DESIGNING FOR COMPLEX SYSTEM UNDERSTANDING IN THE HIGH SCHOOL BIOLOGY CLASSROOM

Complex systems are a prevalent phenomenon across a wide variety of disciplines and are important from both scientific and educational perspectives. These systems are composed of interrelated, hierarchically organized levels; an important aspect of reasoning about them is to account for the interactions within and between these levels. Such reasoning is notoriously difficult for students of all ages. An educationally prominent domain featuring complex systems is molecular genetics. Current instructional strategies tend not to emphasize the system-oriented nature of genetics and leave students with a disconnected, incomplete and often inaccurate understanding of genetic concepts and phenomena. In our research we designed a novel instructional unit aimed at fostering a deeper and more systems-oriented understanding of genetics. We employed several design strategies to make the structure and dynamics of the genetic system a salient and central aspect of the inquiry process and the activities students engaged in. Our findings suggest that the curriculum does, to large extent, engender a deeper understanding of the complex genetic system, its components and dynamics. In particular we observed gains in students’ understanding of the hierarchical structure of the system, the nature of the genetic information and the role of proteins in mediating genetic effects.

Ravit Golan Duncan, Northwestern University
Brian J. Reiser, Northwestern University

Introduction

One of the main goals of science education is creating a scientifically literate public (American Association for Advancement of Science, 1989, 1993). What does this goal entail? According to Science for all Americans, the recent call for reform in science education, the science curriculum should focus on a few overarching themes and fundamental concepts that are essential to science literacy (AAAS, 1989). This approach is contrasted with current curricula that reinforce a superficial and rudimentary understanding of a vast amount of topics. The shortcomings of current instructional approaches in preparing scientifically literate students are particularly evident in the domain of genetics. Prior research on students’ understanding in genetics suggests that students are ill prepared to understand and benefit from technological advances in genetics that they may encounter in their everyday lives such as genetic counseling and screening (Lewis & Wood-Robinson, 2000), nor are students informed enough to understand and participate in current debates involving genetic issues such as genetically modified foods, cloning, gene therapy (Garton, 1992).
Current instruction in genetics fails to help students construct deep and meaningful understandings of concepts and phenomena in the domain. This is a result of both inadequate instructional methodologies and the complexity of the domain itself. Learning genetics is challenging because genetic processes and the resulting outcomes involve many interrelated components that exist at different organization levels comprising a larger system (Dreyfus & Jungwirth, 1999; Horwitz, 1996; Marbach-Ad & Stavy, 2000; Golan & Reiser, 2003). Moreover, these organization levels can span vast time and space scales from molecular processes lasting fractions of a seconds to changes at the organism level which can take decades (Horwitz, 1996; Kapteijn, 1990).

Understanding the genetic system entails reasoning across the different organization levels (such as genes, proteins, cells, tissues, etc) and explaining how changes to one level would affect other levels in the system. Such reasoning is a challenging conceptual task not only in the context of genetic phenomena but also in other scientific context such as: interactions in an ecosystems, behavior of molecules in a gas, dynamics of traffic jams and many others (Penner, 2000; Resnick, 1996; Wilensky & Resnick, 1999).

Given the goal of Science for All Americans (AAAS, 1989), the state of genetics education and the conceptual obstacles learners face in the domain, the need for research-based and effective instructional strategies is more pertinent than ever. We believe that instruction for scientific literacy in genetics should help students construct generative and connected understandings of the domain. By connected we mean that students should come to view genetic phenomena from a more holistic perspective and understand that these phenomena are generated by a system of interrelated components and processes. A more systemic approach to genetics will help students understand the underlying principles and mechanisms that help us explain how our genetic blueprint functions. By generative we mean that students’ understanding should allow them to reason not only about genetic problems taught in class but also about problem and phenomena they are likely to encounter outside of the classroom and throughout their lives. Providing students a conceptual toolkit for the domain of genetics will empower them to understand current and future advances in genetics.

In our research we attempted to address two central questions through the design and implementation of a project-based-based genetics unit for high school students:

1) How can we help students construct such empowering understandings?
2) What are the ways in which project-based instructional environments (Krajcik, Czerniak, & Berger, 2002) can support the goal of connected and generative understandings of genetics?

In this paper we describe the design of a genetics unit to support student engagement with the inquiry practices and conceptual model in the domain and to help students construct more systemic and connected understandings of genetic phenomena. We also provide findings from the pilot enactment and evaluation of this unit in an urban public high school biology classroom.

In the following section we will describe in greater detail the design of the genetics unit in terms of: 1) the problem/needs-analysis, 2) the specific design solutions
we employed in addressing the needs we identified, and 3) the design frameworks and cognitive theories that informed the our design. We will then turn to a discussion of an empirical study of the pilot enactment of the unit in a Chicago public high school exploring the ways in which the unit supported student learning of the target content and skills. In the final section of the paper we will discuss the lessons learned and the implications for design both specific to this unit and more broadly for the design of project-based science curricula.

**The Design of a Project-based Genetics Unit**

In his argument for the contribution of design research to theory and practice in education, Edelson (2002) describes design as “a sequence of decisions made to balance goals and constraints” (p.108). He further postulates three collections of decisions made as part of the design process. Decisions about the process and people involved in the design- design procedure; decisions about the goals and needs the design is intended to address- problem analysis, and decisions about the specific implementation and resulting design- design solution (Edelson, 2002). We will describe our design in terms of these three decision spaces focusing on the problem analysis and design solution. We will not discuss the design procedure in much detail since in our research we were primarily concerned with developing research-based and effective instruction as opposed to understanding and studying the design process itself.

**Design Procedure**

We designed the unit in the context of a collaborative team of teachers and researcher. Four secondary science teachers participated in the design, two were veteran teachers with over ten years of experience teaching and two were relatively new to teaching with less than five years experience. One of the teachers, also the second author on this paper, enacted the curriculum in her science classroom. The researcher on the team, first author of this paper, is a graduate student in a learning sciences program. The design team developed the core of the unit over the course of 3 months and further revisions and refinements took place over the course of another two months prior to the enactment.

Overall the design proceeded in several steps the specific of which will be discussed in later sections. First, we established the learning objectives and performances in terms of the target concepts and skills. In this phase we relied on Illinois State Standards and the AAAS benchmarks (AAAS, 1993) as well as the teachers’ experience content goals of the high school curriculum. Second, we chose a context for the investigation. Third, we constructed the backbone of the unit, delineating the instructional sequence. Fourth, we wrote the specific lesson plans explicating the activities that the students will engage in at all stages of the unit. Lastly, we refined the lessons plans creating coherent curriculum materials written in a consistent and clear format.
There were several types of expertise available in the composition of the design team. The teachers as practitioners brought with them practical experience in classroom management as well as pedagogical knowledge and pedagogical content knowledge. The researcher had previously studied student cognition in genetics and thus brought expert understanding of the conceptual barriers to reasoning in the domain, as well as experience in instructional design and learning theory. In addition the researcher and teachers had varying degrees of content expertise in genetics. The researcher held an M.Sc degree in molecular biology/genetics and the teachers had extensive experiences learning and teaching genetics. Thus, overall the team was well poised to engage in the design of an innovative, research-based instruction in genetics.

Problem Analysis

The problem analysis defines the design solution space in that it defines specific needs, goals and challenges the design is expected to address. In our mapping of this space we informed by prior research on genetic learning and the conceptual model in the domain (the canonical understandings in the domain). In this section we shall discuss each of these aspects. We begin, for the reader’s benefit, with a discussion of the conceptual model in genetics- the canonical scientific understandings of the genetic system and what explanations in this domain entail. We then turn to a discussion of relevant findings from prior research regarding student cognition in genetics.

The conceptual model is the representation of scientific models generated to explain phenomena in the domain. In the case of molecular genetics these models strive to explain how the genetic code brings about effects at higher organization levels. Explanations in this domain are essentially depictions of the causal mechanisms by which genes bring about observable phenomena at the organism level, such as the symptoms of a genetic disease (for an example see Collins, 1992). Investigations of genetic phenomena such as inherited diseases usually begin at the organism level, with the identification of the symptoms of the disease and then attempt to explain these symptoms and the underlying biological cause by sequential delving from higher organization levels to lower one. Thus, the next step in such an investigation would be to identify the organ and tissue that is affected followed by an investigation of the affected cells within those tissues. At this point the investigation would center on proteins, key bio-molecules responsible for many of the cell’s internal structures and functions. This is because genes, which are the ultimate cause of genetic diseases, code for proteins. Genes are essentially codes/recipes for the construction of proteins, from their building blocks- amino acids. The genetic information stored in genes specifies the type and order of amino acids in a protein (Watson & Crick, 1953). The protein’s properties and function depend on its three dimensional structure which is generated by the folding of the amino acid chain. This folding depends on the properties and interactions of the amino acids in the chain. Thus, the specific sequence of amino acids in the chain determine the three dimensional structure of the protein. Understanding how a change to the genetic code would affect the structure and function of a particular protein and subsequently the function of the cells,
tissues and organ involved in the disease, is at the crux of explanations in genetics and represent the conceptual model in this domain.

Prior research on student understanding in genetics presents several important findings that have instructional implications. As noted earlier a major obstacle to reasoning about and explaining genetic phenomena is the existence of multiple levels of organization within the genetic system. Research by Marbach-Ad & Stavy (2000) has shown that students struggle to provide explanations of a phenomenon at one organization level using concepts and terms from another level, for example explaining the inheritance of traits by refereeing to molecular mechanisms such as meiosis and recombination. This may in part be due to the unfamiliarity and invisibility of the entities and processes in the genetic system such as cells, proteins, amino acids and genes (Driver, Squires, Rushworth & Wood-Robinson, 1994; Dreyfus & Jungwirth, 1990, Fisher, 1985; Golan & Reiser, 2003; Marbach-Ad & Stavy, 2000).

This state of affairs is farther complicated by nature of levels in the genetic system. The genetic system contains ontologically distinct levels: an information level containing the genetic information and a physical level containing physical elements (such as proteins, cells, tissues, etc) that bring about the effects of the information (Golan & Reiser, 2002). The physical elements are hierarchically organized such that tissues are composed of cells that are composed of proteins, etc. Understanding genetics entails understanding how genes (containing the genetic instructions) bring about their effects. To do this the learner must understand the content of the information level (what genes code for) and how that information is mapped onto or manifested in physical elements in the system. Research by Venville & Treagust (1998) on students’ conceptions of genes suggests that students’ understanding undergoes conceptual change involving ontological shifts of a similar nature to those attributed to conceptual change in physics (Chi, 1992). They describe a trajectory of conceptual change from a conception of genes as passive particles correlated with a trait to a more fruitful conception of genes as sequences of instructions (Venville & Treagust, 1998). A conception of genes as sequences of instructions is essentially an acknowledgement that the system contains an information level, which is a key step towards developing a more expert-like understanding of the nature of the genetic information. However, in and of itself such an understanding is not sufficient since it does not define the content of the genetic instructions, that is, exactly what is coded for by genes.

Prior research suggests that students do not conceive of the genetic information as being limited to the make up of proteins rather, they attribute to genes the ability to code for more than just proteins both in terms of the structure and the function of entities at higher organization levels (Golan & Reiser, 2002). Students’ view of the genetic instructions affords the direct coding of both the structure and behavior of biological entities at many organization levels. The assumption that genes can code for phenomena at multiple organizational levels is problematic because it eliminates the need for a complete mechanistic explanation of molecular genetics phenomena. If a student assumes an inherited lung disease is caused because the responsible gene “tells” the lungs to...
malfunction there is no need for further explanation. However, the focus of scientific explanations of genetic phenomena is on the molecular mechanisms that generate the observable effect. A conception of genes as coding for the phenomenon at the observable level limits the student’s ability to both comprehend the need for a mechanistic explanation and to generate such an explanation.

Compounding the problem generated by students’ broad view of the information system is their lack of familiarity with the physical entities that are directly coded for by the genetic information (proteins) and the role they play in the mediation of the genetic instructions (Golan & Reiser, 2003). Students are often unaware of the centrality of proteins in explanations of molecular genetics phenomena and do not automatically assume that a protein must be involved in any genetically mediated phenomena. This is somewhat of a bootstrapping issue, since students possess a broad concept of the nature of the information system they are not constrained to assume that proteins must be involved since genes only code for proteins. Conversely since students are not aware of the central roles proteins play in biological processes they do not assume that any biological phenomena that is genetically based is likely mediated by proteins. Overall this hampers their ability to construct causal/mechanistic explanations of genetic phenomena that are aligned with the conceptual model of the domain.

Interestingly, Golan & Reiser (2003) suggest that students are able to reason, at least to some extent, about the mechanisms by which effects propagate through a hierarchical system. That is, they are able to explain how effects at one level of organization (for example proteins) can generate consequent effects at a higher level of organization (cells). In their study students were able to provide speculative causal mechanisms for how a change to a protein’s structure might cause it to malfunction, and how a change at the protein level would affect the entire cell and subsequently the organ in question. Thus, students were able to reason, to some extent, about structure-function relationship in molecular entities (proteins) and reason about the interactions between hierarchical organization levels.

Another important contribution to our understanding of student cognition in genetics came from a prior study conducted by one of the authors (Duncan & Reiser, 2005) that examined the role of domain-specific knowledge in understanding genetic phenomena. This work resulted in a cognitive model of reasoning in genetics that highlighted two types of domain-specific knowledge forms- heuristics and explanatory schemas, which were particularly critical for successful reasoning in genetics. The domain-specific heuristics define key components and dynamics of the system, for example the heuristic genes-code-for-proteins functions much like a rule of thumb that allows the reasoner to speculate the existence of a protein as a mediator of genetic effects. Thus, this heuristics defines genes and proteins as key components of the genetic system and the relationship of genes coding for proteins as a key dynamic in the system (in essence this heuristic helps the reasoner map the information level onto the physical level in the system). Explanatory schemas, in turn, define the possible mechanisms/interactions that are possible in the system. For example the explanatory schema inhibit defines a
prevalent and important molecular mechanism known as inhibition. This mechanism is involved in the regulation of protein activity and many medical drugs exert their effects by inhibiting (or activating) proteins in our body. Taken together the heuristics and schemas are generative and powerful domain-specific knowledge forms that enable the construction of domain-appropriate (aligned with canonical understandings in the domain) explanations in genetics.

In summary, prior research highlights several aspects of students’ reasoning in genetics that are relevant to the problem analysis at hand. First, students’ conceptions of the informational content of genes are too broad and unspecific (coding for more than just proteins). Second, students are unaware of the central (most if not all genetic phenomena are mediated by proteins) and robust (proteins carry out a multitude of functions) role of proteins in the genetic system. Third, students struggle to provide causal/mechanistic explanations of genetic phenomena that link the information in genes to observable effects at the organism level. Fourth, students are somewhat capable of reasoning across hierarchical levels in a system and are also reasonably apt at inferring structure-function correlations in molecular structures. Overall it seems that the challenges to reasoning about genetics are due to inadequate understandings of the genetic system, in terms of its key components and the roles they play (genes and proteins) as well as the relationships between the information and physical level (how genes bring about their effects). Moreover, we have some sense of what are the important domain-specific understandings that are crucial to successful reasoning in this domain and therefore should strive to organize instruction around these powerful and generative ideas. In the following section we describe the design solution we constructed to address the learning obstacles raised in the problem analysis.

Design Solution

As noted by Edelson (2002) the design solution often entails decomposing a complex design problem into more manageable components. Tarnoff & Duncan (2004) suggest that there are four facets to the design solution: 1) identifying learning objective, 2) selecting a problem/inquiry context, 3) constructing the learning sequence, and 4) constructing the learning activities in that sequence. We discuss our design solution in terms of these four facets.

Our design was also informed by three existing design-frameworks: 1) Backward Design (Wiggins & McTighe, 1998), 2) Learning for Use (Edelson, 2001), and 3) Sacffolding Design Framework (Quintana, Reiser, Davis, Fretz, Duncan, Kyza, Edelson & Soloway, 2004). We provide a brief description of each of the frameworks as it pertained to informing our design.

Backward Design (Wiggins & McTighe, 1998) provides guidelines for the design of instructional activities that foster meaningful understandings- understanding by design. There are three key elements of this framework that were useful in our design process. First, as denoted in the name, the design process should start by focusing on the learning
outcomes or goals and progress backwards to constructing learning activities to promote the desired learning outcome (hence backward design). Wiggins & McTighe (1998) emphasize the importance of defining the learning goals up front and designing with them in mind. Second, Wiggins and McTighe (1998) further suggest that to help students construct deep and meaningful understandings we need to identify a small set of enduring understandings, key ideas in the domain, and center the curriculum on them. The emphasis here is on limiting the learning objectives to a few major understandings that are most important and forgoing breadth of coverage for the sake of depth of understanding. Third, the Backward Design Framework advocates maintaining a constant and coherent alignment between the identified learning goals and the designed learning activities such that what students do directly supports the construction of the enduring understandings.

The Learning for Use framework (Edelson, 2001) provides guidelines for the design of environments that engender the construction of knowledge in context and in a way that will allow its use in relevant situations in the future. Edelson (2001) suggests that to accomplish this the design solution should consist of three parts: motivate, construct and refine. The motivate stage fosters a “need to know” on part of the learner for the target concepts/skills. That is, the activities in this stage (which often set the context of the investigation) pique students’ interest in the problem and motivate a need to understand the target concepts/skills in order to solve or understand the presented problem. It is important to note that motivate in this sense connotes more than just a motivational hook to a curriculum or “cool” problem context, rather it entails structuring the problem context such that it requires the learner to engage with the target content in order to solve/explain the problem thus creating a need to know for that content. In the subsequent construct stage students engage with learning activities that help them build the understandings already motivated by the prior stage. The final stage of refine allows students to reflect and apply their understandings such that they are refined in ways that make them more readily available for future use.

The Scaffolding Design Framework is a set of guidelines for the design of scaffolded computer-based learning environments (Quintana et al., 2004). The framework delineates guidelines for scaffolding along with the underlying rational for the guidelines and specific strategies for implementing the guidelines in software (as well as examples from existing software). While the framework was crafted with computer environments in mind we feel that it still provides useful guidelines for designing scaffolded environments that do not rely on more advanced technologies. We will discuss the specific guidelines that informed our design in relevant sections of the paper.

Learning objectives

There were two types of understandings or learning objectives we wanted students to gain from our unit: conceptual understandings about the genetic system, and understandings of the disciplinary practices in genetics. The conceptual understandings we chose were shaped by both our understanding of student cognition in genetics as
explicated in the problem analysis and by the AAAS benchmarks for scientific literacy as well as some current process and content state standards. These two sources (problem analysis and standards/benchmarks) provided a pool of potential concepts in genetics from which we selected the specific conceptual understandings we wanted students to acquire. In this selection process we carefully to limit the learning objectives to a few, key, enduring ideas as suggested by Wiggins & McTighe (1998).

More specifically we chose three enduring understandings of the genetic system as our conceptual learning objectives. These understandings reflect domain-specific knowledge forms (heuristics and explanatory schemas) that enable generative reasoning in the domain. Conceptual learning objective 1: the understanding that the genetic information specifies only the structure of proteins. This goal emerged predominantly from the finding, identified in the problem analysis, that students’ conceptions of the genetic information are too broad and that this hampers their ability to provide causal/mechanistic explanations of genetic phenomena. The idea that genes code for proteins is one of the important domain-specific heuristics characterized by Duncan & Reiser (2005). This concept is also identified as a key ideas in the AAAS benchmarks on Cells (5c) “The genetic information encoded in DNA molecules provides instructions for assembling protein molecules” (AAAS, 1993, pp 114); as well as the Illinois State Goal 12A (I) which states that students should be able to explain the molecular nature of the genetic code.

Conceptual learning objective 2: the understanding that proteins are a central and robust element in the system that mediates genetic effects. Similarly to the first objective this goals also stemmed from our problem analysis and students lack of familiarity with the important role that proteins play in the genetic system and is another critical domain-specific heuristic (Duncan & Reiser, 2005). The importance of proteins is also reflected in the AAAS benchmarks (it receives less emphasis in the state goals): “The idea that protein molecules assembled by cells conduct the work of that goes on inside and outside the cell in an organism can be learned without going into biochemical details”. (AAAS, 1993, pp 112).

Conceptual objective 3: the understanding that the sequence of amino acids in a protein affect the overall structure and properties of the protein that affords and constrains its function. The second half of the AAAS guideline regarding the role of proteins highlights the connection between the structure and function of proteins: “It is sufficient for students to know that the molecules involved are different configurations of a relatively few kinds of amino acids, and that the different shapes of the molecules influence what they do” (AAAS, 1993, pp 113). As noted in the problem analysis students’ already possess a budding understanding of structure-function correlations in molecular entities. In our third learning objective we seek to refine students intuitive ideas by providing them with an understanding of some important structural elements of proteins and their influence on protein function. It is important to note that this idea is not emphasized in traditional curricula and was mostly implicit in the state goals.
The final enduring idea we selected, conceptual learning objective 4: the understanding that genetic effects propagate through a hierarchically organized system such that changes at one level bring about changes to subsequent levels, highlights the connected and systemic nature of genetic phenomena. This idea focuses on applying a more systemic approach to understanding genetics. Systems, is one of four central themes for literacy identified in Science for all Americans (AAAS, 1989). We believe that it is an underrepresented theme in genetic instruction (and some other topics in science) that holds the key to a more coherent and meaningful understanding of genetic concepts. The shift towards a more system-oriented understanding of scientific phenomena has been advocated by many other educational researchers (Hmelo, Holton & Kolodner, 2000; Jackson, Stratford, Krajcik & Soloway, 1995; Jacobson, 2000; Resnick, 1996; Wilensky & Resnick, 1999) and scientists alike (Casti, 1994; Gell-Mann, 1994; Simon, 1996).

Along with the four conceptual learning goals described above there were two disciplinary practices we wanted students to engage with. Conducting scientific inquiry, a notion that is at the crux of current science education reforms, entails engaging in authentic problem solving and investigation in a domain. Thus students are not merely learning the science concepts in a domain they are also apprenticing in its specific investigational strategies and ways of thinking becoming more fluent in the specific practices as unit progresses (Collins, Brown & Newman, 1989; Lave & Wenger, 1991). While it is important to help students experience science as science is practiced it is a challenging endeavor since students lack tacit knowledge of domain-specific strategies possessed by experts in the domain (Reiser, Tabak, Sandoval, Smith, Steinmuller & Leone, 2001; Tabak, 1999, Tabak, Smith, Sandoval & Agganis, 1996). Moreover, traditional instruction often fails to make domain-specific inquiry practices an explicit focus of learning (Collins & Brown, 1988; Merrill, Reiser, Ranney & Trafon, 1992).

We therefore chose to make three important disciplinary practices explicit learning objectives for our unit: 1) comparing normal to abnormal, 2) continuous delving into subsequent organization levels, and 3) providing causal/mechanistic explanations that trace genetic effects throughout the various organization levels all the way to the organism level. Comparing normal to abnormal is a relatively straightforward strategy that is not unique to genetics and we suspected that use of this strategy would be relatively intuitive to students. In genetics inquiry comparison is coupled to a progressive exploration of the causal mechanism at continuously lower organization levels, moving form the organism, to the organ, tissue, cell, protein and gene. In many cases genetic diseases are first identified by their symptoms (organism level) and subsequent research reveals the underlying biological cause, thus featuring a top-down analysis of the phenomenon. Thus the disciplinary game is to explain the abnormalities in each level by postulating changes at a subsequent (lower) level. For example, a cell may not be able to take in a particular nutrient if it is lacking a protein responsible for transporting the molecule into the cell, this the observation at the cellular level (inability to take in x) can be explained by a change to the protein level (missing protein transporter for x).

This continual delving into progressively lower levels of organization is also reflects on the nature of explanations in this domain. Such explanations attempt to link...
the genetic information with observable effects at the organism level (such as symptoms). This linking essentially traces the underlying biological mechanism through the organization levels. A complete and appropriate explanation will provide a coherent “story” that addresses the changes to each organization level and how those changes affect subsequent levels. Such a story could be structured as either a top down (organism to gene) or bottom up (gene to organism) description however it must account for all the levels.

In summary, we chose a few foundational ideas in the domain as conceptual learning objectives. These ideas primarily reflect the structure and dynamics of the genetic system in terms of the key components, their roles and relationships between them. We also wanted to explicitly engage students with important investigational and argumentation practices in the domain. Given these goals, we were next faced with the choice of a problem or investigation context that would motivate the enduring ideas we would like students to learn and afford engagement with the disciplinary strategies we want them to practice.

Problem context

In choosing a context we were concerned with two instructionally important properties: motivating and authentic. We have already alluded to the importance of engendering a motivation or a need to know for the particular learning objectives as a critical component of Learning for Use (Edelson, 2001). Thus we sought a problem context that would require students to tackle the conceptual and disciplinary objectives en route to solving the problem. Such problems are often authentic to the domain of inquiry, that is, solving real-world problems naturally involves engagement with domain-specific concepts and strategies (Collins, 1996; Collins et al., 1989; Hmelo, 1998; Hmelo-Silver, 2004; Williams, 1992). Moreover, such problems help students link their classroom learning to real-world experiences making learning more meaningful (Collins, 1996; Singer, Marx & Krajcik, 2000).

We thus chose an authentic context in genetics- an inherited disease and asked students to first explain the underlying biological mechanisms and then suggest a treatment. Suggesting a fix for a genetic disease (or anything else for that matter) entails understanding the underlying cause of the problem and in genetics that would require students to explore the disease by delving into subsequently lower organization levels in an attempt to link the observable symptoms to the genetic information. Therefore this context would likely engender a need to understand/know the conceptual objectives we adopted. In addition this context allowed us to introduce students to an authentic investigation, much like the investigations of genetic diseases that scientist would conduct, that would allow them to engage in important domain-specific practices.

The specific disease we chose for the investigation was familial Hypercholestrolemia (FH) an inherited predisposition to extremely high blood cholesterol levels that elevated the patient’s risk of heart disease and death due to a heart attack.
Patients with FH have elevated blood cholesterol levels because their liver cells do not take up excess cholesterol due to a missing protein receptor for cholesterol (LDL receptor). This receptor is involved in a rather complex process by which cholesterol, packaged as LDL particles, binds to the receptor and is internalized into the cell. FH patients have a mutation in the gene that codes for the LDL receptor, the mutation can result in either a missing or non-functional receptor (due to an altered shape). We chose this disease in particular for two reasons: heart disease is a serious health problem in the US and a relatively familiar condition to students, 2) there was already an existing unit on the non-genetic risk factors for heart disease and we wanted to capitalize on students’ exiting knowledge of heat disease as well as extend the content coverage of the existing unit.

The choice to focus on one genetic phenomenon presented a problematic tradeoff. On the one hand it afforded a coherent, authentic and motivating context. On the other hand it was only one example or case that afforded limited generativity and generalizability if any. As science educators we want students to learn concepts and be able to transfer them to novel context (Bransford, Brown & Cocking, 1999). Since learning if often situated in a particular context and dependent on properties of that context (Brown, Collins & Duguid, 1989; Carraher, Carraher & Schliemann, 1985; Greeno, Collins & Resnick, 1996; Scribner, 1984), as designers we run the risk of helping students learn the specific details of a specific case without understanding the underlying concept and how it would apply to other contexts (Cognition and Technology Group at Vanderbilt, 1997). However, we rejected the idea of providing several examples to be investigated in parallel due to logistical constraints, the burden this would put on the teacher and the risk of ending up with an incoherent and confusing unit. We attempted to reconcile this dilemma by having students examine FH as the main context for the unit and then explore, in a more cursory manner, several other genetic diseases in order to suggest a potential treatment that would work for all the diseases. This compromise allowed us to maintain the coherence of the investigation by focusing on one disease, but then capitalize on multiple examples for the treatment aspect of the problem. We shall discuss the ramifications of this decision in the results sections.

Learning sequence

Constructing the learning sequence was essentially designing the backbone of the unit in terms of the order and nature of the activities students would engage in. The choice of problem context: investigating and treating a genetic disease contributed to defining the space of potential designs, clearly students would have to engage in an investigation of the disease and only then in trying to find a potential cure. Aside from the problem context there were two design frameworks that informed and constrained the sequencing: Scaffolding Design and the Learning for Use, as well as our learning objectives.

One of the scaffolding strategies suggested by the Scaffolding Design framework is “Make disciplinary strategies explicit in learners’ interaction with the tool” (Quintana
et al., 2004, pp 345). While this strategy is referring to software tools it can clearly be expanded to include curriculum materials. As noted earlier we had several learning objectives that centered on disciplinary practices in the domain. In particular we wanted students to understand that inquiry in genetics involved sequential delving into lower organization levels. We chose to make this strategy explicit in the curriculum materials by structuring the sequence of activities to reflect this notion of sequential exploration of organization levels. Thus students begin the investigation at the organism level by identifying the symptoms of the disease in twelve patient cases that they continue to work with throughout the investigation. They then try to understand the underlying cause for those symptoms at the tissue level (blood), then at the cellular level (liver cell), protein level (LDL receptor) and finally at the DNA level. We hoped that this top down and by level sequencing would make explicit the nature of investigations in this domain.

The sequencing of the unit also reflects our assumptions about what is important in understanding the genetic system- the role of proteins and the nature of the genetic code. The problem analysis suggested that students are unaware of the important role of proteins in the system and tend to ascribe to genes the ability to code for both structures and functions at multiple organization level (Golan & Reiser, 2003). Golan & Reiser (2003) suggest that this is a bootstrapping problem in that students do not have a good understanding of the molecular mechanisms (mainly at the protein level) that mediate genetic effects and thus assume that genes code for the effects themselves, at what ever level they observe them. Traditional instruction does little to alleviate this problem in its disconnected approach to genetics. Often such instruction addresses the structure of proteins and the nature of the genetic information in different parts of the curriculum. Moreover, there is very little emphasis on the function of proteins and the role they play in the genetic system. Rather, the focus is on the molecular structure of DNA and detailed understandings of the processes involved in translating the genetic code into amino acids without providing a more general understanding of the genetic phenomenon as a whole (Venville & Treagust, 1998).

In our sequencing of the curriculum we help students see how effects at higher order levels such as the organism can be explained and affected by smaller entities at lower organization levels, thereby illustrating the molecular mechanisms at play. We also refrain from explaining the role of genes and the ways in which the genetic code is interpreted until there is a clear need for an information level/layer in the system. That is, once students have explored the role of proteins in the genetic disease and how the specific structure of proteins affects their function, only then do we raise the question of what determines the protein’s structure. At this stage the link between the physical level and the information level is relatively straightforward - the protein structure is the lowest physical organization level, the only aspect left unexplained, and hence is the point of interface between the genetic information and the physical level. The need to assume that genes code for structures and functions at higher order levels is alleviated because we have already explained how effects at those levels come about.
Another important characteristic of the learning sequence is the way in which it motivates a need to understand the biological mechanism. The Learning for Use framework (Edelson, 2001) emphasizes the importance of motivating the need to learn the target concepts, however, the framework does not specify a grain-size at which this motivating should occur. That is, is it enough to choose an overarching context that will require the learning and applying of target concepts or should each lesson or activity be motivated in and of itself? We believe that, as designers, we should be motivating at multiple grain-sized levels both the overall context and specific activities should motivate learning. The overarching context motivates the need to engage with the particular content addressed by the instructional materials. However, we also need to engender a need to know for the each of the component concepts in the curriculum.

In our unit we chose to implement this idea by having questioning be the driving force for progression through the learning sequence. That is, uncovering the causal mechanism at one particular level leads to questions that to be answered require delving into the subsequent organization level. For example, figuring out that the liver cells of FH patients do not take up cholesterol (thus explaining the biological cause at the cellular level) raises the question of what in the cell is responsible for allowing cholesterol uptake. To answer this question one needs to explore the protein level since the molecule responsible for cholesterol uptake is the protein receptor for LDL. This constant questioning-leading-to-more-questioning typifies the sequencing of our unit and imparts a narrative feel to the whole curriculum. Figure 1 and 2 provide an overview of the unit, and show the driving questions for each lesson, respectively.

Lesson 1: Why do some people continue to have high cholesterol levels even though they exhibit none of the risk factors for coronary artery disease? Can these people be helped?
Students will discover an additional risk factor for Coronary Artery Disease (CAD)– the genetic risk factor. They will examine data from cases who have elevated cholesterol levels, but who do not have any other risk factors. Some cases have attempted to modify their lifestyle to eliminate the risk factor. Students will examine data regarding these patients’ lifestyles, medical information, and family trees. Their goal is to determine for each case why the individual’s cholesterol levels remain high.

Lesson 2: How is cholesterol made and controlled in healthy people and what can that tell us about people with high cholesterol?
Students will learn about the metabolism and use of cholesterol in the body. The investigation of cholesterol is driven by a student generated KWL. Thus, students will identify a need to understand the role of cholesterol in the body as part of understanding what might be wrong with people that have chronically elevated cholesterol levels. Students investigate where the cholesterol in our body comes from (synthesis and intake), what it is used for in the body (hormones, membranes, bile salts) and how it gets from food or where it is made to where it is used (LDL and HDL in the blood). Based on this knowledge students will create a model diagram of cholesterol in the body. This knowledge will help them further explore what is wrong in people who have a genetic risk factor for CAD.

Lesson 3: Why do FH patients have such elevated blood cholesterol levels? What is different about them compared to healthy individuals?
In this section students compare normal individuals with FH individuals at subsequent levels of organization: cells, organelles, proteins. Students will explore why FH patients have such high cholesterol levels by exploring the mechanism of cholesterol up-take in normal and FH individuals.

National Association for Research in Science Teaching (NARST) April 4-7, 2005
Lesson 4: What determines the number and sequence of amino acids in a protein?
In the prior lesson students learned that the LDL receptor protein in FH has an altered protein structure and that the protein’s structure is determined by the amino acids that make it up. In this lesson students will learn how the amino acid sequence of a protein, such as the LDL receptor, is determined by the genetic code. They will review DNA structure and function and learn the central dogma by associating the amino acids of a protein with genes of the DNA. Students explore how DNA is transcribed into mRNA and then translated into a protein. Students will determine that changes in the genetic code can alter the amino acid sequence and that this in turn can alter the structure of a protein.

Lesson 5: What is the biological basis for people who continue to have high cholesterol levels even though they exhibit none of the risk factors?
Students will explain the biological basis for the elevated cholesterol levels of the cases from Village Park. They will write a letter to the cases explaining the problem at all relevant levels.

Lesson 6: Is the biological basis of other inherited diseases similar, could we find a common method of treatment for all of these diseases?
In previous lessons students learned about the biological basis of a particular inherited disease. In this lesson students explore the biological basis of three other diseases using the provided resources (readings). Students apply this knowledge by creating a brochure explaining the disease they explored. In addition, students try to figure out if there is some common form of treatment that could be used to cure all the inherited diseases they have explored in the unit thus far.

Figure 1: An overview of the unit lessons. Lessons 1-4 deal with the top-down investigation of FH. Lesson 5 is a culmination of the previous lessons in which students are required to synthesize everything they have learned thus far and provide a complete explanation for the biological basis of FH. In lesson 6 students explore three other genetic diseases and suggest a treatment method that would work for all of them.

We feel that this question-driven structure is particularly beneficial for learning in two ways. First, it affords constant motivation of the target objectives, students are constantly answering and asking questions that focus on key conceptual understandings. Second, it mimics a narrative structure in that each phase, much like a chapter in a novel, sets the stage for the next part of the story. The narrative structure enhances the coherence and flow of the unit while motivating the learners’ progression through the learning sequence.
Lesson 1: Why do some people continue to have high cholesterol? Can these people be helped?
Lesson 2: How is cholesterol made and controlled in healthy people and what can that tell us about people with high cholesterol?
Lesson 3: Why do FH patients have such elevated blood cholesterol levels? What is different about them compared to healthy individuals?
  - Is the problem in FH over-synthesis of cholesterol or lack of up-take?
  - What is the difference between FH and normal cells that may explain the deficient cholesterol up-take?
  - How does cholesterol get into the cells and what is the role of the receptor protein?
  - Why do some of the LDL receptors fail to bind LDL or clathrin?
Lesson 4: What determines the number and sequence of amino acids in a protein?
  - Where is the information for determining the structure (amino acid sequence) of a protein?
  - How is the genetic information read and interpreted by the cell?
  - How is the genetic information for making proteins passed on from generation to generation?
Lesson 5: What is the biological basis for people who have inherited heart disease?
Lesson 6: Is the biological basis of other inherited diseases similar, could we find a common method of treatment for all of these diseases?

Figure 2: The driving questions for the unit. Each lesson, 1-5, has a main driving question that the students explore throughout that lesson. In addition, lessons 3 and 4 have sub-questions that guide the progression within the lesson.

In general students begin the unit by examining case files form twelve individuals all of whom have symptoms associated with heart disease. Of those cases eight represent patients with FH and four represent individual that have non-genetic risk factors for heart disease and can thus serve as controls (normal comparisons) for examining FH. Students identify the cases that have a genetic risk factor (have FH) and then proceed over the next four lessons to explore the biological basis for FH. The culmination of this investigation (lesson 5) entails writing a letter to one of the patients (at this point each students is assigned a case) and explaining to them the nature of their disease. In the last activity of the unit students explore three novel genetic diseases (lesson 6). Essentially they are provided with readings about those diseases that touch upon the manifestations of this disease at each organization level. The students have to piece together the story and drawing inferences that link the different levels and generate a coherent and comprehensive explanation of the biological basis for each disease. Finally examine all the diseases and drawing on their commonalities suggest a hypothetical method of treatment that has the potential to work for all genetic diseases. In the next section we will provide illustrative examples at a more specific level for some of the unit’s lessons.
Learning activities

Before we describes some of the specific activities designed for this unit we feel that it would be beneficial to define what we are referring to as an activity. The unit is divided into six lessons; however, these lessons do not necessarily correspond to a single class period. Lesson 3 and 4 for example, are designed to encompass five (50 minute) class periods. Within each lesson there are several activities that students engage in to uncover the biological cause of FH at a particular organization level.

The activities often begin with a brainstorming discussion in which students reflect on their findings from the previous activity and suggest paths for further investigation in terms of which questions and relevant data should be examined next. The teacher, based on her knowledge of the following activities, guides these discussions steering the students in the direction taken by the next activity. Our goal was to create the illusion, for the students, that they are driving the inquiry by letting them figure out what the next step should be. Yet the teacher navigates these discussions so that they converge on the next step in the designed sequence in terms of which type data should subsequently be explored. Next students work in groups to analyze and compare biological data from the 12 case files (which include both normal and FH affected individuals). The activity typically ends with a brief discussion that summarizes the findings from students’ analysis of the data and what they mean in regard to the biological cause of FH at that organization level. Thus an activity can be an entire lesson or part of it and is a coherent data-centered exploration of the phenomenon at a particular organization level.

To give the reader a sense for what such an activity entails and what were some of the design considerations involved we will focus our discussion of some of the learning activities in lesson 3 and 5. We chose to focus on lesson 3 because it deals with some really challenging ideas in the domain and is one of the longest lessons. We chose lesson 5 because it illustrates the type of performances we wanted students to achieve as a consequence of using this curriculum. We believe that activities in these lessons are good examples of the ideas and practices we wanted students to grapple with in our unit and the ways in which we scaffolded their engagement with these ideas and practices.

One of the main learning objectives of the unit was to help students understand the role of proteins in the genetic phenomena (conceptual objective 2). In lesson 3 three of the four activities explore the role of proteins in FH. During these activities students discover, by exploring membrane composition data, that FH cells have an altered protein in their cell membrane. They then learn through a reading and subsequent class discussion about the structure and function of the missing protein and its role in cholesterol up-take. Lastly, they examine composition of this protein and uncover critical amino acids that, when altered, can result in a malfunctioning protein.

There are several important aspects to these activities. First they all feature a comparison between normal individuals and FH patients. Understanding that investigations in this domain entail comparisons of normal to abnormal was one of the
disciplinary practices we wanted to students to learn. In accordance with the Scaffolding Design strategy that advocates making disciplinary strategies explicit in learners’ interactions with the tool/curriculum (Quintana et al., 2004), every data set that students receive includes data from normal individuals and FH patients. Thus, in each of the activities described above students engage in comparisons of normal and abnormal in an attempt to discern important differences that may suggest an underlying cause. An example of such a data set is shown in Figure 3 depicting the protein composition data of cell membranes from FH and normal individuals. By inspecting this data students are expected to notice that the lower bands differ in their relative location and that in some cases are missing entirely. The significance of these findings is later discussed with the whole class.

In addition to helping students understand the role of proteins in the genetic system we also wanted to refine students understanding of structure-function correlations in these molecules (conceptual objective 4). In particular we wanted students to understand that protein functions are mediated through physical interactions with other molecules (for example binding to LDL affords bringing it into the cell). Moreover, the protein’s ability to interact with other molecules depends on the three-dimensional configuration of the amino acid chain.

![Figure 3: Data taken from a protein electrophoresis experiment in which the sizes of proteins from the cell membranes of normal and FH derived cells are compared. Each band represents a protein of a particular size, bands that are higher (towards top) of the gel represent larger proteins and bands at the bottom represent smaller ones. Vertical displacement of bands indicates a change in a given protein’s size. Note the differences in alignment of the lower bands in the gel (in rectangle) as well as the absence of a protein band (missing protein) in the middle two cases (in oval).](image)

We therefore designed an activity in which students compared the structure and function (in terms of activity level) of the LDL receptor protein (the protein involved in...
cholesterol uptake). There are many ways of representing proteins. In figure 3 proteins are represented as bands on a gel; Figure 4 illustrates five additional ways in which proteins are represented.

Given these different representations of essentially the same element we were faced with a choice regarding which representation to use. Our choice in this case was informed by one of the strategies presented by the Scaffolding Design framework which states” provide representations that can be inspected to reveal underlying properties of data” (Quintana et al., 2004, pp 345). Each of the representations in figure 4 highlights a different aspect of the proteins structure, some of which were important to us and some of which were not. For our purposes we wanted students to understand that proteins are chains of amino acids and that these chains fold in certain ways to create the protein’s three-dimensional structure that in turn affects its function (AAAS, 1989).

Figure 4: Five different representations of protein structure found by searching Google/images using the key terms “protein structure”. All of these images depict a protein and highlight some aspect of its structure. The lock and key model (top left) presents the protein as a filled square with a jagged end, whereas the representation below it (bottom right) presents the protein in a ribbon-like structure. The representation on the top-right is somewhat of a mix of the previous two representations as it shows the amino acid chain with an overlay of the filled globular structure.

Students had learned in the previous activity (a reading and class discussion of the function of the LDL receptor) that the LDL receptor binds two molecules: the LDL particle containing cholesterol which is outside of the cell, and clathrin- a protein inside the cell that is involved in the internalization of the receptor. We therefore wanted our representation to readily reveal these functions and the structures that afford them.
Figure 5 depicts the representation we used in the protein structure data set. For each case students were provided with information regarding the structure of the protein as well as its ability to bind LDL and clathrin (binding activity). Note the inherent comparison structured into the data set (by providing data of normal and FH derived proteins). The representation of reveals the chain-like structure of the protein by depicting it as a continuous line that is folded onto itself creating a particular 3 dimensional structure. This structure affords a particular function, for example the two indentations at the top of the protein afford binding to the tooth-like structure of LDL particle allowing a puzzle-like fit between them. Moreover, the representation also shows the cell membrane thus reminding students that the protein is embedded in the membrane a part of it sticking out of cell and another part sticking into the cell (in the cytoplasm), this affords binding to extra-cellular material (LDL) and “pulling” it in.

Our choice of a particular representation of protein structure represents design decisions about the epistemic fidelity of the scientific knowledge we wanted students to engage with. Roschelle (1990) discusses the tradeoffs involved in engaging students with scientific models (simulations and representations) that vary in their complexity and accuracy. He suggests that while science experts have at hand powerful, theory-based models that can accurately represent a set of phenomena in a domain these models are often confusing and too complicated for students to work with. Roschelle (1990) concludes that it is sometimes educationally prudent to forgo some of the complexity and accuracy inherent in the expert model in lieu of simpler, learner-friendly models that will allow students to engage with important conceptual issues in the domain without over burdening them with difficult to interpret models. We agree with Roschelle’s perspective and therefore chose depictions of proteins not generated by state-of-the-art renderings of protein structure but rather simpler caricature-like depictions that are less accurate but easier for learners to interpret and analyze.
Figure 5: case data for protein structure and function. Data for two of the 12 cases are shown. For each case we provided students with information regarding the protein activity in terms of LDL and clathrin binding, two important aspects of the LDL receptor’s function. We also provided a representation of the protein showing the structural element/part responsible for binding the LDL particle (top circle) and clathrin binding (bottom circle). The middle section of the protein (middle circle) transverses the membrane (labeled).

Scaffolding the types of data representations students work with is not enough to support analysis and sense making. Quintana et al., (2004) suggest that as designers we should “provide reminders and guidance to facilitate articulation during sense-making” (pp. 345). We implemented this scaffolding strategy primarily in the form of worksheets that accompanied the data sets students received during the investigation. These worksheets prompted students to identify patterns in the data by having them categorize the cases based on the differences between them, to reflect on the significance of the identified patterns in terms of the underlying biological cause of FH, and to connect their
current findings to the overall “story” of FH. Figure 6 illustrates three questions from the worksheet accompanying the protein-structure data set shown in figure 5.

1) For each case provide the following information:
   - Patient’s name
   - Is there LDL binding activity normal: yes or no
   - Is their clathrin binding activity normal: yes or no
   - Describe any structural abnormalities in the LDL receptor protein

2) Sort the cases into categories based on the data you have about their LDL receptors. List the categories and which cases belong to each category.

3) What are the differences between the LDL receptor (structure and function) of patient Luis Cortez compared with patient Greg Williams. Which patient has more severe FH? Why?

Figure 6: Questions from the worksheet associated with the proteins structure data set. The first question asks students to note the main aspects of the data provided. In the second question students are asked to sort the cases thereby beginning the process of identifying important patterns in the data. The third question asks students to compare two of the cases and entails explaining how a difference in the structures of the proteins from the two would affect the function and overall severity of FH. This question requires students to apply their understandings of protein-structure and function in explaining a genetic phenomenon (differential severity of symptoms).

Lastly, we want to briefly discuss the design of lesson 5, the culmination of the FH investigation. We have already shown how we scaffolded students’ engagement with the disciplinary practices of comparison and sequential delving into subsequent organization levels by making these ideas explicit in both the sequencing of the activities and the activities themselves. There was a third disciplinary practice we wanted students to learn, namely that explanations in this domain are causal mechanistic depictions that stitch the events in each of the organization levels into a coherent story. This practice was mainly scaffolded in lessons 5 and 6, in which students were required to generate causal/mechanistic explanations of both FH (lesson 5) and several other disease (lesson 6). We supported student engagement with this practice by implementing yet another scaffolding strategy from the Scaffolding Design framework (Quintana et al., 2004). The strategy we used was “make disciplinary strategies explicit in the artifacts students create” (pp 345). We have already discussed in this paper a similar strategy emphasizing explicating disciplinary strategies in students’ interactions with the curriculum. These scaffolding strategies complement each other in that students are “surrounded” with explicit attention to disciplinary practices in every aspect of their learning. We made disciplinary practices explicit in the sequencing of the unit, as well as in specific activity
structures. In lesson 5 we also made these practices explicit in the product (artifact) we asked students to create, namely a letter to one of the patient cases detailing the biological basis for their condition. In writing this letter we explicitly asked students to address each organization level as well as explain the differences between normal and FH at each level. Figure 7 shows the instructions students were given for this activity.

Note to student:
In this letter, you need to use your case’s background information to explain why they have high cholesterol. In your explanation, make sure you include a description of the problem at all the relevant levels. For each level explain what is different about normal versus sick individuals for that level:

- Trait/Symptom level
- Tissue/Organ level
- Cell level
- Protein level
- And Gene/DNA level.

Figure 7: The instructions students received regarding the culminating activity in lesson 5. The worksheet included these instructions as well as a template for the letter with specific slots/boxes for each organization level.

The worksheet also included a template for the letter in which space was allocated for the explanation of the mechanism at each organization level. In this way we hoped to make salient to students the nature, in terms of form and content, of explanations in this domain.

Taken together the instructional sequence and the structure of the unit’s activities elucidate the important disciplinary practices in the domain, namely- comparisons of normal and abnormal, sequential delving into lower organization levels and causal/mechanistic explanations that transverse all the organization levels. Moreover, the instructional sequence deals with the genetic system as a whole. The genetic phenomenon of inherited diseases is explored at all levels of the system in a connected and coherent way that emphasizes the role of proteins and the nature of the genetic information.

Up to this point we have discussed our design of the genetic unit in terms of the problems we attempted to address, the learning objectives we identified and the design solution we implemented. We have shown our rational for choosing an investigation context, an instructional sequence and some of the particularities of specific activities. In the remainder of the paper we will discuss the pilot implementation of this unit in a high school biology classroom and the findings from the evaluation study of this enactment.
Methods

Study Context

The curriculum was piloted by one of the teachers on the design team (second author) in her biology classroom at a science magnet public high school in an urban setting. The teacher taught biology to four classes: one regular 9th grade biology and three honors 8th grade biology. The curriculum was identical for all classes. The enactment lasted 5 weeks during late fall and early winter of 2003-4, and closely followed our intended design. The first author was also present during most of the lessons to support and videotape the instruction.

Data Collected

We focused our data collection efforts on the 9th grade regular biology class since the curriculum was designed for an introductory regular biology class. While the 8th grade honor students were also taking biology for the first time these students represented a different audience than originally intended by the design. We collected four types of data: 1) pre and post written assessments, 2) pre and post clinical interviews (Ginsburg, 1996), 3) artifacts of students work, and 4) video of classroom instruction documenting the enactment of the curriculum. Pre and post interviews were only conducted with a subset of the 9th grade students. The artifacts and classroom video were collected from the 9th grade class as well as one of the 8th grade classes (we chose the 8th grade class that was taught immediately prior to the 9th grade class for logistical simplicity), and pre-post written assessments were collected from all 4 classes.

We conducted pre and post interviews with 11 of the 9th grade students who were selected based on the teacher’s recommendations and were representative of the 9th grade class in terms of gender, ethnic background and a range of abilities. Interviews took place during a 2-week period before and after the five-week instruction period. Interviews lasted about 30-45 minutes and were composed of three tasks aimed at eliciting students’ reasoning about molecular genetic phenomena particularly as it pertains to their understandings of concepts targeted in the learning objectives (understanding of the genetic system).

The first interview task was similar to concept-mapping in that students were given cards each with one of the following terms: cell, organ, genetic disease, cellular respiration, trait, gene, DNA, nucleotide, amino acid, hydrophobic, enzyme, proteins, digestion; and were asked to make as many connections as they could between these terms. Our goal in this task was to see what students understood about these terms, the level of connectedness of students’ understandings and whether their understandings changed post instruction. During the second task we asked students to provide a hypothetical mechanism to explain an observable genetic phenomenon described in a brief written scenario. The scenario referred to a phenomenon called chemotaxis in which...
bacteria can sense their environment and respond. We wanted to determine the extent to which students could offer hypothetical explanations of how this genetic process is mediated. In the third task students were asked to put together a “story-in-pieces” using nine statements that describe the processes and interactions of information and physical elements at the different organization levels of the system. While much of the biological details were provided in these statements students still had to compile the story by inferring the causal connections between the different levels of the system. The tasks were structured such that students were required to provide an underlying causal mechanism for a phenomenon either hypothetically or using partial information that was provided to them as part of the task. This structuring of the tasks allowed us to ascertain students’ abilities to reason about and explain genetic phenomena.

In addition to conducting comprehensive interviews with a subset of students we also administered pre and post written assessments composed of short answer and multiple-choice questions. The questions were aimed at eliciting students’ understanding of the components and interactions/mechanisms of the genetic system. For example students were asked to explain what genes are, what is the connection if any between genes and proteins, what proteins do in our bodies, how genes cause genetic diseases etc.

Data Analysis

Our general analysis approach is predominantly content analysis of qualitative data. However, we also conducted quantitative analysis of a subset of the data sources to obtain a broader view of prevalent trends and patterns in the data as well as to triangulate findings from data sources with a smaller subject pool (such as interview data). To enhance the trustworthiness (Guba & Lincoln, 1989, Lincoln & Guba, 2000) and quality of our analysis we employed several methodological and presentation strategies: 1) collection of multiple data sources and triangulation of findings, 2) establishing inter-coder reliability for written assessment and artifact data, and 3) providing detailed descriptions of our instructional design, methodology, and data corpus. By presenting multiple illustrative examples of the “raw” data we make it possible for the reader to make their own judgments regarding the credibility of the analysis and transferability of the findings (Geertz, 1973; Guba & Lincoln, 1989).

In our analysis we differentially used the data sources to determine different aspects of students reasoning (about the genetic system and the disciplinary practices involved in investigating it) and to triangulate our findings. We looked at the video of classroom instruction and the students’ artifacts (explanation of the biological basis of FH-culminating activity of lesson 5) to examine student engagement with disciplinary practices of comparison and explanation of genetic phenomena. In the analysis of the video data we looked at group and whole class discussions in several key lessons of the unit and searched for episodes in which students were engaged with the disciplinary strategies as was evident by their talk. We present some of these episodes as illustrative examples in later sections of the paper. We analyzed the artifacts to determine students understanding of the nature of explanations in this domain, that is, we looked to see
whether students’ explanations referred to all the organization levels of the system, connected events at one level to their effects at a subsequent level and were an accurate depiction of the mechanism at each level.

To analyze the pre-post interview data we used qualitative methods for analysis and quantification of verbal (oral and written) data (Chi, 1997). In this analysis we attempted to reveal the ways in which students reason about: 1) the information level and content of the genetic code, 2) the physical level and centrality and robustness of proteins, and 3) the mechanisms that mediate interactions across organization levels in the system (mainly interview data). We began by identifying reasoning episodes in the interview data (this was also applied to students’ responses to the short answer questions in the survey data). Episodes were then coded based on the ideas/topic of reference in the episode, coding categories reflected relevant ideas regarding the structure and dynamics of the genetic system, for example ideas about the information level (genes), the various physical components of the system such as proteins and interactions between levels. Following the content analysis we attempted to characterize the types of reasoning and conceptions expressed for each topic/idea; as part of this characterization process we also compared the notions students expressed to canonical understandings in the domain.

The analysis was an iterative process in that we began by using coding categories of student conception that were identified in prior research and refined, as well as added, categories based on emerging themes in the data.

Analysis of the written assessment data proceeded in a similar fashion. Students’ responses to test questions were categorized using coding schemes that were informed by prior literature (Golan & Reiser, 2003; Venville & Treagust, 1998) and reflected emerging themes in students reasoning. To establish inter-coder reliability a fifth of the written assessment (a mix of both pre and post assessments) were independently coded by an additional researcher. The inter-coder reliability rate obtained in this process was between 83-87%, and any differences in the coding were resolved. To assess the statistical significance of the observed changes in response frequencies (across categories) between the pre and post assessments we conducted a one-tailed z test of proportions, significance at the p<0.05 level is noted in the results section where relevant.

Results and Discussion

We present our results in two sections. In the first we provide a brief description of the curriculum implementation as it unfolded in the real-class context. In the second section we evaluate what students’ learned in the unit particularly in terms of the central learning objectives of the unit. Lastly we discuss the implications of our findings both specifically in terms of the design of this unit, and as broader lessons we learned about designing for systemic understanding from this design-research presented herein.

The Implementation of the Genetics Unit

Overall the curriculum implementation followed the intended design closely both in spirit and in terms of carrying out activities following the instructional sequence. There
were, however, several instances during the enactment in which activities did not unfold as planned or changes were to the lesson plan (changes were agreed upon by both the teacher and researcher). The first instance occurred during the opening discussion in lesson 2. In this brainstorming discussion students were asked to suggest the next step for the investigation given what they have found in the previous lesson—there is a genetic risk factor for high cholesterol (for some cases heart disease runs in the family). In our design we anticipated that students would suggest two potential moves: investigating genes/DNA of the patients, and investigating the cause for the persistent high cholesterol. The latter is the course of action we wanted students to pursue in the unit, we hoped that during this discussion both options would be raised and the teacher would steer students towards investigating the cholesterol connection first. What actually occurred in class was that students did not suggest investigating the high cholesterol and the teacher had to “push” the students in that direction by suggesting that there may be an easier solution than fixing the genes but that we won’t know it unless we figure out what is wrong with the patients’ bodies. Eventually after such prompting the students suggested investigating why the FH patients have such high cholesterol. Thus, despite our assumptions, students did not perceive the need to explore the high cholesterol rather were more interested in “fixing” the gene even when they didn’t know what genes do exactly.

The second instance occurred during lesson 3, we had originally planned to use a software tool for modeling protein folding called the Molecular Modeler (Tinker, 2000a; Tinker 200b). However, the computer room in the school turned out to be non-conducive to small group work around a computer. We thus designed an alternate paper and pencil activity to help students understand the ways in which the properties of amino acids in a protein affect its three dimensional shape. In this activity students read a short description of a phenomenon in involving a receptor protein and were asked to design the protein based on given specs for its structure and function (for example the protein must transverse the membrane and bind a positively charged molecule inside the cell, etc). We provided students with a list of amino acids divided into a few categories based on their properties (charged, hydrophobic, hydrophilic, etc) and asked them to construct a schematic of the protein showing the specific amino acids in the chain (amino acids were marked as small circles). This activity was actually rather successful in engaging students and helping students understanding structure-function correlations in proteins. However, it was somewhat of a break in the flow of the curriculum since the activity did not relate to FH.

The third instance occurred during the final discussion of the unit, the one in which students suggest a fix for FH and other genetic diseases. This discussion ended up being a rather anticlimactic end to the unit. Students readily suggested that one could alter the genes of affected individuals or give them the correct protein however these conclusions seemed rather obvious and mundane and since students did not know how one could apply these solutions in practice there was not enough material/content for discussion. Perhaps a better culmination of the unit would be an activity in which students manipulate DNA (there are several commercial educational kits for extracting...
and manipulating DNA) and experience some of the tool scientists could potentially use to implement their hypothetical solutions.

Despite these inadequacies in the curriculum and enactment we feel that overall the curriculum promoted a motivating and fruitful learning environment. While we do not have specific measures of engagement and interest, it was our impression based on our observations of students’ work and participation, students found the context and problem motivating and were attentive during class. A large proportion of the students participated in whole class discussions throughout the unit and students were collaborative and reasonably on task during group work.

We were particularly impressed with the quality of participation in both the group work and whole class discussions in Lesson 6, the final lesson in the unit. In this lesson students were given readings about three other genetic diseases the readings included descriptions of the diseases at each organization level but none of the connecting inferences were provided (thus students had to infer how changes described at one level could bring about the changes described for the subsequent level). At first students worked in groups to explain, based on the provided readings, the underlying biological mechanism for one of the diseases (each group received a different disease), then student groups were split and new groups were formed (jigsaw arrangement) that included an “expert” for each disease from the previous groupings. Students shared their explanations of diseases such that each group member understood all three diseases and finally these diseases were discussed in the context of the whole class. In their discussions, both in the group and whole class contexts, students were engaging with the key ideas of the unit and the task at hand demonstrating an ability to provide domain-appropriate explanations of genetic phenomena.

The quality of students’ engagement with the task and the target content is exemplified in the following two episodes taken from the 8th and 9th grade classes. In the first episode four 8th grade students are working in a jigsaw group, each taking a turn at explaining one of the diseases to the rest of the group, here Amber is explaining the genetic disease Cystic Fibrosis to her group. Note that her explanation accounts for all the levels of organization and that it is a coherent story of how a change to the protein (that originated at the with a point mutation at the DNA level) can bring about the symptoms (coughing) of the disease. The other students in her group are actively participating in this activity and even ask Amber to elaborate on her initial explanation.

Amber: At the organ/tissue is lungs and the cells are the lung cells and the membranes of those cells have a messed up protein called CFTR
Geoff: Wait its called CFTR
Amber: Abbreviated. And then the protein is missing an amino acid called phe
Geoff: Phenylalanine (make fun of name)
Tamara: And the DNA
Geoff: DNA is deletion obviously
Amber: No,
Geoff: Really?
Amber: Its point cause it only changes one amino acid
Geoff: You are right
Tamara: OK, umm symptoms
Geoff: Wait hold on, Amber you have to teach us a bit about it so I can describe it if she (teacher) asks me
Amber: This thing is the CFTR (shows diagram of proteins) and it goes on membrane and wait lets start over. Um the cells have to let salt in because when the salt comes in water come has to come in to dilute it right and while it dilutes the salt it dilutes the mucous and the mucous is not sticky, this thing is what lets the salt in but since it has the wrong amino acid it doesn’t let the salt in so water cant get in to the cell to dilute the mucous and the mucous is all sticky and gets all in there and you (makes coughing noises)

The second episode is taken from the 9th grade whole class discussion of Duchene’s Muscular Dystrophy (DMD) another inherited disease the students investigated. The discussion lasted approximately 15 minutes and 14 of the 25 students verbally contributed to the discussion. In the excerpt shown below five different students participate (some are called on by the teacher but most volunteer) in the course of less than two minutes; with the teacher’s guidance they weave together the explanation underlying DMD paying particular attention to the role of proteins in mediating this genetic disease.

Teacher: Ok could be something like that or a change in charge. OK let’s talk about DMD, Janet you had it so I will ask Bill to see how good of a teacher you were. Bill what’s DMD like
Bill: Umm, problems walking your muscles keep breaking down you have muscle loss
Teacher: What tissue is it then
Students: Muscle tissue
Teacher: Jesus could you tell us a bit about the cells of people with DMD
Jesus: Yeah, The DMD muscles tear up because of cracks in the membrane
Teacher: OK good, so you have a cell and it tears forms a break, how come it breaks with DMD
Rick: Because dystrophin it’s the binding protein, no it’s the shock absorber protein so if you run or do anything the muscles tear
Morris: Its like a spring, its like as if you jump on a bed then the cushion protects you
Teacher: Right we have a dystrophin spring in our muscles so we do so our muscles are OK there is the shock absorber but if that shock absorber…. what happens to the protein?
Gwen: Its mutated, the [amino acid] sequence is wrong
Teacher: So is it the same size
Students: No
Teacher: So it’s either missing or broken or smaller so what happens when these people use their muscles? They crack there is no cushion. Right.
Morris: There is a leak
Teacher: Right there is a leak in the membrane and then what happens to the muscle cell
Morris: It dies

Thus, overall we feel that the curriculum was relatively effective at engaging students with the problem context and contributed to the construction of a classroom environment in which students participated and were active agents in their learning. We next turn to a discussion of our findings regarding the ways in which the curriculum supported students’ learning of the target content.

What did Students Learn?

Our goal for this curriculum was to support student construction of generative and systemic understandings in genetics. Towards this end we identified as our learning objectives several domain-specific understandings and practices that are at the crux of reasoning about and explaining genetic phenomena. In this section we examine students’ understandings of these key ideas and practices and their ability to reason generatively about genetic phenomena.

Students’ conceptions of the genetic information

A key aspect of reasoning in genetics entails understanding what the genetic information specifies, what genes code for. Helping students understand that the genetic information merely specifies the structure of proteins was a central objective of the unit. In our design we motivated this understanding by having students first investigate the underlying cause of the genetic disease down to the protein level, such that the only question remaining was what determines the structure of proteins, at which point we introduced the genetic information into the mix. Thus, providing the molecular mechanism onto which the information can be mapped bootstrapped the understanding that genes only code for proteins.

We analyzed the written assessments and interviews to gain an understanding of students’ conceptions of the genetic information before and after instruction. Two of the written assessment questions were particularly informative in this respect. The first question was the simple prompt “what are genes?” and in the second we asked students to explain how a gene can cause a disease. Students’ responses to the first question before and after instruction are shown in Table 1. In our analysis of this question we wanted to ascertain whether students perceived genes as having information of any sort and if so what did they think the information was about. Venville & Treagust (1998) describe a trajectory of conceptual change for the concept of gene (as perceived by high school students) from conceiving of genes as passive particles that are associated with traits to
perceiving of them as active particles that control traits and finally to conceiving of genes as instructions for traits (genes carrying information).

We found a similar trend of conceptual change in the data, prior to instruction a large proportion of the students (40%) conceived of genes as passive particles associated with traits and a smaller proportion of students (24%) conceived of genes as containing information (instructions view of genes). However, after instruction this distribution changed, there was a significant decrease in the percentage of students who conceived of genes as passive particles (14%) and a large increase in the percentage of students who perceived of genes as carrying information, moreover the majority (68%) of the students claimed that the information specified/coded for proteins (these changes were statistically significant at the p<0.05 level based on a one-tailed z test of proportions). Thus, not only did students come to perceive genes as containing information, the most sophisticated view of genes according to Venville & Treagust (1998), they made significant gains in terms of what precisely the genetic information specifies.

**Table 1: Students' responses to the written assessment question: What are genes?**

<table>
<thead>
<tr>
<th>Response Categories (What are genes?)</th>
<th>Pre (%)</th>
<th>Post (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inherited: we get genes for disease from our parents, they are a unique combination we get, etc</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>DNA segment: genes are a segment/piece/section of DNA.</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Genes code for proteins: genes code for proteins, etc</td>
<td>1</td>
<td>68</td>
</tr>
<tr>
<td>Instructions: # genes say what we will be like. Genes are like a program/map/blueprint of our traits/features, etc</td>
<td>24</td>
<td>6</td>
</tr>
<tr>
<td>Active: # genes determine/control/command our traits</td>
<td>22</td>
<td>7</td>
</tr>
<tr>
<td>Passive: # genes make up our traits/features, etc</td>
<td>40</td>
<td>14</td>
</tr>
<tr>
<td>Total student responses</td>
<td>N=105</td>
<td>N=111</td>
</tr>
</tbody>
</table>

#- Based on Treagust & Venville (1998) conceptual change trajectory for the concept of gene

Students’ responses to the second written assessment question (Table 2) fell into six categories. Prior to instruction the majority (61%) of the students provided explanations that emphasized the inherited nature of genetic diseases claiming that if the parents have a genetic disease the offspring will likely get it too. This sort of response doesn’t not actually answer the question at hand since it merely reiterates information given in the prompt itself (the full question was: Assume a person inherited a gene for a particular disease from his parents. This person is now also sick with the same disease his parents had, explain how a gene can cause a disease?). Some (19%) of the students went further to suggest that the gene has information about the disease, in a way it “tells” the body to be afflicted. However, there was no elaboration regarding how the information is carried out, moreover it seemed that students who wrote such responses did not conceive of genes as coding for proteins or understand that diseases are due to a mistake (mutation0 in the genetic information. A smaller proportion (12%) of students did seem to understand that genetic diseases are due to abnormalities or mutations in the gene (malfunctioning or wrong information), yet again these students did not elaborate on the mediating mechanisms. This distribution also changed significantly after instruction. A
much smaller proportion of the students explained the disease as being inherited from the parents (23%) and the most common response category (44%) was that a mutation in a gene results in a mutation of the protein leading to the disease (these differences are also statistically significant according to a $z$ test of proportions). This type of response suggests that students understand the nature of the genetic information as coding for proteins and that changes at the gene level result in direct changes to the protein level.

**Table 2: Students responses to the written assessment question: How can a gene cause a disease?:**

<table>
<thead>
<tr>
<th>Response Categories (How can a gene cause a disease?)</th>
<th>Pre (%)</th>
<th>Post (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inherited: we get genes for disease from our parents, our parents’ genetics predisposes us to the disease, etc</td>
<td>61</td>
<td>23</td>
</tr>
<tr>
<td>Genes have information about the disease: the gene tells/says you will get the disease, etc</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td>There is a mutation/change in the gene: gene is deformed/wrong/malfunctioning</td>
<td>12</td>
<td>24</td>
</tr>
<tr>
<td>The mutation/change in the gene cause a change in the protein: relate a change of the genetic information to a change at the protein level</td>
<td>0</td>
<td>44</td>
</tr>
<tr>
<td>Mutation in protein: mentions changes to a protein’ structure and/or function but does not mention genes as the cause of the change</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Other: gene has a bacteria, the disease alters the gene, some structure in body doesn’t work, genes don’t cause diseases, etc</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>Total student responses</td>
<td>N=107</td>
<td>N=111</td>
</tr>
</tbody>
</table>

Similar trends in understanding were also evident in the interview data and were particularly prevalent in the first interview task. In this task we asked students to make as many connections as they can between 12 terms presented on small cards (such as genes, proteins, DNA, genetic disease, amino acids, etc). In the pre-interview only 2 of the 11 students connected genes to proteins and in these few cases the explanations were that “genes say what enzymes you need” and “DNA tells the body how to use proteins”. In the former case we can assume that the student understands genes as having information about enzymes (a type of proteins) presumably information required for the synthesis of these entities. However, in the latter case the student was referring to proteins we eat, thus in his mind the genetic information specified how we should use the proteins we obtain from food, in this case there does not seem to be an understanding of genes as coding for the making of novel proteins. The state of affairs was different for the post-interviews in which 7 of the 11 students connected the terms genes and proteins and their explanations of this connection were aligned with the scientific understanding of genes coding for proteins.

Taken together these data suggest students’ conceptions of the nature of the genetic information changed over the course of instruction. Prior to instruction students tended to conceive of genes as either non-information based or as having information about entities at multiple organization levels (information about cells, tissues, organs,
etc). Moreover, students reasoned about the connection between genes and genetic phenomena as one of inheritance (genes from our parents predispose us to traits and diseases) and were unable to explain how genes bring about their effects. Post instruction, however, a much larger proportion of the students reasoned about genes as containing information about proteins and were able to link genes to proteins as a mediating mechanism for genetic phenomena. The ability to reason about genes as coding for proteins is critical to generative understandings because this idea holds true for a multitude of genetic phenomena. Such an understanding allows the learner to assume that if a gene is involved in a particular phenomenon a proteins must be the product of that gene and is therefore also involved, hence any explanation of the phenomenon needs to account for proteins as a mediating mechanism.

**Students’ understandings of proteins as central mediators of genetic effects**

Congruent with the understanding that genes code for proteins is the equally critical understanding that proteins play a central and robust role in the genetic system and essentially mediate almost all genetic effects. The understandings that proteins are central and robust entities, and that the amino acid composition of the protein determines its function were reflected in our second and third learning objectives respectively. There were several aspects of the curriculum design that reflect our attempts to scaffold these understandings: 1) students explore the protein mechanisms that are involved in FH in great detail there are several activities devoted to the exploration of the protein level, 2) students examine data regarding the structure of the protein involved in FH (LDL receptor) to ascertain the ways in which differences in structure affect function, 3) students examine the involvement of proteins in three other genetic disease in lesson 6.

Several of the data sources we collected proved useful in ascertaining students’ understandings of these key ideas. In the written assessments we asked students to explain the connection, if any, between proteins and genetic diseases. We wanted to see whether students conceived of proteins as central mediators of genetic effects. We also asked students to list as many functions of proteins as they can recall, this question was designed to tap into students’ familiarity with proteins and their robustness in the system (understanding of the various different roles they play in the system). We found that there were changes to the pattern of students’ responses to the first question (what if any is the connection between proteins and genetic diseases?) before and after instruction (Table 3).

Prior to instruction a large proportion of the students (33%) claimed that there was a connection and that proteins were involved in at least some genetic diseases either because the diseases involved anatomical structures that students knew were made of proteins (for example diseases of the hair, nails and muscle) or because the disease involved inadequate use or processing of proteins acquired through food. While these responses clearly indicate that students did perceive some involvement of proteins in some genetic diseases, it was not clear that students understood that since genes code for proteins that genetic phenomena, such as disease, are inherently mediated by proteins.
Moreover, a significant portion (19%) of the students responded that there is no connection between proteins and genetic diseases.

Students’ responses after instruction demonstrated a much deeper and more generative understanding of the role of proteins in genetic phenomena. The majority of the students (67%) were able to explain that genetic diseases are due to a mutation in a gene that results in a mutation to a protein thereby causing the disease. This type of response illustrates an understanding of proteins as mediators of genetic disease. Moreover, students were able to construct their explanations without needing to know the specifics of the disease in question. Thus, their understanding is generative in nature and applies to a multitude of genetic diseases and other phenomena. In addition, the proportion of students who responded that there was no connection between proteins and genetic diseases (2%), or that proteins were somewhat involved in some genetic diseases (6%), decreased significantly (statistically significant based on a z test of proportions).

**Table 3: Students' responses to the written assessment question: What, if any, is the connection between proteins and genetic disease?**

<table>
<thead>
<tr>
<th>Response Categories (What is the connection, if any, between proteins and genetic diseases?)</th>
<th>Pre (%)</th>
<th>Post (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No connection</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td><strong>Mutation in proteins:</strong> mention changes to a protein’ structure and/or function but do not mention genes as the cause of the change</td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td><strong>Mutation in gene causes a mutation in the protein:</strong> relate a change in the genetic information to a change at the protein level</td>
<td>4</td>
<td>67</td>
</tr>
<tr>
<td><strong>Proteins involved:</strong> suggest protein involvement in some diseases (muscle, hair and nail diseases, predisposition to not digesting/eating enough proteins)</td>
<td>33</td>
<td>6</td>
</tr>
<tr>
<td><strong>Genes are made of proteins</strong></td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td><strong>Other:</strong> both genes and proteins are in us/our cells, both are passed down, our parent’s genetics predispose us to disease</td>
<td>17</td>
<td>4</td>
</tr>
<tr>
<td><strong>No response</strong></td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total student responses</strong></td>
<td>N=108</td>
<td>N=111</td>
</tr>
</tbody>
</table>

In order to get a sense of students’ understanding of the diversity of functions proteins have in the system we looked at students responses to the written assessment question asking them to list as many function of proteins as they could (Table 4). We categorized students’ ideas of protein functions, such that each response could include several propositions. Table 4 lists these categories as well as the number of time they were mentioned by students. There were two main changes between the pre and post responses: the total number of categories and the total number of propositions, both of these parameters increased post instruction. Prior to instruction the majority of students’ responses listed proteins as important for the making of hair, nails and muscle as well as enzymes and overall there were a total of nine different categories of protein functions mention. After instruction there were 16 categories of protein function some were functions specific to a particular example of a protein discussed in class (bringing in LDL and absorbing shock in muscle cells) or a type of protein function extensively studied in
class (receptors, transport and unzipping/copying DNA). Overall the total number of ideas mentioned increased from 207 to 296 thus per-response students mentioned more ideas of protein functions after instruction.

### Table 4: Students conceptions of protein function

<table>
<thead>
<tr>
<th>Students propositions regarding protein functions</th>
<th>Frequency of proposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteins are used as energy/fuel</td>
<td>Pre: 7 Post: 9</td>
</tr>
<tr>
<td>Proteins are involved in making energy/ATP</td>
<td>Pre: 6 Post: 5</td>
</tr>
<tr>
<td>Proteins help make/build other proteins</td>
<td>Pre: 13 Post: 22</td>
</tr>
<tr>
<td>Proteins build/make: nails, hair, muscles</td>
<td>Pre: 97 Post: 76</td>
</tr>
<tr>
<td>Proteins are/or make enzymes</td>
<td>Pre: 72 Post: 66</td>
</tr>
<tr>
<td>Proteins build/make structures (other than hair, nails, muscle)</td>
<td>Pre: 3 Post: 7</td>
</tr>
<tr>
<td>Proteins help genes or hold the genetic code</td>
<td>Pre: 1 Post: 5</td>
</tr>
<tr>
<td>Proteins help grow or repair tissue</td>
<td>Pre: 7 Post: 4</td>
</tr>
<tr>
<td>Proteins are involved in digestion of food</td>
<td>Pre: 1 Post: 10</td>
</tr>
<tr>
<td>Proteins regulate body/cell functions</td>
<td>Pre: 0 Post: 5</td>
</tr>
<tr>
<td>Proteins are receptors, or proteins receive substances (nonspecific)</td>
<td>Pre: 0 Post: 26</td>
</tr>
<tr>
<td>Proteins are involved in metabolic cycles</td>
<td>Pre: 0 Post: 3</td>
</tr>
<tr>
<td>Proteins help transport substances</td>
<td>Pre: 0 Post: 13</td>
</tr>
<tr>
<td>Proteins unzip and/or copy DNA</td>
<td>Pre: 0 Post: 6</td>
</tr>
<tr>
<td>Protect cell membrane of muscle from shock or shock absorber</td>
<td>Pre: 0 Post: 18</td>
</tr>
<tr>
<td>Bring cholesterol/LDL into the cell</td>
<td>Pre: 0 Post: 22</td>
</tr>
<tr>
<td>Total number of propositions</td>
<td>Pre: 207 Post: 296</td>
</tr>
</tbody>
</table>

These data are encouraging, however, we had hoped to observe larger shifts between the pre and post measures. While students did mention more types of protein function post instruction, the majority of students’ responses were still limited to conceiving of proteins as important for making hair, nails, enzymes and muscles. Some of the functions of proteins mentioned in the unit (in lesson 6 and in readings throughout) such as transport, signaling, regulation, structural proteins were not mentioned by many of the students in the post assessment. Thus, it seems that students were still somewhat unfamiliar with the robustness of proteins in the system, they seemed to know that proteins are important entities but were not sure as to what roles they play.

Conceiving of proteins as central entities is, as we noted earlier, important for generative reasoning in genetics. We therefore turned to the interview data; specifically the second task in which we asked students to provide a hypothetical mechanism to explain a novel genetic phenomenon, to examine students’ ability to reason generatively. We were very excited to find that students were relatively apt at providing at least parts of the underlying mechanism despite never having learned any of the details of the phenomenon at hand. In this task students were asked to read a short excerpt describing a genetic phenomenon at the cellular level, for example: bacterial cells can sense...
substances in their environment and move towards or away from them, some cells have a mutation and can no longer sense (there were two versions of the task, comparable in terms of familiarity and difficulty level but different in the specific context). We then asked students to speculate what is happening inside the cell that allows it to sense and move towards or away from the substances. Prior to instruction students’ mechanisms were rather vague, based on instinct, multi-cellular capabilities like smelling, or human-like abilities to sense and respond to the environment. The excerpts below, taken from Alma’s and Ben’s and Jesus’s pre interview, illustrate typical responses to this task.

Alma

So [bacteria] can sense it maybe like by um, by touch kind of. Like maybe it feels different.

Interviewer: Ok, so if this food touches here… How would they know?

Alma: Maybe the substance, the substance has like the food you mean, has some kind of substance that gives off like a smell or a touch or something, or um, yeah. And then the bacteria feels or smells or does whatever it does, needs, I guess. //

Interviewer: Ok, and if let’s say the food touches this side of the membrane and it needs to move that way. Let’s say we have some sort of moving mechanism maybe it’s like a propeller thing. How does it know to move now?

Alma: Like it eats the food, and then it knows that there’s more, so it kind of gives off like a signal kind of to the bacteria to like make it move.

Interviewer: So there’s a signal, where would the signal be in this picture?

Alma: I’d say it was inside, like…

Interviewer: Moving from here to here?

Alma: No, yeah, yeah, kind of like the body when like, when you pinch yourself or something and you start to bleed, your body tells you, like ouch that hurts or something, like your brain I mean tells you that. So maybe, maybe that’s what does that, like it maybe has a little brain or something like that. And then it tells it what to do.

Ben: I’d say [bacteria] can probably smell [the substance].

Interviewer: What do you mean by smell it?

Ben: It probably has some sort of like gland or something that can like smell the stuff, probably.

Interviewer: Ok and what happens when they sense the food, then what happens?

Ben: The uh, bacteria starts moving towards the food.

Interviewer: Ok, um how does it sort of know where the food is?

Ben: It tells the direction where the smell is coming from.

Interviewer: So these glands can convey information about the direction of the food?

Ben: Yes

Interviewer: Ok, but how do they get that information into the cell? So let’s say they
know it’s here, right, like this food is right here, but how does the cell know to move there?

Ben: Probably the nucleus can actually find out somehow that it’s over there, and it’s just telling the whole thing to go over there.

Jesus: Well [bacteria] has to have some kind of… it has to have like some nucleus in there that senses the food

Interviewer: Okay so there’s a nucleus in there, where would it be?
Jesus: In the middle and it will sense the food
Interviewer: How would it sense the food?
Jesus: Uh….by…oh I don’t know really
Interviewer: You can guess…
Jesus: By maybe smelling it or something
Interviewer: Okay let’s say that little dot is a piece of food, how would it sort of know it’s out there, it can’t really see it, right?
Jesus: No, it might smell it, or like feel it or something so it will go toward it if it knows its food but it will stay away from it if it’s poison
Interviewer: Okay, how does it know it is food?
Jesus: I guess it already knows how food is and how poisonous stuff is
Interviewer: But it doesn’t really have a brain or anything it’s not that sophisticated. Somehow it would…
Jesus: …Just know

None of the students we interviewed (pre-interview) mentioned proteins as being involved in the sensing or signaling process, rather, students speculated mechanisms that were based on a human-like sense of smell (Jose and Ben) or touch (Alma) and an often unexplained or nerve-like ability to signal. Moreover, when asked to explain what is happening in the mutated bacteria students would often claim that the bacteria would lose its ability to sense but could not provide more detailed mechanisms for how this occurred.

Post instruction students responses were much detailed, protein-based and plausible. That is, students were able to hypothesis mechanisms that are aligned with canonical understandings in the domain. The episodes below, taken from Cortez’s, Bill’s and Morris’s post interviews illustrate the changes in the quality of students reasoning about this task (in some of these cases the scenario described involved a human cell being able to sense a growth hormone and divide).

Interviewer: How do they know it’s there?
Cortez: Okay I’m thinking we’re gonna go with the idea of the receptors //
Interviewer: What is the result of the hormone getting into the cell and going in to the nucleus?
Cortez: Duplication
Interviewer: Do you have any idea how a hormone could cause the cell to divide?
Cortez: Activating enzyme
Interviewer: What kind of enzyme might that be?
Cortez: A growth enzyme ha ha! I don’t know. An enzyme that tells the nucleus, hey it’s time to divide
Interviewer: Okay there might be some enzyme in here
Cortez: A signal enzyme //
Interviewer: Now it says here that some cells have a mutation, now what does that mean to you here in this context?
Cortez: That there’s a problem either here (points to receptor)… or inside the same nucleus, something to do with the protein that doesn’t let the growth hormone trigger whatever it is that it’s triggering so it could grow. And this problem would come from the DNA. There’s a mutation in the DNA, more so in the gene which makes up the DNA and these genes probably codes for the protein and the mutation is in the gene means that the protein doesn’t grow the way it’s supposed to be.

Note that Cortez is able to speculate the existence of a receptor and an activating enzyme (both of which are types of proteins) and provides an explanation of the mutated cells that includes reference to a change in the genetic code that might alter the protein. This sort of explanation is plausible and is derived from an understanding of proteins as central entities that mediate genetic effects.

In the next episode Bill is able to suggest a receptor as the sensing mechanism but does not suggests a protein for the signaling/messaging mechanism, rather he suggest an nerve-impulse like signal (similarly to students’ suggestions in the pre-interviews). We conjecture that Bill may understand that proteins are central elements in the system and can provide a protein-based mechanism for at least some aspect of the system (sensing) since he is familiar with protein receptors as entities that can “sense” substances (due to the diseases studied in the curriculum). However, Bill does not seem to be familiar with a protein-based mechanism for signaling and thus he provides one that he is familiar with (a nerve-impulse). Bill is able to provide a rather detailed explanation of what happens in the mutated bacteria in terms of precisely how the shape of the receptor would change (from v-shape to a flat line) and why (a change to a specific amino acid, cystein, that allows the v-like bend to occur). Thus, Bill has a deep and generative understanding of structure-function relationships in proteins.

Interviewer: How can it sense the food?
Bill: With like a receptor.
Interviewer: Could you sort of draw what might be a hypothetical receptor might look like.
Bill: It might like be… like that (draws receptor on membrane of cell). And then, or like if it touches it then it sends a message to the nucleus telling it that there’s food.
Interviewer: Ok and what sort of message might that be? How is this receptor signaling to this, the nucleus?
Bill: I think um an impulse or something. //
Interviewer: Now it says here that some bacteria have a mutation, what does that mean to...
Bill: Like if this was made of protein right here (referring to receptor), then the protein would be messed up, so it might not be sensing that there’s like sugar here. So then it won’t ever get the message and then it’s like a chain, so it’ll never get the food.

Interviewer: Ok and when you say the protein might be messed up, what about it might be messed up?

Bill: The like, the um amino acids and nucleotides that like code for what kind of protein it is, are like out of order or they’re missing some. So when it’s coded for, it codes the wrong type of protein. So it might be like a smaller amino acid than it was supposed to be.

Interviewer: So, this protein, could you draw what a messed up, like one option for what a messed up receptor might look like.

Bill: Well if it’s supposed to be a V-shape like that then it might just be straight, like it didn’t have the cystein (type of amino acid) right there, it might just be straight or… yeah.

Similarly to Bill, Morris speculates the existence of a protein receptor that can sense the growth hormone in this case. He is also able to explain that the mutated receptor differs in its shape due to a change in the amino acid composition (missing the proline amino acid) resulting from a change to the DNA codons (codons are the coding aspect of the DNA molecule). Like Bill he does not provide a protein-based mechanism for the signaling, but rather assumes the cell membrane some how knows to tell the cell to divide once the hormone touches the receptor.

Interviewer: Ok, so um, let’s see, how do you think the cell senses the growth hormone?

Morris: It probably has a receptor.

Interviewer: For?

Morris: The hormone. //

Interviewer: Ok, but I don’t really get how this signals anything to the cell, is it just the fact that it touched that hormone?

Morris: Yeah

Interviewer: Then it automatically divides?

Morris: The growth hormone when into there, so it has to signal for it to divide I guess. Since it touched the receptor.

Interviewer: Is there anything in the cell that needs to get the signal to divide?

Morris: Probably the membrane.

Interviewer: Just the cell membrane.

Morris: Yeah

Interviewer: So as long as it touches the cell membrane, it’s like the membrane knows.

Morris: Once it signals to tell the cell to divide.

Interviewer: It that says here that some cells have a mutation, like this one that you drew, what do you think causes it to look different than this one?

Morris: Because this one probably has a proline or some other um thing that makes it bend and this one doesn’t so it’s just flat.

Interviewer: So is this thing then made of…

National Association for Research in Science Teaching (NARST) April 4-7, 2005
Overall 10 out of 11 students we interviewed (post interviews) suggested a receptor protein as a sensing mechanism and two were further able to suggest a protein-based signaling mechanism. Moreover, 8 out of 10 students (one student did not complete the second half of the task dealing with the mutation) were able to explain the mutation as arising from a change to the genetic material that resulted in a change to the structure and thus function of the receptor, and five of those 8 students were able to speculate about specific changes to the protein’s amino acid composition and how those would affect its structure and function (such as Morris’s and Bill’s reference to the changes in the proline and cystein amino acids in the receptor and their effects on the structure of the molecule). These findings are very encouraging since they signify a marked improvement in students understanding of the genetic system and their ability to transfer these understandings to a novel context. It is precisely these sorts of understandings that afford generative reasoning and engender scientific literacy in genetics.

Students’ understandings of the nature of explanations in the domain of genetics

We have thus far discussed students’ understandings of the key components of the genetic system and their interactions. Another important goal of the curriculum was to help students learn the nature of explanations in the domain. More specifically, we wanted students to understand that the genetic system is a hierarchical one such that changes at one level have consequences that affect subsequent levels, and therefore explanations in genetics are depictions of the casual mechanisms underlying the phenomena that transverse all the organization levels, from the gene to its observable effects. As discussed in the design section we scaffolded the practice of domain-specific explanations in several ways: 1) having the instructional sequence reflect a process of investigating and explain genetic phenomena by continual delving into lower organization levels, 2) having students construct such an explanation as the culminating activity of the investigation of FH (lesson 5), and 3) having students engage in “piecing” together such explanations by reading and compiling discrete descriptions of the underlying mechanism at each organization level (lesson 6).

To gage students understanding of the nature of explanations in this domain we analyzed two sources of data: whole-class discussions during lesson 6 (discussions of the biological basis of genetic diseases) and students’ artifacts from lesson 5 (a letter to a patient explaining the biological basis of FH).
We began by analyzing students’ explanations as manifested in their artifacts for lesson 5. This lesson marked the culmination of the investigation of FH and as its conclusion students were asked to write a letter to one of the case patients and explain to them the biological basis for their affliction beginning with the symptom level and delving all the way to the DNA level. We analyzed both the completeness of students’ explanations (did they mention all five organization levels—symptoms, tissues, cell, protein, DNA), and the coherency of these explanations (did students make appropriate connections between levels). Figure 8 depicts the number of levels (out of a total of 5) mentioned in students’ explanations. The majority of the students (67%) included all five levels of organization in their explanations.

Figure 8- The number of organization levels mentioned in students explanations in the culminating activity of lesson 5.

It is not too surprising that most students mentioned all levels since the instructions for the task (Figure 7) specifically states that the explanations should refer to all five levels. We therefore turned to examining another aspect of students’ explanations, their coherency in terms of whether students made appropriate connections between levels and could reason about the propagation of effects through the different levels.
In this analysis we counted the number of connections made such that each level (except the lowest and highest) can be connected both to the level above and the level below except for the lowest and highest levels, which can only be connected one-way. Thus there are a total of 8 possible connections. We found that the majority of the students (78%) made 6 or more connections out of a possible 8 in their explanations (75-100% of the connections). Since the essence of explanations in genetics this is the ability to link genes to their observable effects, these data suggest that students are capable of constructing such holistic explanations.

![Connections Between Organization Levels](image)

Figure 9- Percent of connections made between levels in students’ explanations in the artifact for lesson 5.

Students’ ability to construct holistic and domain-appropriate explanations of genetic phenomena was also evident in classroom discussions. We present two episodes taken from the whole-class discussions during lesson 6 of the 9th and 8th grade classes. In both examples the class is discussing the genetic disease Duchenne’s Muscular Dystrophy (DMD), an inherited muscle disease. Prior to this lesson students had worked in groups to “piece” together explanations of the biological basis of three inherited diseases (Duchenne’s Muscular Dystrophy, Cystic Fibrosis, and Phenylketonuria) from sets of readings that provided descriptions of the diseases at each organization level. The work of piecing together an explanation entailed drawing the appropriate inferences to explain how the links between the various organization levels. Once student groups compiled their explanations the class reconvened for a whole-class discussion of each of the diseases. These discussions were guided by the teacher who prompted students to...
proceed the explanation one level at a time. In both classes students were able to explain
the vents at each level and how changes at one level affected subsequent levels.

9th grade class, whole-class discussion lesson 6:
Teacher: Ok could be something like that or a change in charge. OK let’s talk about DMD,
Janet you had it so I will ask Bill to see how good of a teacher you were. Bill
what’s DMD like
Bill: Umm, problems walking your muscles keep breaking down you have muscle loss
Teacher: What tissue is it then
Students: Muscle tissue
Teacher: Jesus could you tell us a bit about the cells of people with DMD
Jesus: Yeah, The DMD muscles tear up because of cracks in the membrane
Teacher: OK good, so you have a cell and it tears forms a break, how come it breaks with
DMD
Rick: Because dystrophin it’s the binding protein, no it’s the shock absorber protein so
if you run or do anything the muscles tear
Morris: Its like a spring, its like as if you jump on a bed then the cushion protects you
Teacher: Right we have a dystrophin spring in our muscles so we do so our muscles are
OK there is the shock absorber but if that shock absorber.... what happens to the
protein?
Gwen: Its mutated, the [amino acid] sequence is wrong
Teacher: So is it the same size
Students: No
Teacher: So it’s either missing or broken or smaller so what happens when these people
use their muscles? They crack there is no cushion. Right.
Morris: There is a leak
Teacher: Right there is a leak in the membrane and then what happens to the muscle cell
Morris: It dies (end of video-tape)

8th grade class, whole-class discussion lesson 6:
Teacher: OK let’s start, who had DMD? Can somebody tell me, somebody who didn’t
study the disease, what are the symptoms?
George: Some form of mental retardation
Students: What No, that’s PKU
George: Sorry, sorry my bad, it happens in boys, what happens is they become clumsy
they are lame you are weaker, you have a lot of muscle fatigue you walk on your
toes calf muscles enlarged and something with the belly
Amber: They have respiratory failure
Teacher: They have respiratory failure
Amber: Which means they die cause they cant breathe
Teacher: What’s going on… what tissue is this?
Simon: Muscle tissue
Teacher: What’s going on at the cell level, what’s inside the cell what’s going on that’s
making these things happen?
Amber: There is nothing to absorb the shock
Teacher: There is nothing to absorb the shock, so what happens if the shock doesn’t get
absorbed, why is there shock to begin with?
Patrick: Because it is contracting and there is shock on the cell and well if there is no
shock absorbed then tiny cracks form in the cell membrane.

National Association for Research in Science Teaching (NARST) April 4-7, 2005
The ability to understand genetic phenomena in this way is at the crux of scientific literacy in genetics. The most likely way in which students will have to use their genetics knowledge in contexts outside of the classroom is to understand descriptions/explanations of genetic phenomena presented in the general media or by a health care provider. Often such presentations include some but maybe not all the relevant information and the reader/listener has to process the provided information, draw inferences and fill in the blanks in order to make sense of the phenomenon. The activities of lesson 6 emulate such a situation. Our findings suggest that students in both classes (based on participation in these discussions) were relatively proficient at constructing explanations of genetic phenomena from somewhat incomplete and fragmented accounts of the phenomena. Moreover, these explanations were systemic in that they accounted for all levels and detailed the mechanisms by which genes bring about their effects.

Conclusions and Implications
In this paper we presented a design solution, in the form of a curriculum, to support student construction of systemic and generative understandings in genetics. We began by conducting a problem analysis that characterized the conceptual obstacles students face when learning genetics as well as key domain-specific understandings and practices that are crucial for generative reasoning in genetics. We then described in a detailed manner the ways in which we designed the curriculum to scaffold students learning of these key ideas and practices. In the previous section we presented our findings from the pilot enactment of the unit in an introductory biology course at the high school level. We now turn to a more general discussion of three important outcomes of design research: 1) lessons learned about the specific design solution described herein, 2) lessons learned about students cognition in genetics (contribution to cognitive theory), and 3) broader lessons/issue for design of project-based curricula and instructional support for generative reasoning.

Lessons learned about our specific design solution

We set several conceptual and practice-based learning objectives for our unit. More specifically, we wanted students to understand the nature of the genetic information, the role of proteins in mediating genetic effects, the hierarchical nature of the genetic system, and the nature of investigations and explanations in the domain. Our data suggest that after instruction the majority of the students conceived of genes as coding for proteins, recognized the centrality of proteins in genetic phenomena and were capable of providing holistic and coherent explanations of genetic phenomena that reflect the structure and content of canonical explanations in the domain. Moreover, the understandings students constructed allowed them to reason generatively about novel genetic phenomena and to generate plausible, systemic and domain-appropriate explanations of such phenomena. Thus, overall we feel that the curriculum was relatively successful in supporting systemic and generative reasoning in genetics.

There was, however, one aspect of students’ reasoning that we feel did not greatly improve from pre to post and that limited students’ ability for generative reasoning. Students did not seem to develop a robust understanding of the various functions proteins can have in the system. Thus, while they did conceive of proteins as central mediators of genetic effects they had limited understanding of what sort of functions proteins can take on in order to mediate those effects. Pre instruction students conceived of proteins predominantly as enzymes and as builders of muscle, hair and nail tissues. This state did not change much post instruction, students gained insight of only a few more functions that proteins play (receptors, making other proteins and manipulating DNA).

Supporting the construction of a robust understanding of protein function was an important goal of the curriculum, yet it was also at the center of a problematic trade-off issue that is endemic to project-based instruction. We chose to focus instruction on the investigation of a single genetic disease, we made this choice because we wanted to provide students with an authentic context for inquiry and allow them to engage in an in-depth and systemic analysis of genetic phenomena. Due to time and logistical constraints...
of the classroom we chose not to pursue multiple line of investigation simultaneously. However, in making this choice we ran the risk of supporting students in the construction of very contextualized and non-transferable knowledge. To mitigate this problem to some extent we designed lesson 6 in which students briefly investigate (through readings) three other genetic diseases, we also included various readings about other contexts other than FH throughout the unit. Our findings suggest that students may have constructed more generalized understandings of the centrality of proteins but that the examples of other diseases provided in lesson 6 and in scattered readings may not have been enough to support the construction of a robust understanding of protein function. Perhaps what is needed are additional examples and a more targeted focus on the different functions proteins have and the different mechanism in which proteins are involved.

Lessons learned about students’ cognition in genetics

As discussed above students did not form robust understandings of the various functions of proteins and the mechanisms they are involved in. The difficulty in helping students gain insights into the functions and behaviors of a system has been documented in prior research. Hmelo et al., (2000) also attempted to scaffold students’ understanding of a biological complex system (respiratory system) through a design-based curriculum (the final project entailed designing an artificial lung). Their findings suggest that after instruction students seemed to understand the structures involved in the system but still had difficulties reasoning about functions and behaviors in the system.

Why is reasoning about functions/mechanisms so difficult for students? We conjecture that students in our study did not develop domain-specific understandings about potential interactions and mechanisms in the genetic system. In related research the first author proposed a cognitive model of student reasoning in genetics, this model depicts key domain-specific understandings that are critical for generative and domain-appropriate reasoning in genetics (Duncan & Reiser, 2005). Two important types of domain specific knowledge were identified in this work: domain-specific heuristics and domain-specific explanatory schemas. The domain-specific heuristics the define key entities and dynamics of the system, for example the understandings that genes code for proteins and that proteins are central mediators of genetic effects are examples of heuristic knowledge. The domain-specific explanatory schemas further define possible interactions and mechanisms in the system, for example the understandings the different roles proteins play in the system are captured this form of knowledge.

This work informed the identification of learning objectives for the curriculum described herein. We conjecture based on our findings, that students may have only formed understandings that correspond to the domain-specific heuristics and a very limited subset of explanatory schemas. Our conjecture explains the observation that students in this study seemed to understand the central role of proteins in the system but were often unable to provide specific mechanisms for how proteins mediated genetic effects. The findings from this work are suggestive of the importance of domain-specific explanatory schemas in reasoning.
Broader lessons/issue for design

We would like to conclude this paper by considering some broader issues regarding the design of learning environments that are pertinent to the design of the genetics unit. The first issue pertains to the tradeoff between contextualization and generalization. As we discussed earlier, there is a common tension in the design of project and inquiry-based curricula between providing students with an authentic and motivating context for investigation, and helping students construct more generalized understandings of the target content. In choosing a specific context, we allow students to experience an in-depth and sustained investigation of the chosen context. However, in doing so, we run the risk of supporting the construction of very contextualized knowledge that is not amenable to use in other contexts because it has not been sufficiently abstracted from the context of study.

Our findings from this study suggest that students in our study were able to generalize some of the key ideas in genetics but did not form robust understandings of the various functions of proteins since they may not have engaged with sufficient examples of different protein functions. We do not pretend to know of a solution for this problem; however, we believe that this problem is a common plague of project and inquiry-based learning environments and merits research. It seems that our design solution of focusing on a single case for the majority of the unit and then providing a more limited context in which students investigated several other cases did not suffice, at least in the implementation we studied. There are likely several potential design solutions for this problem, and comparative studies of different solutions may yield important insight into balancing this intricate tradeoff.

The second issue we wish to raise involves yet another common tradeoff endemic to science curricula—the tension between accuracy and fidelity and learnability. We have already made reference to this tradeoff in the design section of this paper when we discussed our choice of representation of protein structure (lesson 3). In our design approach, we placed higher value on using data and scientific representations that are accessible and salient to students as opposed to more accurate and extensive depictions. Arguably, students resulting understandings of the data and underlying mechanisms are more simplistic and rudimentary. However, given our goal of generative and systemic understandings that foster scientific literacy in genetics, we believe that such simpler understandings will suffice and our findings regarding students’ ability to reason productively about novel phenomena support this assumption.

There has been relatively little research on the tradeoff between fidelity and accessibility to learners. In his work, Roschelle (1990) provides a set of guidelines for navigating this tradeoff space in the context of simulation models in physics. These guidelines are informative and highlight some important dimensions of the tradeoff space, however, their transferability beyond physics simulation models is somewhat limited. The problem of negotiating a balance between providing accurate and often complex models...
of current scientific understandings, and providing models that students can engage with and make sense of, is a thorny issue in science instruction and merits more extensive research. Particularly in respect to the influences of choices in this tradeoff space on students’ scientific literacy.

Lastly, we wish to highlight a feature of our curriculum that we believe is a successful aspect of our design and may be transferable to other contexts. Edelson (2000) in his Learning for Use framework emphasizes the importance of motivating the construction of knowledge. In our design we implemented this aspect of the Learning for Use framework by structuring the curriculum as cycles of inquiry each motivated by findings from the previous cycle, much like the unfolding narrative of a detective story. In our unit students investigate FH at each organization level, the identification of the underlying cause at one organization level instigates questions that require investigation of the subsequent lower organization level. For example, students’ investigation of the cell level shows that cells do not take up cholesterol from the blood, this raises the question of what allows cells to take up cholesterol and what is the problem with this mechanism in FH cells, these questions form the basis of the investigation at the subsequent protein level. The curriculum that precedes the genetics unit is also structured as an unfolding narrative and provided the template for our design (Kanter, Schwille, MacKenzie, & Reiser, 2003).

This narrative-like structure of the curriculum is favorable in two ways: it affords a more coherent and cohesive learning experience for the student, and it maintains a constant motivation or need to know of the target content. The narrative structure also mitigates against the temptation to teach the content first and only then have students engage in project work, as opposed to having the project itself be the context for learning. Through the narrative structure students are always engaged in the project or investigation, beginning with more familiar aspects of the problem/project (in our case starting the investigation at the more familiar organism level) and moving towards less familiar and more complicated aspects. In this ways students experience the curriculum as an unfolding narrative in which they investigate science questions and work towards solving a problem or designing a product in a seamless manner. In future work we hope to elaborate on this idea of narrative structured curricula and provide a design framework to inform the design of similarly structured curricula in science.

References:


Tinker, R. (2000a). Atomic-scale models: the key to science education reform. @Concord 4:1 p4-5.