# Good problems within and across disciplines

Daniel Reinholz, Tara Slominski, Timothy A. French, Sam Pazicni, Chris Rasmussen, Bradley McCoy

*This paper focuses on the question of what makes a good disciplinary or interdisciplinary problem. We draw from literature across the STEM disciplines and two conference sessions to provide insight into what makes a good problem within a specific discipline and across the disciplines. We use various frameworks to analyze a variety of problems that were nominated as exemplars by STEM education research experts. Common features identified include real-world connections, reinforcement of conceptual understanding, a low floor and high ceiling, multiple solutions paths, and building dispositions of professionals in the discipline. While a good problem need not have all of these features, in general, good problems have more of these features. We also recognize that these problems are context-specific, as what may be considered a problem for one learner could be a trivial exercise for another. We discuss some of the challenges of designing good interdisciplinary problems and identify some features that can make a problem interdisciplinary, including use of cross-cutting concepts and drawing on the specific expertise of each discipline.*

## Introduction

STEM professionals are problem solvers. Engineers design solutions – bridges, airplanes, computers – that transform the way we live and work. Computer scientists develop applications that take our productivity to new levels. Physicists use fundamental laws to build models of increasingly complex real-world phenomena. Biologists develop new techniques for taming

microscopic hazards that attack the human body. Chemists transform and analyze matter, from drug therapies to designer materials, to better our quality of life. All of these solutions arise from addressing basic problems of human existence, and these solutions transform our society.

Given the fundamental role of problem solving across the STEM disciplines, problem solving is a highly-valued practice in STEM education. But what types of problems should students solve? Which problems will help them develop the relevant skills that later allow them to solve problems like minimizing the effects of climate change, treatments for diseases, or engineering solutions? How do instructors support students to engage with such problems?

This paper addresses the question: what makes a problem good? Thinking beyond problems generally, what makes a problem good for biology, physics, math, engineering, computer science, or chemistry? What are the implications for interdisciplinary problems? In this paper, we discuss a variety of frameworks for thinking about good problems, both within and across disciplines. Drawing from two conference sessions focused on good problems, we analyze a variety of problems nominated by Discipline Based Education Research (DBER) scholars as exemplars. The authors of this paper span a variety of disciplines, so the individuals analyzing each of the problems given have expertise in that particular discipline. We close by discussing the challenges and opportunities for interdisciplinary problems.

## Theoretical Background

### What makes a problem?

We take the view that problem solving is relative. In this way, what constitutes a problem is a relationship between a problem solver and the task itself, *not* an inherent property of the task (Schoenfeld, 1985). Thus, a problem is something that is difficult for someone who is engaging

with it. A good problem is typically more than just a technical or computational challenge, but something that poses an impasse in conceptualizing or implementing a solution. We recognize that some great unsolved problems in the STEM disciplines are *also* computational challenges, like image recognition, but this is different from a computational challenge such as multiplying two four-digit numbers mentally. In other words, if someone has a readily available solution schema, that they could in principle implement, it is an exercise, not a problem (Schoenfeld, 1985). Further, problems can be conceptualized as seeds to disciplinary thinking (Schoenfeld, 1991). This means that problems serve as a platform for students to engage in the practices of a discipline. For our own work, we focus on tasks that are *likely* to serve as problems for students in our courses, recognizing that they would not be problems in this sense for disciplinary experts.

**How do problems support learning?**

Educators have developed a variety of metaphors for learning. Two prominent models are learning as knowledge acquisition and learning by apprenticeship (Sfard, 1998). From the knowledge acquisition perspective, we can think of learning in the STEM disciplines as a process of developing more normative conceptions of scientific (and mathematical) phenomena (Linn, Clark, & Slotta, 2003; Smith, diSessa, & Roschelle, 1993). From the apprenticeship perspective, learning is about being able to "do" the same things that STEM professionals do, in increasingly sophisticated ways (Lave, 1996). Taken together, we can think about learning as a process of changing what it means to "know and do" within the STEM disciplines (Lampert, 2003).

Across the STEM disciplines, scholars have paid great attention to features of problems that can support learning. We consider these as the pedagogical aspects of good problems. Our

goal here is not to provide an exhaustive list of such features. Indeed, a recent dissertation focused on good problems analyzed a number of problems along over 20 dimensions, or mathematical problem features (Quarfoot, 2015). Aggregating these dimensions across disciplines, the list soon becomes intractable. Rather, we provide a brief look at some of these literatures simply to make the case that problems are good both because of *what* they help students learn, and *how* they help them learn it (i.e., the pedagogical features).

In mathematics, Schoenfeld (1991) describes a problem aesthetic, favoring problems that have some or all of the following four features: (a) accessibility, (b) multiple solution paths, (c) insight into mathematical concepts or practices, and (d) generalizability (to promote future exploration). Principles (a) and (d) are often described as having a "low floor and high ceiling," meaning that a problem is easy for any student to start, but allows for a great deal of challenge and exploration. The emphasis on multiple solution paths is pervasive in mathematics education (e.g., National Council of Teachers of Mathematics, 2000), because it allows students to compare and contrast their solutions to develop deeper knowledge. Taken together, these problem features are valuable because they support learning across a variety of students in heterogeneous environments (Cohen & Lotan, 1997), which is a real pedagogical challenge.

In computer science, the notion of "nifty" problems adds some other aspects of good problems (Layman, Williams, & Slaten, 2007). Such problems also meet four criteria, they are: (a) playful, fun, or interesting; (b) topical (fitting into standard curricula); (c) scalable (i.e. with a low floor and high ceiling); and (d) easily adoptable (by other instructors). This work emphasizes that good problems should be enjoyable for students, having some sort of fun factor that motivates students to want to work on the problem. Also, fitting in with standard curricula and having materials (e.g., handouts, source code) that can be easily used by others means that they

are problems that can have utility to instructors in a variety of situations. Like the problem aesthetic described above, much of this niftiness is that the problems are designed in a way that they can work well for a variety of students in a variety of situations.

In physics, several novel problem formats have been developed: context-rich problems (Heller & Hollabaugh, 1992), ranking tasks (O'Kuma, Maloney, & Hieggelke, 1999), jeopardy problems (Van Heuvelen & Maloney, 1999), and more. As an example, context-rich problems are designed so that the problem is situated in a real-world context, often with extraneous or missing information, and with some level of solver choice in defining the aim of the problem, selecting an appropriate physical model of a situation, and evaluating the results (Heller, Keith, & Anderson, 1992). The real-world context can serve as a low floor because students may have experience with the setting, while also providing a high ceiling when simplifying approximations to the problem are replaced with more real-world practicalities. In addition, the context supports the disposition of professionals to apply physical thinking outside the classroom. Emphasis on development of physical models, estimation, and approximation supports problem solving practices used by professional physicists and allows for multiple solution paths. Context-rich problems and other common formats of physics problems attempt to include pedagogical incentives for expert-like problem solving practices, with varying success. While professional physicists classify problems in terms of the relevant underlying physics principles, novices are more likely to focus on surface features of a problem (Chi, Feltovich, & Glaser, 1981; Docktor, Mestre, & Ross, 2012). To move the students towards move expert-like practices, problems have been structured to emphasize the key physics practices of blending conceptual and mathematical reasoning (Leonard, Dufresne, & Mestre, 1996; Hull, Kuo, Gupta, & Elby, 2013) and employing multiple representations (Kohl, Rosegrant, & Finkelstein, 2007). While context-rich problems

provide many useful pedagogical opportunities for students, we also acknowledge that there are times when creating a fantasy or theoretical situation can also be beneficial, because it allows an instructor to emphasize particular concepts without the messiness of a real-world situation.

**Does it matter how we use problems?**

As emphasized above, it is helpful when problems can be used by a variety of instructors and students, in a variety of settings. In other words, problems do not exist in a vacuum. They serve as a starting point for learning, whether that takes place in the classroom, at home, in a lab, or out in the field. However, there is the whole process of teaching and learning that must take place, which can either enhance or inhibit the value of a task.

The first question is how many people should work on a given task. One way that this has been conceptualized is the notion of a "group worthy problem" (Cohen & Lotan, 1997; Featherstone et al., 2011; Heller, Keith, & Anderson, 1992). Simply put, any given task is appropriate for a certain number of people. Assigning a one-person task (e.g., procedural computations) to a group results in disengagement, because it is not group worthy. Similarly, assigning a task made for a team to a single person will result in frustration. So, when thinking about whether a problem is good or not, it is also a matter of thinking about how many people it is good for.

Another issue relates to how an instructor uses a task. In other words, what sort of scaffolding does the instructor provide? In mathematics, this issue has been conceptualized in terms of the math tasks framework with the idea of cognitive demand (Henningsen & Stein, 1997). When using a task in the classroom, there is the task as given, the task as framed by the instructor, and the task as implemented by students. These "tasks" may differ, because a teacher

can provide scaffolding (e.g., support towards completing the task) or extensions (e.g., asking new and harder questions) during the enactment of a task. These techniques modify the cognitive demand of the task, influencing what students learn. It is common for instructors to over-scaffold tasks, which may remove their learning potential by only leaving the trivial aspects of the task for the students to complete. So, even if a task is good to begin with, it might be trivialized in implementation. Similarly, skilled teachers can take relatively mundane tasks and transform them into something worthwhile.

Related to the idea of cognitive demand is *Bloom's Taxonomy of Cognitive Domains* (Bloom et al., 1956), which has been taken up extensively by biology educators. Bloom's taxonomy organizes the type of thinking required by a task into six levels. These six levels can be further grouped as lower-order cognitive skills (knowledge, comprehension) and higher-order cognitive skills (application, analysis, synthesis, evaluation) (Crowe et al., 2008). Just as with cognitive demand, the way a teacher enacts a task can be used to engage higher-order skills, which would help make something into a "good problem." That said, recent inventory of end-of-chapter problems from popular General Chemistry textbooks noted a serious lack of problems in the higher cognitive categories, a situation that could very well limit both instructor and student access to examples of more meaningful learning opportunities (Dávila & Talanqer, 2010).

Another issue at hand in problem solving is working memory. As previously stated, "expert" problem solvers in a discipline are much better able to pay attention to the relevant features of a problem, while ignoring other aspects that are more peripheral (Chi, Feltovich, & Glaser). This relates to working memory, insofar that better problem solvers are able to better utilize their working memory by paying attention to the relevant features at hand (Passolunghi & Siegel, 2001). In contrast, less effective problem solvers may become bogged down because they

are trying to attend to too many things at once. Through practice problem solving, students can better develop these metacognitive skills, and thus become better at problem solving as a practice (e.g., Schoenfeld, 1985).

Finally, one might consider the way that an exploration around a problem is sequenced. In science education, the Knowledge Integration (KI) framework describes a process for developing conceptual understanding (Linn et al., 2003). KI involves a four-step process: (1) elicit ideas, (2) add new ideas, (3) develop criteria, and (4) sort out ideas (Slotta & Linn, 2009). The developers of KI then embed this sequence into a variety of different project types (e.g., investigation, controversy, critique, design). If a particular part of the sequence is removed from how one engages with a task, less learning results (Slotta & Linn, 2009). This emphasizes that the curriculum surrounding a problem matters, and can either set up a problem for success or failure.

**Are the STEM disciplines all the same?**

Up to this point, one might simply think that good problems are good for any discipline. Indeed, certain pedagogical features are relatively standard across disciplines. For instance, it is hard to imagine a discipline in which good pedagogical practice would mean making problems too difficult for most of the students, or trivializing them to the point where there is nothing left for students to do. At the same time, we acknowledge that the STEM disciplines are actually quite diverse, and simply referring to them simultaneously in a single acronym obscures this diversity. One of our goals with this paper is to highlight such diversity. We begin to do so by highlighting that the major problem types in disciplines are quite different, owing to the disciplines themselves.

In the biosciences, the use of case studies is a powerful tool for learning (Allchin, 2013). Such problems are particularly valuable because they provide students with an opportunity to use evidence to make predictions about authentic problems that highlight complexity of living organisms and their environments. Chemistry education often focuses on coordinating macroscopic, submicroscopic, and symbolic representations of chemical phenomena (Talanquer, 2011). This coordination models the reasoning patterns of experts, who tend to think about chemistry using the three domains simultaneously. In the geosciences, spatial navigation and mapping are key practices for field work (Riggs, Lieder, & Balliet, 2009). In physics, there is a huge emphasis on building simplified models from which theoretical generalizations can be made (Hestenes, 1987; Brewe, 2008). Computer science centers on developing programs, so debugging and programming are key parts of the problem-solving process (Layman et al., 2007). In engineering, building solutions to real-world problems is the goal, so there is a focus on designing, prototyping, testing, and refining (Ikonen, Piironen, Saurén, & Lankinen, 2009). Finally, (pure) mathematics revolves around conjecturing, argumentation, and proof; it is deductive rather than empirical (Harel & Sowder, 2007). A summary of these problem types is given in Table 1.

Table 1. Canonical problem types by discipline.

| Discipline | Problem Type | Values / Implications |
|---|---|---|
| Biosciences | Case Studies | Reasoning from evidence |
| Chemistry | Representation | Coordinating macro, submicro, and symbolic |

| | | |
|---|---|---|
| Geosciences | Spatial Navigation | Guiding field work |
| Physics | Model Building | Articulation of physical laws |
| Computer Science | Programming/Debugging | Creating tools |
| Engineering | Design Challenges | Real-world solutions |
| (Pure) Mathematics | Proof/Argumentation | Theory building |

This is not to say that each discipline has only one core practice or type of problem. Rather, we argue that certain problem types are emblematic of certain disciplines and rarely used in others. This is due to fundamental differences in the nature of the disciplines and their valued practices. For instance, while it is possible to have mathematical proofs that speak to the science disciplines, it is rare for a proof to afford opportunities to develop disciplinary knowledge in science disciplines, because they are empirical while pure mathematics is not. This means that developing good problems across the disciplines has an extra level of challenge involved.

**What is disciplinary learning?**

Now that we have considered some of the nuances of problems across STEM disciplines, we introduce a framework for thinking about learning in the disciplines. This gives us another way to think about which problems are good – they are the ones that result in more learning. Within the STEM disciplines, there are a variety of guiding documents that set out standards for what students should know and be able to do. Here we draw on two frameworks from the National Research Council: *Adding it Up* in mathematics (National Research Council, 2001), and the *Science Education Framework* for science and engineering disciplines (National Research

Council, 2011). Taken together, we can draw out a few key areas of focus across all STEM disciplines.

The *Science Education Framework* has three major areas of focus for student understanding: disciplinary practices, crosscutting concepts, and core disciplinary ideas (National Research Council, 2011). Practices are things such as developing and using models, analyzing and interpreting data, or explaining and justifying one's ideas, that are key to doing the work of STEM professionals. Crosscutting concepts are big conceptual ideas like patterns, systems, and conservation, that are relevant to many disciplines such as physics, biology, and chemistry. Finally, core disciplinary ideas are concepts such as forces (physics), evolution (biology), and the particulate nature of matter (chemistry), that are critical to understanding particular disciplines. In biology, these concepts and practices are clearly described by the 3D-LAP (Laverty, et al., 2016) and Vision and Change (AAAS, 2011), which makes biology exemplary in this regard of describing the desired disciplinary knowledge at the undergraduate level. In chemistry, the discipline's "anchoring concepts" (as they are operationalized in an undergraduate curriculum) are articulated in the American Chemical Society Examinations Institute's Anchoring Concepts Content Maps (ACCM) (Murphy, et al., 2012; Holme, et al., 2015), which have been established by the Chemistry community at large, with input from participants of several workshop and focus group sessions. Other disciplines have varying specificity in documents describing the core ideas of the discipline, but this is mostly at the K-12 level.

Drawing from *Adding it Up*, there is a conceptualization of mathematical proficiency as having five strands: conceptual understanding, procedural fluency, strategic competence, adaptive reasoning, and productive dispositions (National Research Council, 2001). Conceptual

understanding relates to math concepts, operations, and relations. Procedural fluency is the ability to carry out procedures with flexibility, efficiency, and accuracy. Procedures are things like performing integration by parts, balancing a chemical equation, or constructing a Fourier series. Strategic competence relates to solving mathematical problems and adaptive reasoning relates to interpreting and explaining mathematical ideas. Finally, productive dispositions relate to how one sees and identifies with the discipline. Ideas like "steady effort in learning mathematics pays off ," "deep thinking, not speed is valued," or "math is useful in the real world" would be considered productive dispositions.

Taking these two frameworks together, we offer four main categories: concepts, practices, procedures, and dispositions. We condense crosscutting concepts and core disciplinary ideas together in the category of concepts, recognizing that some of these may be specific to a discipline while others are cross-disciplinary. We also condense strategic competence and adaptive reasoning together in the category of practices, seeing as these are both core disciplinary practices in mathematics. Of the four areas, we de-emphasize procedures compared to the others. Procedures are not unuseful, but in most typical curricula, they receive far more attention than the other three categories. Thus, part of the work of good problems is to counteract this narrowing of the curriculum.

From the perspective of our framework, a problem can be considered "good" when it helps students learn along these four dimensions. The pedagogical and other features described above all make it easier for a student to engage meaningfully with a problem, but there must also be something meaningful to learn from the problem, which these dimensions describe. A good *disciplinary* problem supports growth along these dimensions with respect to that particular discipline (e.g., a good physics problem would help students develop the practices and

dispositions of a physicist). A good *multidisciplinary* problem (Choi & Pak, 2006) does this for multiple disciplines (e.g., the practices and concepts of both mathematics and physics). One can also think of *interdisciplinary problems* (ones that build concepts in multiple disciplines and help integrate them) or *transdisciplinary problems* (those that also include a humanities context). For this paper we focus on disciplinary and interdisciplinary problems.

According to these classifications, it is possible to have a physics problem that includes mathematics, which is a good physics problem but not a good mathematics problem. In this sense, it could be a *good disciplinary* problem, and a multidisciplinary problem, but not a *good multidisciplinary problem* because it does not help build core ideas or practices in multiple disciplines.

## Method

This paper builds on conversations from two interdisciplinary STEM education conferences: *Breaking the Boundaries in STEM Education,* hosted on April 7 at Loyola Marymount University; and *Transforming Research in Undergraduate STEM Education (TRUSE)*, hosted from July 5-9, 2017, at the University of St. Thomas. There were approximately 20 participants throughout the day in the Breaking Boundaries problem solving strand, and approximately 40 in a follow-up breakout session on disciplinary problems at TRUSE. We acknowledge the conversations in these sessions as the inspiration for this paper. The participants in these sessions were Discipline-Based Education Researcher (DBER) scholars who spanned the disciplines of mathematics, physics, chemistry, and biology, with minimal representation of geology, computer science, and engineering.

During the TRUSE session, participants were first split up at different tables, by discipline, so that they could brainstorm features of good problems and exemplary problems

within their disciplines. We draw from these problems in our analyses below. Afterwards, participants were split up into cross-disciplinary tables, and tasked to generate good interdisciplinary problems. We also offer some of these problems below with commentary on why this was a very difficult task.

## Good Disciplinary Problems

### Mathematics: Filling Bottles

The first problem was a mathematics problem that was suggested by the presenters (see Figure 1). This problem focuses on core ideas in calculus, notably function and rate of change. This problem also has core practices of representation (e.g., generating graphs from some realistic situations). As it relates to everyday life, the problem can promote productive dispositions that math is supposed to make sense and that functions do not necessarily need to be defined by a single algebraic formula. The problem de-emphasizes procedures, but there are some procedural aspects of labeling axes and putting together graphs.
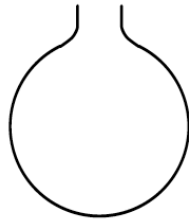
Figure 1. Filling bottles.

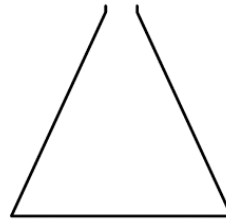You are filling up bottles with liquid coming from a tap at a constant rate.

1. For each bottle, sketch a graph of the height of liquid in the bottle as a function of time.
2. For each bottle, sketch a graph of the rate of change of the height of liquid in the bottle as a function of time.
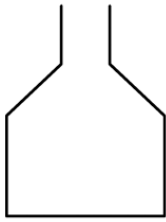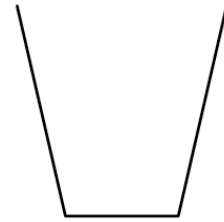
Ink bottle   Evaporating flask   Conical flask   Bucket

This task is a nice example of a problem with a low floor and a high ceiling. For the first part of the task students can readily reason that in each case a graph of height versus time would be increasing and hence everyone can make some progress. Issues of concavity, average rate of change, and instantaneous rate of change, and the connection between the graph of a function and a graph of its rate of change can also emerge for students, making the ceiling quite high. In particular, the second bottle shape change point in the ink bottle offers an opportunity for students to informally and intuitively reason about instantaneous rate of change, even if this concept has not formally been defined in their calculus class. For example, Marrongelle and Rasmussen (2008) describe how offering alternatives for the shape of the graph at the point where the second bottle transition from spherical to cylindrical (see Figure 2) can be generative for eliciting justifications that foreground instantaneous rate of change.

Figure 2. Which graph most accurately reflects what happens at the neck of the bottle?

height    C    B

A

neck

volume
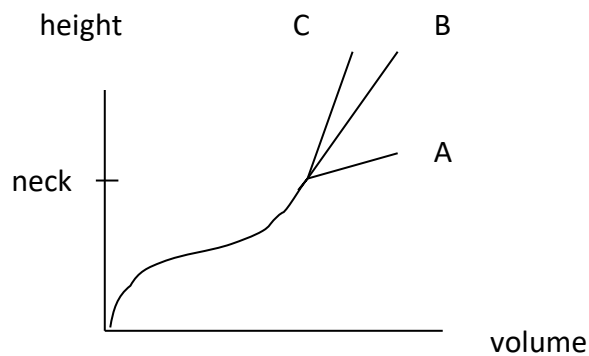
The problem also has multiple solution pathways. For example, we have seen some students approach this problem by imagining a cup of water being poured in the bottle, marking on the image of the bottle how high they conjecture that the water rose, adding another cup of water and marking again on the bottle how high the water rose, etc. This cup by cup process would then be graphically re-represented and students can justify why it would make sense to connect the points with either straight line or non-straight line segments. Graphs of height versus time can also be generated by dynamically imaging how fast the water is rising as water flows into the flask at a constant rate. This type of reasoning is emblematic of what Carlson et al (2002) refer to as covariational reasoning, which is a foundational disciplinary way of conceptualizing how two quantities simultaneously change. While this task does not lead to formal deductive proofs, it does lend itself to conjectures, argumentation, and justification - key disciplinary practices in proving.

**Physics: Superman: The Ride**

The second problem is a context-rich physics problem written by one of the authors, who was actively involved in the Breaking Boundaries sessions. The problem statement is in Figure 3.

As with many context-rich problems, this problem requires students to identify a calculable target variable that can be used to answer a non-quantitative question via comparison. The problem could be framed with a more explicit and more scaffolded question, such as "Assuming the brakes do not fail, is the safety track long enough?" However, we prefer this more open-ended question because it gives students valuable practice in identifying and justifying their assumptions and in determining which variables are calculable and relevant. As this problem is solved by groups in class, the instructor can reframe the question as necessary if students oversimplify the problem or have trouble identifying a calculable approach.

This question can be answered by calculating the distance that the car needs to stop under the influence of friction and comparing to the length of the safety track or calculating the speed after travelling the length of the safety track. Multiple solutions paths are available; this problem can be effectively solved using forces and kinematics, conservation and the work-kinetic energy theorem, or the impulse-momentum theorem. The instructor can have students explore the alternate solution paths as extensions to the problem. Each solution path requires a different physical model of the situation. Logically complete solutions require the students to articulate and justify approximations as part of their model, such as assuming negligible air resistance.

Figure 3. Superman: The Ride

I once rode Superman: The Ride at Magic Mountain. The ride accelerates a car through a launching station, then the track turns vertically upward. The car rises up the track to a height of 415 ft, then falls back down under the influence of gravity and stops again back at the starting point. The braking mechanism at the end of the ride malfunctioned so the car continued through the launching station onto a safety track where it was stopped by friction. I

estimate that the coefficient of friction was 0.5. I can see that the safety track is 750 feet long. The ride specs say that the car weighs 6 tons. I also noticed that the brakes did partially engage. The ride attendant apologized for the malfunction and offered to let us ride again without waiting in line. Should I get off?

The physics concepts used in this problem vary by choice of solution path. Regardless of the path, the solution requires employing at least two fundamental physics concepts (e.g. energy and forces, forces and kinematics, or energy and momentum). Any solution is likely to include at least two representations: a sketch of the ride and a motion diagram, free body diagram, or energy bar chart. Physics practices supported by this problem include defining a good question, identifying an appropriate physics principle, and application of physical laws. Once a physics principle has been chosen and translated into an equation using the appropriate physical law, the calculational procedures are simple algebra. Dispositions of physics experts encouraged by this problem include examining everyday situations for underlying physics, seeing physics as relevant to everyday life, modelling by making simplifying assumptions and approximations, and interpreting the conceptual meaning of calculated results.

**Chemistry: Structure-Property Relationships**

The third problem was suggested by a group of chemists; versions of this problem are well-known and published as exemplars (Laverty, et al., 2016). The problem stem presents two substances that possess the same molecular formula (dimethyl ether and ethanol, $C_2H_6O$) along with representations of their molecular structures (i.e. Lewis structures, Figure 4). The problem consisted of six prompts related to these two substances: (1) what intermolecular forces (IMFs)

would be present in a container of dimethyl ether? (2) what IMFs would be present in a container of ethanol? (3) draw a collection of dimethyl ether molecules, showing the strongest IMF for this compound, labeling the location and type of IMF you are drawing; (4) draw a collection of ethanol molecules showing the strongest IMF for this compound, labeling the location and type of IMF you are drawing; (5) using the relative strengths of IMFs, predict which compound would have the higher boiling point and explain why; and (6) suggest a way in which a mixture of ethanol and diethyl ether could be separated and use your previous responses to support your suggestion.

Figure 4. Data for structure-properties relationship problem.



dimethyl ether          ethanol
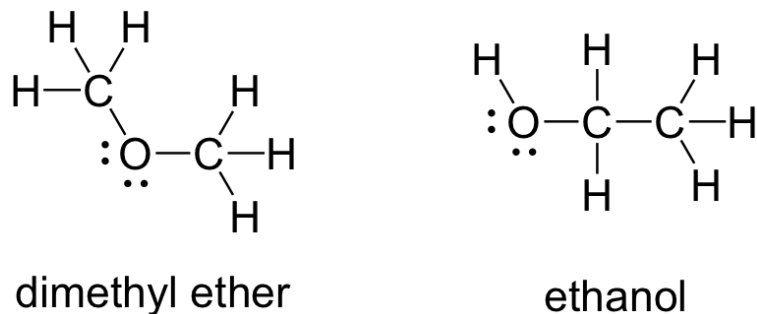
The chemists identified a number of reasons why this represents a "good problem." First, the problem centers on two core chemistry disciplinary ideas: (1) chemical compounds have geometric structures that influence their chemical and physical behaviors; and (2) intermolecular forces (electrostatic forces between molecules) dictate the physical behavior of matter. While chemistry's core ideas are articulated in a variety of ways, language from the ACCM was used here. Alignment with articulated disciplinary core ideas was considered a hallmark of "good problems" by the chemists because such organization mirrors how experts think about and reason within their own disciplines (National Research Council, 2000). Moreover, this problem is aligned with well-documented difficulties that students have when attempting to connect

macroscopic properties to representations of molecular-level structure (Cooper et al., 2012).

Thus, if implemented properly, this problem could aid students in overcoming a common struggle encountered when working with one of chemistry's core ideas.

Second, this problem requires students to engage in practices central to science in general and to chemistry specifically. Integral to Coppola's appeal to "Do Real Work, Not Homework" is the perception that students are only learning *about* science if they do not engage in the practices that are central to *doing* science (Coppola, 2015). The result of the former is, of course, the many shortcomings experienced by those who engage in rote versus meaningful learning (Mayer, 2002). Eight scientific and engineering practices were articulated in *A Framework for K-12 Science Education* (National Research Council, 2011) and there is little reason to believe that these practices should not be applied to the college classroom (Cooper, et al., 2015). An analysis of the "good problem" above using the Three-Dimensional Learning Assessment Protocol (3D-LAP, Laverty, et al., 2016) demonstrates that it is well-aligned to the scientific practices of Developing and Using Models—students are asked to construct models of the strongest IMFs present in samples of dimethyl ether and ethanol, make a prediction about a phenomenon (boiling points) using these models, and provide reasoning that links the model to the prediction. Moreover, this problem is aligned with a disciplinary practice articulated in Sevian and Talaquer's Chemical Thinking Learning Progression: *analysis*, the development and application of strategies for detecting, separating, and quantifying chemical substances (Enke, 2001; Sevian and Talaquer, 2014).

Lastly, this "good problem" is well-aligned (or has the potential to be, given how it is implemented in the classroom) with attributes of the aforementioned "Real Work" framework (Coppola, 2015). The problem balances convergent and divergent tasks; it demands that students

construct responses that (a) can be identified as "right" or "wrong" (e.g. prompts 1, 2, and 5) and (b) that are more generative in nature, whose accuracy is defined within a set of common guidelines (e.g. prompts 3, 4, and 6). Absent from the problem articulation is how it is to be operationalized in a real learning environment; such choices could further align the "good problem" with the ideals of Real Work. For example, the problem could incorporate peer presentations, review, and critique so as to allow learners to develop explanatory knowledge; additionally, the problem could incorporate technology to permit learners to develop models using multiple modes.

**Biology: The 'Dino Problem'**

Variations of the 'Dino Problem' (Bray Speth et al., 2009; Grant 2008, cartoon adapted from Frank Hauser, Jr., published in Science, vol. 250, pp. 1103, 1990) have been used in biology education long before the TRUSE conference, but this problem was identified as a "good" by some of the biology education researchers that attended the conference. While this cartoon appears to be deceptively simple, it can actually pose real challenges for undergraduate biology students.

Figure 5. The Dino Problem

**Evolutionary Theory Made Simple**
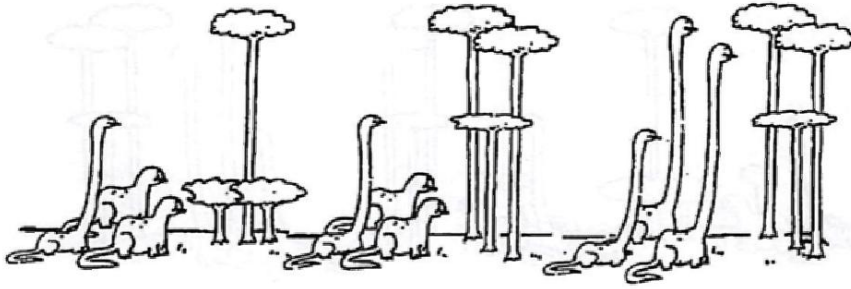
The cartoon above represents change that has occurred in a population of animals and a population of plants over thousands of years (time is read from left to right). Use your current understanding of evolution by natural selection to explain how the changes came about.

We consider the 'Dino Problem' to be a "good problem" for a number of reasons. First, this problem asks students to reason about evolution by natural selection, a core biology concept as recognized by reports from the NRC (2011) and AAAS (2011). A good response from a student needs to be quite complex and make use of all the principles that govern evolution by means of natural selection (i.e. variation within a population, genetic basis of that variation, trait inheritance, differential survival and reproductive success, and changes in phenotype frequency over time). To generate an exemplary response, students have to 1) recall the multiple principles that govern natural selection, 2) apply these principles to this hypothetical scenario, and 3) integrate the changes occurring in both of these populations by connecting with those governing principles. Even after instruction, biology students struggle to successfully incorporate all of the necessary principles when presented with this assessment.

Second, this problem explicitly asks students to generate a reason for why change came about, and therefore, students are primed to present their explanation in terms of cause and effect (a crosscutting concept). Also, this problem asks students to reason about two different populations (animals and plants) over time. By presenting these organisms in tandem (both pictorially and textually), students are encouraged to consider the interactions occurring between

the organisms. So, we would argue that the 'Dino Problem' supports at least two crosscutting concepts from the NRC framework and the resulting 3D-LAP: a) cause and effect and b) systems and system models.

Third, this problem asks students to make inferences on the hypothetical evolutionary scenario depicted in the cartoon. Students are expected to generate an explanation for populations' change over time using their understanding of evolution by natural selection (Bray Speth et al., 2009). Because this question is situated in a hypothetical scenario, students would not be able to simply recall knowledge they have previously learned. Instead, students would need to apply the principles of natural selection in this new scenario, connecting those principles with the inferences they make from the figure to construct an accurate explanation of the evolutionary event, a task that is known to be very difficult for students. As a result, we would classify this as a higher-order cognitive question along Bloom's taxonomy (Crowe et al., 2008) that supports the scientific practices of a) developing and using models and b) constructing explanations that have been identified by the NRC (2011) and the resulting 3D-LAP (Laverty, et al., 2016).

## Good Interdisciplinary Problems

When participants in the TRUSE session were placed into interdisciplinary tables, they had a much harder time coming up with "good" interdisciplinary problems. Some groups were more successful than others (see below), but the fact that groups consisting of experts of multiple STEM fields had difficulty generating problems highlights the challenge of generating good, truly interdisciplinary problems.

Although participants struggled to generate many problems, the process itself elucidated some salient features of good interdisciplinary problems. Good interdisciplinary problems must have a contextual intersection that is interesting and relevant to multiple disciplines. A problem with a very narrow, discipline-specific context does not easily allow for the integration of multiple disciplines. For example, having students from biology and chemistry balance a chemical equation gives them practice of an important procedure used in both fields, but it does not allow for the biology students to bring in their domain-specific knowledge to add to the solution. Oftentimes instructors (and textbook authors) try to rectify this issue by making the surface features of the problem relate to a second discipline, while the deep feature remains discipline-specific—asking students to balance a chemical equation related to glycolysis does not automatically make it a biology problem as well.

However, a rich context is not a sufficient condition for a good interdisciplinary problem. Good problems should generate knowledge in multiple disciplines. One way to do this is through the application of a crosscutting concept. Every scientific discipline is unique, but many of them use the same tools and concepts to tackle problems, such as energy (Watkins et al., 2012) and conservation laws, which can be applied in a large range of situations. A bonus consequence of the universality of these crosscutting concepts is that members of each discipline think about, focus on, and invoke different aspects of the underlying concept when solving problems. This leads to a more creative and dynamic problem-solving process as students from different disciplines can unlock different parts of the problem, which then uncover new entry points for the other disciplines to add to a possible solution.

The application of crosscutting concepts emphasizes the interconnectedness of the various disciplines and helps fight the compartmentalization of concepts by discipline. Science is

not done in a vacuum; big, important scientific problems facing humanity like curing cancer or minimizing the effects of climate change require creative solutions and the cooperation of many scientists from different disciplines. Group solving of good interdisciplinary problems mimics how science is done "in real life." This is a way to promote dispositions that are valued across multiple disciplines. It provides students—who will eventually be the ones working on these big, important scientific problems—an opportunity for them to act as proto-practitioners in a classroom setting. Here are two examples of possible interdisciplinary problems created during the TRUSE session.

**Climate Modeling.**

Keeping in mind that a good interdisciplinary problem should have broad contextual relevance, one group suggested climate modeling as a candidate for a suite of problems but did not agree upon a specific problem statement (again highlighting the difficulty of writing problems like this). Such a large and complex topic allows for the integration of many different STEM disciplines, but can quickly become overwhelming if the focus is not narrowed enough that the beginning, or at least a toehold, of a solution can be seen. One such question could be "What effect would a 10% decrease in the amount of carbon dioxide in the Earth's atmosphere have on the oceans and their ecosystem?" Geologists, environmental scientists, and biologists would bring their knowledge of the climate and Earth as a system to the solution. This base knowledge would provide a starting point for the group to rally around, and others could build on it in various ways. Testing possible models would necessitate some level of computational modeling and, thus, writing computer code, in which mathematicians, computational physicists, and computer scientists all have expertise. An understanding of the chemistry of the oceans and

the atmosphere and the effect temperature has on relevant reactions and chemical equilibria, such as acidification of the oceans due to increased dissolved carbon dioxide, is crucial for an accurate and complete model of the planet's climate. A broad topic like climate modeling requires input from practitioners of many different STEM fields, each of whom would bring valuable discipline-specific practices and core ideas necessary to completely understand the problem at hand.

**Fat Problem.**

This problem is simply stated as: **"If you were to lose five pounds of fat, where does it go?"** The most obvious applications of this problem relate to biology and chemistry: What is fat? What is its chemical structure? How and why is fat created in the body? How and why is fat broken down in the body? How is the broken-down fat expelled from the body? Biologists might take a big-picture, system-level view of the problem by looking at the various metabolic processes, while chemists might focus on the products and energetics of the specific reactions involved.

Although this problem is clearly relevant to biology and chemistry, members of other disciplines can provide valuable insight into its solution. Applied mathematicians and physicists, who generally value modeling and theory building, can provide a simple model from which generalization can be made. They could also work on various scaling arguments to see how the size of the organism affects the conversion process itself, as well as the relative amounts of product and heat generated. For example, how might the solution found initially for a human change for a mouse? A cat? An elephant? A blue whale?

In addition to a rich and relevant context, this problem requires the use of crosscutting concepts like energy and conservation of mass. A chemist would be interested in the thermodynamics and energetics of the individual chemical reactions, reactants, and products. A biologist might take a broader view of these reactions and focus on the energetic coupling of one reaction to another, say, or how the release of energy due to a specific reaction affects the organism as a whole. Conservation laws require the defining of a system, which changes depending upon what is being solved for. Mass must be conserved in a given chemical reaction, which is shown by a balanced chemical equation; an organism, on the other hand, can "lose weight" due to interactions and transfers between it and the surroundings.

This problem is also an example of a "low floor, high ceiling" problem. The problem is simple on its face—"Where does the fat go?"—and is very approachable for scientists and non-scientists, alike. This initial simplicity belies a possibly complex and intricate solution that can be extended to further applications or other related problems. For example, students could propose and then perform experiments to test the predictions generated during the problem-solving process. They could burn a sample of fat in an enclosed space and measure the amount of carbon dioxide produced or any resulting temperature changes. This then leads to another good interdisciplinary problem relevant to many disciplines—the interpretation of actual scientific data with respect to a system of interest.

## Discussion

This paper tackles the problem of what makes a good problem. STEM DBER scholars across their disciplines have spent a considerable amount of time thinking about problems, but there is certainly no consensus across the disciplines. Nevertheless, scholars have developed a

variety of different, yet complementary, frameworks that can be used to think across the disciplines. For instance, mathematics educators use the idea of cognitive demand to describe what type of thinking a task requires, whereas biologists have taken up Bloom's taxonomy. Here we benefit from the interdisciplinarity of the authors, as it provides us access to literatures and concepts that we otherwise would not be familiar with.

Although we took different lenses to the problems above, we note common features including real-world connections, reinforcement of conceptual understanding, multiple solutions paths, a low floor and high ceiling, and building dispositions of professionals in the discipline. We propose that good problems in any STEM discipline will include at least some of these features, though any individual problem need not contain all of these features. While these are similar features, they are also discipline-specific. Each discipline favors certain concepts, dispositions, and real-world connections (e.g., engineering design vs. theory-building). Indeed, as we reflect on our collaborative work through these sessions, we recognize that for an individual who does not have disciplinary expertise, it may be very difficult to tell whether or not a problem is good.

We also reflect on the extent to which students across the disciplines actually have the opportunity to engage in good problems. In physics, for example, despite wide consensus among physics education faculty on the benefits of the features of good problems, many instructors persist in assigning problems that are not well aligned with their learning goals (Yerushalmi, Cohen, Heller, Heller, & Henderson, 2010). Textbook analyses in mathematics also points to the impoverished nature of tasks for instructors to choose from. For example, Lithner (2004) found that in a mainstream calculus textbook, 90% of the tasks could be solved by superficial reasoning (e.g., relying on recall, keywords, pattern matching to worked out examples). Thus even if an

instructor wanted to assign good problems, either for homework or for collaborative work in class, finding or developing such problems can be a challenge.

Finally, creating interdisciplinary problems was a challenge for members of our sessions. Because a good deal of disciplinary knowledge is required to engage with many problems, interdisciplinary problems require deep understanding of multiple disciplines. This is difficult to achieve unless one is teaching in a truly interdisciplinary environment. Another possibility that we see is using different problem types across disciplines. For instance, programming is ubiquitous in computer science, but it could also be used in physics, integrating the two subjects better. Similarly, mathematical biology is a new area that has emerged in mathematics. While it is challenging, we remain optimistic for possibilities of truly interdisciplinary problems.

## Acknowledgements

## References

AAAS. (2011). *Vision and change in undergraduate biology education: a call to action*. Washington, D.C.: American Association for the Advancement of Science.

Allchin, D. (2013). Problem- and Case-Based Learning in Science: An Introduction to Distinctions, Values, and Outcomes. *CBE-Life Sciences Education*, *12*(3), 364–372. https://doi.org/10.1187/cbe.12-11-0190

Bloom B. S., Krathwohl D. R., Masia B. B. (1956). Taxonomy of Educational Objectives: The Classification of Educational Goals. New York, NY: D. McKay.

Bray Speth, E. Long, T. M., Pennock, R. T., & Ebert-May, D. (2009). Using Avida-ED for

    Teaching and Learning About Evolution in Undergraduate Introductory Biology Courses.

    *Evolution Education Outreach*, 2(3), 415-428.

Brewe, E. (2008). Modeling theory applied: Modeling Instruction in introductory physics.

    *American Journal of Physics*, *76*(12), 1155–1160. https://doi.org/10.1119/1.2983148

Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning

    while modeling dynamic events: A framework and a study. *Journal for Research in*

    *Mathematics Education, 33*(5), 352–378.

Chi, M. T. H, Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics

    problems by experts and novices. *Cognitive Science, 5,* 121.

Choi, B. C. K., & Pak, A. W. P. (2006). Multidisciplinarity, interdisciplinarity and

    transdisciplinarity in health research, services, education and policy: 1. Definitions,

    objectives, and evidence of effectiveness. *Clinical and Investigative Medicine. Medecine*

    *Clinique Et Experimentale*, *29*(6), 351–364.

Cohen, E. G., & Lotan, R. A. (1997). *Working for equity in heterogeneous classrooms:*

    *Sociological theory into practice*. New York, NY: Teachers College Press.

Cooper, M. M., Underwood, S. M., Hilley, C. Z., Klymkowsky, M. W. (2012). Development and

    assessment of a molecular structure and properties learning progression. *J. Chem. Educ*.,

    *89*, 1351-1357.

Cooper, M. M., Caballero, M. D., Ebert-May, D., Fata-Hartley, C. L., Jardeleza, S. E., Krajcik, J.

    S., Laverty, J. T., Matz, R. L., Posey, L.A., & Underwood, S. M. (2015). Challenge

    faculty to transform STEM learning. *Science, 350*, 281-282.

Coppola, B. P. "Do Real Work, Not Homework" In, Garcia-Martinez, J.; Serrano-Torregrosa, E., Eds. Chemistry Education: Best Practices, Opportunities and Trends. Weinheim, Germany: Wiley-VCH, 2015; pp. 203-257.

Crowe, A., Dirks, C., & Wenderoth, M. P. (2008). Biology in bloom: implementing Bloom's taxonomy to enhance student learning in biology. *CBE-Life Sciences Education*, 7(4), 368-381.

Dávila, K., & Talanquer, V. (2010). Classifying end-of-chapter questions and problems for selected general chemistry textbooks used in the United States. *J. Chem. Educ*., *87*, 97-101.

Docktor, J, Mestre, J., & Ross, B. (2012). Impact of a short intervention on novices' categorization criteria. *Physical Review Special Topics - Physics Education Research, 8*, 020102.

Enke C. G. (2001). *The Art and Science of Chemical Analysis*. New York, NY: Wiley.

Featherstone, H., Crespo, S., Jilk, L. M., Oslund, J. A., Parks, A. N., & Wood, M. B. (2011). *Smarter together! Collaboration and equity in the elementary math classroom*. Reston, VA: National Council of Teachers of Mathematics.

Grant, B. W. 2008. Practitioner Research as a Way of Knowing: A Case Study of Teacher Learning in Improving Undergraduates' Concept Acquisition of Evolution by Natural Selection. Commissioned Paper from the National Research Council Board On Science Education Workshop "Linking Evidence and Promising Practices in STEM Undergraduate Education," 30 June 2008, at the National Academies, Washington, DC

Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. Lester (Ed.), *Second handbook of research on mathematics*

*teaching and learning* (pp. 805–842). Charlotte, NC: National Council of Teachers of Mathematics.

Heller, P., Keith, R.,  Anderson, S. (1991). Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving. *American Journal of Physics, 60(7)*, 627-636.

Heller, P. & Hollabaugh, M. (1991). Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups. *American Journal of Physics, 60(7),* 637-644.

Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, *28*(5), 524–549.

Hestenes, D. (1987). Toward a modeling theory of physics instruction. *American Journal of Physics, 55(5),* 440-454.

Holme, T., Luxford, C., & Murphy, K. (2015). Updating the General Chemistry Anchoring Concepts Content Map. *J. Chem. Educ.*, *82*, 1115-1116.

Hull, M., Kuo, E., Gupta, A. & Elby, A. (2013). Problem-solving rubrics revisited: Attending to the blending of conceptual and formal mathematical reasoning. *Physical Review Special Topics - Physics Education Research, 9*, 010105.

Ikonen, A., Piironen, A., Saurén, K., & Lankinen, P. (2009). CDIO Concept in Challenge Based Learning. In *Proceedings of the 2009 Workshop on Embedded Systems Education* (pp. 27–32). New York, NY, USA: ACM. https://doi.org/10.1145/1719010.1719016

Kohl, P., Rosengrant, D. & Finkelstein, N. (2007). Strongly and weakly directed approaches to teaching multiple representation use in physics. *Physical Review Special Topics - Physics Education Research, 3,* 010108.

Lampert, M. (2003). *Teaching Problems and the Problems of Teaching*. New Haven, CT: Yale University Press.

Lave, J. (1996). Teaching as learning, in practice. *Mind, Culture, & Activity*, *3*(3), 149–164.

Layman, L., Williams, L., & Slaten, K. (2007). Note to Self: Make Assignments Meaningful. In *Proceedings of the 38th SIGCSE Technical Symposium on Computer Science Education* (pp. 459–463). New York, NY, USA: ACM. https://doi.org/10.1145/1227310.1227466

Laverty, J. T., Underwood, S. M., Matz, R. L., Posery, L. A., Carmel, J. H., Caballero, M. D., Fata-Hrtley, C. L., Ebert-May, D., Jardeleza, S. E., & Cooper, M. M. (2016). Characterizing college science assessments: The three-dimensional learning assessment protocol. *PLoS ONE* 11(9): e0162333. https://doi.org/10.1371/journal.pone.0162333

Leonard, W., Dufresne, R., & Mestre, J. (1996). Using qualitative problem-solving strategies to highlight the role of conceptual knowledge in problem solving. *American Journal of Physics, 64(12),* 1495-1503.

Linn, M. C., Clark, D., & Slotta, J. D. (2003). WISE design for knowledge integration. *Science Education*, *87*(4), 517–538. https://doi.org/10.1002/sce.10086

Lithner, J. (2004). Mathematical reasoning in calculus textbook exercises. *Journal of Mathematical Behavior*, *23*, 405–427.

Marrongelle, K., & Rasmussen, C. (2008). Meeting new teaching challenges: Teaching strategies that mediate between all lecture and all student discovery. In M. Carlson, & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate*

*mathematics education* (pp. 167-178). Washington, DC: The Mathematical Association of America.

Mayer, R.E. (2002). Rote versus meaningful learning. *Theory into Pract*ice, *41*, 226–232.

Murphy, K., Holme, T., Zenisky, A., Caruthers, H., & Knaus, K. (2012). Building the ACS Anchoring Concepts Content Map for undergraduate chemistry. *J. Chem. Educ.*, *89*, 715-720.

National Research Council. (2000). *How People Learn: Brain, Mind, Experience, and School: Expanded Edition*. Washington, D.C.: National Academies Press.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: The National Council of Teachers of Mathematics.

National Research Council. (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

National Research Council. (2011). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.

O'Kuma, T., Maloney, D, & Hieggelke, C. (1999). *Ranking Task Exercises in Physics*. Prentice Hall.

Passolunghi, M. C., & Siegel, L. S. (2001). Short-Term Memory, Working Memory, and Inhibitory Control in Children with Difficulties in Arithmetic Problem Solving. *Journal of Experimental Child Psychology*, *80*(1), 44–57. https://doi.org/10.1006/jecp.2000.2626

Quarfoot, D. (2015). The Music of Mathematics: Toward a New Problem Typology. *eScholarship*. Retrieved from http://escholarship.org/uc/item/68f901n3

Riggs, E. M., Lieder, C. C., & Balliet, R. (2009). Geologic Problem Solving in the Field:

Analysis of Field Navigation and Mapping by Advanced Undergraduates. *Journal of Geoscience Education*, *57*(1), 48–63. https://doi.org/10.5408/1.3559525

Schoenfeld, A. H. (1985). *Mathematical problem solving*. New York: Academy Press.

Schoenfeld, A. H. (1991). What's all the fuss about problem solving. *ZDM*, *91*(1), 4–8.

Sevian, H., & Talanquer, V. (2014). Rethinking chemistry: A learning progression on chemical thinking. *Chem. Educ. Res. Pract*, *1*5, 10-23.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, *27*(2), 4–13.

Slotta, J. D., & Linn, M. C. (2009). *WISE science: Web-based inquiry in the classroom*. New York, NY: Teachers College Press.

Smith, J., diSessa, A. A., & Roschelle, J. (1993). Misconceptions Reconceived: A Constructivist Analysis of Knowledge in Transition. *Journal of the Learning Sciences*, *3*(2), 115–163.

Speth, E. B., Long, T. M., Pennock, R. T., & Ebert-May, D. (2009). Using Avida-ED for teaching and learning about evolution in undergraduate introductory biology courses. *Evolution: Education and Outreach*, *2*(3), 415-428.

Talanquer, V. (2011). Macro, Submicro, and Symbolic: The many faces of the chemistry "triplet." *International Journal of Science Education*, *33*(2), 179–195. https://doi.org/10.1080/09500690903386435

Van Heuvelen, A. & Maloney, D. (1999). Playing Physics Jeopardy. *American Journal of Physics, 67,* 252-256.

Watkins, J., Coffey, J., Redish, E., & Cooke, T. (2012). Disciplinary authenticity: Enriching the reforms of introductory physics courses for life-science students. *Physical Review Special Topics - Physics Education Research, 8,* 010112.

Yerushalmi, E., Cohen, E., Heller, K., Heller, P., & Henderson, C (2010). Instructors' reasons for choosing problem features in a calculus-based introductory physics course. *Physical Review Special Topics - Physics Education Research, 6,* 020108.