Daniel Z. Meyer, Leanne M. Avery

Why Inquiry is Inherently Difficult… and Some Ways to Make it Easier

The authors offer a framework that identifies two critical problems in designing inquiry-based instruction and suggest models for developing instruction that overcomes those problems.

"I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description; and perhaps I could never succeed in intelligibly doing so. But I know it when I see it …"

—Potter Stewart
Jacobellis v. Ohio

Justice Stewart's statement regarding pornography would seem to be applicable to the current state of inquiry in the science education field. We have numerous rich descriptors of inquiry in action (National Research Council, 2000; Minstrell & van Zee, 2000), as well as robust rubrics designating levels of inquiry (Herron, 1971; Wheeler, 2000; Beerer & Bodzin, 2003). In other words, we know it when we see it. But these fall short in providing teachers with the tools for how to develop inquiry-based activities. Much of the research investigating this has focused on the structural barriers (e.g. time, resources, teacher knowledge, etc.) (Anderson, 2002; Minstrell & van Zee, 2000). This research suggests areas for policy makers and teacher educators to work on, but these descriptions fall short of actually providing guidance to teachers.

While we do not aim to produce a straightforward, "cookbook" process for generating inquiry activities, we do aim to push beyond "I know it when I see it". We feel this can be done by considering the design of inquiry activities as a problem space. By exploring what makes inquiry inherently difficult, as well as three potential models that overcome these challenges, we aim to build a framework that has heuristic power. That is, it has the potential to suggest further solutions to the particular problem of designing inquiry activities. We are aiming for a middle ground between the two current extremes: something that is more general than good examples of inquiry activities, but more specific and oriented towards creating activities than outcome descriptions of inquiry in action.

Our analysis is informed both by conceptual frameworks from science studies and by the experience of facilitating a variety of science educators in developing inquiry-based instruction. This paper can be seen as a formalization of the advice we find ourselves regularly giving science educators. Our framework can be divided into two broad sections. First, we outline a problem space component—an articulation of the challenges in designing inquiry activities. Second, we provide a solution component—a series of activity types that have the potential to resolve the challenges of the problem space.

By exploring what makes inquiry inherently difficult, as well as three potential models that overcome these challenges, we aim to build a framework that has heuristic power.

Our View of Inquiry

The National Science Education Standards describe inquiry as "the diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work" (NRC, 1996). Therefore, we draw heavily on studies of scientific practice to form our approach. Two concepts in particular have been useful and guide our further discussion.

The first concept is the notion that context matters. This is probably best encapsulated in Kuhn's (1970) principle of a paradigm: scientists operate in an existing framework that guides aspects of their work, such as

what counts as evidence. This effects how participants react to empirical evidence. Scientists from different fields that have points of overlap will approach common topics in different manners. For example, when results from neutrino experiments differed from current theory, different types of involved researchers questioned different parts of the theoretical framework (Pinch, 1981, 1985). But the paradigm is more than just a gauge by which to judge new work. It provides the impetuous and purposefulness that motivate researchers to take on new work in the first place. Individual pieces of scientific work (to the extent one can even define an individual piece) only have meaning in their specific context.

The second concept—interpretive flexibility—comes from sociological studies of the work of done to develop scientific and technological knowledge (Collins, 1981a, 1981b). This refers the situation in which differing conclusions can be made from the same set of empirical data. These situations occur frequently at the cutting edge of scientific work. In Collins' research on early gravitational wave detection, odd results could be attributed to a variety of sources, because there was, by definition, no universally accepted interpretation (Collins, 1975, 1981a, 1985). In particular, data conflicting with current theory could indicate either counterevidence to that theory or a flaw in experimental technique. The formation of scientific knowledge involves social interactions to reduce this variability in interpretation to the point that one conception wins out and becomes accepted as fact. This concept can also be applied to the development of technology. Different actors will have different conceptions of what existing technologies are, the nature of current problems, and what should be valued in potential future solutions (Pinch & Bijker, 1987).

With regard to both of these concepts, we argue that in inquiry in general, the role of argumentation is central (Bricker & Bell, 2008). The development of scientific and technological knowledge involves making substantive arguments using empirical and theoretical warrants. Indeed, we use this as a litmus test of inquiry instruction. It must require and enable students to make non-deterministic, empirically supported arguments at some point in the experience. By non-deterministic, we mean to exclude cases (sometimes found on standardized tests) where evidence points (and is often designed to point) in a clear, predictable direction.

Why is this hard?

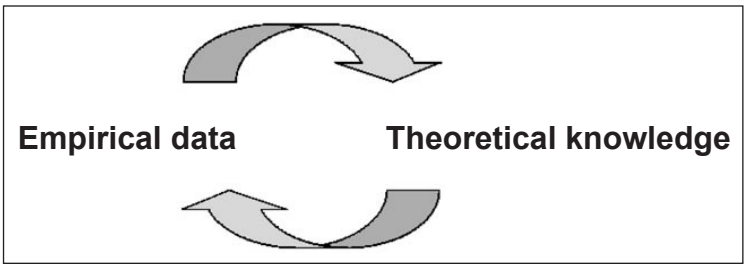
Creating the circumstances in which students can make these types of arguments often runs into two problems, which we will term the Getting on Board Problem and the Variability Problem.

The Getting on Board Problem can be illustrated by considering the simple diagram in Figure 1.

This is a very generic diagram of the development of knowledge in scientific communities. Phenomena are observed, and this generates the need to describe, organize, and explain it. This results in the development of new theoretical knowledge. This new framework suggests opportunities for new empirical observations. The new observations lead to new explanations, and so on. This is a cycle, and as member of a community, one never needs to consider a beginning. In the case of individuals, new participants begin by piggybacking on the work in progress of others, wherever in the cycle that might be. (For example, new Ph.D. students begin by working on the ongoing work of their mentors.) Once on board, they are part of the cycle, and are never working outside of a historical context.

This works in the general science community. However, in the classroom science setting, particularly at the pre-college level, things become problematic. Very few science classrooms are structured to allow junior students to benefit from interactions with more advanced students. Rather, students are grouped to all be at the same level.[1] Furthermore, students enter into scientific investigations with little or no background. Hence, no matter what part of the cycle we

Figure 1: A simple cycle of scientific work



1. We should note that typical heterogeneous classrooms do not do enough to remedy this problem. Mirroring science communities requires the interaction of individuals that are not just at different abilities but have different experiences, roles, and objectives. Having students with a variety of aptitudes but all of whom are engaging in a particular activity for the first time does not overcome the problem.

choose to inject students into, it is very difficult for it to have any meaning to them.

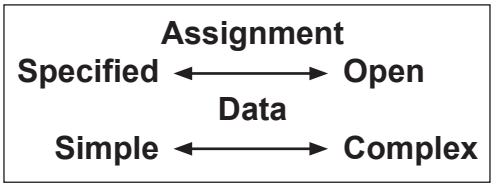
The Variability Problem stems from the need for a real argument. Making an argument means making statements about relationships. If students are only determining isolated attributes—say, the solubility of a particular chemical—there is no argument to have. Some sort of relationship among variables[2] is needed to create the tension that makes investigations meaningful. Furthermore, as noted above, we need the arguments to not be deterministic. This means having a degree of messiness to the data. There must be something there to argue over!

This requirement is also not easily achieved in the pre-college science classroom. First, the greater the technical resources of a classroom, the greater the opportunity to have data over which arguments can be had. This threshold can often be beyond the capabilities of classrooms. Second, the content most pre-college classrooms focus on is often very well established. The arguments have simply already happened. Lastly, understanding that there *is* tension or ambiguity in data generally depends on prior knowledge, leading us back to the Getting on Board Problem.

Balancing Acts

Designers of inquiry-based instruction can re-conceptualize these two problems as two balancing acts as shown in Figure 2.

Figure 2: The balancing acts of inquiry design



The first balancing act concerns the challenge or task given to students. This can range from very specific and rote to very open ended and ill-defined. Each end of the spectrum has problems. The specific end is the traditional cookbook lab, with all of its well-deserved criticism. There will be no variation in data (if done correctly). The goal is not to make a data-supported argument but to follow directions accurately in order to achieve the predetermined outcome (Amerine & Bilmes, 1990). There is nothing to argue over. But the open end of the spectrum also has problems. There is a limit on how open a task a given set of students can handle. The question is not merely whether or not the students can accomplish the task. It is conceivable that a task that students might technically not complete could still enable them to learn a tremendous amount through the effort. The problem arises when students are unable to determine how to make any movement whatsoever on a task. In other words, the task is so unintelligible to students that they cannot even proceed in a wrong direction, and therefore, also have nothing to argue over.

The role of the inquiry designer is to create a challenge/task/question that is understandable by the student as a challenge/task/question but not as a solution. Understanding an assignment as a solution means that there will be no argumentation over what is the proper solution. It is worth also noting that the right balancing point depends tremendously on the students. The challenge "build a car" is not an appropriate task for most middle school students. They wouldn't know where to start. However, this challenge might be appropriate for an advanced undergraduate mechanical engineering class. (There are other problems with this assignment that we will discuss a bit later.)

The second balancing act is on the data space that students will work with. Again, it is the extremes that are the problem, but also easy to fall into. The simplistic end is the traditional confirmation lab. The simple data is one-dimensional and conforms to the expected pattern, which students dutifully report. Since the data is straightforward, any arguments and their outcomes are predetermined.

The role of the inquiry designer is to create a challenge/task/question that is understandable by the student as a challenge/task/question but not as a solution.

Any deviation is automatically treated as a sign of poor experimental technique rather than possible support for alternative conclusions. On the complex end, you have data that is beyond the abilities of the students to collect and/or evaluate. This might be for conceptual reasons or for technical reasons. The data that can be collected with a particle collider offers plenty of opportunities for multiple interpretations—and therefore arguments—but such equipment is beyond most secondary schools!

So the task of the inquiry designer is to find (or create) a data space that

2. This does not need, however, to be quantitative in nature.

is approachable by students but also has some work for students to do. In other words, the data must be usable in terms of making arguments but not so usable that there is only one obvious argument available. Again, the right balance point depends on the particular students.

Frameworks

So what to do about these problems? We do not believe there is an easy, step-by-step way to produce inquiry instruction. Creativity and context will always be an essential element. However, in facilitating pre-service and in-service teachers' development of inquiry activities, we have noticed a pattern in instructional plans that seem to overcome the barriers we have discussed. We propose three models that inquiry designers can use to produce more inquiry instruction.

We must make two important caveats. First, we do not claim that these models are anything new. There are numerous examples of these models already in existence. Rather, we aim to put a label on them and, thus, identify how they are instances of a common phenomenon. By identifying a common pattern, we hope to provide guidance for generating new activities. Second, we do not claim that this list is exhaustive. There are certainly other sound inquiry-based activities that do not fall neatly within these forms.

Protocol Model

The Protocol Model has its origins in the Environmental Inquiry Project at Cornell University (ei.cornell.edu). A *protocol* is a well-defined procedure for collecting data. In terms of definition and clarity of steps, it is quite similar to a traditional cookbook lab. However, it is clearly portrayed as being just a tool—as opposed to the entirety of the lab experience. More importantly, a protocol can be applied to a wide variety of situations—not just the situation in which it is introduced and learned. (Hence, some cookbook labs can be adapted to form protocols but others cannot.) Once the students learn the protocol in an initial circumstance, they can then apply it to further research. This research can be more varied and more student-directed.

The prototypical case of a protocol is the lettuce seed bioassay (Trautmann, 2001a, 2001b). Students are given fairly clear directions for producing a serial dilution of a salt solution, setting up a bioassay using lettuce seeds, and evaluating the results. Once they have had that experience, they can now engage in further, more varied research: other concentration ranges, other toxins, and even other biological indicators. At the most sophisticated end of the spectrum, the bioassay can become a moderate piece in a larger extensive research endeavor.[3]

Another example of a protocol is the Watershed Habitat Evaluation and Biotic Integrity Protocol (WHEBIP) (Carlsen and Trautmann, 2004). This protocol was created to allow scientists to use models to predict aquatic biodiversity in watersheds. In this protocol, stream integrity ratings are assigned using land use criteria and can be accomplished using aerial photographs or remote sensing without requiring ground truthing (although in some instances, it is appropriate). Ratings are based on information students assess, including size of riparian belt, type of land use near stream, gradient, pollution, and conservation activity. Students can use this protocol to make a preliminary assessment of a habitat and, if desirable, make comparisons to data gleaned from ground truthing. This tool enables students to obtain data for one or multiple sites within watersheds or comparative studies between watersheds and make recommendations for remediation.

Learning a protocol is not just a question of now having a new technical skill. The student has also been introduced to a way of looking at the natural world. The dataset they produce in the initial learning round is also significant. It can be an indicator of what aspects of the phenomenon merits investigation next, just as with science at large. Hence, the student has been brought on board the knowledge development cycle.

A counter example can help illuminate the nature of an effective protocol. A common physics cookbook lab is to measure the period of a pendulum with various lengths and masses. Unlike some cookbook labs, this is not easily configured into a protocol. It fails to overcome both problems. The data produced is not likely to have any ambiguity—and any that does occur will be attributed to practitioner error. In addition, once the initial data is collected, then what? The experience will not introduce students to a new empirical realm.

Design Challenge Model

The Design Challenge Model has had more common use. Design Challenges are centered on an explicit

3. This is one way we have seen where (with a lot of work) the Getting-on-Board problem can be overcome. Seniors carrying out an extensive research project "contract out" their bioassay needs to lower grade students, providing them the opportunity to learn the basic procedures (Avery, 2003).

production task. Often the task will motivate the practical need to acquire certain knowledge bases. Sometimes inquiry designers will use a jigsaw arrangement in which students are divided into specialty groups to learn one of the applicable knowledge bases, then rearranged into design teams made up of representatives from each specialty group.

Forming the explicit charge that is given to the students is the critical and creative focal point of designing Design Challenges. Accomplishing this goal can determine if the balancing acts have been achieved. As mentioned above, a way of framing the problem is to give students a challenge that is understood as a question but not as a solution. A question for which students already have a single, preconceived solution will not generate the argument opportunity necessary for inquiry. At the other end of the spectrum, a task for which students have no conception or ability to proceed is equally unfruitful. However, it should also be noted that there is another way in which Design Challenges can be too open-ended. Consider the challenge for middle school students "design and build a paper airplane." This avoids both of the problems noted so far. Students understand what a paper airplane is, they have the intellectual and material resources to meet this challenge, and they are likely to propose multiple solutions. But then what? The litmus test we described requires *warrant*-based arguments—essentially saying "this is better because of such and such". As stated, this design challenge does not include any means to defend why one design is better than another. The challenge is *too* open—not in term of students' cognitive or technical ability to achieve it, but in terms of it being a meaningful competition. Design Challenges must include pressures that require student designers to make judgments and back up those judgments with arguments. Hence, the "build a car" challenge noted above would be an ineffective design challenge even for a group of students that could build a car. More pressure is needed.

We do not believe there is an easy, step-by-step way to produce inquiry instruction.

Design challenges often result in tangible products. One example of this is the stormwater treatment design challenge (Carlsen and Trautmann, 2004). This activity models how cities develop systems for collecting and draining runoff from storms. Using simple materials such as plastic soda bottles, tape, coffee filters, cat litter, sand, gravel, and plastic tubing, students are given the task of creating a filtering system that can handle a simulated storm event over a relative period of time. They need to take into account the various types of substances (such as chemicals, dirt, oils, etc) found in runoff, the volume of the storm event, the time between events and the extent to which the stormwater needs to be filtered. Like engineers in the real world, they are also constrained by materials, guidelines, budget, time, and design. From a curriculum design point of view, the specifics of the design constraints (size, materials, etc.), evaluation measures (ph, DO, etc.) and simulated runoff (particulate matter, oil, etc) will be what determine how the balancing acts have been achieved.

Just making something, however, does not make an effective design challenge. Construction activities can be the Design Challenge equivalent of a cookbook lab. Consider the common example of students in physics classes designing roller coaster rides. The details of the assignment are crucial in determining whether this is an effective Design Challenge. Often, students design the ride in a fairly arbitrary way, and then post facto apply physics principles to determine elements like speed. The laws of physics do provide limitations on the design (e.g. a hill can not be too high that a car will not have the energy to reach the top), but there are not competing constraints that provide for points of debate. Once a student stays within the bounds of physics, any choice is an arbitrary preference, and hence, there is no opportunity for argumentation.[4] This illustrates how the details of an assignment can have a profound effect.

Although producing a tangible object is perhaps the most common type of design challenge, we should not limit our students or our own thinking to this format. Consider the following example (Meyer, 2003). Students are given a scenario in which a community that is experiencing pollution in a local waterway. The students are divided into different constituency groups: farmers, homeowners, industry, and municipal authorities. They are given a variety of information resources—some common and some specific. They then

4. We should note that we are not arguing that such an activity is not worthwhile, but simple that it does not work as an effective design challenge.

have a variety of meetings—some in homogenous groups and some in heterogeneous groups. The task and final outcome of those meetings is to develop a restoration plan. This example creates opportunities for debate without being too open-ended, but it results in a plan of action rather than a physical product.

Product Testing Model

In general, the Protocol Model and Design Challenge Model can be seen as corresponding to scientific work and engineering work respectively. We have used these frameworks with pre-service and in-service science teachers and feel they genuinely represent general frameworks that can be utilized to inspire and guide further design of inquiry instruction. We end by proposing a third framework. It will take further work to flesh out its details and legitimacy.

The Product Testing Model is inspired in part by the Discovery Channel show Mythbusters (Rees, 2003). A common thread through much of the work on the show, and product testing in general, is the challenge to reproduce natural phenomena under lab conditions—i.e. in an intentional, controllable, measurable, and reproducible manner. In this sense, it is much like a sub-set of design challenges. But there is also a second point of contention: once results are obtained, how should they be evaluated? Consider the task of determining the best paper towel. What makes the best paper towel? How can a desired characteristic like durability be measured (in order to make clear that brand A is more durable than brand B)? And once that is done, how should durability be related to other characteristics, like price? Hence the Product Testing Model operates in two problem spaces: physically performing the relevant tests and determining criteria for success and failure. In a way, it is the combination of the Protocol and Design Challenge Models. A task generates the needs for various knowledge domains and the development of data collection routines.

Conclusion

We have put forward a framework that identifies two critical problems in designing inquiry-based instruction and suggests three models for developing instruction that overcomes those problems. The Protocol Model overcomes the Getting on Board Problem by providing students an initial experience through clearly delineated steps with a data collection technique that can be applied to a wide variety of further settings. It not only gives students a new tool, but also suggests questions to which it can be applied. It overcomes the Variability Problem by being applicable to a wide variety of settings and utilizing messy data. The Design Challenge Model overcomes the Getting on Board Problem by presenting a practical need to acquire certain knowledge bases. It asks students to understand it as a question before understanding a solution. It overcomes the Variability Problem by imposing a variety of pressures that allow students to balance competing needs in a variety of ways. Lastly, the Product Testing Model overcomes the Getting on Board Problem by centering on everyday phenomena. It overcomes the Variability Problem both through the challenge of producing the phenomena in the lab setting and through competing values.

References

Amerine, R., & Bilmes, J. (1990). Following instructions. In M. Lynch & S. Woolgar (Eds.), *Representation in scientific practice* (pp. 323-336). Cambridge, Mass.: The MIT Press.

Anderson, R. D. (2002). Reforming Science Teaching: What Research says about Inquiry. *Journal of Science Teacher Education, 13*(1), 1-12.

Avery, L. M. (2003). Knowledge, Identity, and Teachers' Communities of Practice. Unpublished Doctoral Dissertation, Cornell University, Ithaca.

Beerer, K. and Bodzin, A. (2003). Promoting inquiry-based science instruction: The validation of the Science Teacher Inquiry Rubric (STIR). Journal of Elementary Science Education, 15(2), 39-49.

Bricker, L. A., & Bell, P. (2008). Conceptualizations of Argumentation From Science Studies and the Learning Sciences and Their Implications for the Practices of Science. *Science Education, 92*, 473-498.

Carlson, W.S., & Trautmann, N.M. (2004), *Watershed Dynamics*, Arlington, VA: NSTA Press.

Collins, H. M. (1975). The seven sexes: A study in the sociology of a phenomenon, or the replication of experiments in physics. *Sociology, 9*, 205-224.

Collins, H.M. (1981a). Son of seven sexes: The social destruction of a physical phenomenon. *Social Studies of Science, 11*, 33-62.

Collins, H. M. (1981b). Stages in the empirical programme of relativism. *Social Studies of Science, 11*, 3-10.

Collins, H. M. (1985). *Changing order: Replication and induction in scientific practice*. Chicago: The University of Chicago Press.

Herron, M.D. (1971). The nature of scientific inquiry. *School Review, 79,* 171-212.

Kuhn, T. S. (1970). *The structure of scientific revolutions*. Chicago: University of Chicago.

Meyer, D. Z. (2003). Social Engagement in Curriculum Design. Unpublished Doctoral Dissertation, Cornell University, Ithaca, NY.Minstrell, J., & van Zee, E. (Eds.). (2000). *Inquiring into inquiry learning and teaching in science*. Washington, D.C.: American Association for the Advancement of Science.

National Research Council. (1996). *National science education standards*. Washington: National Academy Press.

National Research Council. (2000). *Inquiry and the National Science Education Standards*. Washington, DC: National Academy Press.

Pinch, T. (1981). The Sun-set: The presentation of certainty in scientific life. *Social Studies of Science, 11*, 131-158.

Pinch, T. (1985). Towards an analysis of scientific observation: The externality and evidential significance of observational reports in physics. *Social Studies of Science, 15*, 3-36.

Pinch, T., & Bijker, W. (1987). The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. In W. E. Bijker, T. P. Hughes & T. J. Pinch (Eds.), *The social construction of technological systems* (pp. 17-50). Cambridge, Mass: MIT Press.

Rees, P. (Writer) (2003). Mythbusters: Discovery Channel.

Trautmann, N. M. (2001a). *Assessing Toxic Risk: Student Edition*. Arlington, VA: NSTA Press.

Trautmann, N. M. (2001b). *Assessing Toxic Risk: Teacher's Guide*. Arlington, VA: NSTA Press.

Wheeler, G. F. (2000). The three faces of inquiry. In J. Minstrell & E. H. van Zee (Eds.), *Inquiring into Inquiry Learning and Teaching in Science* (pp. 14-19). Washington, DC: American Association for the Advancement of Science.

Daniel Meyers is an assistant professor, Department of Mathematics and Science Education, Illinois Institute of Technology, Chicago, IL 60616. Correspondence concerning this article may be sent to <meyerd@iit.edu>.

Leanne M. Avery is associate professor of science education, Department of Elementary Education and Reading, State University of New York College at Oneonta.