Oxford, England: Grossman.

Postmes, T., Spears, R., & Cihangir, S. (2001). Quality of decision making and group norms. *Journal of Personality and Social Psychology, 80(6),* 918-930.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, *93*(3), 223-231, doi:10.1002/j.2168-9830.2004.tb00809.x

Richland, L. E., Stigler, J. W., & Holyoak, K. J. (2012). Teaching the conceptual structure of mathematics. *Educational Psychologist*, *47*(3), 189-203, doi:10.1080/00461520.2012.667065.

Renkl, A. (1999). Learning mathematics from worked-out examples: Analyzing and fostering self-explanations. *European Journal of Psychology of Education 14*(4), 477–488.

Renninger, K.A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist*, *46*(3), 168-184.

Rittle-Johnson, B., & Koedinger, K. (2009). Iterating between lessons on concepts and procedures can improve mathematics knowledge. *British Journal of Educational Psychology*, *79*(3), 483-500.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93*(2), 346-362.

Rotgans, J.I., and Schmidt, H.G. (2011). Situational interest and academic achievement in the active-learning classroom. *Learning and Instruction*, *21(1)*, 58-67.

Ruiz-Primo, M. A., Briggs, D., Iverson, H., Talbot, R., & Shepard, L. (2011). Impact of undergraduate science course innovations on learning. *Science*, *331*, 1269-1270, doi:10.1126/science.1198976

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43*(3), 450-461. doi: 10.1037/0022-3514.43.3.450

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, *3*(4), 207-217, doi:10.1111/j.1467-9280.1992.tb00029.x.

Schwartz, D. L. (1995). The emergence of abstract representations in dyad problem solving. *Journal of the Learning Sciences*, *4*(3), 321-354, doi:10.1207/s15327809jls0403_3

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*(4), 475-522, doi:10.2307/3233709

Schwartz, D. L., Bransford, J. D., & Sears, D. (2005). Efficiency and innovation in transfer. In J. P. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective (pp. 1-51).* Greenwich, CT: Information Age Publishing.

Schwartz, D. L., Lindgren, R., & Lewis, S. (2009). Constructivism in an age of non-constructivist assessments. In S. Tobias & T. M. Duffy (Eds.), *Constructivist instruction: Success or failure?* (pp. 34-61). New York, NY: Routledge/Taylor & Francis Group.

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction, 22*(2), 129-184. doi:10.1207/s1532690xci2202_1

Schwartz, D. L., Sears, D., & Chang, J. (2007). Reconsidering prior knowledge. In M. C. Lovett, P. Shah, M. C. Lovett, & P. Shah (Eds.), *Thinking with data* (pp. 319-344, Carnegie Mellon symposia on cognition). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Sears, D. A. (2006). Effects of innovation versus efficiency tasks on collaboration and learning. Stanford University, U.S.: ProQuest Information & Learning.

Silvia, P. J. (2008). Interest—The curious emotion. *Current Directions in Psychological Science, 17(1),* 57-60.

Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21*(1), 43-69. doi:http://dx.doi.org/10.1006/ceps.1996.0004

Smith, K. A., Johnson, R. T., & Johnson, D. W. (1981). Can conflict be constructive? Controversy versus concurrence seeking in learning groups. *Journal of Educational Psychology, 73*(5), 651-663.

Smith, K. A., Sheppard, S. D., Johnson, D. W. and Johnson, R. T. (2005). Pedagogies of engagement: Classroom-based practices. *Journal of Engineering Education, 94*, 87–101. doi:10.1002/j.2168-9830.2005.tb00831.x

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research, 69*(1), 21-51, doi:10.2307/1170643

Sweller, J. (2009). Response to Schwartz, Lindgren, and Lewis. In S. Tobias & T. Duffy, (Eds.), *Constructivist instruction: Success or failure?* (p. 53). New York, NY: Routledge/Taylor & Francis Group.

Sweller, J., Kirschner, P. A., & Clark, R. E. (2007). Why minimally guided teaching techniques do not work: A reply to commentaries. *Educational Psychologist, 42*(2), 115-121, doi:10.1080/00461520701263426

Sweller, J., van Merrienboer, J.J.G. & Paas, F.G.W.C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296. https://doi.org/10.1023/A:1022193728205

Vansteenkiste, M., Niemiec, C.P. and Soenens, B. (2010). The development of the five mini-theories of self-determination theory: An historical overview, emerging trends, and future directions. *Advances in motivation and achievement*, *16*, 105-166.

Vygotsky, L. S. (1978). *Mind in society* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.), Cambridge, MA: Harvard University Press.

Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education*, *22*(5), 366-389, doi:10.2307/749186

Westermann, K. & Rummel, N. (2012). Delaying instruction: Evidence from a study in a university relearning setting. *Instructional Science*, *40*, 673-689. https://doi.org/10.1007/s11251-012-9207-8

Wiedmann, M., Leach, R.C., Rummel, N., and Wiley, J. (2012). Does group composition affect learning by invention? *Instructional Science, 40(4),* 711-730. https://doi.org/10.1007/s11251-012-9204-y

Wiedmann, M., Leach, R.C., Rummel, N., and Wiley, J. (2015). Mathematical skills and learning by invention in small groups. In Y. H. Cho, S. I. Caleon, & M. Kapur, (Eds.), *Authentic Problem Solving and Learning in the 21st Century: Perspectives from Singapore and Beyond* (pp. 249-265). Singapore: Springer Science + Business Media.

Wirkala, C., & Kuhn, D. (2011). Problem-based learning in K-12 education: Is it effective and how does it achieve its effects? *American Educational Research Journal, 48*(5), 1157-1186. Retrieved from http://www.jstor.org/stable/41306381

Wise, A. F., & O'Neill, K. (2009). Beyond more versus less: A reframing of the debate on instructional guidance. In S. Tobias & T. M. Duffy (Eds.), *Constructivist instruction: Success or failure?* (pp. 82-105). New York, NY: Routledge/Taylor & Francis Group.

Wittenbaum, G. M., Hollingshead, A. B., & Botero, I. C. (2004). From cooperative to motivated information sharing in groups: Moving beyond the hidden profile paradigm. *Communication Monographs, 71(3),* 286-310.

Yetter, G., Gutkin, T. B., Saunders, A., Galloway, A. M., Sobansky, R. R., & Young, S. Y. (2006). Unstructured collaboration versus individual practice for complex problem solving: A cautionary tale. *The Journal of Experimental Education*, *74,* 137-159.