Geological models for Explorations Of Dynamic Earth (GEODE)

Supporting middle school students' learning through geodynamic modeling

IMPORTANCE

Understanding geoscience is among the most relevant challenges of our age. On a daily basis, humans are affected by Earth's processes, reliant on Earth's resources, and awed by Earth's beauty. A multitude of our everyday needs—water for drinking, industry, and crop irrigation; energy to power cars and to warm homes; and minerals for the manufacturing of today's pervasive electronics—are centrally dependent upon Earth's dynamic processes and embedded resources. At the same time volcanic eruptions, earthquakes, and landslides disrupt and endanger human lives past, present, and future. Two hundred years ago, the Indonesian volcano Tambora destroyed an entire tribe of people, cooled Earth's atmosphere, and caused famines around the world [1]. Only a few years ago, ash from an Icelandic volcano grounded much of Europe's air travel for nearly a month, crippling the continent's transportation amid a period crucial for financial recovery [2]. The number of earthquakes in the central and Eastern United States have increased dramatically in correlation with wastewater disposal from hydraulic fracturing [3]. And a historically overdue, "when, not if" earthquake of staggering magnitude threatens to destroy a sizable portion of the coastal Northwest [4]. Still, hundreds of millions of people around the world live in areas prone to natural disasters, and humans continue to deplete natural resources at increasing rates despite adverse effects on the economy and the environment's sustainability [5]. Although scientific knowledge of geoscience is more sophisticated and expanding more quickly than ever before, formal education in the geosciences is minimal for the large majority of students [6].

Geoscience demands an understanding of complex, invisible, dynamic processes. Teaching geoscience poses significant challenges. Experiments with the Earth's geology are impossible, and the natural processes that shape the Earth take place out of sight, over unimaginably long times [7–9]. Students have difficulty imagining the magnitude of geologic time and reasoning about the frequency and rate of geologically and environmentally important events [10–11]. And separating geological events from this temporal scale adds an additional layer of cognitive difficulty [12–13]. Although understanding geoscience is quite possibly one of the most complex conceptual arenas, it is undoubtedly crucial to our continued overall livelihood and safety as humans.

Yet geoscience education today is locked into descriptive and non-dynamic approaches. Despite the current calls for engaging students actively and promoting explanation building in science learning [14–16], introductory geoscience is taught largely in a descriptive manner, with a heavy focus on naming and classifying eras and periods, rocks and minerals, the Earth's structure, landforms, and archaic landmasses and oceans. Curricula rely on static illustrations and images, limiting students' understanding of Earth as a dynamic system driven by energy within it [17]. Lesson plans rely on static illustrations or non-manipulative animations to show changes in geologic systems, limiting students in developing understanding of the emergent nature of the inner workings in the Earth's geosphere. Hands-on activities use analogies [18] to demonstrate Earth systems, but generally employ materials such as styrofoam to represent Earth's crust or play dough to explore metamorphism in ways that may gloss over or underemphasize important aspects of these processes. At a time when we understand better than ever the complexity of learning geoscience concepts and the importance of developing explanatory models of phenomena, approaches to teaching geoscience focus are still largely non-dynamic and descriptive.

Technology holds huge potential for transforming geoscience pedagogy and understanding. The possibilities current technology offers for supporting students' understanding of complex, invisible, dynamic systems are unparalleled. Geodynamic models today are transforming geology research by providing ways of understanding the mechanisms and physical processes that shape Earth's surface [19]. Streamlined versions of research-grade geodynamic models also have the potential to revolutionize geology education in grades 6-12. Such models make it possible for students to experiment with otherwise inaccessible concepts, setting starting conditions and controlling parameters during a simulation run. The models can also help students gain insights about the causal mechanisms responsible for changes in Earth's surface because the behavior of these models emerge from properties built into them and students can quickly see results of their choices.

We must develop active, technology-enabled systems modeling approaches to geoscience pedagogy. Current state of the art technology and foundational learning sciences research make this an ideal time for new instructional approaches. The design and development of geodynamic models holds transformational potential for middle school teaching and learning in Earth Science. The *Geological models for Exploration Of Dynamic Earth* (GEODE) project takes full advantage of this opportunity. GEODE will create model-based curriculum designed to explore how Earth's surface features are currently changing. It will explore how plate tectonics helps explain past and current movements of rock layers at Earth's surface, and how plate movements are responsible for most features on Earth's surface. In GEODE, students will experiment with single or combined geological processes that play out over time, setting parameters in simulation-based models to gain understanding of central concepts in geodynamics, such as how forces on Earth's crust caused by plate tectonic movements change Earth's surface features. This approach, vastly different from approaches with non-manipulative geological animations that lock students into a single predetermined scenario, will allow students to simulate a wide variety of possible scenarios across timescales and engage in authentic scientific practices, exploring the limitations and uncertainties of these models as they compare outcomes to real-world complexities.

Goals and Objectives

The Concord Consortium (CC), in partnership with Pennsylvania State University (PSU), proposes an Early Stage Design and Development proposal addressing Strand 2, Learning. The goal of the GEODE project is to design transformational geodynamic modeling software and curriculum modules for middle school Earth science classes. We will conduct design-based research to study the development of these materials for supporting student learning around plate tectonics and associated geologic processes, as well as modeling practices and uncertainty-infused scientific argumentation practices. In addition, we will co-develop teacher practices with participating teachers in order to support student learning with model-based curriculum. In parallel, we will develop an evidence-based understanding of how Next Generation Science Standards (NGSS) science practices and related teaching practices must be adapted to implement complex modeling tools and uncertainty-infused argumentation in the classroom. This project builds on CC's significant experience in developing online curriculum materials with computer-based dynamic modeling environments and PSU's experience in teacher professional development and learning progressions research related to Earth and Space Science in the middle grades.

Objective 1: Build simulations enabling students to explore geodynamic features of the Earth. We will develop an online plate tectonics modeling tool and an online geodynamic modeling tool. For the plate tectonics software we will add features to a model developed at CC (see Technology Development). With the geodynamic modeling software, we will simulate key geologic processes and their interactions over geologic time. The geodynamic modeling software will permit students to "program" a series of geologic events, gather evidence from the emergent phenomena, revise the model, and use their models to explain the causal mechanisms related to a particular arrangement of rock layers. The two combined tools will be referred to as the "GEODE software," which will allow the background recording of every transaction students make for later analysis to investigate the process of student learning with the models.

Objective 2: Develop three learning progression-inspired curriculum modules with the simulation models. The project will create curriculum materials bridging traditional hands-on classroom activities with the GEODE software as a way to explore the dynamic system of the Earth's surface and discover how real-world surface features emerge from underlying geological processes. The modules will be informed by the plate tectonics, modeling, and argumentation learning progressions as described in the research section in this proposal. Using models embedded in the curriculum, students will set up models, run them under a range of input conditions, and compare the model output with real-world examples.

Working closely with our teacher partners—who will field test the GEODE software and curriculum modules—we will develop teacher practices necessary for connecting the classroom materials to the modules.

Objective 3: Research. Research will focus on curriculum modules, assessment materials and professional development materials. The model-based curriculum modules will be developed through three cycles of design-based research. In the first phase, the GEODE models will be tested with a small set of students and teachers in think-aloud settings. In the second phase, three cycles of design-based research will be conducted with the model-based curriculum. In the third phase, the revised curriculum materials with the GEODE software will be pilot tested by a larger number of teachers in diverse classroom settings (teacher N = 30). The assessment and professional development will be researched and continually revised throughout the project. This final pilot test will allow for a comparison between partner teachers who have been involved in face-to-face (F2F) professional development for the use of the modules and tools and "maverick" teachers who are using the curriculum and online professional development materials. Findings of this research will inform learning progressions related to student understanding in plate tectonics and associated Earth processes, modeling, and argumentation.

Objective 4: Disseminate through publications, presentations, and teacher partnerships and networks. We will produce revised and polished materials ready for wide dissemination to Earth science teachers through the Pennsylvania Earth and Science Partnership and the Pennsylvania Earth and Space Science Teachers Association, as well as to a national audience through web resources at CC and through the National Earth Science Teachers Association (NESTA) network. We will publish research results in peer-reviewed journals for teachers as well as for researchers and developers in science education and learning sciences.. We will also present at conferences and disseminate curriculum and assessment materials through teacher networks.

Results from Prior NSF support

The **High-Adventure Science (HAS)** and **High-Adventure Science: Earth Systems and Sustainability** projects *(PI: Pallant; DRL-0929774; $695,075; 9/15/09 – 8/31/12; PI: Pallant with Co-PIs Lee and Larson; DRL-1220756; $2,328,593; 10/1/12 – 12/31/16)* **Summary of Project Results:** The HAS projects have developed modules for Earth and environmental science classes to test the hypothesis that students who use computational models of complex Earth systems, analyze real-world data, and engage in scientific reasoning and argumentation practices will be better able to understand core ideas about Earth systems science and the impact humans can have on these systems. HAS projects also created and validated two assessment frameworks to measure students' formulation of uncertainty-infused scientific arguments and students' explanations of Earth systems in terms of stocks and flows. Analysis based on pre- and post-tests administered before and after each HAS module showed that students improved on argumentation and systems thinking explanation practices across diverse school settings. **Intellectual Merit:** The six modules successfully manifested five design principles to address how to incorporate scientists' current empirical research and modeling practices into short-duration, inquiry-based curriculum modules. This work contributed to the scientific education research field on argumentation by addressing how students incorporate uncertainty in formulating scientific arguments and describing how to assess uncertainty-infused scientific argumentation [20]. In addition, findings from the project are illuminating how to promote understanding of Earth's systems and how human interactions might impact them. **Broader Impacts:** These emanate from High-Adventure Science's six short-duration online modules addressing climate change, fresh water availability, land management, energy resources, air pollution, and the search for life in space, which are freely available to teachers through the National Geographic Society (NGS) and Concord Consortium websites. A total of 49 teachers and roughly 5,100 students have used the online curriculum modules for field testing purposes during the High-Adventure Science projects, and a much larger community of teachers and students have already begun using them independently; these numbers will continue to grow as NGS promotes the materials. **Publications:** Research has led to publications that describe a theoretical framework for assessing students' uncertainty-infused scientific arguments [20], an analysis of student articulation of uncertainty in argumentation tasks [21], and a methodology for promoting scientific argumentation using computational models [22]. In addition, one paper is currently in review [23]; five peer-reviewed papers were published in teacher journals [24–27]; three in-house newsletter articles were published [28–30]; and 11 conference presentations were made [27, 31–40]. Note that HAS:ESS is currently ongoing and is at the start of the fourth project year.

Investigating How to Enhance Scientific Argumentation through Automated Feedback in the Context of Two High School Earth Science Curriculum Units (ESAAF). (*PI: Liu with Co-PIs Pallant and Lee; DRL-1418019; $1,309,991; 9/1/14-3/31/18*) **Summary of Project Results:** This project is responding to the need for technology-enhanced, formative science assessments that promote argumentation practice through inquiry-based science teaching and learning. Guided by comprehensive argumentation theories and drawing on advanced automated scoring technologies with proven validity [41–42], we are applying automated scoring tools to facilitate immediate feedback to formative, constructed-response to uncertainty-infused argumentation assessment items in High-Adventure Science modules. We have completed three iterative design cycles integrating the automated scoring and feedback technology to an online HAS curriculum module. Preliminary findings indicate that automated scoring and feedback can be seamlessly integrated and with it students improve their arguments through revision. We are in the early stages and will continue testing in classrooms with the goal of establishing the impact of this technology on student learning. **Intellectual Merit:** ESAAF has great potential to enhance the use of constructed-response items to elicit deep understanding. ESAAF will study the conditions of feedback and identify the feedback associated with effective learning. ESAAF feedback is anticipated to provide customized opportunities for students to monitor progress and improve learning. **Broader Impacts:** ESAAF's implementation of automated scoring and feedback could produce transformative impacts on the teaching and learning of Earth science content and scientific argumentation. When automated scoring is empowered by quality feedback, it provides unique opportunities for students to develop reflection and metacognitive skills that contribute to lifelong learning. Ultimately, ESAAF can provide a model of formative assessments enhanced by automated scoring and feedback. **Publications:** One paper describing the validation of the automated scoring system has been submitted [43]; another paper highlighting the design cycle has been submitted to a learning sciences conference [44]; and several presentations at conferences have been made [45–48]

Pennsylvania Earth and Space Science Partnership (ESSP). (*PI: Furman, Co-PI: McDonald, Educational Research and Professional Development Workshop Lead; DUE-0962792; $9,181,723; 10/01/10 - 06/30/16)* **Summary of Project Results:** The ESSP is a partnership focused on improving and promoting Earth and Space Science (ESS) in the middle grades (4-9). We have completed five years of summer workshops and academic year professional development focusing on plate tectonics, solar system astronomy, energy, water, and climate. We have also been involved in initiatives to transform key ESS related introductory courses in higher education. **Intellectual Merit:** The primary project research focuses on developing learning progressions in plate tectonics and solar system formation. Both learning progressions focus on students' understanding of the causal mechanism underlying the phenomena. **Broader Impacts:** 115 teachers from our Pennsylvania Research Practice Partnership districts of Philadelphia, Reading, York, State College, Bellefonte, and Bald Eagle Area have participated in our summer workshops. The project is in the process of developing a book outlining the philosophy and structure of the workshops, with details from all our workshops for use by partnerships schools (and beyond) to support the local delivery of professional development based on the project model. We also established the Pennsylvania Earth and Space Science Teachers Association, which has since become one of the most successful state-level NESTA associations, with over 600 members. **Publications:** Resulting research on learning progressions has been published [17], with additional manuscripts in preparation. There have been two papers in practitioner journals [49–50], and a large number of presentations at state, national, and international conferences [51–66].

Research and Development Design

Disciplinary Core Ideas. The three proposed modules address the following two sets of core ideas in the Next Generation Science Standards [67] related to Earth and Space Science (ESS): Core Idea ESS1: Earth's Place in the Universe (ESS1.C: The History of Planet Earth) and Core Idea ESS2: Earth's Systems (ESS2.A: Earth Materials and Systems; ESS2.B: Plate Tectonics and Large-Scale System Interactions).

Science Practices: Developing and Using Models and Formulating Arguments. Because students will learn content by modeling Earth systems, the curriculum will focus on Science Practice 2: Developing and using models as well as Practice 7: Engaging in argument from evidence. The curriculum will "use provided computer simulations … as a tool for understanding and investigating aspects of a system, particularly those not readily visible to the naked eye" (p. 58) [15]. Students will be able to use outcomes from their exploration as evidence in their arguments about how different geological processes and forces will create emergent phenomena under specified circumstances (model-based argumentation).

Crosscutting Concepts Related to Systems and Systems Models. All seven crosscutting concepts identified in the NGSS [67] will be present in the curricula to some extent, but it would be a distraction to emphasize each. Because student learning in GEODE will be based on designing and exploring models to identify causes and effects in Earth systems, we will focus on Systems and system models, paying particular attention to students "developing models, component parts, exploring their interactions, in terms of inputs, outputs, and processes and to generating questions about aspects of their model representation of the system, and, eventually refining their models" (p. 93) [15]..

Proposed Modules

Below we describe concepts covered in the three modules, which are intended to be completed in order. Modules will vary in duration between three and five days and will be designed to be easily interwoven into traditional classroom sequences. Online teacher guides will be developed based on insights from our partner teachers and will include an overview, learning goals, and strategies for integrating and implementing materials in class. In addition, teachers will be supplied with answers to embedded questions and a variety of potential student responses. Each module will be designed around a framing question, a model-based challenge, and uncertainty-infused scientific argumentation to conclude the challenge.

Module 1: Plate tectonics and geologic processes along plate boundaries. In this module, students spend three days using the plate tectonics software to compare the characteristic distribution of earthquakes, their depth, magnitude, frequency, and location along the different plate types of plate boundaries—convergent, divergent, and transform. Students then explore a 3D visualization of plate boundaries to determine how interactions of the plates result in the emergent pattern of earthquakes, volcanoes, and mountains along the plate boundaries. Students begin to connect the subduction, divergent, and transform boundaries to prominent geologic processes and landforms. This module is critical for framing the ways in which Earth's landforms are related to how plates move over geologic time.

Module 2: Earth's dynamic processes creating landforms. In this module students will learn how Earth's surface features can be used to infer the history of a landscape and how specific sequences of Earth's processes create landforms. Students will use the geodynamic modeling software to vary sequences of events to understand how landforms develop, weather, and erode. On the first day students will explore the different depositional environments as represented in the model (e.g., deep ocean, continental shelf, and river deposition), and how time affects the thickness of layers. Next, students will spend a day exploring volcanic processes. They will see how intrusions and volcanic activity cross through and move between pre-existing layers. On the third day, students will explore erosion represented by rivers or wind and learn about the principles of horizontality and laws of superposition. Students will use models to explore how geologic processes are represented in the rock layers. Finally, on the fourth and final day of the module, students will be challenged to "program the model" to match real-world examples (photographs); in so doing students will explore the interactions of the geologic processes over time, generate questions related to the visualization, and refine their models. When the model and real-world examples do not match, there will be scaffolding to help students take this as evidence that some aspects of the model are incorrect or incomplete. Students will be encouraged to modify, rerun, and re-compare until the model matches the real-world example fairy closely.

Module 3: Internal forces and processes changing Earth's features. This five-day module will enable students to add compression, tension, and uplift stresses on the layers in the geodynamic model. Students will spend two days exploring several different processes including folding (deformation of rock), faulting (stresses that result in fracture of rock), uplift (increase in elevation of surface), and subsidence (motion of Earth's surface as it shifts down). On the third day, as students use the geodynamic model they will also focus on how the geologic processes are associated with plate motion and how locations represent very different environments over geologic time. For example, one location may have started in the deep ocean, and as the tectonic plates moved, the location may be part of a collision between plates, which could then uplift the formerly deep-ocean location to the top of a mountain. Finally, on the last two days of the module, students will be challenged to again "program the model" to match more complex real-world examples than in Module 2, and will be asked to explain the events that led to a particular arrangement of layers and then connect them with causal mechanisms (e.g., plate collision causing mountain building, river erosion, etc.).

An example of a day in a class using GEODE curriculum

Figure 1: Mountain showing folded rock layers. ***http://disc.sci.gsfc.nasa.gov/geomorphology/GEO_2/geo_images_T-42/FigT-42.7.jpeg***

Josiah is working with two other students in his 7th grade science class to try to connect plate motion to the location of mountains. He has been fascinated by the recent earthquake in Nepal, which was described as being so violent it moved mountains, so he is excited to use the GEODE plate tectonics software and toggle on plate boundary markings and plate motion arrows. He has zoomed in and out all over the map. The group is attempting to answer questions like "Where do mountains form?" and "Are some boundaries more likely to produce mountains than others?" Josiah and his group explore for a while, then make a claim related to mountain locations, using evidence from the model to support their claim. They explain their uncertainties and share their findings in a class discussion. Josiah brings up the earthquake in Nepal and talks about its location on the map and how the Himalayas are located on a plate boundary.

The group then launches the geodynamic modeling software. They can't wait to "program" a sequence of geological processes and see how the model shows change over time when it is played. They try at least three scenarios where they see mountains form, each time trying to create even bigger mountains. The teacher, Mrs. O'Brien, asks them to create a sequence of events similar to the folded mountains in the photograph in Figure 1.

Using controllers in the software, they take turns placing geologic processes—like deposition and uplift forces—on the timelines, describing what they think will happen, and then playing the model. They set up, run, and refine the model multiple times looking at cause and effect. They create several distinct rock layers by adjusting sedimentation duration and number of sedimentation events and explore compression *and uplift forces on those layers. Through a series of embedded argumentation activities with models and discussions facilitated by the teacher they are asked to consider how these models are related to the hands-on activities they did earlier, the role of plate tectonics in creating the sequence, and the limitations of the model in replicating real-world examples. Just before class ends, Josiah is pleased that their model looks just like the photograph and he takes a snapshot to show they did it!*

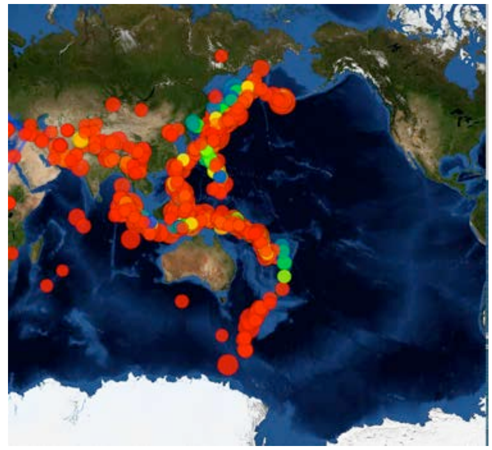

Figure 2: Map showing magnitude, depth, and frequency of earthquakes from 1960 to present.

TECHNOLOGY DEVELOPMENT

The GEODE computer-based software will run in any browser that supports HTML5, including all school computers, including Google Chromebooks, plus many tablets and other portable devices.

The plate tectonics modeling software developed at CC for a prior project currently displays hypocenter data obtained from the U.S. Geological Survey for the world from 1960 to the present of magnitude 4.0 and greater, and location and type of plate boundaries on a satellite image of Earth (see Figure 2). We will add additional data related to rate and direction of plate motion, volcanic eruptions, the ability to update data to include current events, and the ability to isolate topographic features such as mountains, river valleys etc. Students will be able to modify how the data are displayed and to draw specific cross-sections in order to look at the depth, magnitude, and location of earthquakes.

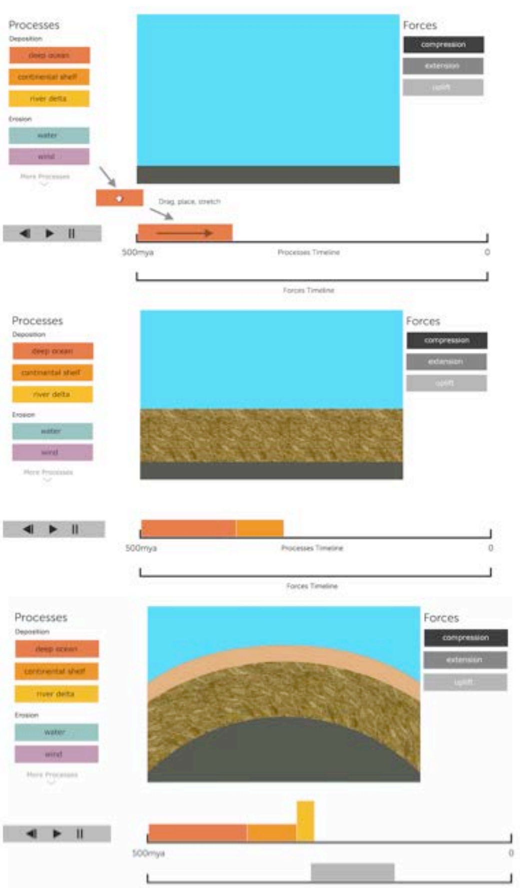

Figure 3: Students drag a geologic process to the Processes timeline (top), run the model, then drag a second process (middle), add an additional layer and drag an uplift force to the Forces timeline (bottom).

The geodynamic modeling software will depict a two-dimensional cross-section (cutaway profile view) of a sample of Earth's crust. The model will be a canvas where students explore scenarios that include sedimentary accumulation, uplift, subsidence, folding, faulting, eruption, and erosion. Students will see young rock layers overlie old rock layers. They will see how gaps in layers occur. The model will always begin with slowly subsiding bedrock. Three levels of interactivity will accommodate more challenging modeling as the curriculum modules progress.

Level 1: Basic principles. The model will show a small region approximately 4 km deep to help students learn about near-surface interactions. The timescale and processes will be exaggerated to help students visualize the events. Learners will "program" the model by selecting geologic process blocks from a process "palette" (see Figure 3, top). Student interaction with the model happens via the timeline. Students will drag rectangles from the geologic process palette to the timeline. Each geologic process rectangle is resizable on the timeline. Stretching the rectangle horizontally changes the length of time over which that process acts. Stretching the rectangle vertically changes the rate at which the process acts. The combination of rate and time will determine the effect of the process on the area (e.g., the thickness of a sediment deposit). Students can explore changing the process rectangle parameters and determining how their changes affect the outcomes in the model.

Level 2: Surface layers' response to forces. Learning how layers respond to regional forces is an important part of interpreting geologic history. In Level 2, a second timeline will be added to the model. After students have created a few rock layers they will then be offered the option of introducing compression, extension, and uplift forces (see Figure 3, bottom) by placing force blocks on the second timeline and experimenting with strength of force (height) and duration (length).

Level 3: Cross-section of tectonic plates and landforms. Because the cross-section represents one location that is moving over time (due to plate motion), the model will introduce a new toggle that will locate an event programmed on the timeline onto a schematic cross-section of tectonic plates and the geologic processes associated with specific locations. For example, if a student's cross-section begins with volcanic eruptions, the toggle will zoom into the divergent plate boundary. When the next layer is added via sedimentation, the toggle will zoom into the ocean basin caused by the diverging plates.

PROFESSIONAL DEVELOPMENT

Our teacher partners will help with the development and field testing of the GEODE software and curriculum, giving feedback on how students engage with the materials and software during their learning. Teachers will be drawn from an ongoing Research and Practice Partnership (RPP) established as part the ESSP (see Results from Prior NSF Support). Since 2010 the ESSP has established professional development experiences for teachers in partner districts that are co-developed and co-delivered by science education and disciplinary faculty at PSU. These professional development experiences are designed to seamlessly blend disciplinary content and reform-based pedagogical practices in science. These teachers will lend their experiences to provide a context for online materials that will be utilized by teachers who will use the modules without the benefit of face-to-face professional development.

To support the initial development of GEODE modeling software, we will recruit five teachers from two middle schools in the State College Area School District (SCASD) as participants in Years 1 and 2. These teachers have Earth Science as a required part of their curriculum and constitute the complete set of seventh grade science teachers in the district. SCASD also is a one-to-one laptop district, making it ideal for pilot testing the GEODE project's model-based online curriculum. In Year 1, we will hold five targeted meetings with the teachers—three in the fall, one in February, and one in April—to get feedback and design suggestions on the software and curriculum.

During Year 2, we will organize up to four Studio Days focused on implementing the second iteration of project materials in teachers' classrooms. Studio Days are based on a Japanese Lesson Study Model (e.g., [68]) and involve all teachers co-planning a lesson, and then all attending one teacher's classes as they teach the lesson. The team meets as soon as the lesson is over and re-plans the lesson, which is then retaught to another group of students. The lesson is iterated at least three times with planning in between and then there is a final debrief at the end of the day. The iterative nature of the process lends itself to the development of new curricular materials.

In the summer at the beginning of Year 3 we will host a professional development workshop at PSU for approximately 20 teachers drawn from the RPP districts. Teachers will spend one week engaged with the GEODE materials in preparation for the teachers' larger scale implementation during Year 3.

The professional development workshop will focus on the following: understanding geoscience content, using geodynamic models with students, developing student argumentation skills, and implementing the curriculum in the upcoming year. This workshop will be organized on a pattern used for the last five years as part of the ESSP project. Teachers will engage with GEODE modules as learners, followed by an in-depth debrief to help teachers understand key aspects of the lessons. Then teachers will work in small groups to consider how to integrate the modules into curricular context in their schools. ESSP teachers have been working with a Claims, Evidence, and Reasoning framework [69], which means they will be familiar with supporting students' scientific argumentation; however, introducing uncertainty-infused argumentation tasks based on modeling is new and requires new teacher learning. Also, teachers will need support in developing practices to help students with their new modeling practices. During the school Years 2 and 3 teachers will participate in one Studio Day per district during enactment of the GEODE materials. This will provide design and development feedback for the GEODE team. Thus, while the ESSP teachers are an excellent pool to draw from in terms of their teaching expertise, they will also be new to the foci of the GEODE project and the professional development will be geared to preparing them for a full pilot enactment in Year 4.

Project Research

Theoretical Framework

Geoscience focused learning progressions. Prior work in students' understanding of geoscience concepts has focused primarily on geocognition, that is, how students reason about large scale (both temporal and spatial), and how they reason within and across 2D and 3D representations of geodynamic phenomena. In contrast, McDonald, et al. [51] have been investigating a learning progression (LP) focused on students' ability to develop understandings of plate tectonics in terms of causal mechanisms and systems-level understanding of plate interactions. The results characterize three main progress variables: plate movement, plate dynamics, and intra-plate interactions [54]. In addition to developing the first LP focused on the big idea of plate tectonics, the research has also yielded findings that can inform the development of the GEODE modules specifically.

To be able to understand plate movements as a way to explain the distribution and patterns of landforms on Earth's surface, students must link the phenomenon on the surface of the Earth together to create a pattern (plates) and then connect them with one overall causal mechanism (convection based on heat transfer). Fundamentally, this also means they must transition from viewing the Earth as a static body with landforms to a dynamic system driven by Earth's internal energy. One of the most significant findings from the LP work is that students in middle grades are capable of understanding the Earth as a dynamic system, but this requires moving them past a boundary/events focus. The initial findings also indicate that some of the traditional historical data associated with teaching plate tectonics (Pangea, magnetic sea-floor striping, etc.) can cause problematic misunderstandings and unproductive models of Earth systems with students who do not have a mechanistic understanding of Earth systems [51]. GEODE will be explicitly designed to take into consideration our current understandings from the LP work. In addition, the project will focus on expanding the LP from a relatively narrow focus on plate tectonics and its underlying mechanism to a broader set of Earth system phenomena linked to plate tectonics, such as sedimentation, volcanic eruptions, earthquakes, deformation of strata, etc. This will expand the scope of the LP and connect it to concepts in the NGSS middle grades set of disciplinary core ideas (e.g. ESS2.A, ESS2.B, ESS1.C).

Modeling learning progression. The Earth is a system of many interacting parts; matter and energy flow in and out of the system; the system is maintained and controlled by feedback loop mechanisms [70–72]. It is difficult to isolate one variable in Earth's system [73], which makes experimentation with the controlled variable strategy less feasible in teaching Earth systems concepts. A substantial body of research shows that computational models and simulations allow students to understand through exploration the behavior of systems that are difficult to understand by other means [74–76]. For this, agent-based computational models [77] are ideal for exploring geologic systems at large space and temporal scales [15]. Much as geoscientists do, students in the GEODE project will be able to experiment with models by controlling the parameters, the starting conditions, and conditions during a run. The models will have vivid graphics and run quickly, so that students can experiment, observe cause and effect, and gain insights about the system.

Students' difficulty with imagining the magnitude of geologic time and reasoning about the frequency and rate of geologically and environmentally important events hinders student learning [10]. This difficulty stems from the fact that people think in terms of events, not the interval scale of linear time. Most often in education, developing an understanding of geologic time and its unfamiliar timescale is done through analogies (e.g., imagining the time since Earth was formed represented by a football field) [78]. The partitioning of the spatial field with the temporal scale can add another layer of cognitive difficulty. Dodick and Orion's [12, 13, 79] research suggests that temporal organization (the sequence of events) and the principles of change (the present is the key to the past) help people to understand that the static sequences of rock strata were created dynamically over time.

In accordance with the scientific modeling literature, GEODE models will be designed to create representations that abstract and simplify Earth systems by focusing on key geologic forces and processes to explain and predict geodynamic phenomena. Schwarz, et al. [80] recognized students' learning progression associated with modeling in two directions: (1) getting more and more sophisticated in each of four activities essential to the modeling practice such as construct, use, evaluate, and revise models, and (2) improving metamodeling knowledge such as considering models as tools to predict and explain scientific phenomena and models as capturing and reflecting latest developments in understanding. Considering the fact that "all models are wrong" [81] (p. 501), another metamodeling knowledge to include would be analyzing sources of uncertainty stemming from conceptual, technological, and methodological limitations embodied in what models at hand can depict [82].

Argumentation learning and progression. The added benefits of engaging students in scientific argumentation from modeling is that scientific argumentation deepens science concept learning [83–84], altering student views of science [85–86], and supporting student investigation [87–88]. Research on scientific argumentation has grown substantially in the last 10 years [89]. One aspect that has been overlooked, however, is how students treat uncertainty in formulating their arguments [90]. Uncertainty can play two roles when students construct an argument. One type of uncertainty represents students' confidence in their own knowledge and ability [91]. The other type is inherent in scientific inquiry due to measurement errors, lack of conclusive theories or models, and limitations associated with current equipment and technologies [82]. Because all the GEODE topics involve using models to interpret Earth's history, issues about the reliability of models and the kinds of conclusions that can be justified from the models make student understanding of uncertainty a central focus of the project [21–22]. As argumentation is a central scientific practice in the discourse of science, it will give students insight into how scientists construct knowledge [15], and the role of modeling in knowledge construction [22].

From empirical and theoretical work available in the literature, we can infer several ways to describe students' learning progress. For example, research on assessment of scientific argumentation [89, 92] indicates that students' understanding of scientific argumentation generally moves towards (1) including necessary argumentation elements such as claim, evidence, reasoning, qualifiers, and conditions of rebuttal [85, 87, 93], (2) including a larger number of warrants [84, 92], (3) tightly coordinating between theory and evidence [94–96], (4) better grasping epistemological and epistemic aspects of argumentation [97–99], and (5) including both scientific reasoning and critical thinking [20]. GEODE curriculum will be designed to engage students in uncertainty-infused scientific argumentation based on evidence construed from models and from classroom experiences.

RESEARCH QUESTIONS

RQ1: **(students' development of understanding and learning progression)** How and to what extent does students' understanding develop through three modules: from applying the basics of plate tectonics knowledge to discovering how diverse Earth's processes and internal forces create Earth's surface features? How does this development compare with existing plate tectonics learning progression models?

RQ2: **(students' development of modeling practice and learning progression)** What patterns emerge from students' modeling practices with geodynamic models? How do the modeling patterns relate to ways in which students describe and explain plate tectonics and associated geological processes that shape Earth's landforms? How and to what extent do students' modeling practices develop over three modules? How does the development of students' modeling practices compare with existing learning progression models?

RQ3: **(students' development of uncertainty-infused argumentation with evidence from modeling and learning progression)** How do students use evidence generated from models in formulating arguments about geological phenomena? How do students frame uncertainty in accepting or refuting evidence from the models? How does the development of students' uncertainty-infused argumentation practice compare with existing learning progression models?

RQ4: **(teacher practice)** What practices are productive in supporting students' use of dynamic computer models of complex phenomena? What is the repertoire of practices developed by teachers around the use of modeling that include an explicit focus on uncertainty-infused argumentation?

RESEARCH PLAN

Research will focus on curriculum modules, assessment materials, and professional development materials. Curriculum modules will undergo three design cycles from Year 1 to Year 3; assessment materials will be developed and continuously revised from Year 1 to 3; professional development materials will be outlined in Year 1, developed in Year 2, and used and revised in Years 3 and 4. Year 4 will be a pilot of GEODE software and curriculum modules, as well as supporting professional development materials.

Design-Based Research on Model-Based Curriculum Modules (Year 1 to Year 3)

During development of curriculum materials, we will follow the design-based research paradigm where established learning theories, available research results, and prior designers' experiences inform initial design choices and iterative redesign provides data to modify curriculum materials and theories [100].

Model-in-Activity Testing in Small Student Groups (Design Cycle 1). Prototype geodynamic models will be developed and tested at CC through several "clinical" trials and think-alouds with 20 students (a few at each trial) from local middle schools. The goal of the trials and think-alouds is to focus on the user experience and interaction with the modeling environment in order to improve the user interface and examine whether the learning goals are attained by students.

Curriculum Prototyping with Lead Teachers (Year 1). Informed by research literature, NGSS, and other documents, research staff at CC and PSU, scientists on the advisory board, and lead teachers will work together to develop performance expectations for the GEODE modules for content understanding, modeling practices, and uncertainty-infused scientific argumentation for each module. Based on the performance expectations, we will prototype the first GEODE module and implement it in one teacher's classrooms local to CC in order to see how the module functions in a real classroom setting.

Curriculum Testing by 5 Lead Teachers in Year 2 (Design Cycle 2) and by 20 Teachers in Year 3 (Design Cycle 3). We will work with five lead teachers and their students (approximately 400 students) in Year 2 and 20 teachers drawn from the ESSP RPP districts and their students (approximately 1,600 students) in Year 3, to implement the three online GEODE modules along with pre-tests and post-tests.

Pre/Post-Test Data Collection and Analysis. Every time GEODE modules are implemented, the teacher will administer two instruments related to (1) core content understandings and (2) uncertainty-infused scientific argumentation along with students' demographic information such as gender, ethnicity, language, and computer familiarity, grade, prior knowledge, and ELL status. Student learning outcome variables will be created for content understanding and for scientific argumentation. On each learning outcome variable, we will apply repeated measures ANCOVA with the teacher as a fixed effect and other student demographic variables as covariates. This will allow us to examine how variations in implementation impact student learning gains as well as for whom GEODE modules are most beneficial.

Module Data Collection and Analysis. Each GEODE module will have at least five modeling tasks combined with uncertainty-infused scientific argumentation prompts. Data on modeling practices will be collected through students' responses to modeling prompts and their actions recorded automatically by the server. Since log data track all student activities in the curriculum authoring system, log data provides valuable information on students' model uses and students' navigation in the module. Log data will provide information on how much time students spend on the construction, running, and revision processes with models, and the number of iterations students go through when creating each model. Students' uncertainty-infused scientific argumentation tasks will be analyzed by the established protocol developed by Lee et al. [20]. The connection between students' modeling practices and their argumentation performances will be investigated.

Logging and Screencast Data Collection and Analysis. The curriculum server will automatically collect students written responses to prompts in the curriculum modules and log all students' interactions with the module and the models. Using screencast technology, we will also collect video data for a sample of two student groups each from five lead teachers (10 student groups) in Year 2 and five teachers (10 student groups) in Year 3. Screencast data include videos of computer screens showing user interactions overlaid with students' voices. In selecting student groups for the screencast data collection, we will consider grouping based on science achievement determined by pre-test scores as well as in consultation with the teachers. Three groups will be represented: one high performing, two medium performing, and one low performing. This data will create a comprehensive picture of student activity, including dialogue between partners and modeling decisions made in real time. We will use screencast data to validate log data analyses methods and results.

Assessment Research and Development (Year 1 to Year 3)

We will examine students' performances related to (1) understanding plate tectonics and related Earth processes, (2) modeling practices, and (3) uncertainty-infused argumentation practices.

Students' understanding of core ideas related to plate tectonics and related Earth processes will be measured by a multiple-choice plate tectonics instrument being developed as part of learning progression work by McDonald et al. [51]. The current version (being revised) has two reliable and valid subscales (Cronbach Alpha values of 0.72 and 0.69). A new round of data will complete reliability and validity for the assessment in spring 2016. Since this instrument focuses exclusively on plate tectonics, it will need to be broadened for use with all three modules. First, we will add items related to the application of mechanisms involved in plate tectonics to the creation of a diverse set of geological landforms through geological processes and forces. Second, in order to access students' reasoning behind students' choices, we will add an open-ended explanation prompt (i.e., "Explain your choice") at the end of some multiple-choice items. Students' open-ended responses will be scored using knowledge integration scoring (KI) rubrics [101–102], which can measure the depth of student understanding related to how and why scientific phenomena occur. The KI scoring rubrics were designed to measure students' abilities to elicit and connect relevant ideas in a scientific context [103]. Items in the knowledge integration framework have been validated in classroom-based trials for psychometric rigor [104], instructional sensitivity to instruction that fosters integrated understanding of science across physical, biological, and Earth sciences [105], and learning progression in energy [106].

Students' modeling practice will be examined as part of students' responses to open-ended prompts embedded throughout the GEODE modules. Their actions while interacting with the models will be automatically collected as log data in the background. For each GEODE module, we will develop modeling tasks that consist of setting up challenges, asking questions, planning, testing, and revising models. In defining performance expectations within the context of GEODE modules, we will use the literature on general modeling practice such as Schwartz et al. [80] template of key modeling aspects: construct, use, evaluate, and revise models as well as systems modeling in geology [72–73]. Students will be asked to explain causal relationships and the mechanisms by which they are related.

Uncertainty-infused scientific argumentation items will measure the extent to which students make reliable claims based on available evidence from their models backed by accepted scientific knowledge or theory, and particularly the extent to which students recognize limitations associated with their claim and evidence-based justification (uncertainty). The item format consists of four parts: multiple-choice structured claim, open-ended explanation, five-point Likert scale uncertainty rating, and uncertainty rationale [20, 107–108]. The uncertainty-infused scientific argumentation construct was developed for the HAS project and validated using Rasch Analysis based on Partial Credit Model [20]. Rasch PCM results indicate that students' claims, explanations, and uncertainty rationales formulate a unidimensional construct and had an acceptable model fit. This unidimensional scientific argumentation construct had a reliability of 0.91 Cronbach Alpha. The item format was used as embedded assessments in the HAS online curriculum modules and was shown to be effective in eliciting students' use of evidence from models in their argumentation [20, 107]. It will be used after students manipulate GEODE models. Finally, we will develop three uncertainty-infused scientific argumentation tasks for the pre- and post-tests.

Assessment Data Collection and Analysis. For each implementation, we will use pre- and post-tests to measure students' gains in understanding plate tectonics and related Earth processes and uncertainty-infused scientific argumentation. Since the plate tectonics instrument and the uncertainty-infused scientific argumentation instrument will be created or modified for the GEODE project, we will administer them to students in the five lead teachers' classrooms (approximate students N = 400) in Year 1. In order to establish construct validity of pre/post-tests we will be developing, we will analyze students' responses to the tests using the construct modeling approach suggested by Wilson (2005) where (1) a construct map is theorized from the relevant literature on the target construct; (2) items that elicit various levels on the construct map are selected for an instrument; (3) student responses are collected on the instrument; and (4) appropriate item response models are applied to student response data. We will use Rasch Modeling of student responses [109–110] to establish the validity of each instrument [111]. To ensure Rasch Modeling is appropriate, we will test for multi-dimensionality with exploratory factor analysis (EFA) using principal axis factoring with a promax rotation. EFA is an exploratory statistical method used to determine the underlying structure of an assessment [112]. We will also test for local dependence [113] as the items can be clustered. We will use psychometric properties to validate the construct underlying the instrument and the items to establish a measurement scale for each construct. In Years 2 and 3, the instruments with revised items will be used before and after modules to examine the items' sensitivity to GEODE modules. The instruments will be continuously improved based on psychometric and instructional sensitivity results throughout the design cycles so they will be ready for the pilot in Year 4.

Year 4: Pilot Testing

By the end of Year 3, we will have refined versions of curriculum modules, assessment materials, and professional development materials. In Year 4, these materials will be used by 20 ESSP RPP teachers. They will also be made available to the community of teachers who have been field test teachers in HAS, HAS:ESS, and ESAAF projects (a pool of 100-plus teachers spread across 20 states in the U.S.). We will solicit participation and select 10 HAS teachers that represent diverse school settings in terms of students' language and ethnic minority status, school locale (suburban, urban, rural), and student population socioeconomic status to implement the materials. As a result, 30 teachers will participate with an approximate student sample size of 3,000. The HAS teachers will receive the GEODE online professional development materials and minimal online professional development from the GEODE staff (2-3 hours). The ESSP teachers will have had professional development in the summer before Year 3.

Participating teachers will implement all three GEODE modules in their classrooms. Before and after the modules, they will administer identical pre-tests and post-tests content understanding and uncertainty-infused scientific argumentation instruments. We will create two learning outcome variables from these instruments. On the curriculum server, all students' actions with the modules will be automatically recorded. Using the log data, we will characterize teacher enactment in terms of the extent to which the modules are completed and the extent to which models are used. We will pool students from both ESSP and HAS teachers, and conduct analyses on the two student learning outcome variables such as repeated measures ANCOVAs to investigate how student demographics, teacher training (F2F vs. online professional development), and school settings impact student learning outcomes.

Repertoire of Teacher Practices

We need to understand the kinds of practices teachers develop around the use of complex geodynamic models, as well as how they support students' uncertainty-infused argumentation practices. To do this, we will focus data collection on the professional development contexts and also on their classroom enactment with students. All of the professional development contexts will be video and audio recorded, and artifacts produced during professional development will be collected. The corpus of data will include: video recordings of all Studio Days in Years 2 and 3 (up to 10 days and approximately 50 hours total) and all activities during the weeklong professional development at the beginning of Year 3 (approximately 40 hours); in Year 4 we will collect case study data for two teachers' enactment of all three modules with video recordings of all instruction, as well as a collection of all student classroom artifacts, such as in-class assignments and tests. The case study data will be the primary focus of the analysis for teacher practices, as Year 4 will be the most mature version of the GEODE modules and the teachers' practices. We will focus initially on the video recordings, using a video analysis tool, Studiocode [114]. Analysis will begin using the NGSS student practices as a framework, with particular attention to identifying classroom segments focused on *developing and using models* (Science Practice 2) and *formulating arguments* (Science Practice 7). Once these segments have been identified the analysis will focus on how these practices play out in the context of complex geodynamic models and with uncertainty-infused argumentation. Analysis of these segments will be used to create a framework for characterizing teacher and student practices to serve as a focus for analysis of the remaining data. The results will be a clearer characterization of both NGSS practices and newly developed teaching practices that can inform the creation and study of curricula, especially those involving complex modeling tools.

Mechanisms to Assess Success of the Project

Dissemination

The project will create online curriculum modules, teacher guides, and research. All curriculum materials will be available and ready for dissemination on CC's website, as well as being linked to the Earth and Space Science Partnership and Pennsylvania Earth and Space Science Teachers Association website. In addition, we will promote the material through the NESTA network. The Concord Consortium will produce a GEODE project website, which will distribute the curriculum and will include results from our research and evaluation. CC will also promote the materials through various social media channels. Additionally, the partners will promote the project through presentations and workshops at conferences (NSTA, NECC, and NARST) and through articles published in peer-reviewed journals in science education (*Journal of Research in Science Teaching, International Journal of Science Education,* and *Science Education*), learning sciences (*Journal of the Learning Sciences*), and for teachers (*The Science Teacher*, etc.). In addition, the *@Concord* biannual newsletter, distributed for free to over 29,000 digital and print subscribers, will be another communication venue.

Expertise

Amy Pallant will serve as Principal Investigator. She will direct the development of the models and curriculum materials and be responsible for the overall coordination and budgeting of the project.
Scott McDonald, Ph.D., (Co-PI) will oversee the PSU portion of the budget, lead professional development activities, data collection, and analysis in the PA schools.

Hee-Sun Lee, Ph.D., (Co-PI), will lead research on assessment development and validation as well as student learning of modeling practice and uncertainty-infused scientific argumentation captured before, during, and after the modules.

Advisory Board

We intend to have one advisory panel meeting each year.

Bryan Brightbill is a middle school teacher in Park Forest Middle School in State College, PA. He has been part of the ESSP RPP for more than five years and has participated in the learning progressions research around Plate Tectonics. He holds B.S. and M.S. degrees in geosciences.

Rick Duschl is a professor of science education and the Waterbury Chair for Secondary STEM Education at Penn State University. He is a former editor of *Science Education*. His research explores issues of the epistemology and practice of science.

Tanya Furman is a professor of geosciences and an associate vice president and associate dean for undergraduate Education at Penn State University. She is the lead investigator on the NSF-funded Earth and Space Science Partnership.

Greg Kelly is a professor of science education and the Associate Dean for Research, Outreach, and Technology in the College of Education at Penn State University. He is a former editor of *Science Education*. His research explores issues of knowledge and discourse in science education settings.

Ann Rivet is an Associate Professor of Science Education at Teachers College, Columbia. She is interested in researching student learning in inquiry and project-based learning environments; urban middle school science reform; content literacy in secondary science; Earth Science teaching and learning.

External Evaluator

Xiufeng Liu will assume the role of external evaluator. Dr. Liu has extensive experience in science assessment, science instruction, and measurement, and is currently co-editor of the *Journal of Research in Science Teaching*. Dr. Liu will independently monitor project progress and evaluate the quality of curriculum, models, and assessment materials from both psychometric and content perspectives. He will evaluate the reports, presentations, and materials produced from this project. Dr. Liu will meet with the Co-PIs in teleconference meetings to learn about project updates and provide feedback. The evaluator will also participate in the annual face-to-face advisory board meetings and work with the advisory board members in providing recommendations to the team regarding project progress.

Broader Impacts

The models, curriculum, and online teacher professional development materials developed in GEODE have great potential for improving Earth science education. The materials will be made available for free to all future learners, teachers, and researchers beyond the participants outlined in the proposal. The GEODE curriculum implementation itself will involve over 5,000 students and approximately 30 teachers from diverse school systems serving minorities and students from families with diverse socioeconomic backgrounds. Changes made in teachers' practices as a result of the professional development will also impact new students for years to come. The online materials will be promoted and distributed to a national audience primarily through CC's website that attracted over 38,000 Earth science visitors between 2013 and 2015, through the Pennsylvania partner districts, the ESSP and Pennsylvania Earth and Space Science Teachers Association websites, and at national teacher conferences. The project will provide new knowledge on how to teach Earth science phenomena, and develop modeling and scientific argumentation practices. The GEODE project materials and results can provide a model for how geodynamic models can be integrated into learning opportunities and could serve as a template for other curriculum and technology developers. Finally, students will be exposed to geoscience concepts that underlie important challenges facing humanity at this time.

References

[1] J. Lee, "Volcanic Eruption That Changed World Marks 200th Anniversary," *National Geographic*, 2015. [Online]. Available: http://news.nationalgeographic.com/2015/04/150410-tambora-volcano-eruption-climate-change-famine-earth-science/.

[2] S. Erlanger and J. Ewing, "Air Travel Crisis Deepens as Europe Fears Wider Impact," *New York Times*, New York, NY, 2010.

[3] J. L. Rubinstein and A. B. Mahani, "Myths and Facts on Wastewater Injection, Hydraulic Fracturing, Enhanced Oil Recovery, and Induced Seismicity," *Seismol. Res. Lett.*, vol. 86, no. 4, pp. 1060–1067, 2015.

[4] K. Schulz, "The Really Big One," *New Yorker*, 2015.

[5] "Visitor Use Statistics," *National Park Service*, 2015. [Online]. Available: http://www.nature.nps.gov/socialscience/stats.cfm. [Accessed: 01-Jan-2015].

[6] "Promoting Earth Science Literacy for Public Decision Making," *The Geological Society of America*, 2013.

[7] B. E. Herbert, "Student Understanding of Complex Earth Systems," in *Special Paper 395: Geologist Think and Learn about the Earth*, C. Manduca and D. W. Mogk, Eds. Boulder, CO: Geological Society of America, 2005.

[8] K. M. Kortz and D. P. Murray, "Barriers to College Students Learning How Rocks Form," *Journal of Geoscience. Educ.*, vol. 57, no. 4, p. 300, 2009.

[9] C. King, "Geoscience education: an overview," *Studies Science. Education*, vol. 44, no. 2, pp. 187–222, 2008.

[10] C. Cervato, "The significance of geologic time: Cultural Education and economic frameworks," in *Earth and Mind II: A synthesis of Research on thinking and Learning in the Geosciences*, K. Kastens and C. A. Manduca, Eds. Geological Society of America, 2012, pp. 19–28.

[11] I. Resnick, K. Atit, and T. Shipley, "Teaching geologic events to understand geologic time," in *Earth and Mind II: A synthesis of Research on thinking and Learning in the Geosciences*, K. A. Kastens and C. A. Manduca, Eds. Colorado: The Geological Society of America, 2012, pp. 41–44.

[12] J. Dodick and N. Orion, "Geology as an historical science: Its perception within science and the education system," *Science Education*, vol. 12, pp. 197–211, 2003.

[13] J. Dodick and N. Orion, "Measuring student understanding of geological time," *Science Education*, vol. 87, pp. 708–731, 2003.

[14] National Research Council, *Inquiry and the National Science Education Standards: A Guide for Teaching and Learning*. Washington, DC: The National Academies Press, 2000.

[15] National Research Council, *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press, 2012.

[16] A. W. Duschl, Richard A.; Schweingruber, Heidi A.; & Shouse, Ed., *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies Press, 2007.

[17] J. D. Plummer, C. Palma, A. Flarend, K. Rubin, Y. S. Ong, B. Botzer, S. McDonald, and T. Furman, "Development of a Learning Progression for the Formation of the Solar System," *International Journal of Science Education*, vol. 37, no. 9, pp. 1381–1401, 2015.

[18] E. J. Pyle, “A model of inquiry for teaching earth science,” *Journal of Science Education.*, vol. 12, 2008.

[19] M. Gurnis, M. Aivazis, J. Tromp, E. Tan, P. Thoutireddy, Q. Liu, E. Choi, M. Dicaprio, M. Chen, S. Simons, B. Quenette, B. Appelbe, C. Aagaard, L. Williams, L. Lavier, Moresi, and H. Law, “GeoFramework: A Modeling Framework for Solid Earth Geophysics,” in *American Geophysical Union*, 2003.

[20] H. S. Lee, O. L. Liu, A. Pallant, K. Crotts, S. Pryputniewicz, and Z. Buck, “Assessment of uncertainty-infused scientific argumentation,” *Journal of Research in Science Teaching.*, vol. 51, no. 5, pp. 581–605, 2014.

[21] Z. Buck, H.-S. Lee, and J. Flores, “I’m sure there may be a planet: Student articulation of uncertainty in argumentation tasks,” *International Journal of Science Education*, vol. 36, no. 4, pp. 2391–2420, 2014.

[22] A. Pallant and H.-S. Lee, “Constructing scientific arguments using evidence from computational models,” *Journal of Science Education and Technology*, vol. 24, no. 2/3, pp. 378–395, 2015.

[23] W. Xing, H.-S. Lee, and A. Shibani, “Identifying patterns in students’ use of data and reasoning in scientific argumentation: An automated text mining approach. .,” *Science Education. submitted.*

[24] A. Pallant, D. Damelin, and S. Pryputniewicz, “Deep Space Detectives,” *The Science Teacher*, vol. 80, no. 2, pp. 45–50, 2013.

[25] A. Pallant, H. Lee, and S. Pryputniewicz, “Modeling Earth’s Climate,” *The Science Teacher.*, vol. 79, no. 7, pp. 32–36, 2012.

[26] A. Pallant, “Encouraging Students to Think Critically About Earth’s Systems and Sustainability,” *The Earth Scientist.*, vol. 29, no. 4, pp. 13–17, 2013.

[27] A. Pallant, “No simple answers-How models and data reveal the science behind environmental topics,” *MEES Obs.*, 2013.

[28] A. Pallant, “The future of fracking: Exploring human energy use,” *@Concord*, vol. 17, no. 1, pp. 10–11, 2013.

[29] A. Pallant, “Modeling the unknown is high adventure,” *@Concord*, vol. 14, no. 1, pp. 6–7, 2010.

[30] A. Pallant, “Looking at the evidence-What we know. How certain are we?,” *@Concord*, vol. 15, no. 1. pp. 4–6, 2011.

[31] Z. Buck, H.-S. Lee, and J. Flores, “Students’ uncertainty during exoplanet detection tasks.,” in *National Association for Research in Science Teaching*, 2014.

[32] H.-S. Guo, Y., Xing, W., & Lee, “Identifying students’ mechanistic explanations in textual responses to science questions with association rule mining,” in *Data Mining for Educational Assessment and Feedback I during nternational Conference on Data Mining 2015.*, 2015.

[33] H. Lee, A. Pallant, S. Pryputniewicz, and O. L. Liu, “Measuring Students’ Scientific Argumentation Associated with Uncertain Current Science,” in *National Association for Research in Science Teaching*, 2013.

[34] H.-S. Lee, A. Pallant, and S. Pryputniewicz, “Articulating uncertainty as part of scientific argumentation during model-based exoplanet detection tasks.” in *The 29th International Astronomical Union General Assembly*, 2015.

[35] H.-S. Lee and A. Pallant, “Building an intelligent online curriculum system: High Adventure Science example,” in *Learning Analytics Summer Institute*, 2015.

[36] A. Pallant and H.-S. Lee, “Characterizing uncertainty associated with middle school students’ scientific argumentation,” *National Association for Research in Science Teaching*. Orland, FL, 2011.

[37] A. Pallant and S. Pryputniewicz, “High-Adventure Science: Making and Defending Scientific Claims in the Face of Uncertainty,” *Massachusetts Environmental Education Society.*, 2013.

[38] A. Pallant, “Scientific seminar-looking at the evidence: How certain are we?,” in *Cary Institute for Ecosystem Studies*, 2012.

[39] A. Pallant, “Developing integrated understanding of global climate using dynamic visualizations.,” in *American Association for the Advancement of Science Conference,* 2012.

[40] A. Pallant, “High Adventure Science: Free simulations exploring Earth’s systems and sustainability.” in *National Science Teacher Association*, 2014.

[41] J. Z. Sukkarieh and J. Blackmore, “c-rater: Automatic content scoring for short constructed responses.,” in *22nd International FLAIRS Conference*, 2009, pp. 290–295.

[42] J. Z. Sukkarieh and S. G. Pulman, “Automatic short answer marking,” in *Second Workshop on Building Educational Applications Using NLP*, 2005, pp. 9–16.

[43] L. Mao, O. L. Liu, K. Roohr, V. Belur, M. Mulholland, H. –S. Lee, and A. Pallant, “Validation of automated scoring for formative assessment of students’ scientific argumentation in climate change,” *Educational Assessment.*

[44] H.-S. Lee, A. Pallant, O. L. Liu, T. Lord, and M. Pryputniewicz, S. Mulholland, “Evidence-based iterative design cycles: Developing automated feedback for written scientific argumentation,” in *International Conference of the Learning Sciences, submitted*.

[45] V. K. Belur, L. Mao, O. L. Liu, and H.-S. Lee, “Student and teacher perceptions of automated feedback in an earth science assessment,” in *Annual Northeastern Educational Research Association Conference*, 2015.

[46] K. Roohr, L. Mao, and O. L. Liu, “Psychometric properties of an earth science formative assessment.,” in *Annual Northeastern Educational Research Association Conference*, 2015.

[47] M. Mulholland, “C-rater-ML: Model building and evaluation,” in *Annual Northeastern Educational Research Association Conference*, 2015.

[48] M. Zhu, L. Mao, O. L. Liu, and A. Pallant, “Impact of automated scoring and feedback on scientific argumentation in earth science: Through the log data analysis,” in *The IEEE International Conference on Data Mining*, 2015.

[49] T. Licona, P., McDonald, S. and Furman, “Student Conceptions of the Cause and Locations of Earthquakes,” *The Earth Scientist.*, vol. 29, no. 2, pp. 50–54, 2013.

[50] Y. S. Rubin, R., Plummer, J., Palma, C., Flarend, A., Spotts, H., McDonald, S. & Ong, “Assessing Student Progress Along a Solar System Learning Progression,” *Science Scope*, vol. 38, no. 1, pp. 27–33, 2014.

[51] S. McDonald, J. Plummer, A. Rivet, C. Delgado, K. Kastens, A. Flarend, K. Rubin, M. Pickard, M. Bembenic, and C. Anderson, “Integrating Crosscutting Themes, Practices, and Core Ideas: Learning Progressions in Earth and Space Science Symposium,” in *National Association for Research in Science Teaching*, 2014.

[52] S. McDonald and B. Brightbill, “Science Teacher Leadership: What can you do to lead?,” in *Pennsylvania Science Teachers Association Annual Meeting*, 2013.

[53] T. McDonald, S. & Lewis-King, "Science Teacher Leadership: In and Out of the Classroom," in *Invited Keynote at the annual Pennsylvania Earth and Space Science Teachers Association conference*, 2013.

[54] S. McDonald, K. Bateman, A. Ozcelik, H. Gall, A. Webb, and T. Furman, "Understanding Students' Ideas about Plate Tectonics: A Learning Progressions Approach," in *Annual meeting of the Geological Society of America (GSA)*, 2015.

[55] A. Webb, S. McDonald, T. Furman, H. Gall, K. Bateman, and A. Ozcelik, "A Plate Tectonics Multiple Choice Assessment: A Pilot," in *Annual meeting of the Geological Society of America*, 2015.

[56] S. Neville and L. Guertin, "Sky in Google Earth as a tool for student inquiry and exploration," *Geological Society of America Abstracts with Programs,* vol. 43, no. 1, p. 79, 2011.

[57] L. Guertin, "Middle and High School Teacher Misconceptions of Geologic Time.," *Geological Society of America Abstracts with Programs*,, vol. 43, no. 5, p. 67, 2011.

[58] J. Petula, "Grade 4-12 Student Conceptions of Astronomical Phenomena," *Geological Society of America Abstracts with Programs* vol. 43, no. 5, p. 66, 2011.

[59] M. Pickard, "Upper Elementary through High School and Teacher Misconceptions of the Interior of the Earth," *Geological Society of America Abstracts with Programs,* vol. 43, no. 5, p. 66, 2011.

[60] P. Licona, "Upper Elementary through High School and Teacher Misconceptions of the Causes of the Breakup of Pangaea," *Geological Society of America Abstracts with Programs,* vol. 43, no. 5, p. 67, 2011.

[61] T. Furman, C. Zembal-Saul, R. Crane, and M. Merritt, "Activities using authentic data introduce climate change in the middle grades classroom," *Geological Society of America Abstracts with Programs,* vol. 43, no. 1, p. 79, 2011.

[62] S. Gajeski, N. Miller, and M. Pickard, "Allowing student choice in a middle school solar system unit to increase end-of-the-year engagement," *Geological Society of America Abstracts with Programs,* vol. 43, no. 1, p. 79, 2011.

[63] L. Guertin, "Benefits and strategies for adding scientific audio news stories to the K-12 classroom," *Geological Society of America Abstracts with Programs,* vol. 43, no. 1, p. 79, 2011.

[64] C. Palma, J. Plummer, C. Ghent, T. Gleason, Y. S. Ong, and S. McDonald, "Have Astronomers Been to Neptune? Results of a Study of High School Students' Ideas about How Astronomers Study the Solar System," in *American Astronomical Society Meeting Abstracts*, 2015, vol. 225.

[65] C. Palma, A. Flarend, J. Petula, M. Richards, H. Spotts, S. McDonald, and T. Furman, "Design, Delivery, and Results of the Earth and Space Science Partnership Teacher Professional Development Program," in *American Astronomical Society Meeting*, 2013.

[66] C. Palma, J. Petula, J. Plummer, A. Flarend, and G. Goldsborough, "First Step in Building an Astronomy Learning Progression: Analyzing Student Conceptions of Astronomical Phenomena," in *American Astronomical Society Meeting Abstracts*, 2012.

[67] NGSS Lead States, *The Next Generation Science Standards*. The National Academies Press, 2013.

[68] C. C. Lewis, R. R. Perry, and J. Hurd, "Improving mathematics instruction through lesson study: a theoretical model and North American case," *Journal of Mathematics Teacher Education*, vol. 12, no. 4, pp. 285–304, 2009.

[69] C. L. Zembal-Saul, K. L. McNeill, and K. Hershberger, *What's Your Evidence?: Engaging K-5 Children in Constructing Explanations in Science*. Boston, MA: Pearson, 2012.

[70] V. J. Mayer, “Using the earth system for integrating the science curriculum,” *Science Education*, vol. 79, no. 4, pp. 375–391, 1995.

[71] O. B.-Z. Assaraf and N. Orion, “Development of system thinking skills in the context of earth system education,”*Journal of Research in Science Teaching*, vol. 42, no. 5, pp. 518–610, 2005.

[72] Y. Kali, N. Orion, and B.-S. Eylon, “Effect of knowledge integration activities on students' perception of the Earth's crust as a cyclic system,” *Journal of Research in Science Teaching*, vol. 40, no. 6, pp. 545–565, 2003.

[73] W. Hung, “Enhancing systems-thinking skills with modelling,” *British Journal of Educational Technology*, vol. 39, no. 1099–1120, 2008.

[74] W. Feurtzeig and N. Roberts, *Modeling and simulations in science and mathematics education*. New York: Springer, 1999.

[75] P. Horwitz, “Linking Models to Data: Hypermodels for Science Education,” *The High School Journal.*, vol. 79, no. 2, pp. 148–156, 1996.

[76] T. de Jong and W. R. van Joolingen, “Scientific discovery learning with computer simulations of conceptual domains,” *Review of Educational Research*, vol. 68, no. 2, pp. 179–201, 1998.

[77] M. A. Honey and M. Hilton, *Learning science through computer games and simulations*. Washington D.C.: The National Academies Press, 2011.

[78] B. D. Jee, D. H. Uttal, D. Gentner, C. Manduca, T. Shipley, B. Sageman, C. J. Ormand, and B. Tikoff, “Analogical thinking in geoscience education,” *Journal of Geoscience Education*, vol. 58, no. 1, pp. 2–13, 2010.

[79] J. Dodick and N. Orion, “Cognitive factors affecting student understanding of geologic time,” *Journal of Research in Science Teaching*, vol. 40, no. 4, pp. 415–442, Apr. 2003.

[80] C. V Schwartz, B. J. Reiser, E. A. Davis, L. Kenyon, A. Acher, D. Fortus, Y. Schwartz, B. Hug, and J. Krajcik, “Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners,” *Journal of Research in Science Teaching*, vol. 46, no. 6, pp. 632–654, 2009.

[81] J. D. Sterman, “All models are wrong: Reflections on becoming a systems scientist,” *Systems Dynamics Review*, vol. 18, no. 4, pp. 501–531, 2002.

[82] D. Allchin, “Teaching the nature of science through scientific errors,” *Science Education,*, vol. 96, no. 5, pp. 904–926, 2012.

[83] P. Bell and M. C. Linn, “Scientific arguments as learning artifacts: designing for learning from the web with KIE,” *International Journal of Science Education*, vol. 22, no. 8, pp. 797–817, 2000.

[84] R. A. Duschl and J. Osborne, “Suppoeting and promoting argumentation discourse in science education,” *Studies in Science Education.*, vol. 38, pp. 39–72, 2002.

[85] D. Kuhn, “Teaching and learning science as argument,” *Science Education*, vol. 94, no. 5, pp. 810–824, 2010.

[86] D. Kuhn, “Science as argument: Implications for teaching and learning scientific thinking,” *Science Education,*, vol. 77, no. 3, pp. 319–337, 1993.

[87] S. Erduran, S. Simon, and J. Osborne, “Tapping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse,” *Science Education.*, vol. 88, pp. 915–933, 2004.

[88] J. Osborne, S. Erduran, and S. Simon, "Enhancing the quality of argumentation in school science," *Journal of Rsearch in Science Teaching*, vol. 41, no. 10, pp. 994–1020, 2004.

[89] V. Sampson and D. B. Clark, "Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future directions," *Science Education*, vol. 92, pp. 447–472, 2008.

[90] L. A. Bricker and P. Bell, "Conceptualizations of argumentation from science studies and the learning sciences and their implications for the practices of science education," *Science Education,* vol. 92, pp. 473–493, 2008.

[91] K. E. Metz, "Children's understanding of scientific inquiry: Their conceptualization of uncertainty in investigations of their own design," *Cognition and Instruction*, vol. 22, no. 2, pp. 219–290, 2004.

[92] D. B. Clark and V. Sampson, "Assessing dialogic argumentation in online environments to relate structure, grounds, and conceptual quality," *Journal of Research in Science Teaching.*, vol. 45, pp. 293–321, 2008.

[93] T. D. Sadler and S. R. Fowler, "A threshold model of content knowledge transfer for socioscientific argumentation," *Science Education*, vol. 90, pp. 986–1004, 2006.

[94] L. K. Berland and K. L. McNeill, "A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts," *Science Education*, vol. 94, no. 5, pp. 765–793, 2010.

[95] K. L. McNeill, D. J. Lizotte, J. Krajcik, and R. W. Marx, "Supporting students' construction of scientific explanations by fading scaffolds in instructional materials," *Journal of the Learning Sciences.*, vol. 15, no. 2, pp. 153–191, 2006.

[96] A. Zohar and F. Nemet, "Fostering students' knowledge and argumentation skills through dilemmas in human genetics," *Journal of Research in Science Teaching.*, vol. 39, no. 1, pp. 35–62, 2002.

[97] G. J. Kelly and A. Takao, "Epistemic levels in argument: An analysis of university of oceanography students' use of evidence in writing," *Science Education*, vol. 86, pp. 314–342, 2002.

[98] W. A. Sandoval and B. J. Reiser, "Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry," *Science Education.*, vol. 88, pp. 345–372, 2004.

[99] E. de Vries, K. Lund, and M. Baker, "Computer-mediated epistemic dialogue: Explanation and argumentation as vehicles for understanding scientific notions," *Journal of the Learning Sciences.*, vol. 11, no. 1, pp. 63–103, 2002.

[100] A. Collins, D. Joseph, and K. Bierlaczyc, "Design research: Theoretical and methodological issues," *Journal of the Learning Sciences.*, vol. 13, no. 1, pp. 43–76, 2004.

[101] O. L. Liu, H.-S. Lee, C. Hofstetter, and M. C. Linn, "Assessing knowledge integration in science: Construct, measures, and evidence," *Educational Assessment.*, vol. 13, no. 1, pp. 1–23, 2008.

[102] O. L. Liu, H.-S. Lee, and M. Linn, "Measuring knowledge integration: A four year study," *Journal of Reearch in Science Teaching.*, vol. 48, no. 9, pp. 1079–1107, 2011.

[103] M. C. Linn, "The knowledge integration perspective on learning and instruction," in *The Cambridge handbook of the learning sciences*, K. Sawyer, Ed. New York: Cambridge University Press, 2006, pp. 243–264.

[104] H.-S. Lee, O. L. Liu, and M. C. Linn, "Validating Measurement of Knowledge Integration in Science Using Multiple-Choice and Explanation Items," *Applied Measurement in Education*, vol. 24, no. 2. pp. 115–136, 2011.

[105] H.-S. Lee, M. C. Linn, K. Varma, and O. L. Liu, "How does technology-enhanced inquiry instruction with visualizations impact classroom learning? ," *Journal of Research in Science Teaching*, vol. 47, no. 1, pp. 71–90, 2010.

[106] H.-S. Lee and O. L. Liu, "Assessing learning progression of energy concepts across middle school grades: The knowledge integration perspective," *Science Education*, vol. 94, no. 4, pp. 665–688, 2010.

[107] Z. Buck and H. S. Lee, "I'm sure there may be a planet there: Student Articulation of Uncertainty in ARgumentation Tasks," *International Journal of Science Education*, vol. 36, no. 14, pp. 2391–2420, 2014.

[108] A. Pallant and H.-S. Lee, "Constructing Scientific Arguments Using Evidence from Dynamic Computational Climate Models," *Journal of Science Education and Technology*, Jun. 2014.

[109] G. Rasch, *Probabilistic models for some intelligence and attainment tests*. Chicago, IL: University of Chicago Press, 1966.

[110] G. N. Masters, "A Rasch model for partial credit scoring," *Psychometrika*, vol. 47, pp. 149–173, 1982.

[111] T. G. Bond and C. M. Fox, *Applying the Rasch model*. Mahwah, NJ: Lawrence Erlbaum Associates, 2007.

[112] F. J. Floyd and K. F. Widaman, "Factor analysis in the development and refinement of clinical assessment instruments," *Psychological Assessment*, vol. 7, pp. 286–299, 1995.

[113] W. M. Yen, "Effects of local item dependence on the fit and equating performance of the three-parameter logistic model," *Applied Psychological Measurement.*, vol. 8, no. 2, pp. 125–145, 1984.

[114] S. McDonald and G. J. Kelly, "Understanding the Construction of a Science Storyline in a Chemistry Classroom," *Pedagogies: An International Journal.*, vol. 2, no. 3, pp. 165–177, 2007.