8-2-2019

Mission to Planet Markle: Problem-Based Learning for Teaching Elementary Students Difficult Content and Practices

Melanie Peffer
University of Colorado, Boulder, melanie.peffer@colorado.edu

Maggie Renken
Georgia State University, mrenken@gsu.edu

Patrick Enderle
Georgia State University, penderle@gsu.edu

Jonathan Cohen
Georgia State University, jcohen@gsu.edu

Mission to Planet Markle: Problem-Based Learning for Teaching Elementary Students Difficult Content and Practices

Abstract

Young children can struggle to learn difficult disciplinary content and important skills for practicing science. Problem-based learning (PBL) may be useful for addressing such difficulties, yet evidence to support its usefulness in elementary school-aged children is limited. We considered the role of a PBL unit in improving students' genetics content understanding and their skills specific to creating arguments with coordinated claims, evidence, and reasoning. First- through fifth-grade students participated in a six-week PBL unit about evolution and genetics. Students worked in mixed age groups and were charged with illustrating a fictitious alien species, called markles, based on a series of facts they collected about factors expected to impact markle adaptation. This work was particularly unique in its assessment of student groups' illustrated design solutions as arguments. Although students demonstrated weaknesses in coordinating claims and evidence overall, they were able to demonstrate success in gaining difficult genetics content knowledge and in preparing arguments with, at minimum, two components of well-constructed arguments, in most cases, providing a claim supported by reasoning. This work is informative for understanding student abilities, the potential of PBL, and considerations for its use.

Keywords: Problem-based learning, genetics education, argumentation, elementary science education

Mission to Planet Markle: Problem-Based Learning for Teaching Elementary Students Difficult Content and Practices

Problem-based learning (PBL) is a learner-centered approach to instruction. In a PBL context, the learner develops a solution to a defined problem through research and the application of his or her knowledge and skills (Savery, 2015). PBL has been touted for its effectiveness in garnering a host of education-relevant outcomes across disciplines. By engaging students in practical, experiential learning, a PBL approach in science disciplines is expected to foster students' general science practices and discipline-specific content knowledge in tandem (Dochy, Segers, Van den Bossche, & Gijbels, 2003). Such an approach is in line with *A Framework for K-12 Science Education* and the derivative Next Generation Science Standards (NGSS), which emphasize the value in coordinating students' development of science practices and core content knowledge acquisition (NGSS Lead States, 2013; National Research Council, 2012).

In a thorough review of what and how students learn via PBL, Hmelo-Silver (2004) provides evidence of PBL's effectiveness in advancing students' knowledge base, problem solving skills, and self-directed learning skills following problem-based learning experiences. Unfortunately, a shortage of empirical evidence of what students learn in PBL settings remains despite Hmelo-Silver's acknowledgment of as much over ten years ago. The dearth of evidence is particularly noticeable in K-12 settings (cf. Wirkala & Kuhn, 2011). The domains of learning best supported by PBL in Hmelo-Silver's seminal 2004 review neatly align with three broad domains of learning we consider to serve as *holistic science learning* in the current study: content knowledge, science practices, and 21st century skills. Namely, we examined PBL outcomes related to content knowledge in a genetics domain and the scientific practice of argumentation (Kereluik, Mishra, Fahnoe, & Terry, 2013; Voogt & Roblin, 2012). Prior research has suggested these outcomes may be fostered with PBL. We build on this prior work predominately examining the utility of PBL for older students by considering outcomes of PBL associated with elementary school-aged children. The current study aims to contribute empirical support for incorporating PBL in elementary science curriculum.

PBL for teaching difficult genetics content knowledge

PBL use in the classroom, and especially in the K-12 science classroom, abounds (Walker, Leary, Hmelo-Silver, & Ertmer, 2015). Presenting students with ill-structured problems with no single path of inquiry provides students the opportunity to collaboratively assess data provided, refine their ideas, and experience the tenuous nature of science knowledge by deciding on a solution that may not be correct (Gallagher, Stepien, Sher, & Workman, 1995). The popularity of PBL as a pedagogical approach may be due to its expected ability to engage students in learning difficult content. One such challenging content area for school-aged students is genetics. For instance, an analysis of genetics themed essays written by high school students indicated the prevalence of many misconceptions (Shaw, Van Horne, Zhang, & Boughman, 2008). This is particularly concerning since before the Next Generation Science Standards, genetics education began in middle school (National Research Council, 1996) and now begins in elementary school (NGSS Lead States, 2013). In spite of years of instruction in genetics, these misconceptions persist. Children are introduced to terms such as “genes” and “genetic inheritance” passively from the entertainment industry but are unsure what those terms mean (Venville, Gribble, & Donovan, 2005), and such exposure may foster misconceptions that persist through high school, stunting potential genetic understanding gains (Smith & Williams, 2007). Work by Duncan and Reiser (2007) suggests that part of the difficulty with learning genetics is due to challenges associated with reconciling information across different levels, such as how what is occurring at the DNA level impacts what is occurring at the protein level in a cell. Even at the undergraduate level, students do not understand how features are inherited, for example believing that diseases are inherited from biological parents, rather than from genetic material (Henderson & Maguire, 2000).

Genetics may be learned best in the context of inquiry (Duncan, Rogat, & Yarden, 2009). Previous work demonstrated that genetics-themed PBL units were more effective than direct instruction for increasing eighth-grade students' understanding of genetics content (Araz & Sungur, 2007). Although evidence suggests such instruction is effective for improving students' understanding of genetics, the majority of interventions are designed for high school students, with significantly fewer for middle school and late elementary school (Duncan et al., 2009). Previous work suggests that later elementary-aged students are familiar with the concepts of genetic inheritance (Springer & Keil, 1989), and a learning progression proposed by Elmesky (2013) suggested that curricula should leverage the cognitive abilities of K-5 students to develop a more advanced theory of kinship and genetic inheritability to lay the foundation for understanding gene expression in later grade levels. Similarly, a learning progression by Duncan et al. (2009) proposed fifth- and sixth-grade students should have a basic understanding of inheritance, traits, and DNA. In the 2007 National Research Council Report, *Taking Science to School: Learning and Teaching Science in Grades K-8*, the authors conclude that all children, even very young children, are capable of engaging in complex reasoning about the world. This is reflected in the NGSS as genetics content appears in the third-grade standards (NGSS Lead States, 2013). Consequently, interventions and research geared toward genetics learning for students in elementary school, particularly in early grades, are necessary.

Young students' arguments: Coordination of claims, evidence, and reasoning

Beyond promoting acquisition of difficult content knowledge, PBL has been traditionally employed to foster important science practices (Allen, Duch, & Groh, 1996; Baser, Ozden, & Karaarslan, 2017; Bell, 2010; Ferreira & Trudel, 2012; Kolodner et al., 2003; Kwon et al., 2018). For example, in the context of clinical education, medical students who engaged in PBL exhibited enhanced clinical problem solving ability (Savery, 2015). PBL activities can often conclude with students making a final argument in response to the initial ill structured problem (Belland, Glazewski, Richardson, 2011). Argumentation is widely recognized as a critically important practice in science education (Berland & Reiser, 2009; McNeill & Krajcik, 2006; National Research Council, 2012; Ryu & Sandoval, 2012). The ability to engage in arguing from evidence is one of eight core science and engineering practices described in the NGSS (NGSS Lead States, 2013). According to the NGSS, students in elementary school should be able to identify arguments and what makes an argument "good," particularly with regard to evidence. The standards also state that both early and late elementary school students are expected to be able to use evidence to construct an argument. Teaching students about and through argumentation alongside science content accomplishes multiple curricular goals. When engaged in argumentation and the creation of arguments, students apply content knowledge, engage productively in written and verbal discourse, and begin to understand that science is not a discrete collection of facts, but rather a body of knowledge generated through various discursive and cognitive activities (Manz, 2015). Argumentation encourages science learning as a vehicle for making sense of the world, as opposed to a passive student experience (Berland, Schwarz, Krist, Kenyon, Lo, & Reiser, 2015). Research has shown, though, that argumentation and the creation of arguments is challenging for elementary school students (McNeill, 2011; Ryu &

Sandoval, 2012) and middle school students (Belland, Glazewski, Richardson, 2011). In particular, teaching students to support claims with evidence—a key component of scientific arguments—has proven to be a challenge for science educators (Berland & Reiser, 2009). Limitations associated with students' writing and discourse abilities contribute to challenges in helping students learn to support claims with evidence and in assessing students' ability to do so (Felton, 2004; Sampson, Enderle, Grooms, & Witte, 2013).

Despite its status as an important and widely accepted part of science education, creating high quality arguments remains difficult for students to learn. Prior research has examined the nature of some of students' difficulties and the impact of curriculum on them (e.g., McNeill, 2009; Osborne, Erduran, & Simon, 2004). McNeill (2011) analyzed fifth-grade students' written arguments over the course of a school year. Student writing was analyzed to determine if it was argumentative in nature and if it contained arguments in a claim, evidence, reasoning format. Overall, students' argument construction improved over the course of the school year, but when given challenging content, students struggled to make accurate and appropriate arguments. Interviews with participating students revealed that while their overall ability to write scientific arguments had improved, the students still lacked an understanding of the importance of using evidence to support claims (McNeill, 2011). Prior work has demonstrated that, at least by middle school, curriculum that highlights argument components (i.e., claim, evidence, and reasoning) aids students in their ability to ground arguments in evidence (Berland & Reiser, 2009; McNeill & Krajcik, 2006). Berland and Reiser's (2009) analysis of middle school students' written arguments revealed two main categories: arguments that explicitly reference evidence and arguments that implicitly reference evidence. The overall ability of students to make high quality arguments explicitly incorporating evidence requires more attention and guidance (Berland & Reiser, 2009).

Similarly, Ryu and Sandoval (2012) examined third- and fourth-grade students' argument construction over a period of a school year. Students were engaged in a science curriculum in which their teacher prompted students working in groups to justify how they know something or how they would convince others of what they know. Ryu and Sandoval (2012) assessed arguments based on students' use of causal claims, the coherence of claims, citation of evidence, and whether or not the student explicitly justified their argument. The authors noted improvement over the school year in the students' ability to relate claims to each other in a coherent manner, their ability to cite evidence, and the use of explicit justification. These developments were attributed to explicit and consistent guidance offered to students through expectations for arguments that were communally established among teachers and students. Although improved, students' ability to explicitly justify their claims with data was still lower than the other aspects of argumentation examined in the study (Ryu & Sandoval, 2012). These studies, as well as others in the argumentation literature (Sampson & Clark, 2008), demonstrate students' challenges with respect to explicitly supporting claims with evidence in written arguments, especially in classroom settings. Further, challenges persist in the use of curricular interventions to help students with connecting claim and evidence (McNeill & Berland, 2017).

PBL activities are one type of curricular intervention that can foster productive argumentation interactions (Belland, 2010). PBL frames instruction with a context, or issue, requiring students to know and use relevant scientific information (Hmelo-Silver, 2004). Students can use scientific information to lead an investigation or to design a solution. For example, to create a PBL context that fosters productive argumentation, some PBL units revolve around a broad, investigable question that can lead to multiple productive investigations. Other PBL units employ a problematic, ill-structured context, usually drawn from real world circumstances, that necessitates the design of potential solutions (Householder & Hailey, 2012). Real world problems are typically messy, lacking the type of well-defined nature that often mitigates students' motivation or engagement. The real-world context requires students to set parameters and pull resources from a variety of disciplines (Savery, 2015). When framing a PBL unit in the real-world context, students' endeavors change in nature from one focused on conducting scientific investigations to the designing and refining of problem solutions. Although both are emphasized in national standards (NGSS Lead States, 2013), the integration of design challenges for the goal of enhancing science instruction remains problematic (Berland, 2013).

The argument products developed during design challenges must reflect the unique purpose and kinds of reasoning used for designing a problem solution (Berland, 2013). Theoretical frameworks of argumentation identify several commonly accepted or related structural elements inherent in high quality arguments (Grooms, Enderle, & Sampson, 2015; McNeill & Krajcik, 2006; Osborne et al., 2004; Sampson & Clark, 2008). These frameworks have mostly focused on the production of arguments from scientific investigations and not those arguments that result from design challenges. The nature and characteristics of traditional argument elements (e.g., claim, evidence, and reasoning) must shift to reflect the different activity goal of achieving a problem solution that meets certain criteria and specifications.

The first argument element to consider is the *claim*, which is typically an assertion that directly answers the question guiding a scientific investigation (McNeill, 2011). In design challenges, the claim would comprise either a proposed concept from the initial stages that is ideal for developing a prototype. Rather than the assertive statement typical for scientific arguments, the claim in a design argument can encompass a hypothetical schematic or a physical model. When considering *evidence* in a design argument, a design activity shifts the type of information analyzed to support the claim, focusing mainly on the constraints and criteria set forth in the design problem (McNeill & Krajcik, 2006). Finally, the *reasoning* element in a design argument would involve statements that explain how the evidence used for a particular stage supports the claim being made in that stage (Sampson & Clark, 2008). For design arguments, these statements would emphasize how elements of the prototype design (claim) align with the constraints and criteria stemming from the design problem (evidence). In the current study, we examined students' illustrated products at the end of a problem-based learning unit for indication of three component parts of an argument: claim, evidence, and reasoning. One of our research objectives was to determine if elementary students were able to support illustrated claims with evidence.

Illustrated design solutions as arguments. When focusing on the ability to construct high quality arguments, many studies have analyzed written text created through students' activities (Berland & Reiser, 2009; McNeill, 2011; Sampson et al., 2013). By assessing text-based products, such research has shed light on students' understanding of specific science content and the structural elements that make up scientific arguments. Further, written arguments serve as proxies, representing students' proficiency with engaging in the process of arguing from evidence (Sampson et al., 2013). However, written text is not the only manner available for students to express an argument composed of claims, evidence, and reasoning.

Drawings and graphic representations, compared to text-based forms, are considered to be equally valid products for assessing students' understanding of complex systems and phenomena (Bowker, 2007; Dentzau, in press; Lewis & Greene, 1983; White & Gunstone, 1992). Scholars have argued that using drawings can make students' conceptions more accessible due to the perception of drawing as being a less intimidating, and often enjoyable, activity compared to writing, particularly for younger students (Chang, 2012; Thomas & Silk, 1990). Chang (2012) contends that drawings are also applicable for assessment purposes with small groups of children as well as with individuals. Thus, allowing young students to generate drawings to represent conceptual understanding can also provide a way to decrease apprehension related to learning relatively complex science content, such as genetics. Also, having students draw the representations of different argument components could provide another vehicle for conveying the importance of individual components and relationships between them.

Current Study

The driving question of the current study was whether PBL supports gaining content knowledge and science practices. The PBL unit utilized for this study introduced students to the concepts of genetics and evolution and concluded with a capstone argument design project. Given that we know very little about PBL to teach genetics and develop arguments at the elementary school level, we sought to preliminarily examine each individually. We assessed student genetics understanding at the beginning and end of the unit using a pre/posttest format. To assess students' ability to create arguments, we asked student groups to create a capstone argument design project to complete a PBL unit that removed writing demands, the potential impact of prior belief bias, and the possibility of a correct answer. Students were instructed to generate a creative, illustrated product rather than a written argument. We hypothesized that asking students to generate drawings rather than written arguments would make generating arguments supported by evidence more accessible to elementary school students. In other words, we expected to observe evidence-based claims in student groups' collaborative illustrations.

Methods

Participants

The participants in this study were 80 elementary students in grades 1 through 5 at a Title I elementary school located in a suburb of a large southeastern city (Table 1). Title I schools contain a significant proportion of children from low socioeconomic status families and receive federal funding to help children meet state and academic standards. 57% of the students at the participating elementary school were below the poverty line. At the time of this study, schools implemented grade-appropriate Georgia Performance Standards (GPS) as the measure of benchmarks for all academic subjects, but in 2016 the Georgia State Board of Education approved a new set of standards titled the Georgia Standards of Excellence (GSE) for schools to adopt during the 2017-2018 school year (Georgia Department of Education, 2015).

Table 1. Grade distribution of Participating Students

	n	*% of Total*
First Grade	10	12.5%
Second Grade	16	20%
Third Grade	21	26.3%
Fourth Grade	11	13.8%
Fifth Grade	22	28%

All 80 students participated in the PBL unit and design argument assessment. Data used for the genetics knowledge assessment involved only the 67 students who were present in school during both the pre and posttest. These 67 students were comprised of 9 first graders (13%), 10 second graders (15%), 17 third graders (25%), 11 fourth graders (16%) and 20 fifth graders

(30%). Students participating in this study were a mix of talented-and-gifted (TAG) and able learners (i.e., students above grade level on a measure of math or reading ability, or both). During the daily enrichment period, (see intervention procedure below), TAG and able learners engaged in PBL with the school gifted coordinator and support staff including the media specialist, art teacher, math coach, and counselor. Participants were 52.6% female and 46.2% male. One student declined to identify gender. We did not collect race/ethnicity data, but the school district of the participating school is 48% African-American, 37% Caucasian, 8% Hispanic, 4% Multi-racial and 3% Asian.

Unit design and intervention procedure

The intervention took place over six weeks during the daily enrichment period at the research site school. The enrichment period took place from Tuesday to Thursday for 45 minutes a day in the school media center. During this window of time, students school-wide were moved to classrooms or school sites that were not their homerooms for remedial or enrichment time. Students who participated in the study were assigned to enrichment rather than remediation support. This was the third PBL unit students in this group had participated in during the school year, but it was the first unit focused on biology learning. Enrichment period PBLs were designed to align with specific standards that would be addressed in the students' homeroom class for a deeper level of engagement. During all PBL units, students worked collaboratively around a central theme or question. The first PBL unit that students engaged in was a school courtyard redesign project. Students worked in groups to design the space and create a budget for the redesign. Their final projects were presented to a board of community stakeholders and school staff. The second PBL unit was focused on students using scientific reasoning skills to solve a mystery, namely why a farmer's chickens stopped laying eggs. Students were divided into 16 mixed-age groups, with 6-7 students in each group. Students from each grade level were represented within each group, but distribution varied. Each day, the student groups worked through a PBL activity related to genetics and evolution, specifically on animal adaptations. Under the Georgia Performance Standards, approved in 2004, animal adaptations in relation to the environment are introduced as a framework in the first grade; each grade builds on that foundational understanding, but genetics is not included as a distinct unit until the fifth grade. Additionally, these standards identify science communication through writing and drawing as a key competency starting in the first grade and introduce scientific argumentation in the third grade (Georgia Department of Education, 2015). Under the new Georgia Standards of Excellence, the content areas are identical to the previous standards, but the focus has shifted to more inquiry-based instruction.

To start the PBL unit, students were introduced to a fictional scenario in which a group of scientists were dispatched to study organisms called markles, who live on a planet far out of the solar system, Planet Markle. However, the scientists in the scenario are concerned that the markles may evolve and consequently change their appearance during the extended time period of travel from Earth to Planet Markle. Therefore, students were placed in the role of Mission Planners. As Mission Planners, they hypothesized about the appearance of the markles and devised plans for capturing the creatures, based on information on basic genetics principles presented earlier at a series of stations as part of the PBL unit, information provided about the markles' home planet, and the students' own claims about the markles' appearance, respectively. Although fictional, the problem space mimicked that of a real world context. It lacked a rigid and clearly-defined problem space and required students to use information from a variety of sources.

At the beginning of the unit, students were instructed to visit stations to learn how animals on Earth change their appearance over time. Each of the four stations had its own learning objectives, questions to answer, and after completing the station, students earned a factoid (Table 2) related to both their station and fictional markle biology. The first station was titled “Genes, Environment, and Phenotypes: Why do we look the way we do?” and introduced students to the concept of cells containing DNA, the hereditary nature of DNA, and how both DNA and the environment can influence our phenotype, or how we look. Students completed the Dragon Genetics simulation to model how genes influence phenotype, and how this information is passed between parents and offspring. Students also learned about how heat influences the coat color of Siamese cats. The second station, “Adaptations,” introduced students to the concept of adaptations, gave examples of adaptations, (e.g., long necks on giraffes), and described how adaptations can form over time due to evolutionary forces (e.g., Galapagos finches). The third station was titled “Mutations and Survival” and introduced students to mutations (changes in DNA), and how mutation can lead to new adaptations and evolution. Students explored the relationship between mutation and color in pepper moths using an online lesson and accompanying game (peppermoths.weebly.com). The final station, “Genetic Drift and Natural Selection,” explored how different processes influence the number of genes that are available impact a species. Students learned about extinction, non-random selection of genes (natural and artificial selection), and random removal of genes via genetic drift. Students played a game developed by the first author to demonstrate natural selection and genetic drift. For the natural selection demonstration, students took several butterflies and were told that they represented the total population of butterflies in a backyard. Some butterflies contained lots of decoration whereas others were plain. Students then selected half of the butterflies to go into their private collection, and then discussed how removal of those particular butterflies changed the diversity of phenotypes, and consequently genes present in the population. During the second part of the demonstration, students removed 20 numbered beetles from a bag, half green and half orange. Students first counted how many of each type of beetle was present and wrote it in their notebook. Then students rolled a 20-sided die five times. If a beetle's number was selected, it was turned over and considered dead. Students then counted how many orange and green beetles were left after some were randomly removed from the population. This process was repeated at least once more and numbers compared to demonstrate the random impact on the population. Upon completion of each station's activities, students earned a sticker with a factoid relevant to markle adaptations. A list of factoids collected at each station is in Table 2.

After students completed all of the stations, they were randomly assigned features of a particular region of Planet Markle. Examples of these features are shown in Table 3. Based on these features, students drew what they thought the markles in their region would look like. Students also were told to design a trap to capture the markle based on their predictions of the creatures' appearance and behavior. Each student group prepared a presentation during which students described their markle and their plan for trapping the markle. To support their presentations, students prepared notecards with details about the markle and trap design. Students received no instruction on argumentation or scientific explanation.

Unscripted scaffolds were provided by the five instructors present during instructional focus time and by the series of stations at which students learned genetics principles. The first author assessed fidelity of curriculum implementation by visiting the school site once a week to observe the unit. The first and second authors attended the students' concluding presentations to make observations and to collect student artifacts, including the groups' illustrations and notes.

Table 2. *The factoids students earned upon completing each station. The students integrated these factoids into their final markle drawings.*

Station	Factoid
Genes, Environment, and Phenotype	Markles inherit different genes from their parents. Since these genes lead to certain phenotypes, if a markle has one parent with spotted fur (and the gene for spotted fur), they will also have spotted fur. When a markle is born, it is light colored. If the markle grows up in a cold climate, it will eventually be blue. Their blue color helps them blend in with the water, so blue markles tend to live near water and eat fish. Markles that grew up when it is warmer are green and live in the trees where they can blend in better with their surroundings. Markles want to blend in with their surroundings so other animals do not eat them.
Adaptations	Markles have different adaptations that they use to find food, escape predators, and have babies. These adaptations include big noses, long legs, long hair to help camouflage and keep the markle warm, wings for flying, snorkels, and suction cup hands. Big noses and snorkels help the markles find food. Without the snorkels, markles can't go fishing for food. Their big noses help the markles sniff out food and without them, they would be limited to only using their eyes to look for food. Giant bears can eat markles. The giant bears are slow and because of their size, can't climb trees to chase after the markles. Therefore, if a markle has adaptations such as wings, long legs or suction cup hands for climbing the markle is more likely to get away. Without these adaptations, the markle relies on camouflaging behind its long hair, or its color. Remember, much like the finches we learned about in this station, although all markles are similar, they have different adaptations depending on their specific environment. As the markles' environment changes over the time it takes for our astronauts to arrive, the markles' adaptations will also change.
Mutations and Survival	Some markles will be albino (all white) because of a random mutation. Being albino can be good or bad depending on whether there is lots of snow on the ground or not since the markle can hide better on a white background. If there is no snow on the ground, the albino markle can't hide easily and is more likely to be eaten by predators.
Genetic Drift and Natural Selection	There are many volcanoes on Planet Markle. These volcanoes can kill markles that get stuck in the lava flows. The volcanoes can also spread ash everywhere turning all of the trees from green to white. Some markles have spots on their fur. Next to Planet Markle is a planet where a species of aliens called narps. Narps love to wear spotted markle fur and hunt spotted markles for their fur.

Table 3. *Sample features of Planet Markle geography that were distributed to students after completing all of the unit stations. Students based their choices regarding the markles' morphology and the features of traps designed to catch the markles on the random combination of Planet Markle features they received.*

Sample Features of Planet Markle Geography

Your region is very cold.

Your region of Planet Markle is very dark and it is difficult to find food using eyesight alone.

Your region of Planet Markle has large cliff faces. Birds like to nest high on these cliff faces, and markles love to eat their eggs.

Your region is susceptible to tidal waves and markles are at risk for drowning without certain adaptations.

Your region is in a desert with lots of quicksand.

Data collection and analysis

We collected data associated with our primary research questions, including students' genetics content knowledge and illustrated design arguments. Content knowledge was collected at the individual level, while illustrated design arguments were completed in groups.

Genetics assessment. Previous work identified common genetics misconceptions held by students in high school (Shaw et al., 2008). Since these misconceptions persist and may begin in the early elementary grades (Smith & Williams, 2007) we decided to focus our assessment on these common misconceptions, rather than a comprehensive assessment of everything learned during the PBL unit. In order to determine if engagement in PBL resulted in enhanced understanding of these particularly difficult to understand genetics concepts, we asked students to do a pre/posttest assessment of their understanding of two of domains of these common misconceptions: (1) the deterministic nature of genes and (2) the nature of genes and genetic material. We asked students two questions related to the deterministic nature of genes: Why do people look different from each other? Why don't all people have the same hair color? We asked students one question about the nature of genes and genetic material: What are genes? The three items were open ended. Students answered the items at the beginning and end of the PBL unit. We removed data from students who were absent during either the pre or posttest. Sixty-seven students participated in both the pretest and posttest.

Student responses were blinded and then coded by two researchers on a scale of 0-4, where 0 = no response and 4 = mature understanding. A summary of the rubrics used is shown in Table 4. When coders disagreed, a third researcher served as the tiebreaker. Interrater reliability for this coding was acceptable for all three questions (Question 1, Cohen's *kappa* = 0.92; Question 2, Cohen's *kappa* = 0.93; Question 3, Cohen's *kappa* = 0.81).

Table 4. Rubric for coding students' responses to genetics assessment items.

Score	**Questions 1 and 2**	**Example Answer**	**Question 3**	**Example Answer**
0	No response, illegible, un-intelligible or "I don't know"		No response, illegible, un-intelligible or "I don't know"	
1	No attempt at explanation, restates the question	People weren't born the same	Mentions characteristics controlled by genes such as skin or hair color. No mention of parents, inheritance, or DNA.	Genes are a type of animals like cats. All cats are in the same genes
2	Non-biological explanation (God, ethnicity, culture)	Because God made them that way	Describes genes as cells or traits. No mention of heritability.	Genes are things in your body that give you characteristics
3	Biological explanation that includes understanding of family, parents, or	Our parents did not look the same so we are most likely not	Mentions concepts of inheritance, but does not fully explicate what genes are. Describes genes as	Genes are things that are passed on from your parents

	inheritance. Does not mention genes	to look the same as our parents	traits or generic "things" that come from parents.	
4	Explicitly mentions genes, genetics, inheritance	People look different from one another depending on the genes they receive from their parents	Explicitly mentions DNA. Makes connection between DNA and a trait or relationship to parent.	Genes are the DNA in your body that determines who you are and what you look like

Design argument assessment. We collected group-generated design drawings and supporting presentation notecards for data analysis. Notecards contained details about illustrated markles and traps that students felt were most important to their presentations. Students received no direct instruction regarding what information to include on their notecards. Brooks (2009) contended that one of the strengths of using drawings for students to express their understanding lies in the ongoing facilitation of dialoguing in other modes (such as writing) that help students explore complex ideas. In light of this connection, we chose to include the presentation notecards in the argument analysis in an attempt to fully capture student groups' use of evidence and reasoning in their design solution. Final products were evaluated by assessing the presence of claims, evidence, and reasoning elements in the drawings and notecards.

Being a design challenge, we analyzed students' work using adapted conceptualizations of argument components (claim, evidence, reasoning) (McNeill & Krajcik, 2006) better suited for design challenges as opposed to scientific investigations. Broadly speaking, these three elements of written arguments provide different kinds of information to a reader. The claim is an answer to the problem or question posed in a particular context (Sampson, Enderle, & Grooms, 2013). Although, structurally, the claim is typically provided first, these answers are ultimately derived from the evidence generated during an investigation or development of a problem solution (Sampson et al, 2013a). Evidence in scientific arguments entails the data and information that has been analyzed and interpreted, involving the identifications of trends, patterns, comparisons, and contrasts. The reasoning component of an argument involves making connections between the claim and evidence using design principles and scientific concepts (Sampson et al, 2013a).

With respect to *claim* in the Markle design challenge, the groups' markle drawing was considered to be the overarching claim as it represented the actual solution to the design challenge embedded in the PBL unit. Specific features of the markle highlighted in the drawings or notecards were also considered as claims. We characterized *evidence* as the environmental constraints, in the form of factoids or Planet Markle features, provided to the student groups during the unit. For evidence to be considered present in the groups' design arguments, either explicit drawings of the environmental elements described in the factoids or features, or explicit statements describing them in the notecards had to be included. Although we do see instances where students choose features directly from the factoids, this is not always the case. It is likely that reasoning occurred at different levels, with some creative ventures pursued while the factoids are utilized generally. Finally, *reasoning* elements included drawings or statements that connected particular features of the markle to the environmental constraints identified by each group. Typically, reasoning statements involved descriptions of the function or purpose of a certain feature in relation to a specific environmental constraint.

The drawings and notecards for sixteen student groups were analyzed using the design argument framework described. The first and third author collaborated on this analysis, reaching agreement on coding through continuous negotiation until agreement was reached. The third author was not part of the team who administered the PBL unit to students, so was not biased as to how students generated drawn arguments. The analysis focused only on the design solutions generated for the markle organism, as these solutions required the consideration of environmental constraint information collected through student activity and markle traps did not. Thus, the trap designs are not discussed further here. The use of environmental constraint information as evidence required students to build on their understanding of the genetic and environmental concepts taught during the station activities at the beginning of the PBL unit. The analysis quantitatively analyzed the presence and frequency of various combinations of claim, evidence, and reasoning components present in each groups' set of drawings and notecards. The analysis of groups' drawings included counting drawn and written components present and their nature, while the notecards were analyzed for the presence of various combinations of written argument components. Our focus for this analysis was to describe the variation of argument structures produced by students who engaged in the Markle PBL unit. We did not engage in an evaluation of the quality of these elements, as our research focus pertained more to understanding what kinds of elements emerged in students' arguments when afforded opportunities to draw them instead of just writing text for them. Following this analysis, we identified several groups to further describe using small case studies (Stake, 2006) to demonstrate and compare some of the variations in the design solutions generated.

Results

Students' understanding of genetics after engaging in PBL

Since little is known about genetics learning in elementary school students and very few interventions have been developed to improve genetics understanding in elementary school students (Duncan et al., 2009), we first determined if students engaged in a genetics PBL demonstrated learning gains, particularly gains in traditionally difficult to understand genetics concepts. To assess changes in genetics understanding, specifically the relationship of genotype to phenotype and the nature of genes, students ($n = 67$) completed pre and posttest items. Performance on the both the pretest and posttest was normally distributed. First, we conducted a paired sample t-test to compare the pretest and posttest means. Responses on the posttest (*mean* = 3.05, $SD = .45$) were significantly different than responses on the pretest (*mean* = 2.68, $SD = .66$; $t = -2.32$, $df = 19$, $p = .03$, *Cohen's d* = .66). Next, we examined if student performance on all three of the genetics test items improved between the pre and posttest. For the first two items, Why do people look different from each other?; Why don't all people have the same hair color? within-subjects paired t-tests revealed a statistically significant increase in conceptual understanding from pretest to posttest (Question 1: pretest $M = 2.89$, $SD = 1.14$; posttest $M = 3.25$, $SD = 1.13$; $t(66) = -2.20$, $p = .03$, Cohen's $d = 0.32$; Question 2: pretest $M = 2.97$, $SD = 1.18$; posttest $M = 3.35$, $SD = 1.06$; $t(66) = -2.14$, $p = .04$, Cohen's $d = 0.34$; Figure 1). Using a third within-subjects paired t-test, we did not observe a statistically significant difference in students' understanding of the nature of genes between the pretest and posttest (Question 3: pretest $M = 2.22$, $SD = 1.13$; posttest $M = 2.17$, $SD = 1.26$; $t(66) = 0.26$, p = .80, Cohen's $d = 0.04$; Figure 1)[1].

Place Figure 1 Here

---

[1] We noted that second graders performed more poorly on the posttest for Question 3 than did other grade levels (Figure 2). However, when we omitted second graders' data from the analysis, there were still no statistically significant gains between pretest and posttest for Question 3. Figure 1 illustrates pretest and posttest scores for Question 3 with second graders included.

Next, we sought to determine whether improvements in overall performance on the genetics assessment differed for each grade level. We conducted paired samples *t*-tests comparing pretest mean score and posttest mean score for each grade level—1st, 2nd, 3rd, 4th, and 5th. Scores improved from pretest to posttest for all grade levels, except for 2nd grade. The only significant difference from pretest to posttest was for 3rd graders (*mean at pretest* = 2.57, *SD* = .74; *mean at posttest* = 3.08, *SD* = .64; *t* = -2.82, *df* = 16, *p* = .01., *Cohen's d* = .74). (See Figure 2.)

Given our observation above that as a whole, students improved on their posttest scores on questions 1 and 2, but not 3, we next sought to determine differences in change scores from pretest to posttest across the grade levels, first through fifth (Figure 2). For Questions 1 and 2, a one-way analysis of variance (ANOVA) revealed no statistically significant differences in score change (posttest score – pretest score) among the grade levels. For Question 1, on average all students in each grade level demonstrated gains in conceptual understanding from pretest to posttest (Figure 2). For Question 2, overall gains were seen for third-, fourth-, and fifth-grade students, but not for first- and second-grade students (Figure 2). For Question 3, there was a marginally significant difference in change scores across grade level ($F(4,62) = 2.37$, $p = 0.62$). Post-hoc analysis with Fisher's least significant difference test indicated that second graders were different than all other grade levels. Specifically, second graders performed more poorly at the posttest than at pretest for Question 3; whereas first and fourth graders demonstrated modest improvement and third and fifth graders no improvement (Figure 2).

Place Figure 2 Here

To determine whether or not this discrepancy was the result of differences in baseline understanding, we compared pretest scores among all grade levels. Students did not perform differently at pretest on Questions 2 or 3. However we did observe a statistically significant difference on pretest performance for Question 1 ($F(4,62) = 4.70$, $p = 0.00$). Post-hoc analysis with Fisher's least significant difference test indicated that fifth graders' pre-test scores (M = 3.17, SD = 1.23) were statistically higher than all of the other students' scores (first grade, (M = 2.30, SD = 1.23); second grade (M = 2.50, SD = 1.22); third grade (M = 2.57, SD = 1.17); fourth grade, (M = 2.54, SD = 0.97).

Characteristics of student-generated design solution arguments

Reviewing the products generated by student groups during the markle design activity revealed interesting trends. Overall, the different groups were able to incorporate the environmental constraint information in their designs of a markle. Across all groups, features were included in each markle design that could be reasonably connected to the specific environmental factoids each group received during the first part of the unit. Thus, at least anecdotally, evidence exists that all student groups were able to process the environmental constraint factoids and develop designs responsive to them. However, the analysis described here does not include any inferred or anecdotal connections between students' designs and relevant environmental constraints. The following analysis first focuses on general trends in explicit elements, either drawn or written, included in the markle design presented to the other groups. As these designs represent complete argument products emerging from the unit activities, they were the primary unit of analysis. We expanded that unit to also include the notecards that were prepared for the presentation in an effort to capture as many connections as possible being made by students between design features (claims), environmental constraints (evidence), and the relevant functions of those features (reasoning).

The analysis of all groups' arguments involved both the actual drawing presented and the presentation notecards prepared by the students. Table 5 provides an overview of how student groups incorporated different argument elements in the drawing they presented. Recall, two primary aims of this study were to investigate whether elementary students are able to coordinate claims, evidence, and reasoning and how the creation of illustrated arguments relate to this ability. All groups developed a drawn design, so at minimum, they all produced a drawn claim. However, a claim alone is not an argument. Six out of 16 groups (38% of the groups) only provided a drawn design/claim with no other supporting information, thus not providing an argument in their final design solution. Three more groups only included written labels with their drawings to highlight certain features. This means the majority, approximately 56%, of the groups did not provide an argument in their illustrated design solutions. The remaining groups did incorporate explicitly some combination of argument elements in their presented design solution. Only one group explicitly incorporated constraint information through drawing, using illustrated evidence about the environment to support markle features they developed. This group also included written reasoning statements to describe the functions of their design relative to the environment. Another group only included written reasoning statements with their drawing that described the functions of specific design features. Finally, five student groups included several written statements on their presented designs that incorporated combinations of the different argument elements, with some statements including both evidence and reasoning and some with only reasoning included. It is worth noting that all of the illustrated arguments included text on the illustration, with the exception of a single drawn claim plus drawn evidence argument. However, even in this case, the group included a written reasoning statement on their illustration. We describe this group in more detail as a case study later.

Table 5. Quantitative Comparison of Claim-Evidence-Reasoning (CER) Elements in Drawings

	CER Elements in Drawings	**# of Groups**	**% of Groups**
No illustrated argument present	Drawn C Only	6	38%
	Drawn C with Written Label Only	3	19%
Illustrated argument present	Drawn C with Drawn E & Written R	1	6%
	Drawn C with Written C, E & R elements	5	31%
	Drawn C with Written R Only	1	6%

Note: C = claim, E = evidence, R = reasoning.

The analysis of the notecards provided a more complex view of how each group incorporated different elements of evidence and reasoning in the final design solution they presented. Table 6 provides a quantitative description of the different combinations of argument components present on the notecards for each group. Each statement on a group's notecard was analyzed individually for argument components. In some instances, pairs of sentences were analyzed together as they comprised one coherent unit of argument elements. Although more complex in the distribution of combinations, broad trends are readily apparent in the table. First, the majority of the statements students wrote on their notecards emphasized connections between certain features of their markle design (claim) and their function (reasoning). Yet, it is notable that student groups were more explicit in being sure to call out specific environmental constraint evidence to support the features of their markle design for their presentation, either separately or in combination with elements of reasoning, rather than paired with their claim. Thus, in the overall analysis of the groups' final design solutions, student groups' ability to incorporate argument components was more prevalent in their written notations as opposed to in their illustrated forms. Nevertheless, it is important to note that the illustration provided a point from which claims, evidence, and reasoning could emerge.

Table 6. Quantitative Comparison of CER Elements in Notecard Statements

Group	**C Only**	**E Only**	**C + E**	**C + R**	**C+E+R**
A				6	
B				1	2
C	1			4	2
D				2	3
E				8	
F	1	1		5	1
G			2	1	1
H	4			1	2
I		2	3	1	1
J	1	2	1	5	
K				7	2
L					
M					
N					
O	2			2	2
P			1	6	1

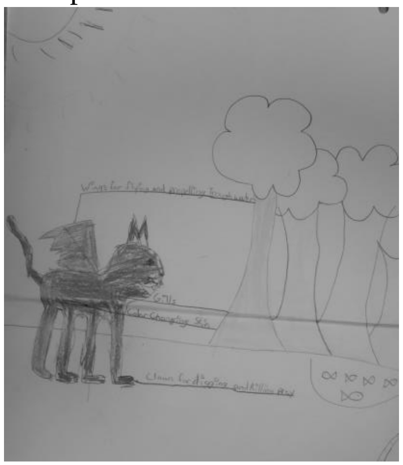
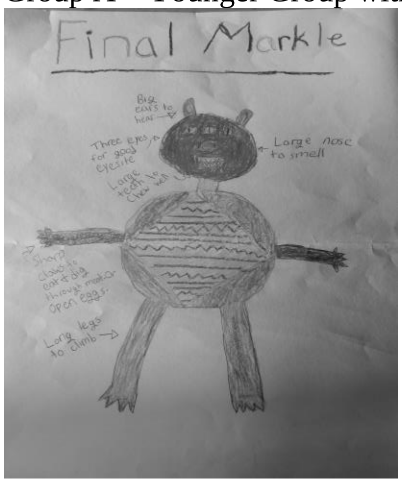
Variations among student groups: Four cases

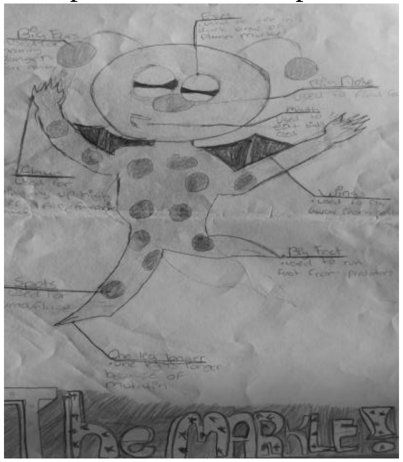
To provide further insight into the kinds of design solutions developed by elementary students in this activity, we provide a more detailed description of four student groups, with each group serving as a case for this part of the analysis. The four cases were selected to represent unique solutions to help the reader see the nature of the design solutions presented. Table 7 provides pictures of each group's markle design and other relevant information.

Group E: A case of drawn evidence. Only one group out of 16 explicitly drew design constraints, or evidence as conceptualized here. The environment constraints related to the presence of lakes and "tall trees" in the markle's environment are prominent elements in Group E's drawing. The few written statements included in the group's drawing focus only on reasoning elements, describing the function of certain features (e.g., "Claws for digging and killing prey", "Wings for flying and propelling through water"). Interestingly, though they were the only group to draw evidence, they were also one of the only groups that did not include evidence in their written notecard. Example statements from their notecards include a string of several claim and reasoning combinations: "The wings are to fly and propel through water. The eyes for night vision. Fangs for biting prey. Gills for swimming. Claws for killing prey and digging wandering around. Changing color of fur."

This group provided one of the largest amounts of written argument combinations on its notecard as well. The group composition was relatively equal in its distribution across grade levels, with two first graders, and one student from every other grade level, second through fifth.

Table 7. Comparison of Student Group Cases

Drawing	Environmental Constraints/Evidence	CER Elements - Drawing	CER Elements - Notecards	Students' Grade Level
Group E– Drawn Evidence				
	• Your region has large cliff faces. Birds like the nest high on these cliff faces, and markles love to eat their eggs. • All of the lakes in your region of Planet Markle contain large markle-eating sea monsters. • Giant snakes that like to eat markles live in your region. These giant snakes can't see markles that are hidden well. • Your region is susceptible to tidal waves and markles are at risk for drowning without certain adaptations. • Your region is full of tall trees	1 Drawn C 2 Drawn E 3 Written C+R	8 C+R Statements	1st – 2 2nd – 1 3rd – 1 4th – 1 5th – 1
Group A – Younger Group with Large Amount of Reasoning on Drawing				
	• Your region is very cold • Your region is very snowy • Your region is full of tall trees • Your region of Planet Markle has large cliff faces. Birds like to nest high on these cliff faces, and markles love to eat their eggs • Your region of Planet Markle is very dark and it is difficult to find food using eyesight alone	1 Drawn C 6 Written C+R	6 C+R Statements	1st – 2 2nd – 2 3rd – 2 4th – 0 5th – 1

Drawing	**Constraints**	**CER-Drawing**	**CER-Notecard**	**Grd. Lvl.**
Group K – Older Group with Large Amount of Reasoning on Drawing				
	• Your region is very dark and it is difficult to find food using eyesight alone • Your region has large cliff faces. Birds like the nest high on these cliff faces, and markles love to eat their eggs. • Your region has very few trees and rives full of fish • Your region is in a desert with lots of quicksand • Your region produces gamma rays that cause mutations in your markles that change their fur color	1 Drawn C 3 Written C+E+R 6 Written C+R	2 C+E+R Statements 7 C+R Statements	1st – 1 2nd – 1 3rd – 2 4th – 2 5th – 1
Group G – Only Claim Drawn, Notable Written Evidence				
	• Your region has large cliff faces. Birds like the nest high on these cliff faces, and markles love to eat their eggs. • Your region if full of tall trees • Your region is susceptible to tidal waves and markles are at risk for drowning without certain adaptations • Your region contains lots of large beaver-like animals that cut down all of the trees to build dams • Aliens from Planet Narp want the markles spotted fur.	1 Drawn Claim	2 C+E Statement 1 C+R Statements 1 C+E+R Statement	1st – 1 2nd – 2 3rd – 1 4th – 1 5th – 2

Groups A & K: The impact of age. Two other groups who both incorporated a larger number of reasoning statements in their drawing and notecards also provide a contrast in group structure. Group A was markedly younger, with most members being in third grade or below. They provided reasoning elements explaining the functions of each design feature they highlighted in their drawing for a total of six, such as “Long legs to climb,” (Written C+R) and “Three eyes for good eyesight.” (Written C+R) Most of these statements were repeated in their presentation notecards.

Similarly, another group, Group K, provided a slightly larger amount of statements with argument components on both their drawing and notecards. The age composition of this group skewed slightly older as the majority of student members were in third grade or older. A notable difference in the statements provided by this group involves their use of explicit references to environmental constraints as evidence to support certain design features. On their drawing, the group wrote statements including, “Claws used for climbing up high cliffs for favorite snack.” (Written C+E+R) Similar statements were included on the group’s notecards, such as, “Eyes: These eyes are used to see in the dark areas of Planet Markle.” (Written C+E+R) Although this group did incorporate evidence in some of their statements, the majority of their written statements involved only claim and reasoning elements. Thus, both groups, although comprised of different age concentrations, demonstrate the ability to incorporate relevant argument components in explaining their design solutions. The presence of more explicit references to constraint information as evidence from the older group of students does suggest a developmental emergence of understanding the need to coordinate multiple argument components.

Group G: Illustrations with no written text. We also analyzed another case where the student group only provided a drawing for their final design solution with no written elements present. Although they did not add any further detail to their drawing, they did provide statements comprised of argument elements on their notecards, although limited to only four. While smaller in amount of statements, the bulk of this group's writing could be considered higher quality due to explicitly incorporating evidence from constraints to defend particular features. Examples of such statements include: "Our markle has fins and gills because our markle is at risk of drowning," (Written C+E) and "Our markle has claws because our markle uses his claws to climb up cliffs to get bird eggs." (Written C+E+R)

The first example provided shows an example of a statement that includes only claim and evidence elements without text describing their reasoning involving function. In comparison, the second statement shows how this group coordinated evidence and reasoning elements for a specific design feature: an ideal argument. As described in Table 7, the grade composition of this group also tended towards being older with over half of the students being in third grade or higher. Indeed, out of the nine groups that only provided a drawing (some with labels) for their design solutions, six of them did incorporate explicit references to constraint evidence in their notecards in a similar manner as the Group G. All six of these groups shared another similarity in that their grade level compositions tended to be a majority of older students in third grade or above. Of the groups that did not provide explicit reference to constraint evidence in the notecards, most of them were more balanced with greater numbers of students in lower grades.

Discussion

PBL for teaching genetics to elementary students

Although genetics is typically taught to older students (Duncan et al., 2009), we observed that the elementary-aged students participating in this PBL unit increased their understanding of genetics principles that are typically misunderstood. Third graders experienced significant gains in their overall genetics understanding after participation in the PBL. More specifically, we noted that students' understanding of the deterministic nature of genes were most likely to improve. Students' understanding of the nature of genes and genetic material did not improve, which may indicate that this is a particularly difficult concept for young students to grasp and likely requires additional instruction beyond what was provided at the unit stations. Regardless of grade level, following the PBL unit, students were more likely to explicitly reference genes, inheritance, or genetic material when asked why people look different from one another. When asked about hair color, however, there was a variation in improvements in conceptual knowledge across grades. Specifically, first graders showed a slight decrease in performance from pretest to posttest and second graders had no change in their performance. The older students all improved their scores on this item. This may indicate that this was a slightly harder question for the younger students, or that the students still required some additional scaffolds to fully understand this concept. The cognitive demands of the questions' open-ended nature may have posed differential demands across development as well. These findings align with prior work that suggests genetics content may be particularly difficult for students, but can be best scaffolded, even for younger children with PBL or inquiry instruction (Araz & Sungur, 2007; Henderson & Maguire, 2000; Shaw, Van Horne, Zhang, & Boughman, 2008; Smith & Williams, 2007; Venville, Gribble, & Donovan, 2005)

Interpreting changes in students' understanding of genes is more difficult, given our findings. Overall, students were no more likely to acknowledge DNA or inheritance in their description of genes following the PBL unit. First and fourth graders demonstrated modest improvements, but we observed no difference in understanding for the third and fifth graders. Interestingly, second graders were less likely to mention heritability following the unit. Notably, this was the case for all second graders rather than being driven by an outlier. Additionally, second graders answered this item similarly to students in other grades before the PBL unit. Since the nature of genes is a common misconception, and since this large drop was only seen among one group of students, it may be the case that instruction outside of the PBL in the second graders' normal classes may have resulted in the presence of a misconception or overemphasis on the connection between genes and traits rather than on heritability. Unfortunately, the research team is unaware of and unsure about what such disruptive instruction may have been. Despite anomalous findings with regard to students' understanding of what genes are, improvements in understanding inheritance following the unit are promising. Even very young students may be capable of learning genetics concepts. We take this as evidence for genetics learning progressions starting earlier than late elementary or middle school. We also take the shift in second grade TAG students – who were engaged in accelerated third grade genetics curriculum outside of the unit – as evidence that students' genetics misconceptions require specific sensitivity in early years.

Arguments constructed in a PBL unit

One aim of this work was to determine if elementary school students are capable of employing reasoning to support claims with evidence in their illustrated products and without explicit instruction about constructing arguments using evidence and reasoning. Students' illustrations and presentation notes as part of a capstone assignment following the PBL unit explicitly incorporated major components of arguments—claims, evidence, and reasoning.

Mixed age groups of high achieving and gifted students successfully made claims about their illustrated design solutions and used evidence to support these claims, all without scaffolding or instruction specific to argument construction. The nature and quantity of different combinations of claims, evidence, and reasoning elements varied noticeably across groups in their drawings and written notation.

The variation noted across these groups offers further demonstration of a developmental trajectory for students' ability to engage in the coordination activities necessary for the development of scientific arguments argued for by others (Kuhn, 1991, 2005). For learners to improve in their ability to construct higher quality arguments, they must also improve their ability to evaluate their knowledge products using metacognitive abilities (Garcia-Mila & Andersen, 2007). Looking across the groups described in Table 6, those that had a larger share of older students produced richer collections of drawn and written argument components. Thus, our findings agree with other scholars who have argued for the importance of developing metacognitive abilities in complement to enhancing their ability to learn through argumentation (Garcia-Mila & Andersen, 2007; Kuhn, 2005). Following this line of thinking, incorporating instructional elements that afforded students opportunities to explicitly reflect on their design solutions could potentially have increased the groups' explicit coordination of their Markles to the environmental constraints they had. Incorporation of such intentional scaffolding has been shown to be helpful in such learning (Felton, 2004; Zohar & Nemet, 2002).

A premise for the current study contends that by allowing students to draw their design solutions, the difference in expectations would facilitate a more accessible venue for elementary students to create high quality arguments. The results developed here do not offer resounding support for this premise in that only one group actually illustrated elements of evidence in its final solution. However, several other groups used their drawings to then generate statements that did incorporate both evidence and reasoning components to explain particular design features. Beyond the drawn design solutions, the presentation notecards also offer further demonstration of student groups explicitly incorporating elements of evidence and reasoning to argue for the design solution they developed. Although students did not express them through drawing, the mode of expression did provide a vehicle for them to incorporate argument components in a coherent manner. Therefore, we agree with scholars who assert that drawings are a valid form for having students express their understanding of complex events (Bowker, 2007; Chang, 2012). The current study adds further texture to this notion by demonstrating that although students may not fully express themselves using this mode of expression, drawing can also facilitate students' use of writing in a more meaningful manner. The student groups were prompted through their drawings to explain at minimum their reasoning for including certain design features as well as evidence (environmental constraints) to support their inclusion through written text, both on the image and in their notecards. The use of drawings in the science classroom can assist in helping students express complex ideas through imagery, but also provide an expressive anchor to ground their writing in as well. Yet, the use of drawings does still present challenges in the science classroom.

For students, the word "argument" may have a negative connotation, which can influence how students engage in discourse (McNeill, 2009). Rather than ask students to create an argument, we asked students to create an illustrated design solution to assess the presence of claims, evidence, and reasoning. Even though this task differs from typical argument construction tasks, our finding that students have difficulties explicitly connecting evidence to claims is consistent with prior research analyzing students' written arguments (Berland & Reiser, 2009; McNeill, 2011; Ryu & Sandoval, 2012). For example, prior to instruction in argumentation, Ryu and Sandoval (2012) rate third- and fourth-grade students' written arguments as having little-to-no evidence cited and lacking explicit justifications. These challenges persisted, although alleviated somewhat, after instruction. In the current study, we also noted students did not explicitly include appropriate evidence in the majority of argument statements they constructed in writing. The persistence of these student challenges provides further support for the emphasis on the scientific and engineering practices identified in the NGSS and state adopted variations of them. Having students engage in these practices is not merely enough, rather we must also help them to understand the nature and role of these practices in science (Ford, 2008; 2015). In light of the results of this study and others noted previously (McNeill, 2011; Ryu & Sandoval, 2012; Venville & Dawson, 2012), we agree with this premise and the importance of incorporating instruction in science classrooms that addresses the practices and their constitutive elements explicitly. Further, if we endeavor to help students gain better understanding and proficiency with arguing from evidence, science educators must also be mindful to help students understand variations in the types of evidence necessary for particular purposes and problems.

The instructional unit and related tasks involved in this study were framed using a PBL approach. Often, PBL uses ill-structured problems to frame the entire unit and contextualize the science content to be learned (Savery, 2015). However, the problems students engage in solving are not always answered through empirical investigation. Rather, the end products for students engaged in some PBL units are more aptly characterized as problem solutions. As such, students must come to understand the difference between an empirical investigation and the development of a designed problem solution (Berland, 2013; Householder & Hailey, 2012; Leonard, 2005).

To develop solutions to design problems, students must engage in the design process, which includes empirical investigations to test prototypes, but also involves the development of potential solutions that can be used to develop prototypes (Berland, 2013). In this study, the main product of students' efforts reflects this first stage of the design process where they developed markle organisms that could potentially survive in the environmental constraints they collected through their factoid finding work at the beginning of the unit. We argue that in this stage of the design process, material regarding constraints is the most plausible source for external information that can be used to assess the appropriateness for a particular claim or design feature. This information is similar to analyzed data collected during an investigation serving as evidence to support a claim answering the question guiding the investigation. These differences in the nature of information needed for evidence for particular types of tasks can also help explain why we did not see as many evidentiary elements incorporated into student groups' drawn and written markle solutions.

The different markle designs from all groups demonstrated that students were mindful of the environmental constraint information they collected, as the features and reasoning statements provided by groups often implied, if not explicitly mentioned, one of the design constraints. It is reasonable to think that if these students had been provided explicit instruction in what the components of a high quality arguments included, particularly in a design solution context, then more groups would have provided explicit connections. The results of this study are promising when interpreted to show that even without such support and guidance intentionally embedded in the PBL unit, several groups did seek out those conceptual connections in the solutions they presented. We concur with other scholars who have also drawn attention to the importance of incorporating explicit teaching in science classrooms that focuses on the unique characteristics and elements of design problems and having students engage in those activities in meaningful ways to help them understand the differences compared to scientific investigations (Berland, 2013; Householder & Hailey, 2012).

Considerations for PBL use

The findings of this study can also inform science teachers who work on implementing PBL instruction in their classrooms. As seen in this study, PBL units do create contexts for students to learn complex science content, such as genetics and evolution, as well as getting them to engage in multiple science and engineering practices, such as arguing from evidence. Indeed, these contexts offer teachers opportunities to have students make personal, affective connections with the content they are trying to teach and provide shared experiences of students' participation in those practices, creating space for students to learn from each other. Yet, to engender these types of learning events through PBL, teachers must be mindful of providing explicit instruction in the fundamental nature of the practices. PBL can be used to frame students' engagement in a practice, but to gain a solid understanding of the practice students must also learn about its elements (Ford, 2008). Thus, with respect to the PBL unit in this study, students could have received instruction about what elements are necessary for a high quality argument once they had completed station activity focusing on genetics content and collection of environmental constraint information. Further, this unit could have been enhanced by also helping students understand the broad goal as a design task and the distinct criteria for what claims, evidence, and reasoning entail in such contexts. Explicitly distinguishing the types of argument producing activities (e.g., investigations and design problems) would also enhance students understanding of the variations of activities that comprise the scientific enterprise.

Limitations & future directions

Despite the benefits of our findings, these are somewhat exploratory and not without limitation. Understanding how illustrations serve as points of assessment and intervention in young children's argumentation skills should be expanded in future work by employing experimental design and adding instructional scaffolds, such as peer reviews or feedback on students' labels of claim, evidence, and reasoning in sample illustrations. Because we did not include any scaffolds or probes of student thinking (e.g., peer review or interviews) in this study, our current findings are limited to coding decisions based on our interpretations. While we expect including additional probes would further validate our findings, we were effortful in our current design in order to observe young students' abilities in constructing arguments without the interference of any instruction. Our findings are further limited in their generalization to a broad population. We examined outcomes with a mixed group of talented-and-gifted and able learner students from one school site. These students participated because they were free to engage in flexible curriculum while other students worked to develop proficiency in math and literacy during a daily enrichment/instructional focus period at the school site. The research questions considered here should be further explored in other populations of students.

Conclusion

Argumentation is an essential part of both science practice and education, but is challenging for students to learn. This work demonstrated that a different type of task, the creation of illustrated fictitious aliens, when assessed as an argument, shared many features and challenges seen when students engage in typical classroom argumentation tasks. The work here proposes an additional method for teaching and studying elementary school students' argumentation practices and provides evidence of its utility for gaining new insights into how children learn and understand arguments, particularly in a design setting. Students were able to demonstrate success in gaining difficult genetics content knowledge as well as in preparing arguments with, at minimum, two components of well-constructed arguments, and in most cases providing a claim supported by reasoning. This work is informative for understanding student abilities, the potential of PBL, and considerations for its use.

References

Allen, D. E., Duch, B. J., & Groh, S. E. (1996). The power of problem-based learning in teaching introductory science courses. *New Directions for Teaching and Learning*, *1996*(68), 43–52. https://doi.org/10.1002/tl.37219966808

Araz, G., & Sungur, S. (2007). Effectiveness of problem-based learning on academic performance in genetics. *Biochemistry and Molecular Biology Education*, *35*(6), 448-451.

Baser, D., Ozden, M. Y., & Karaarslan, H. (2017). Collaborative project-based learning: an integrative science and technological education project. *Research in Science & Technological Education*, *35*(2), 131–148. https://doi.org/10.1080/02635143.2016.1274723

Bell, S. (2010). Project-Based Learning for the 21st Century: Skills for the Future. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *83*(2), 39–43. https://doi.org/10.1080/00098650903505415

Belland, B. R., Glazewski, K. D., & Richardson, J. C. (2011). Problem-based learning and argumentation: Testing a scaffolding framework to support middle school students' creation of evidence-based arguments. *Instructional Science*, *39*(5), 667-694.

Belland, B. R. (2010). Portraits of middle school students constructing evidence-based arguments during problem-based learning: The impact of computer-based scaffolds. *Educational technology research and Development*, *58*(3), 285-309.

Berland, L. K. (2013). Designing for STEM Integration. *Journal of Pre-College Engineering Education Research (J-PEER)*, *3*(1). http://doi.org/10.7771/2157-9288.1078

Berland, L. K., & Reiser, B. J. (2009). Making sense of argumentation and explanation. *Science Education*, *93*(1), 26–55.

Berland, L. K., Schwarz, C. V., Krist, C., Kenyon, L., Lo, A. S., & Reiser, B. J. (2015). Epistemologies in practice: Making scientific practices meaningful for students. *Journal of Research in Science Teaching.*

Bowker, R. (2007). Children's perceptions and learning about tropical rainforests: An analysis of their drawings. *Environmental Education Research*, *13*(1), 75-96.

Brooks, M. (2009). Drawing, visualisation and young children's exploration of "big ideas". *International Journal of Science Education, 31*(3), 319-341. doi:10.1080/09500690802595771

Chang, N. (2012). What are the roles that children's drawings play in inquiry of science concepts?. *Early child development and care*, *182*(5), 621-637.

Dentzau, M. (in press). Students' changing mental models of the longleaf pine ecosystem after involvement in an outdoor environmental education program. *Southeastern Naturalist*.

Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: a meta-analysis. *Learning and Instruction*, *13*(5), 533–568. http://doi.org/10.1016/S0959-4752(02)00025-7

Duncan, R. G., & Reiser, B. J. (2007). Reasoning across ontologically distinct levels: Students' understandings of molecular genetics. *Journal of Research in Science Teaching*, *44*(7), 938-959. doi: 10.1002/tea.20186

Duncan, R. G., Rogat, A. D., & Yarden, A. (2009). A learning progression for deepening students' understandings of modern genetics across the 5th-10th grades. *Journal of Research in Science Teaching*, *46*(6), 655-674. doi:10.1002/tea.20312

Elmesky, R. (2013). Building capacity in understanding foundational biology concepts: A K-12 learning progression in genetics informed by research on children's thinking and learning. *Research in Science Education, 43*(3), 1155-1175. doi:10.1007/s11165-012-9286-1

Felton, M. K. (2004). The development of discourse strategies in adolescent argumentation. *Cognitive Development*, *19*(1), 35–52. http://doi.org/10.1016/j.cogdev.2003.09.001

Ferreira, M. & Trudel, A. R. (2012). The Impact of Problem-Based Learning (PBL) on Student Attitudes Toward Science, Problem-Solving Skills, and Sense of Community in the Classroom. *Journal of Classroom Interaction*, *47*(1), 23–30.

Ford, M. (2008). Disciplinary authority and accountability in scientific practice and learning. *Science Education, 92*(3), 404-423.

Ford, M. (2015). Educational implications of choosing “Practice” to describe science in the Next Generation Science Standards. *Science Education, 99*(6), 1041-1048.

Gallagher, S. A., Stepien, W. J., Sher, B. T., & Workman, D. (1995). Implementing problem-based learning in science classrooms. *School Science and Mathematics, 95*(3), 136-146. doi:10.1111/j.1949-8594.1995.tb15748.x

Georgia Department of Education. (2015). *Georgia performance standards.* Retrieved from https://www.georgiastandards.org.

Grooms, J., Enderle, P., & Sampson, V. (2015). Coordinating Scientific Argumentation and the Next Generation Science Standards through Argument Driven Inquiry. *Science Educator, 24*(1), 45-50.

Henderson, B. J., & Maguire, B. T. (2000). Three lay mental models of disease inheritance. *Social Science & Medicine, 50*(2), 293-301.

Hmelo-Silver, C. E. (2004). Problem-Based Learning: What and How Do Students Learn? *Educational Psychology Review*, *16*(3), 235–266. http://doi.org/10.1023/B:EDPR.0000034022.16470.f3

Householder, D. L., & Hailey, C. E. (2012). Incorporating engineering design challenges into STEM courses.

Kereluik, K., Mishra, P., Fahnoe, C., & Terry, L. (2013). What knowledge is of most worth. *Journal of Digital Learning in Teacher Education*, *29*(4), 127–140. doi:10.1080/21532974.2013.10784716

Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., … Ryan, M. (2003). Problem-Based Learning Meets Case-Based Reasoning in the Middle-School Science Classroom: Putting Learning by Design(tm) Into Practice. *Journal of the Learning Sciences*, *12*(4), 495–547. https://doi.org/10.1207/S15327809JLS1204_2

Kwon, K., Shin, S., Brush, T. A., Glazewski, K. D., Edelberg, T., Park, S. J., … Alangari, H. (2018). Inquiry learning behaviors captured through screencasts in problem-based learning. *Interactive Learning Environments*, *26*(6), 839–855. https://doi.org/10.1080/10494820.2017.1419496

Leonard, M. J. (2005). Examining tensions in a ‘‘design for science’’activity system: science versus engineering goals and knowledge. *Tidskrift for Lararutbildning och Forskning* [Journal of Research in Teacher Education], 3, 132–146.

Lewis, D., & Greene, J. (1983). *Your child’s drawings: Their hidden meaning.* London: Hutchinson & Co. Ltd.

Manz, E. (2015). Representing student argumentation as functionally emergent from scientific activity. *Review of Educational Research, 85*(4), 553-590. doi:10.3102/0034654314558490

McNeill, K. L. (2009). Teachers' use of curriculum to support students in writing scientific arguments to explain phenomena. *Science Education*, *93*(2), 233–268. http://doi.org/10.1002/sce.20294

McNeill, K. L. (2011). Elementary students' views of explanation, argumentation, and evidence, and their abilities to construct arguments over the school year. *Journal of Research in Science Teaching*, *48*(7), 793-823.

McNeill, K. L., & Krajcik, J. (2006, April). Supporting students' construction of scientific explanation through generic versus context-specific written scaffolds. In *Annual Meeting of the American Educational Research Association, San Francisco*.

McNeill, K. L., & Berland, L. (2017). What is (or should be) scientific evidence use in k-12 classrooms?. *Journal of Research in Science Teaching*, *54*(5), 672-689.

National Research Council. (1996). *National Science Education Standards*. Washington, D.C.: The National Academies Press.

National Research Council. (2007). *Taking Science to School: Learning and Teaching Science in Grades K-8.* Washington, D.C.: The National Academies Press. https://doi.org/10.17226/11625.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* Washington D.C.: The National Academies Press.

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States.* Washington, D.C.: The National Academies Press.

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching*, *41*(10), 994–1020. http://doi.org/10.1002/tea.20035

Ryu, S., & Sandoval, W. A. (2012). Improvements to elementary children’s epistemic understanding from sustained argumentation. *Science Education*, *96*(3), 488–526.

Sampson, V., & Clark, D. B. (2008). Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future directions. *Science Education*, *92*(3), 447–472. http://doi.org/10.1002/sce.20276

Sampson, V., Enderle, P., & Grooms, J. (2013a). Argumentation in science education. *The Science Teacher*, *80*(5), 30.

Sampson, V., Enderle, P., Grooms, J., & Witte, S. (2013b). Writing to learn by learning to write during the school science laboratory: Helping middle and high school students develop argumentative writing skills as they learn core ideas. *Science Education*, *97*(5), 643–670. http://doi.org/10.1002/sce.21069

Savery, J. R. (2015). Overview of problem-based learning: Definitions and distinctions. In Walker, A., Leary, H., Hmelo-Silver, C., & Ertmer, P.A. (Eds.), *Essential readings in problem-based learning: Exploring and extending the legacy of Howard S. Barrows* (pp. 5-15). West Layfette, Indiana: Purdue University Press.

Shaw, K. R. M., Van Horne, K., Zhang, H., & Boughman, J. (2008). Essay contest reveals misconceptions of high school students in genetics content. *Genetics*, *178*(3), 1157-1168.

Smith, L. A., & Williams, J. M. (2007). "It's the X and Y thing": Cross-sectional and longitudinal changes in children's understanding of genes. *Research in Science Education, 37*(4), 407-422. doi:10.1007/s11165-006-9033-6

Springer, K., & Keil, F. C. (1989). On the development of biologically specific beliefs: The case of inheritance. *Child Development*, *60*(3), 637-648.

Stake, R. E. (2006). *Multiple case study analysis.* New York: Guilford Press.

Thomas, G. V., & Silk, A. M. (1990). *An introduction to the psychology of children's drawings*. New York University Press.

Venville, G., & Dawson, V. (2012). *The Art of Teaching Science: For middle and secondary school*. Allen & Unwin.

Venville, G., Gribble, S. J., & Donovan, J. (2005). An exploration of young children's understandings of genetics concepts from ontological and epistemological perspectives. *Science Education, 89*(4), 614-633. doi:10.1002/sce.20061

Voogt, J., & Roblin, N. P. (2012). A comparative analysis of international frameworks for 21st century competences: Implications for national curriculum policies. *Journal of Curriculum Studies, 44*(3), 299–321. doi:10.1080/00220272.2012.668938

Walker, A., Leary, H., Hmelo-Silver, C., & Ertmer, P. A. (Eds.). (2015). *Essential Readings in Problem-Based Learning: Exploring and Extending the Legacy of Howard S. Barrows.* West Layfette, Indiana: Purdue University Press.

White, R., & Gunstone, R. (1992). Prediction-observation-explanation. In *Probing understanding* (pp. 44-64). London: Falmer Press.

Wirkala, C., & Kuhn, D. (2011). Problem-based learning in K-12 education: Is it effective and how does it achieve its effects? *American Educational Research Journal*, *48*(5), 1157–1186.

Figure Captions

Fig. 1 Mean scores and standard deviations on genetics understanding Questions 1 – 3 at pretest and at posttest. Scores ranged from 0 = no correct conceptual understanding to 4 = mature conceptual understanding. * indicates differences in pretest and posttest scores are significant at the $p < .05$ level.

Fig. 2 Mean change scores from pretest to posttest and standard errors on genetics understanding Questions 1 – 3 across grades 1 through 5. * indicates differences in pretest and posttest scores are significant at the $p < .05$ level.