A learning analytics approach to scaffolding scientific modeling in the classroom

by

David Philip Quigley

B.S., Georgia Institute of Technology, 2011

M.S., Georgia Institute of Technology, 2013

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Computer Science & Institute of Cognitive Science

2018

This thesis entitled:
A learning analytics approach to scaffolding scientific modeling in the classroom
written by David Philip Quigley
has been approved for the Department of Computer Science & Institute of Cognitive Science

Prof. Tamara Sumner

Prof. William Penuel

Prof. Sidney D'Mello

Prof. Rafael Frongillo

Prof. Tom Yeh

Date ____________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

IRB protocol #12-0685

Quigley, David Philip (Ph.D., Computer Science and Cognitive Science)

A learning analytics approach to scaffolding scientific modeling in the classroom

Thesis directed by Prof. Tamara Sumner

Scientific modeling is increasingly important both in K-12 science education and the broader scientific community, but there are significant gaps in both our understanding of how people learn modeling and how we can support them in this process. This work takes a comprehensive look at how students use digital modeling tools in science classrooms, particularly with EcoSurvey, a tool developed to support students in creating a model of the components and interactions in the local ecosystem. This tool has been developed using an iterative process and deployed in three consecutive school years as part of a design-based implementation research project in high school biology classrooms. During this time, I have developed new techniques for analyzing students' models and modeling activity along with the impact of feedback and recommender systems. These approaches have demonstrated significant power in creating a picture of students' modeling activity in real time. In addition, I have determined the effects of certain design decisions on student tool utilization through iterative deployment, and found that explicit scaffolds can have a significant impact on students' models and modeling practices. Finally, I have begun to map how student activity can be related to their learning of modeling as a science and engineering practices. Through this work, I have demonstrated the power of real-time activity analytics to provide insight on the appropriate level of student support to give. This work advances learning analytics, the study of scientific modeling in the classroom, and modeling tool design.

Dedication

To my parents, Phil and Suzy, for always giving me support in whatever I do.

Acknowledgements

Thank you to my committee, Tamara Sumner, William Penuel, Sidney D'Mello, Rafael Frongillo, and Tom Yeh. Each of you have supported me with ideas and feedback throughout my process, and without you I would not be here now. Thank you especially to Tammy for always being responsive and supportive, even when I had weird questions or lost data.

Thank you to the members of the iHub team, both old and new, especially Jonathan Ostwald, Conor McNamara, and Jennifer Jacobs for writing with me. Thank you also to the rest of the team Sam Severance, Heather Leary, Katie Van Horne, Holly Devaul, Douglas Watkins, Patty Kincaid, Kim Frumin, Brian Reiser, Mike Novak, and Tara McGill for supporting me through my research. Thank you also to our design and implementation teachers for being the real heroes in helping students learn with our tools.

Thank you to the lab team, both old and new, including Ovo Dibie, Arafat Sultan, Soheil Danesh, Rick Parker, Alex Gendreau, Srinjita Bhaduri, Varsha Koushik, Ahmed Fahmy, Peter Gyory, Josh Ladd, Ali Raza, and Yael Kali, for all your feedback and support.

to my friends in the program, for helping me get through both classes and life. To the innumerable friends I have in GCF and at Crestview, thank you for being there for me in fun and in fear, and for always keeping me in your prayers.

Thank you to my Uncle Charlie and Poppop, for showing me what you can do with a life in academia. Thank you to my brother and the rest of my family for always being there for me.

Most importantly, thank you to Kaitlin, my darling partner, for putting up with me.

This material is based in part upon work supported by the National Science Foundation under Grant Numbers 1555550 and 1147590. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. In addition, I would like to thank the Gordon and Betty Moore Foundation for their support.

Contents

Tables

Table

Figures

Figure

Chapter 1

Introduction

My research studies students' scientific modeling. In particular, I focus on how we can automatically characterize students' models and their modeling practices, how the design of digital modeling tools can support student modeling practices, and how my measures of modeling engagement relate to student understanding of modeling as a science and engineering practice. I study these questions through the iterative design and deployment of EcoSurvey, a digital modeling tool for high school biology. Within this iterative process, I have developed and applied a framework for normalizing modeling actions across tools, developed a suite of analytics for characterizing modeling and practices in real time, and explored types of feedback that are useful for teachers and students to support reflection. Through this work, I demonstrate that improved modeling tool design and the incorporation of real-time feedback based on novel analytics can have a positive impact on the student modeling experience.

Scientific models represent ideas, processes, and phenomena by describing important components, their characteristics, and their interactions. Models are constructed across a broad spectrum of scientific disciplines, such as the food web in biology, the water cycle in Earth science, or the structure of the solar system in astronomy. Models are central to the work of scientists for understanding phenomena, and for constructing and communicating theories. Constructing and using models to explain scientific phenomena is also an essential practice in contemporary science classrooms. In *A Framework for K-12 Science Education* [58], developing and using models is one of the eight core practices deemed essential for science learning and instruction. According to the

Framework, "[s]cientists use models... to represent their current understanding of a system (or parts of a system) under study, to aid in the development of questions and explanations, to generate data that can be used to make predictions, and to communicate ideas to others" ([58], Pg 57). Systems models are also considered a cross-cutting concept in the *Framework*, allowing students to understand components, interactions, and mechanisms that influence a system.

Scientific models can take many forms, such as textual descriptions, visual diagrams, computer simulations, and mathematical equations. For instance, in elementary physical science, Schwarz et al.[76] studied the development of students' modeling practices by having students sketch models depicting how light interacts with objects to produce shadows. In another case, Bryce et al.[3] asked students to construct a clay model of a cell, using those models to represent their understanding of the spatial relationship between pieces. Even these simple modeling activities push students to represent their current knowledge and to use this knowledge to explain new phenomena. Models are often more complex, involving visual representations or computer simulations. Such models may focus on the complex interactions between components (e.g. predator-prey interactions in a food web) or depict how a substance changes state over time (e.g., how water changes from liquid to gas as it moves through stages in the water cycle).

However, while it is widely recognized that developing students' ability to create and use models to understand phenomena is important, learning sciences research has documented numerous challenges to implementation in the classroom. These challenges include variations in how teachers approach the topic of modelling [50] and variations in how students engage with the practices [39, 1]. Variations in classroom implementation can lead to differences in students' opportunities to learn these important modeling practices [28].

The context for my research is scientific modeling in high school biology classrooms. This work is part of Inquiry Hub, a research-practice partnership developing inquiry-based biology curriculum for middle and high school classrooms [43]. Within the high school ecology unit, students are tasked with creating a model of their local ecosystem using EcoSurvey, a digital modeling tool designed to represent the organisms and interactions the students encounter as they map a local field site.

Table 1.1: Research Questions.

Research Question	Objective	Data Sources
How can we automatically characterize students' models and their engagement with modeling practices at scale?	Computational methods can automatically characterize both models and modeling practices in real-time.	EcoSurvey Final Models & Clickstreams
What methods can we use to promote successful scientific modeling?	Designs that scaffold modeling activities and real-time feedback can help students build "better" models	Analyses of EcoSurvey Modeling Tool Design, Final Models, & Clickstreams
How do students' modeling practices relate to their understanding of scientific modeling?	Students will generally demonstrate a strong understanding of scientific modeling after the use of digital modeling tools. Students that successfully engage with modeling practices will demonstrate a deeper understanding of modeling.	EcoSurvey Clickstreams & Interviews

Through our partnership with a large urban school district in the midwestern United States, we are collecting data from over a thousand students as they use EcoSurvey in their classrooms.

My work crosses the fields of computer science, cognitive science, and education. I develop and adapt new methods of computational learning analytics to understand student activities in modeling tools, as well as advance understandings of human-centered computing in the area of digital modeling tool design. This work also helps define scientific modeling at the student level, providing quantitative representations of student engagement with modeling practices. This quantitative representation also helps solidify definitions of student scientific modeling in learning sciences. Finally, this work has a real-world impact on teachers and students in the classroom, helping them to teach and understand both ecological concepts of organisms and interactions as well as abstract scientific modeling skills.

1.1 Research Objectives

My research is organized around three key questions, summarized in Table 1.1.

1.1.1 (RQ1) How can we automatically characterize students' models and their engagement with modeling practices at scale?

The purpose of this investigation is to understand the classroom activities of students. Scientific modeling is a new focus in the K-12 classroom; while scientists have been using models for many years and students have been using models as part of their classroom activities, emphasizing modeling as a science and engineering practice with connections across fields is a new approach [58].

My approach looks at modeling as a science-wide practice. I use emerging learning sciences theory related to models and associated activities to create a framework of what makes a good system model, as a cross-cutting science concept. This framework drives approaches for creating a standardized representation of models (normalizing) and the actions students take to create, revise, and use them. I study how this approach applies to the use of digital modeling tools, particularly our EcoSurvey tool. By normalizing models and modeling practices to existing learning sciences theory, we can develop analytic approaches that examine the use of digital modeling tools from a modeling perspective, rather than from the domain-specific perspective of how these models apply to the problem at hand.

First and foremost, this work supports teachers and students. In our preliminary work, teachers expressed a pressing need for support in analyzing and evaluating students' modeling to provide substantive guidance through instruction, discussion, and grading. My methods create a summary view of the students' models and activity that teachers can quickly interpret. This summary can then be leveraged in classroom practice to guide students in creating richer models.

This approach also opens new grounds in learning analytics. Normalizing models and modeling practices allows me to apply computational methods to supporting teachers in classrooms at scale. This approach allows me to generalize findings beyond the particular use case of our teachers using EcoSurvey. I also expand learning analytics to incorporate new techniques that apply to sequential activity with a digital tool.

1.1.2 (RQ2) What methods can we use to promote successful scientific modeling?

This question explores the impact of tool design and analytic feedback, specifically in how well students leverage modeling practices and the strength of their final models (as discussed in RQ1). The characterization of models and modeling practices, while an important theoretical development, is ultimately useful in terms of how it can be used to impact students' learning. I use these measures of modeling to understand differences that arise from the iterative design, development, and deployment of EcoSurvey.

I use two different approaches to promote scientific modeling. First, I examine the design of the tool and workflow of the students' modeling process. This work focuses on how the features of our digital modeling tool map to the modeling practices. The process of mapping actions to practices provides insights into how well those practices are represented. Furthermore, analysis of student use of EcoSurvey can highlight where students do not engage with certain modeling practices.

The other approach examines the students' modeling practices to generate customized feedback that can inform future actions. The idea of providing formative feedback is well developed in digital learning environments and broader learning sciences contexts (e.g. [66]). However, teachers in our pilot deployments have requested greater support in understanding their students' models and modeling activity. If we can develop systems that use an online approach to analyzing models and practices, we can provide support to teachers and students during the modeling process, helping teachers give strong formative feedback and helping students to understand modeling as a series of practices.

The normalized features of models and modeling practices provide an ideal metric for answering this question. Since the normalization approach is designed to work across different modeling scenarios, the process provides a direct comparison between different versions of EcoSurvey.

1.1.3 (RQ3) How do students' modeling practices relate to their understanding of scientific modeling?

This question characterizes the experience from the student perspective. My work seeks to develop a system that drives students to not only engage with modeling practices and build successful models, but also to understand modeling as a science and engineering practice. We measure understanding through short interviews that increasingly prompt students to discuss modeling. These interviews are then coded for the depth and generality of students' explanations and the level of prompting required for them to give their most advanced description. In addition, I find the correlation between students' modeling practices and their responses to our modeling interviews. This helps validate our normalization process, providing insight into how those measures relate to understanding.

1.1.4 Conjecture Map

These questions inform a conjecture map [74] of how my interventions will influence mediating processes to affect change on models and modeling practices seen in figure 1.1. A conjecture map, as the name suggests, visualizes how a proposed change to a system (in this case, a digital modeling environment) can influence the outcomes for your targeted element of the system (in this case, the student experience). The map opens with high-level conjectures about possible changes to the system and links those conjectures to action items that embody them. These action items, in turn, are linked to mediating processes that take place within the system, which then link to expected outcomes.

In order to determine the validity of this set of conjectures, I need to understand my outcomes, that is, what makes a "better" model, and how successful students create such models (i.e. what constitutes the practice of modeling). In addition, these definitions should be as generalizable as possible, in order to apply to both ecosystems students in a variety of ecological contexts as well as modeling activities in other domains. This generalizability also supports the development of

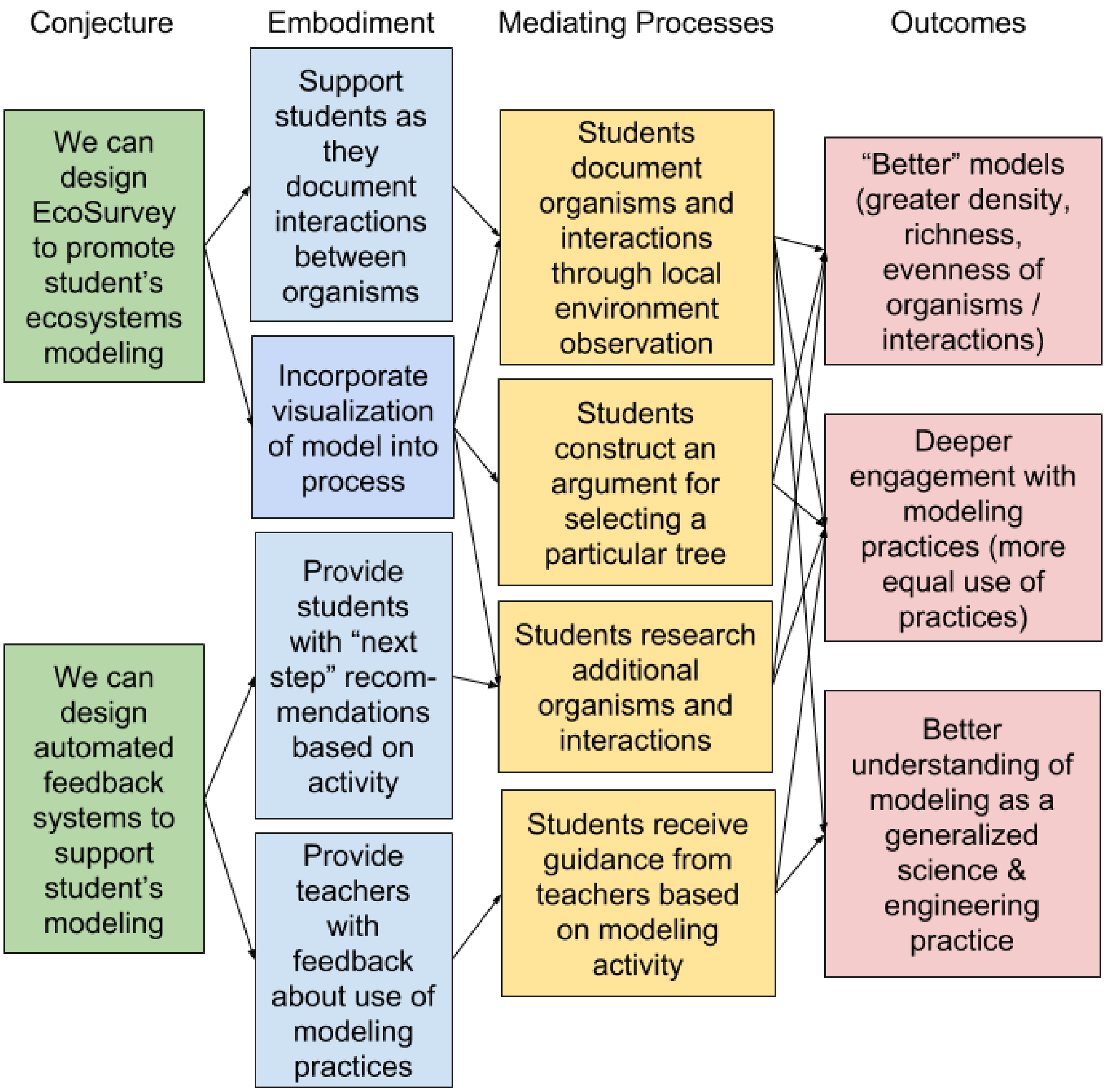


Figure 1.1: Conjecture Map.

analyzing these features in real-time, which can then be presented to users to shape future actions. This leads to RQ1, focused around the development of automated methods for characterizing models and modeling practices.

My next objective is to understand how I can design tools, particularly EcoSurvey, to support students in this modeling process (RQ2). My intervention is focused on designing EcoSurvey to naturally afford [61] engagement with modeling practices and the incorporation of a personalized feedback system that analyzes modeling activity in real time. In particular, my design scaffolds the interaction mapping within the tool and the regular use of a graph visualization to see the model as a whole. Meanwhile, my feedback supports teachers in understanding students' modeling activities within EcoSurvey and gives students direct feedback on their modeling activity.

In addition to examining the student experience within EcoSurvey, I also want to look at how using the tool relates to generalized knowledge about modeling as a science and engineering practice and how students perceive the experience. This leads to RQ3, which examines the outcome measures that evaluate student understanding.

Answering these questions will significantly contribute to the body of knowledge in learning analytics and learning sciences research, as well as provide design guidelines to developers of digital modeling tools and instructional materials. Learning analytics researchers benefit from the exploration of how existing techniques can be extended to understand student scientific modeling (RQ1). Seeing how these methods apply to scientific modeling can connect analytics work to active research in learning sciences and can demonstrate the power of learning analytics to impact learning in real time (RQ2). Learning sciences researchers benefit from an understanding of student and classroom differences in students' scientific models and modeling practices (RQ1), exploration into new ways to guide student modeling (RQ2) and an understanding of the impacts of digital modeling activities on the student experience (RQ3). Instructional tool and materials developers gain insight into the impact of various design decisions on student activity in an open-ended modeling context (RQ2) and ideas on how to support student understanding (RQ3).

My work also helps to support teachers and students in real-world classrooms. Scaffolds guide students through the modeling activity, helping them to "dive in". My analytics make students' models visible to teachers in a summarized, dashboard view that shows how students have engaged with modeling practices. This approach also allows for teachers and students to see what steps might contribute to creating a better model.

1.2 Related Work

My research bridges work and ideas from Learning Sciences and multiple subdisciplines of Computer Science to enhance students' opportunities to learn. The foundations of this work stem from scientific modeling education research, a growing discipline in the Learning Sciences space that seeks to understand the features of classroom scientific models and the processes behind students' modeling. This research demonstrates a growing body of knowledge from a cognitive perspective of student understanding, but there are also opportunities to address issues around addressing equity of access and understanding students' modeling at scale. Machine Learning research provides ideas and methods for understanding these very issues; I leverage normalization and classification techniques to provide a generalized understanding of modeling without expert intervention. Human-Centered Computing techniques bring a unique opportunity to explore both how we can design modeling tools and systems to support strong scientific modeling as well as methods to use our modeling insights to provide targeted and supportive feedback for students during the modeling process.

1.2.1 Scientific Modeling in Education

1.2.1.1 Scientific models that support complete explanations of phenomena

Scientific models are tools for explanation and prediction. A complete scientific explanation should "explain observed relationships between variables and describe the mechanisms that support cause and effect inferences about them" [58]. Thus, to support student explanations, a scientific model of a phenomenon should include important components ("variables"), their interactions ("relationships"), and define the mechanisms involved. When modeling an ecosystem, these correspond to the organisms in the ecosystem (animals, plants, insects, fungi, etc), how these organisms interact with each other and the environment (predator, prey, producer, decomposer, etc), and the involved processes (abiotic, biotic, etc). Professional biologists use this information to measure the biodiversity of an ecosystem in terms of species richness, evenness, and divergence [26, 50, 19]. Implementing modeling in the classroom adds an extra layer of complexity, incorporating teachers' understandings in addition to those of the students [32, 90]. Nevertheless, by building models, students can apply their understanding of ecosystems and gain insight into the components of good models.

In this work, I characterize variation in students models by examining the number of organisms present, the variety of types of organisms present, the number of interactions between organisms that students have identified, and the diversity of these interaction types. I also look at how these features are distributed within a model. These measures are used to understand the complexity of a student model. Interestingly, understanding the complexity of an ecosystem has been shown to support students to develop empathy and other affective stances towards nature [36]. Student understanding the flow of matter and energy through ecosystems has also been shown to vary strongly across cultural boundaries [1], providing further motivation for addressing equity in student models and student modeling practices.

1.2.1.2 Strong student scientific modeling practices

Constructing scientific models is part of the "inquiry" tradition in science education, where students learn scientific concepts through hands-on "doing" [39]. Understanding what students are doing at a fine-grained level can provide teachers with useful insights into learning processes, as well as provide teachers with feedback as to where and when students need additional assistance. Towards this end, several scholars have developed frameworks characterizing effective student modeling practices [76, 3]. Schwarz et al. [76] identify a series of seven practices: (1) identifying the anchoring phenomena to be modeled, (2) constructing a model, (3) testing the model, (4) evaluating a model, (5) comparing the model against other ideas, (6) revising the model, and (7) using the model to predict or explain phenomena. Bryce et al.[3] identify a similar set of practices as important to support student learning during modeling, namely (1) observation (paralleling the anchoring phenomena), (2) model construction, (3) model use, (4) model evaluation, and (5) model revision. Their research suggests that supporting students to engage in these practices can lead to positive learning outcomes [76].

Here, I focus on a subset of these practices - constructing, evaluating, revising, and using models - incorporating them into my analysis framework [10]. I use these four practices as they are directly supported through the EcoSurvey interface and can be readily observed and tracked in the usage log. In addition to these four practices, I examine the degree to which students engaged in iterative design of their models. Iteration occurs when students cycle between the other four modeling practices, where the four practices correspond directly to individual actions in the EcoSurvey interface, such as adding an organism or relationship (construction), editing an organism or relationship (revision), or generating a graph of the entire ecosystem to support a tree choice (using). Iteration is an important modeling practices that is used to both expand the scope of a model and to improve its accuracy [28, 3]. Learning analytic techniques are used to identify the degree to which students used these practices and to examine variations in student modeling practices. While these usage log analysis methods are an excellent passive way to collect data on student practices [65], it is important to note that these methods do not capture information about how students are reasoning with their models. Exploring student reasoning with models and how they generate explanation using models is beyond the scope of this study, and would require deep exploration of students' cognitive processes using think-alouds, cognitive interviews or other learning and cognitive sciences research methods (e.g. [76]).

1.2.1.3 Equity in Student Learning Experiences

Student learning outcomes vary widely across teachers [33, 53]. Students with a top-performing math teacher can be expected to perform .266 standard deviations better on a standardized math test than those with a median teacher [33]. Similarly, McNeill et al.[53] evaluated 22 high school ecology classrooms across the US and found that teacher differences accounted for 34.5% of the variance on scores from a multiple choice assessment and 42.5% of the variance on scores from an open ended assessment. Differences in student learning outcomes can be attributed, in part, to differences in their opportunities to learn different topics [52, 55]. For instance, in a classroom setting, the opportunity for iteration can be driven by the structure of the class: students will not expand or refine their model if they are not given the opportunity to do so.

Differences in student learning can also be attributed to differences in the curriculum being utilized, and differences in how teachers implement curriculum in their specific classroom [33]. Large variations in how teachers implement STEM "inquiry-oriented" curriculum have been routinely observed [40, 75], and curriculum integrating modeling is no exception. Windschitl et al. [90] conducted a series of studies examining how K-12 teachers integrated student modeling into their classrooms and found significant variance in teacher understanding and adoption. For many teachers, the traditional scientific method notion of generating a hypothesis is deeply ingrained in their views of science practices. Subsequently, these teachers had difficulty adopting a scientific practice that required them to ground ideas and predictions in an initial model. In some cases, they found that teachers simply rejected the model-based inquiry approach, citing that providing students with opportunities to engage in iterative practices took too much classroom time and added unnecessary complexity. Similarly, Jordan et al. [32] found that teachers spent most of their time using models as a communication tool rather than a practice for students to engage with in the classroom.

One of the important approaches to defining improvement across learning sciences is to reduce the variance attributable to the sources inherently present in the education system [4]. By reducing the variance attributable to differences in teacher, school, or population demographics, we can provide a more equitable educational experience. Our objective in this work is to reduce the variance in students use of modeling practices, with a particular focus at reducing effects found at the teacher level. This reduction in variance will promote equity in the learning experience by giving all students opportunities to learn [28].

In my analysis, I examine variations in student modeling across classrooms and teachers, analyzing both students' opportunities to learn and variations in the degree to which they engaged in specific modeling practices. For these analyses, I use measures of frequency and variety as features [77]. Frequency characterizes how often students were able to engage in the different modeling practices within EcoSurvey, whereas variety captures the breadth of practices that they engaged in. Frequency and variety have been shown to reliably predict the uptake and adoption of new technologies across different groups of users [77, 51]. These features differentiate patterns of uptake and adoption of modeling practices across classrooms.

I also explore the ways in which an individual student's modeling processes can be indicative of teacher differences. I use sequence classification techniques [90] to detect recurring patterns, called sequential patterns or action sequence features, in student's modeling practices, as they engage in cycles of creating, evaluating, revising, and using their models. I explore the degree to which automatically extracted and optimized action sequence features are able to correctly predict a specific student's teacher. These pattern mining methods have been used by learning analytics researchers to address questions related to course selection trajectories [11] and group work dynamics [68]. Automatic feature optimization is a common technique used in data mining to identify the features that carry predictive value for classification [21]; the resulting features can reveal insights into processes important for differentiating between categories [20]. In this case, I am using these sequences to detect and understand potential differences in modeling curriculum implementation across teachers.

1.2.2 Learning Analytics

1.2.2.1 Understanding Student Activity

Understanding student activity is an active area of Learning Analytics research. There is a large body of work focused on detecting students' skill acquisition using digital tools. One example of this is a study focused on young students' ability to make numeric and fractional estimates in a number line game [46]. In this study, they found significant gains in student accuracy over time.

One important area of research around understanding student activity focuses on detecting variance between students as a predictor of future activity. One example is Jihyun Park et al's [63] work on understanding student use of online classroom resources using features drawn from student clickstreams. These features focus on generalizing beyond the course content, focusing on simple frequency measures (e.g. number of clicks per day) and abstractions of how the content accessed relates to the course schedule, determining whether the content being accessed is being "previewed" or "reviewed". The team used these metrics to categorize whether students' engagement with online resources increased, decreased, or stayed the same, creating three different categories of students. Overall, they found that students who specifically increased their engagement with the tool had a better success rate in the class.

While these aggregate features are useful in understanding activity, another approach to understanding this activity and variance is the use of sequence modeling [91]. This approach focuses on a fine-grained distinction of different activities and analyzes patterns in how activities lead into one another. This approach parallels that used by d'Aquin et al.[11], where they used sequential pattern mining to study student course enrollment patterns.

These innovations inspire my approach to analyzing student engagement with modeling practices. By adapting the approaches of aggregate activity and sequence analysis to scientific modeling in the classroom, I develop new insights into how students participate in this crucial activity.

1.2.2.2 Online Analysis

While Learning Analytics bodies of work demonstrate a strong body of existing work in understanding student activity, this work is often performed in a post-hoc fashion, extracting what improvements, differences, or gaps may exist from a tool deployment. An important area of ongoing research seeks to understand how these methods can be used to predict potential performance. Creating "online" systems (i.e. systems that measure performance and draw conclusions in real time, during use) has the potential to detect critical differences and gaps in learning and engagement. In turn, these predicted differences can be used for intervention, supporting students in useful ways.

Online difference detection and intervention is not new. In general user interface research, this area has driven the development of many different tools, including interventions as ubiquitous as "Clippy", the oft-ridiculed Microsoft Office digital assistant. At the 2017 conference on Learning Analytics & Knowledge, several research teams (e.g. [24, 47, 30]) presented preliminary results in designing and/or deploying predictive systems for education and learning, while many others (e.g. [70, 34]) cited prediction for intervention as an important next step for their research.

There are three main styles of online intervention: alerts, automated task selection, and dashboards. In alert-based interventions, such as the aforementioned Clippy, the system provides some sort of notification to the user of some error or inefficiency in their actions. This approach faces numerous problems, primarily accuracy (and the balance between misses and false positives, e.g [29]) and usability (e.g. Clippy [57]). Automated task selection, or choosing the problem/example to show a student based on features of the problem/object and previous responses of the student, is a common approach in the learning setting (e.g. [73]). However, this approach is uniquely suited to learning tasks focused on repetition, such as image classification. Dashboards serve as a passive learning intervention that focuses on "informating" [92] and guiding only when accessed. This approach, while overcoming the burdens of interruption provided by alert systems and working well in longitudinal tasks, faces issues related to the need to access the information (limiting the access to those with the motivation to seek out feedback) and the possible complexity of interpretation.

In my work, a dashboard-type information view is the most effective method for showing real-time feedback. The motivations of a joint formal classroom activity, such as EcoSurvey, provide a higher level of baseline motivation to improve through the social pressures of collaboration and the external motivation provided by the teacher. The students can also benefit from the metacognitive aspects of interpreting feedback, allowing them to understand both how to improve their models as well as how to engage with modeling more generally. In addition, the teachers have motivation as professionals to both guide students to engaging with feedback as well as scaffold their interpretation.

1.2.3 Human Centered Computing

1.2.3.1 Scaffolds

Scaffolding, or the use of external supports to help a person accomplish a task is a common approach in the learning sciences. Vygotsky [88] discussed such supports when describing the zone of proximal development for apprentices in trade fields. Scaffolds are natural extensions of this idea, tools that support a learner in this process, providing guidance or removing levels of complexity from the problem. These scaffolds can then be removed from the application, setting, or activity when the learner is ready to move on.

This scaffolding approach has been adopted in software-supported learning scenarios (e.g. [71, 16]). One particular line of work has focused on digital applications that support scientific inquiry. Quintana et al.[72] describe three phases of scaffolding scientific inquiry:

> 1) Characterizing the cognitive tasks, social interactions, tools, and artifacts that constitute the scientific practices in which learners are engaged.
>
> 2) Characterizing the aspects of these practices in which learners encounter obstacles.
>
> 3) Characterizing scaffolding guidelines that specify ways that tools can alter the task to address the obstacles by helping make tasks more tractable and productive for learners.

My work naturally builds on these guidelines. By focusing the design of EcoSurvey around the practices outlined in learning science literature on modeling in the classroom, we naturally connect to the real practices of the task. My analytics are designed to automatically measure how successful students are at engaging with these practices and the points at which students run into difficulty. I have used this feedback in design, and the results show that students are using the revised modeling tool to engage more successfully with modeling.

1.2.3.2 Co-Design / User-Centered Design / Iterative Design

User-Centered Design is an important methodology for creating usable and useful tools for real-world use. This process focuses on several key techniques, including iteration [60]. These methods have helped define what makes for a successful design, with a focus on outcome measures related to the task at hand. This work motivates both our design process as well as the need for new approaches to defining successful models and engagement with modeling practices in the classroom.

My design work takes place in the context of a larger Design Based Implementation Research (DBIR; [14]) project. This process builds on the ideals of user-centered design, but expands its commitment to involving end users (in our case, teachers) as co-researchers and designers, testing and refining theoretical assumptions during iterative design cycles. We employ a co-design process [83, 66] to develop and refine our curricula and systems, including EcoSurvey. Co-design is "a highly-facilitated, team-based process in which teachers, researchers, and developers work together in defined roles to design an educational innovation, realize the design in one or more prototypes, and evaluate each prototype's significance for addressing a concrete educational need" ([66], p. 51). We have successfully used this process to design both software tools [83] and new curriculum [77] with teachers.

1.3 Research Context

1.3.1 Inquiry Hub

EcoSurvey was developed as part of a larger collaborative design-based research project called the Inquiry Hub, which is focused on supporting teachers in developing student-centered approaches to curriculum and teaching [77]. Inquiry Hub Biology is a digital high school biology curriculum developed in partnership with Denver Public Schools. Within the ecosystems unit of this curriculum, students are asked to choose a tree to plant on their school grounds or other designated site that will improve their local ecosystem's biodiversity and resilience.

1.3.2 EcoSurvey

To develop an understanding of their local ecosystem, students use EcoSurvey to document and visualize organisms and interactions they encounter. Within EcoSurvey, students join a "survey" set up for their class, take photos and field notes on organisms they find on their school grounds, and create a "card" for each organism. Students add details to each card about the organism including its role in the ecosystem and relationships it has with other organisms. EcoSurvey is designed to support peer-review, allowing for students to provide feedback and make edits to cards quickly and easily. These models are visualized in a graph view, allowing for further review and use as evidence as the proper tree to plant.

1.4 Dissertation Structure

This dissertation is built from three studies that address parts of the research questions discussed above. Chapters 2, 3, and 4 have either been published as a journal or conference article or are being drafted for future submission. Chapter 2 reviews the analysis of our first year of EcoSurvey deployment, and was initially published as a conference paper at the 2017 7th Annual Conference on Learning Analytics & Knowledge in Vancouver. That chapter builds the foundation for answering RQ1, and establishes the motivation and baseline for RQ2. Chapter 3 expands on our first year analysis to perform comparison with data from our second deployment. It was published in the Frontiers in ICT journal special issue on digital education in 2017, and reinforces the impact of my analytics on RQ1 while beginning the exploration of RQ2. The final third year of deployment is captured in Chapter 4 and is targeted for submission to the Journal of Learning Analytics special issue on Human-Centered Analytics. This final study reinforces our findings on RQ1 and RQ2, as well as discovering some boundaries for those conclusions. The chapter also explores RQ3 by examining the correlation between modeling activity and the ability to explain and discuss modeling. All three papers include collaborative work and contributions made by my thesis advisor Tamara Sumner. In addition, my collaborator Jonathan Ostwald made contributions to the work presented in chapters 2 and 3, my undergraduate research assistant Conor McNamara contributed to methods and analyses presented in chapter 3, and my collaborator Jennifer Jacobs contributed to data analysis presented in chapter 4. Finally, Chapter 5 will revisit the research questions, and discuss how each one has been addressed through these three studies.

Chapter 2

Scientific Modeling: Using learning analytics to examine student practices and classroom variation

This chapter was published as a full paper (best paper nominee) at the 2017 Learning Analytics & Knowledge conference in Vancouver (see [70]). This work establishes the foundations of my approach to understanding scientific models and modeling practices at scale (RQ1) and provides a baseline measure of these two facets for exploring the impact of our iterative design (RQ2).

2.1 Introduction

Scientific models represent ideas, processes, and phenomena by describing important components, their characteristics, and their interactions. Models are constructed across a broad spectrum of scientific disciplines, such as the food web in biology, the water cycle in Earth science, or the structure of the solar system in astronomy. Models are central to the work of scientists for understanding phenomena, and for constructing and communicating theories. Constructing and using models to explain scientific phenomena is also an essential practice in contemporary science classrooms. In *A Framework for K-12 Science Education* [58], developing and using models is one of the eight core practices deemed essential for science learning and instruction. According to the *Framework*, "[s]cientists use models... to represent their current understanding of a system (or parts of a system) under study, to aid in the development of questions and explanations, to generate data that can be used to make predictions, and to communicate ideas to others" [58].

Scientific models can take many forms, such as textual descriptions, visual diagrams, computer simulations, and mathematical equations. For instance, in elementary physical science, Schwarz et al [76] studied the development of students' modeling practices by having students sketch models depicting how light interacts with objects to produce shadows. Bryce et al [3] asked students to construct a clay model of a cell. Even these simple modeling activities push students to represent their current knowledge and to use this knowledge to explain new phenomena. Models are often more complex, involving visual representations or computer simulations. Such models may focus on the complex interactions between components (e.g. predator-prey interactions in a food web) or depict how a substance changes state over time (e.g., how water changes from liquid to gas as it moves through stages in the water cycle).

In this research, we study the development of student modeling practices in secondary biology classrooms. In these classrooms, students used a web-based software tool - EcoSurvey - to characterize organisms and their interrelationships found in their local urban ecosystem. Students use EcoSurvey to: (1) photograph, map and characterize local species, (2) document how species interact around shared resources such as food, and (3) identify resources and species that are important to the resilience of their environment. EcoSurvey follows in a rich tradition of computer-based modeling tools [80, 37, 15]. These digital modeling tools provide built-in affordances that foreground important scientific modeling practices, and are explicitly designed to scaffold students' modeling activities, through the careful design of the interface and prompts promoting reflection and appropriate action [71, 15]. As such, they support students to develop more complex models that would be difficult to create using traditional tools and these models can be quickly revised thanks to their digital nature.

Digital modeling tools also provide an opportunity for instrumentation to unobtrusively capture usage. Reflecting contemporary software architectures, EcoSurvey is a cloud-based software tool, where all changes and refinements to student models are centrally captured and stored, providing researchers with a fine-grained record of student modeling practices at scale, across potentially thousands of students in a wide range of classroom settings. These rich data offer opportunities for new learning analytic methods to better characterize student scientific modeling practices and to examine classroom level differences. In this paper, we use learning analytics and machine learning techniques to answer the following questions:

1) What variation do we see in the models created by students to support explanations of scientific phenomena, in our case, ecosystem biodiversity?

2) What variation do we see in student modeling practices across different teachers?

3) Can the action sequences used by students during modeling be used to predict each student's teacher?

We analyzed EcoSurvey usage data collected from over 200 secondary students across ten classrooms. We observed large variations in the completeness and complexity of student models, and large variations in their iterative refinement processes. We also observed large differences in student modeling practices across different classrooms and teachers, and we were able to predict a student's teacher based on the observed modeling practices with a high degree of accuracy without significant tuning of the predictive model. These results highlight the value of this approach for extending our understanding of student engagement with an important contemporary science practice, as well as the potential value of analytics for identifying critical differences in classroom implementation. These results shed light on potential improvements in tools and curricula. Before discussing our approach and results further, we first present the education and learning sciences theories underpinning this work and describe our research context and the EcoSurvey tool in more detail.

2.2 Theory and Related Work

A central goal of our approach is to develop theoretically-grounded analytic methods. Education research and the learning sciences offer insights into three areas critical to our approach: the elements of a "good" student model, how to characterize student modeling practices, and variation in classroom implementation across teachers.

2.2.1 Scientific models that support complete explanations of phenomena

Scientific models are tools for explanation and prediction. A complete scientific explanation should "explain observed relationships between variables and describe the mechanisms that support cause and effect inferences about them" [58]. Thus, to support student explanations, a scientific model of a phenomenon should include important components ("variables"), their interactions ("relationships"), and define the mechanisms involved. When modeling an ecosystem, these correspond to the organisms in the ecosystem (animals, plants, insects, fungi, etc), how these organisms interact with each other and the environment (predator, prey, producer, decomposer, etc), and the involved processes (abiotic, biotic, etc). Professional biologists use this information to measure the biodiversity of an ecosystem in terms of species richness, evenness, and divergence [26, 50, 19].

In this work, we characterize variation in students models by examining the number of organisms present, the variety of types of organisms present, the number of interactions between organisms that students have identified, and the diversity of these interaction types. We also look at how these features are distributed within a model. These measures are used to understand the complexity of a student model. Interestingly, understanding the complexity of an ecosystem has been shown to support students to develop empathy and other affective stances towards nature [36]. Student understanding the flow of matter and energy through ecosystems has also been shown to vary strongly across cultural boundaries [1], providing further motivation for understanding variation in student models and student modeling practices.

2.2.2 Strong student scientific modeling practices

Constructing scientific models is part of the "inquiry" tradition in science education, where students learn scientific concepts through hands-on "doing" [39]. Understanding what students are doing at a fine-grained level can provide teachers with useful insights into learning processes, as well as provide teachers with feedback as to where and when students need additional assistance. Towards this end, several scholars have developed frameworks characterizing effective student modeling practices [76, 3]. Schwarz et al. [76] identify a series of seven practices: (1) identifying the anchoring phenomena to be modeled, (2) constructing a model, (3) testing the model, (4) evaluating a model, (5) comparing the model against other ideas, (6) revising the model, and (7) using the model to predict or explain phenomena. Bryce et al [3] identify a similar set of practices as important to support student learning during modeling, namely (1) observation (paralleling the anchoring phenomena), (2) model construction, (3) model use, (4) model evaluation, and (5) model revision. Their research suggests that supporting students to engage in these practices can lead to positive learning outcomes [76].

Here, we focus on a subset of these practices - constructing, evaluating, revising, and using models - incorporating them into our analysis framework [10]. We focus on these four practices as they are directly supported through the EcoSurvey interface and can be readily observed and tracked in the usage log. In addition to these four practices, we examine the degree to which students engaged in iterative design of their models. Iteration occurs when students cycle between the other four modeling practices, where the four practices correspond directly to individual actions in the EcoSurvey interface, such as adding an organism or relationship (construction), editing an organism or relationship (revision), or generating a graph of the entire ecosystem to support explanations (using). Iteration is an important modeling practices that is used to both expand the scope of a model and to improve its accuracy [28, 3]. Learning analytic techniques are used to identify the degree to which students used these practices and to examine variations in student modeling practices. While these usage log analysis methods are an excellent passive way to collect data on student practices [65], it is important to note that these methods do not capture information about how students are reasoning with their models. Exploring student reasoning with models and how they generate explanation using models is beyond the scope of this study, and would require deep exploration of students' cognitive processes using think-alouds, cognitive interviews or other learning and cognitive sciences research methods (e.g. [76]).

2.2.3 Teacher Differences

Student learning outcomes vary widely across teachers [33, 53]. Students with a top performing math teacher can be expected to perform .266 standard deviations better on a standardized math test than those with a median teacher [33]. Similarly, McNeill et al [53] evaluated 22 high school ecology classrooms across the US and found that teacher differences accounted for 34.5% of the variance on scores from a multiple choice assessment and 42.5% of the variance on scores from an open ended assessment. Differences in student learning outcomes can be attributed, in part, to differences in their opportunities to learn different topics [52, 55]. For instance, in a classroom setting, the opportunity for iteration can be driven by the structure of the class: students will not expand or refine their model if they are not given the opportunity to do so.

Differences in student learning can also be attributed to differences in the curriculum being utilized, and differences in how teachers implement curriculum in their specific classroom [33]. Large variations in how teachers implement STEM "inquiry-oriented" curriculum have been routinely observed [40, 75], and curriculum integrating modeling is no exception. Windschitl et al [90] conducted a series of studies examining how K-12 teachers integrated student modeling into their classrooms and found significant variance in teacher understanding and adoption. For many teachers, the traditional scientific method notion of generating a hypothesis is deeply ingrained in their views of science practices. Subsequently, these teachers had difficulty adopting a scientific practice that required them to ground ideas and predictions in an initial model. In some cases, they found that teachers simply rejected the model-based inquiry approach, citing that providing students with opportunities to engage in iterative practices took too much classroom time and added unnecessary complexity.

In our analysis, we examine variations in student modeling across classrooms and teachers, analyzing both students' opportunities to learn and variations in the degree to which they engaged in specific modeling practices. For these analyses, we use measures of frequency and variety as features [79]. Frequency characterizes how often students were able to engage in the different modeling practices, whereas variety captures the breadth of practices that they engaged in. Frequency and variety have been shown to reliably predict the uptake and adoption of new technologies across different groups of users [79, 51]. Here, we use these features to study the different patterns of uptake and adoption of modeling practices across classrooms.

We also explore the ways in which an individual student's modeling processes can be indicative of teacher differences. We use sequence classification techniques [91] to detect recurring patterns, called sequential patterns or action sequence features, in student's modeling practices, as they engage in cycles of creating, evaluating, revising, and using their models. We explore the degree to which automatically extracted and optimized action sequence features are able to correctly predict a specific student's teacher. These pattern mining methods have been used by learning analytics researchers to address questions related to course selection trajectories [11] and group work dynamics [68]. Automatic feature optimization is a common technique used in data mining to identify the features that carry predictive value for classification [21]; the resulting features can reveal insights into processes important for differentiating between categories [20]. In our case, we are using these sequences to detect and understand potential differences in modeling curriculum implementation across teachers.

2.3 Research Context: Inquiry Hub and EcoSurvey

EcoSurvey was developed as part of a larger collaborative design-based research project called the Inquiry Hub, which is focused on supporting teachers in developing student-centered approaches to curriculum and teaching [77]. Inquiry Hub Biology is a digital high school biology curriculum developed in partnership with a large urban school district in the midwestern United States. Within the ecosystems unit of this curriculum, students are asked to choose a tree to plant on their school grounds or other designated site that will improve their local ecosystem's biodiversity and resilience. Classes use EcoSurvey to create a collective model of their local ecosystem. They use these models to provide evidence and construct arguments to support their choice about the type of tree they choose to plant. The recommended type of tree is then planted on the site, in collaboration with the local Parks and Recreation Department, based on the students' arguments and evidence. Thus, the models students create using EcoSurvey support them to construct arguments with real world consequences. To illustrate the use of EcoSurvey within this context, we follow the experience of "Maria", a fictional student in Ms. Smith's 3rd period class.

2.3.1 Data Collection and Creating the Model

Ms. Smith instructs students to map the ecosystem within a selected site on their school grounds or in the local area, taking pictures and making field notes on the organisms and interactions between organisms that they observe. Maria's group makes observations along the creek that runs next to the school. They find a lady beetle, a honey locust tree, some mushrooms, a gray squirrel, and a few other organisms. Using their smartphones, they take pictures of these organisms and upload them to EcoSurvey, creating a "card" for each organism while out in the field. Each card automatically captures information about the date, time, and location of the observation being recorded. Cards also include a "relations" field to capture interactions between organisms and information about the organism's role in the ecosystem. Students begin entering this information as they observe it in the field, and then continue to augment this information back in the classroom through additional research. In Figure 2.1, we see Maria's lady beetle card under construction. While in the field, she created the card, uploaded a picture, and added details about interactions they saw. At the same time, her team members are also creating cards for other organisms they are observing.

2.3.2 Evaluating the Model

As students create cards, their organisms are added to a shared class "survey". The survey view shows all of the organism cards and their detailed information, ordered by how recently they were edited. Maria can see that her classmates have created many cards, including a Blue Jay card (Figure 2.2).

Ms. Smith organizes the student groups into pairs and asks each group to review the other's

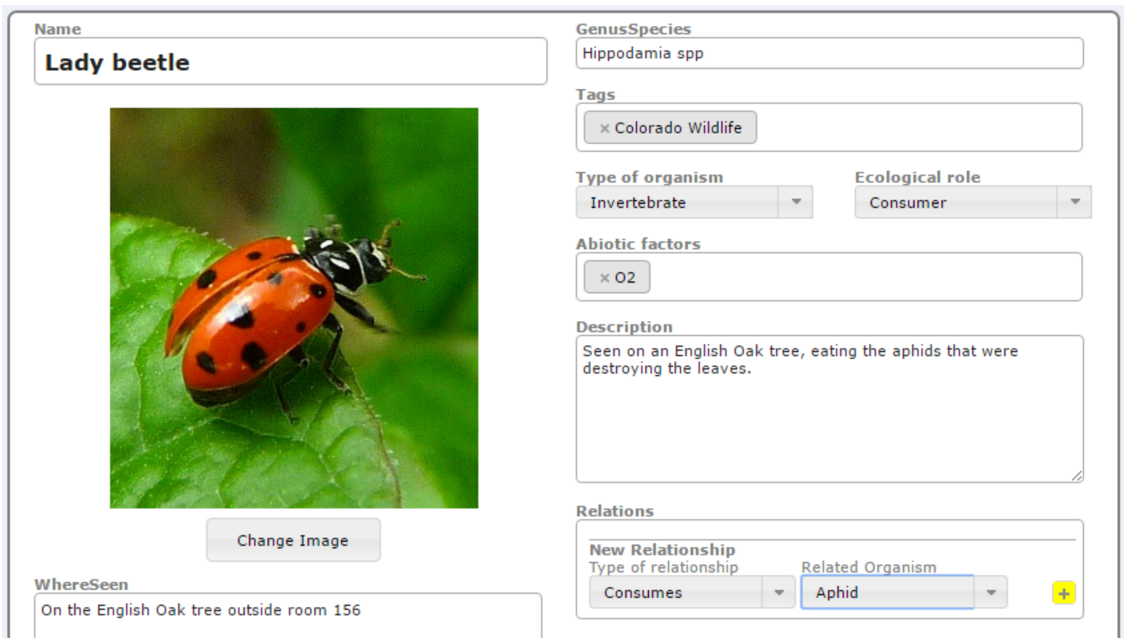


Figure 2.1: The edit view for Maria's Lady Beetle card.

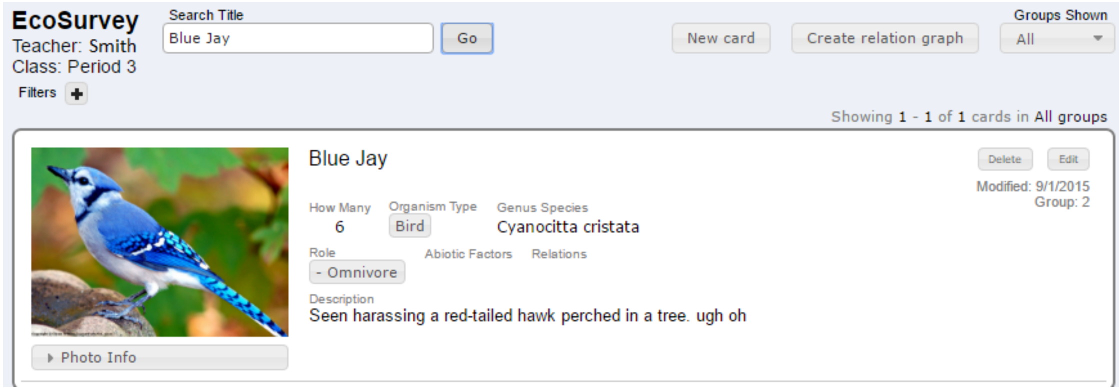


Figure 2.2: The main view of Ms. Smith's class survey.

cards for correctness and completeness. Maria's group is paired with Group 2, who completed several cards. Andre, a member of Group 2, asks Maria to first review the red tailed hawk card he created. Maria uses the search feature of the survey view to quickly find the hawk among the cards. She notices that this card is missing many details, including interactions with other organisms.

2.3.3 Revising the Model

Maria recommends that Group 2 do further research into how the hawk contributes to the local ecosystem. She also takes the chance to update her group's gray squirrel and honey locust cards. She discovered that hawks prey upon squirrels and nest in honey locust trees during her earlier research. She didn't realize that their school ecosystem included hawks until she reviewed the work of her classmates, as her group did not see a hawk. Once Maria has completed editing her group's cards, she continues her review of Group 2's cards. She uses the group select function to view only the cards created by members of Group 2.

Group 2 notices that two people in Maria's group created duplicate lady beetle cards. Maria decides to add her lady beetle information to the other card, since it is more detailed, and uses the delete function to remove her lady beetle card from the model.

2.3.4 Iterating the Model

In reviewing Group 2 cards, Maria sees a card for geese, but notices that the group did not add a predatory relationship to grass, even though she observed geese eat the grass on the soccer field. She uses the search functionality and discovers that no one in class created a card to document grass as an observed organism. Maria adds a new card for grass and includes a predatory-prey relationship with geese. By cycling back through earlier modeling practices (creating new cards), Maria is iteratively improving the class model to be more complete and accurate.

2.3.5 Using the Model

Once the class has created a robust model of their local ecosystem, students use this model to construct arguments for choosing a particular tree to plant. Maria presses the "create relation graph" button, which generates the graph representation of the model and exports it to a digital graphing tool(Figure 2.3). Maria and her team study the resulting diagram that enables them to visualize the relationships (links) between all the organisms (nodes) they have cataloged. It is clear from looking at her graph that the English Oak trees are an important keystone species in their site, involved in a large number of relationships with a wide variety of organisms. The geospatial locations in the observational data indicate that there are only two English Oak trees located in their site; Maria and her group recommend planting an additional tree of this type.

2.3.6 Analyzing EcoSurvey Use

Maria's scenario illustrates how EcoSurvey supports students to engage in the practices of creating, evaluating, revising, iterating, and using models. To use a learning analytics approach to study modeling practices, we must map specific actions, or sequences of actions, taken in the EcoSurvey interface to specific modeling practices. Table 1 describes the mapping between modeling practices and specific EcoSurvey interface actions that we use in our analyses. As students interact with EcoSurvey, the system captures and logs each of the actions shown in Table 2.1. Each log entry includes the time, user, survey, and action type.

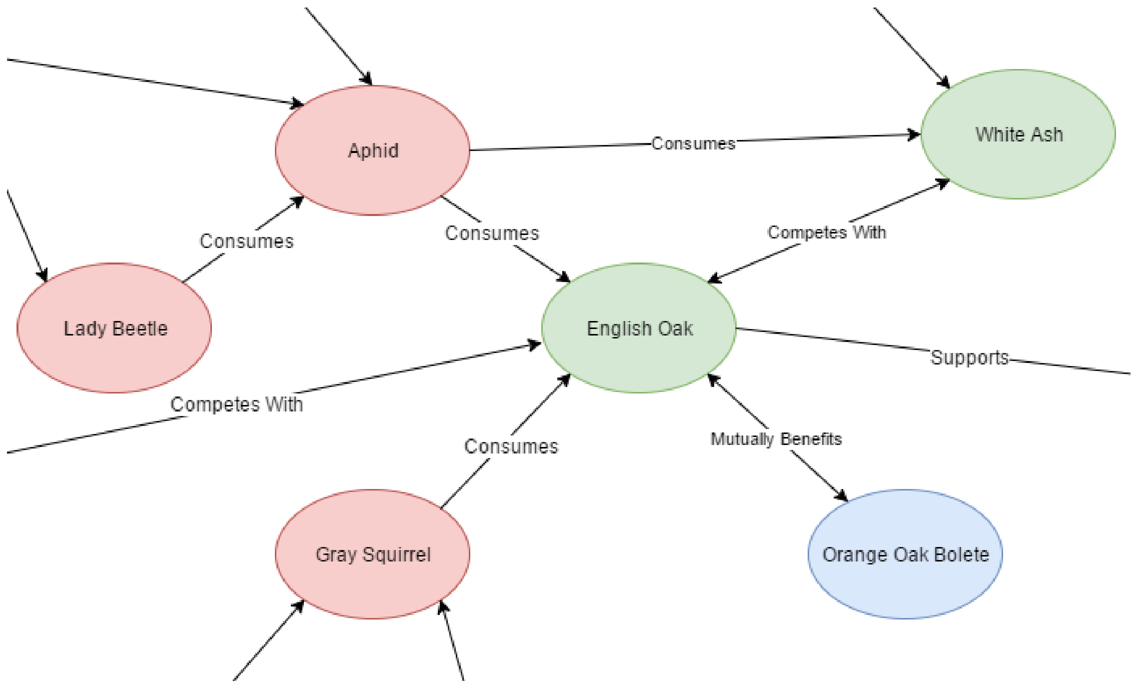


Figure 2.3: A section of Maria's final graph.

Table 2.1: EcoSurvey Actions

Modeling Practice	Description	EcoSurvey Actions
Create Model	Create a new entry in the model	New Card
Evaluate Model	Explore the organisms and interactions in the current model	Group Select, Search
Revise Model	Edit or delete organisms and interactions included in the current model	Edit, Delete
Use Model	Export a representation of the model for use (e.g. constructing an argument)	Generate Graph, Download
Iterate	Cycle between creation, revision, and use practices	New Card, Edit, Delete, Generate Graph, Download

2.4 Methods

Here, we describe data used in our analyses as well as the specific analytic techniques used to answer each of our three research questions. All teachers' names are pseudonyms.

2.4.1 Study Data

EcoSurvey usage log data was collected from 262 students, across 10 high school classrooms, during Fall 2015. These 10 classes were taught by three different teachers: Anderson, Baker, and Chavez. Anderson taught two periods of high school biology, which she elected to combine into one group to produce a single ecosystems model. Baker taught three periods, while Chavez taught five. From the sample, we recorded actions for 204 students, while 58 students did not record any activity. All classrooms in this sample followed a 3:1 device deployment where three students used one laptop together; thus it is not surprising that there are students with no recorded activity. A total of 9 models were created, which included 586 organism cards and 545 interactions, generating 3160 action logs.

2.4.2 Variation in Student Scientific Models

Our first research question examines variation within student models, focusing specifically on the richness of students' models in terms of the number of organisms and their relationships. We analyze the relative number of organisms and interactions within each class survey. We also look at the balance of interactions per organism by evaluating both the average number of interactions per organism and variance in the distribution of interactions. Examining variance allows us to distinguish different patterns in the assignment of interactions to organisms. Some classes may create models where most organisms have a similar number of interactions, while other classes may create models where only a few organisms have been assigned many interactions.

2.4.3 Variation in Modeling Practices

Our second research question examines variation in student modeling practices, focusing on action variety, frequency, and iteration. Action variety refers to the range of actions a student performed. For example, some students may have only created and edited cards, while others may have used the full range of EcoSurvey actions. Frequency refers to the total number of actions completed by an individual student and the number of usage sessions they engaged in. Sessions are defined by a series of actions from a single user without a large break in activity (greater than two hours). Defining a session using a two hour gap allows for any student activity within a long class period to occur within one session; several of our classrooms employ 1.5 hours block periods.

To characterize iteration practices, we look for evidence of design cycles within the log information. Design cycles can be recognized when students engage in multiple sequences of construct-revise-use practices. This focus on a sequence of practices is consistent with Schwarz et al [76], which characterized modeling practices as a series of steps. By extension, a design cycle consists of returning to a previous modeling step after moving on in the sequence (e.g. creating a new card after editing a different card). We counted the number of cycles as a measure of iteration.

Combined, these three metrics - action variety, frequency, and iteration - yield an eight feature vector for each student consisting of total number of EcoSurvey actions, total number of create actions, total number of evaluate actions, total number of revise actions, total number of use actions, total number of EcoSurvey action types taken, number of sessions, and number of iterations. We combined the feature vectors for students with the same teacher, and performed a Kruskal-Wallis H test [42] for each feature to determine differences between teachers. A Kruskal-Wallis H test is a non-parametric adaptation of an ANOVA to compare samples of different sizes, as we have in our groups. We further explored these differences using Tukey's HSD test [84] to test the significance of pairwise differences between teachers.

2.4.4 Predictive Value of Modeling Practices

Our third research question examines the degree to which we can use sequences of student modeling actions to predict that student's teacher. For this prediction task, we use the previously described features of variety, frequency, and iteration as well as automatically extracted sequence patterns. This sequence pattern approach is inspired by the feature-based sequence classification methods summarized by Xing, Pei, and Keogh [91]. In our work, a sequence pattern consists of a series of EcoSurvey actions (e.g. "New Card", "Edit", "Generate Graph") embedded within a student's complete action log. To extract sequence patterns, we used the Colibri Core [86] software package. This software package, originally designed for natural language processing tasks, treats every action as a token and determines the frequency of consecutive token sequences (n-grams) from student usage logs. These token sequences can include wildcard actions (skip-grams). For instance, the software will extract the sequence "New Card", "Edit", "Generate Graph" as either an n-gram or as the skip-gram "New Card", {*}, "Generate Graph". This skip-gram will capture similar sequence patterns, where one action occurs between New Card and Generate Graph actions. This yielded 2,893 unique sequence patterns, that occurred at least three times, across all student usage logs. Once we extracted these sequence patterns, we used them as a new series of features to augment each student's existing feature vector. This approach parallels that used by d'Aquin et al [11], where they used sequential pattern mining to study student course enrollment patterns.

To understand which features that characterize a student's modeling actions are most predictive of his or her teacher, we input subsets of each student's feature vector into four Naive Bayes classifiers using Weka [22]. The first classifier used the eight features related to variety, frequency, and iteration of actions. The second classifier used the full set of sequence pattern extracted by Colibri Core for each student. The third classifier implemented a best-first search [21], which automatically reduced the full set of sequence patterns to the eighteen most predictive features. The last classifier combines the eight variety, frequency, and iteration features with the eighteen most predictive sequence patterns. Each test was run using 10-fold cross validation.

Table 2.2: Final models for each class.

Survey	# Users	Organisms	Interactions	Int. Per Org.	Int. Var.
Anderson 4 & 7	29	155	264	1.7	4.35
Baker 1	28	47	7	0.149	0.297
Baker 2	27	25	5	0.2	0.24
Baker 4	29	19	0	0	0
Chavez 1	27	88	70	0.795	0.663
Chavez 2	29	45	27	0.6	1.31
Chavez 6	30	60	57	0.95	3.78
Chavez 7	31	81	82	1.012	5.72
Chavez 8	32	66	33	0.5	0.826

2.5 Results

Results are presented for each of our three research questions.

2.5.1 (RQ1) What variation do we see in the models created by students?

As shown In Table 2.2, there are substantial variations in the models created by students in different classrooms. We see that Anderson's students documented many more organisms (155) and interactions (264) than all other classes. Though Anderson had both of her classes work together to create one survey, the total number of students contributing to this model is comparable to the number of students contributing in other classrooms. We also see that students in Baker's three classes each documented significantly fewer organisms and interactions. One class only documented 19 organisms (less than one per student) and did not document any interactions. Chavez's classes exhibit wide variation, particularly in the numbers of interactions documented by each class.

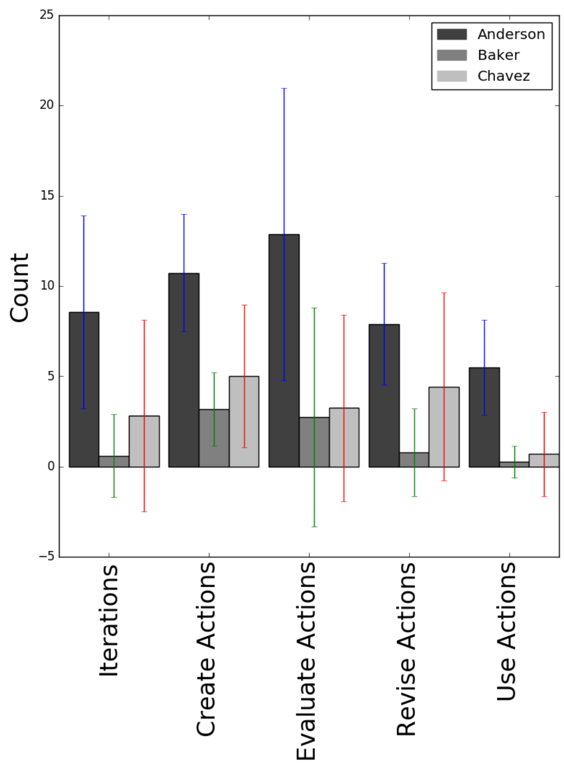
The number of interactions per organism, a broad measure of model complexity, further illustrates apparent classroom differences, with Anderson's class creating more complex models than Baker's and Chavez's classes. To better understand classroom differences, we examine variance in the number of interactions per organism. In Anderson's class, we see a high variance in comparison to the interactions per organism metric, which indicates that there are a small number of organisms with lots of interactions and many organisms with few interactions.

Chavez' P1 and P7 classes provide a particularly interesting case to examine this variation. On reviewing Table 2.2, we see that the variance in the number of interactions assigned to each organism is significantly lower in P1 than in P7, while the actual number of organisms and interactions are comparable. Further analyses reveal that students in Chavez's P1 did not assign any interactions for 39% of their organisms, while students in P7 did not assign interactions to 74% of their organisms. A similar analysis revealed that 42% of the organisms documented in Anderson's model did not include interactions. In most classes, the majority of organisms have no documented interactions. It appears that students engaged significantly more with describing organisms, and spent far less time consistently documenting interactions.

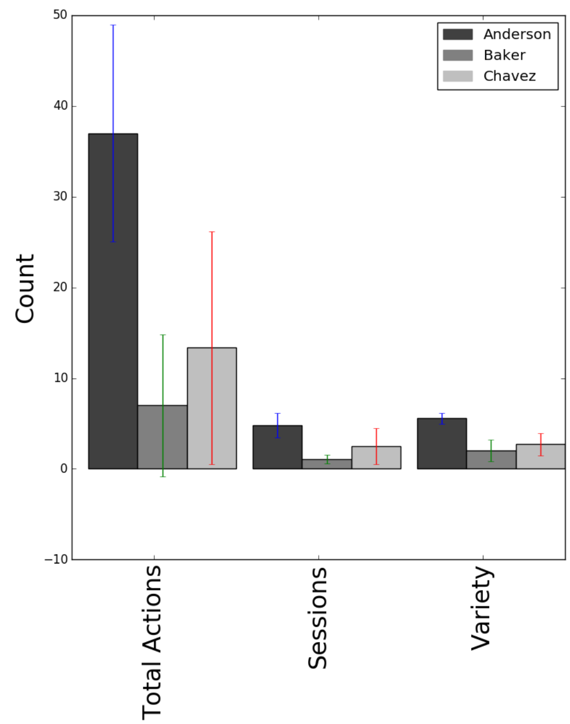
2.5.2 (RQ2) What variation do we see in student modeling practices across different teachers?

There are significant differences between the student action sequences of our three teachers on all eight metrics related to variety, frequency, and iteration ($p < .001$). Our Tukey's HSD test for each feature shows that the three groups are each distinct to a significant degree in Create, Revision, and Iteration frequency (Figure 2.4a, $p < .05$), as well as Overall Actions, Session Count, and Action Variety (Figure 2.4b, $p < .05$). We also see Anderson's students performed significantly more Evaluate and Use actions than the other two teachers' students (Figure 2.4a, $p < .05$), though the differences between Baker's and Chavez's students are not significant. Anderson's class also used EcoSurvey twice as much, as measured by session counts. Overall, Anderson's students engaged in more modeling practices than both of the other two groups, and Chavez's students engaged in more modeling practices than Baker's.

There were also differences in the modeling practices that students employed. Students in Baker's classes rarely engaged in three of the five modeling practices we are studying: revisions, iteration, or use. Chavez's class engaged with four of the five practices, but appeared to rarely use their models.



(a) The average number of actions by modeling practice type.



(b) The average number of actions, types of actions, and action sessions.

Figure 2.4: Student modeling practices for each teacher's students.

Table 2.3: Predictive accuracy of each action sequence feature set

Feature Set	# Attributes	Naive Bayes Acc
Baseline	0	51.96%
All Sequence Patterns	2,893	63.73%
Variety, Frequency, and Iteration Features	4	67.65%
Best Sequence Patterns	18	75.00%
Combined Features	22	80.39%

2.5.3 (RQ3) Can the action sequences used by students during modeling be used to predict each student's teacher?

As shown in Table 2.3, student action sequences can predict their teacher with varying degrees of reliability depending upon the features used. Our baseline assumes that each student is in one of Chavez's classes; almost 52% of the students in this study were in one of his classes. All of the feature sets we studied improved performance over the baseline. Classifying based on all 2,893 sequence patterns improved our classification accuracy by almost 12%, whereas classifying solely based on our variety, frequency, and iteration features improved performance by over 15%. We also trained a model on the best sequence patterns, that is, the 18 most predictive patterns identified by Weka's Attribute Selection tool [21]; this yielded a nearly 25% improvement in performance. The best performing model was one that combined the most predictive sequence patterns with our variety, frequency, and iteration features. This combination resulted in a 30% improvement over baseline, correctly predicting a student's teacher 80% of the time.

The most useful features for classification accuracy are the 18 "best" sequence patterns (Table 2.4). A closer examination reveals that these sequence patterns correspond to our five modeling practices in interesting ways. These patterns prioritize model revision, evaluation, and iteration as distinguishing features, which correspond to the differences in classroom modeling practices discussed under research question 2.

To better understand the types of errors that our best performing model makes, we generated a confusion matrix (Table 2.5). We see that 75% of the errors are due to the misclassification of 30 of Chavez's students as Baker's students. One possible reason for this misclassification is that some

Table 2.4: The most predictive action sequences.

New card, New card, {*}[1], New card, {*}, {*}, New card	New card, {*}, Group Select, {*}, New card	Group Select
Group Select, {*}, Group Select	Group Select, {*}, {*}, {*}, Group Select	Group Select, {*}, New card, {*}, New card
Group Select, Search	Search, {*}, {*}, {*}, Edit	Edit
Edit, Edit	Edit, {*}, Edit	Edit {*} {*} Edit
Edit, Search	Edit, Generate Graph, Download	Edit, Generate Graph, Download, Edit
Generate Graph	Download	Generate Graph, Download

students in Chavez's classes performed very few modeling actions overall, similarly to the majority of students in Baker's classes.

2.6 Discussion

In this study, we demonstrated the utility of learning analytic methods for characterizing variation in students' scientific models and their modeling practices. We also showed that an individual student's modelling action sequences can be used to predict his or her teacher. Our results support Windschitl et al's findings documenting large variations in how teachers implement modeling in their classrooms [90]. While we did not conduct direct classroom observations, our analysis revealed profound, quantifiable differences in the models that students constructed across different classrooms and significant differences in their classroom learning experiences as depicted in the range of modeling practices that they engaged in.

Student models exhibited large variance in the number of organisms and interactions doc-

[1] A {*} refers to a wild card in a skip-gram, which can be compelted with any value.

Table 2.5: Combined features confusion table.

		Classified As		
		Anderson	Baker	Chavez
Correct Class	Anderson	29	0	1
	Baker	1	64	3
	Chavez	5	30	71

umented. These differences could be due to a variety of factors, such as the time allocated to modeling during class, the degree to which modeling practices were incorporated into instruction, or their teacher's dispositions and knowledge about scientific modeling. Our results suggest that such teacher level differences do matter. Another source of variation could be differences in ability and knowledge that individual students bring to the modeling task. In our current work, we are revising the Inquiry Hub curriculum to provide better guidance to teachers to integrate modeling into their classroom, and we are providing more opportunities for students to engage in modeling throughout the unit.

Our analysis of student models also revealed a disturbing similarity across all classrooms and teachers: all the models contained significant percentages of organisms that did not have a single defined interaction with another organism. Thus, these student models are missing critical elements of a complete and sound ecosystem model. It is unlikely that these models can support students to develop comprehensive explanations and predictions as called out in the *Framework* [58]. There are multiple possible explanations for these behaviors, including weaknesses in the Inquiry Hub curriculum, the associated teacher professional development, or the design of the EcoSurvey tool. As a first step, we have made major changes to the design of EcoSurvey version 2 to make it easier for students to establish relationships from multiple parts of the interface, to visualize established relationships through an integrated graph view, and to see which organisms are not connected to others in the model.

The large variance we observed in student modeling practices provides evidence of significant teacher-level differences. Clearly these teachers are implementing EcoSurvey and the corresponding lessons differently in their classrooms, with wildly varying results. When teachers devoted more time to modeling, as measured by sessions, their students' engaged in a richer variety of modeling practices. Prior research suggests that there is a linkage between student engagement in modeling practices and future learning outcomes [76, 3]. Thus, it appears that students in several of our participating classrooms lacked critical opportunities to learn [52, 55], that could ultimately impact their academic performance. In future work, we plan to examine the relationships between student engagement in modeling practices and their learning outcomes as measured by end-of-course school district assessments.

Our predictive analysis provided further evidence of significant teacher-level differences. The feature selection algorithm honed in on the presence or absence of three modeling practices - evaluation, revision, and iteration - as the features that best predicted a student's teacher. This suggests that future professional development and curriculum design should focus on these specific practices, ensuring that all students get an opportunity to participate in these parts of the modeling process. In EcoSurvey version 2, we have expanded features designed to support evaluation, revision, and iteration practices. For instance, we have implemented generating a visual graph of their model directly into the tool, rather than exporting this information into a 3rd party graphic tool. By facilitating students to use (visualize) their models more frequently, we hope that this will prompt them to notice shortcomings and engage in more iterative refinements. The most accurate classifier also benefited from additional features characterizing action variety, frequency (number of actions), and iteration. These features further highlight differences in student engagement, with some students missing the opportunity to explore, develop, and use their models over time.

A core aspect of our analytic approach explicitly linked specific user interface actions in the EcoSurvey tool to individual modeling practices identified through prior research: creating, evaluating, revising, using, and iterating [76, 3, 28]. This approach enabled us to work with theoretically and empirically sound features identified through prior classroom research. And, this approach enabled us to interpret the action sequences identified as salient by our algorithms in a theoretically-informed way, enabling us to link our findings back to instructional concerns, such as curriculum design and professional development. This method of linking interface actions to identified modeling practices could support generalizing this analytic approach to other tools that support scientific modeling.

While this study yielded many results that have informed our partnership design work, there are several limitations that are important to note. First, we are working with a limited data set, containing data from only three teachers and 9 models. While we generated interesting insights into differences between these classrooms, it is difficult to generalize our findings to a broader spectrum of classrooms. Second, we cannot attribute our observed variation in models and modeling practices to student-level differences, due to the shared and collaborative nature of the deployment. All our participating classrooms asked students to work in groups and each group shared a single laptop computer; we are actually observing the collaborative modeling practices of small groups rather than individual students.

2.7 Conclusion

We have demonstrated that learning analytics can be used to study student scientific models and student modeling practices at a scale that has previously been impossible. We used quantitative statistical measures to study variation across models and teachers. We also used methods drawn from data mining and machine learning to identify critical differences in student modeling practices and to explore which features of student modeling sequences are useful for classification.

This work opens the door for a wide variety of further research. Future directions could incorporate student demographics and examine potential differences in the uptake of modeling practices across various populations. Future work could also incorporate student assessment data to look at connections between engagement in modeling practices and student learning outcomes. Other work could further explore teacher-level differences, combining classroom observations with learning analytics to better understand the different approaches teachers take during classroom implementation.

The work presented here has already informed the Inquiry Hub partnership's effort. The design-based research team is making evidence-based changes to our curriculum, professional development, and classroom tools based on these results. Other research groups studying student scientific modeling can apply these theories and analytic techniques in their settings to understand variation in models, modeling practices, and classroom implementation.

2.8 Acknowledgments

This material is based in part upon work supported by the National Science Foundation under Grant Numbers 1555550 and 1147590. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. In addition, the authors would like to thank the Gordon and Betty Moore Foundation for their support.

Chapter 3

Using Learning Analytics to Understand Scientific Modeling in the Classroom

This chapter first appeared as a journal article in the Frontiers in ICT publication special issue on Digital Education (see [69]). This work is a direct extension of the previous publication/chapter, but adds information on the second year of deployment. This chapter reveals some impacts of our iterative design (RQ2), particularly in regards to the distribution of relationship types in models between years. There are also significant null results on the impacts of design presented here, particularly in the coninued persistence of the orphaned card phenomenon and the teacher and classroom variance in size and complexity of models.

3.1 Introduction

Scientific models represent ideas, processes, and phenomena by describing important components, their characteristics, and their interactions. Models are constructed across a broad spectrum of scientific disciplines, such as the food web in biology, the water cycle in Earth science, or the structure of the solar system in astronomy. Models are central to the work of scientists for understanding phenomena, and for constructing and communicating theories. Constructing and using models to explain scientific phenomena is also an essential practice in contemporary science classrooms. In A Framework for K-12 Science Education ([58]), developing and using models is one of the eight core practices deemed essential for science learning and instruction. According to the Framework, [s]cientists use models... to represent their current understanding of a system (or parts of a system) under study, to aid in the development of questions and explanations, to generate data that can be used to make predictions, and to communicate ideas to others ([58]).

Scientific models can take many forms, such as textual descriptions, visual diagrams, computer simulations, and mathematical equations. For instance, in elementary physical science, Schwarz et al [76] studied the development of students modeling practices by having students sketch models depicting how light interacts with objects to produce shadows. Bryce et al [3] asked students to construct a clay model of a cell. Even these simple modeling activities push students to represent their current knowledge and to use this knowledge to explain new phenomena. Models are often more complex, involving visual representations or computer simulations. Such models may focus on the complex interactions between components (e.g. predator-prey interactions in a food web) or depict how a substance changes state over time (e.g., how water changes from liquid to gas as it moves through stages in the water cycle).

However, while it is widely recognized that developing students modelling skills is important, learning sciences research has documented numerous challenges to implementation in the classroom. These challenges include variations in how teachers approach the topic of modelling ([40, 32]) and variations in how students engage with the practices ([76, 3]). Variations in classroom implementation can lead to differences in students opportunities to learn these important modeling practices ([52]).

Learning Analytics can play a valuable role in understanding these differences in opportunities to learn. By focusing on how data streams can be used to characterize learner activity and understanding, researchers have been creating adaptive and responsive systems that leverage new insights to improve the learning experience for those students who need support. This approach has been leveraged in many learning scenarios (e.g. [7, 27]), but has seen limited application in scientific modeling.

In this research, we study the development of student modeling practices using digital modeling tools in secondary biology classrooms. In these classrooms, students used a web-based software tool - EcoSurvey - to characterize organisms and their interrelationships found in their local urban ecosystem. Students use EcoSurvey to: (1) photograph, map and characterize local species, (2) document how species interact around shared resources such as food, and (3) identify resources and species that are important to the resilience of their environment. EcoSurvey follows in a rich tradition of computer-based modeling tools (e.g. [80, 37, 15]). These digital modeling tools provide built-in affordances that foreground important scientific modeling practices, and are explicitly designed to scaffold students' modeling activities, through the careful design of the interface and prompts promoting reflection and appropriate action ([71, 15]). As such, they support students to develop more complex models that would be difficult to create using traditional tools and these models can be quickly revised thanks to their digital nature.

Digital modeling tools also provide an opportunity for instrumentation to unobtrusively capture usage. Reflecting contemporary software architectures, EcoSurvey is a cloud-based software tool, where all changes and refinements to student models are centrally captured and stored, providing researchers with a fine-grained record of student modeling practices at scale, across potentially thousands of students in a wide range of classroom settings. These rich data offer opportunities for new learning analytic methods to better characterize student scientific modeling practices and to examine classroom level differences. In this paper, we use learning analytics and machine learning techniques to answer the following questions:

1) How can we automatically measure the extent to which students scientific models support complete explanations of phenomena?

2) How does the design of student modeling tools influence the complexity and completeness of students models?

3) How do clickstreams reflect and differentiate student engagement with modeling practices?

We analyzed EcoSurvey usage data collected from over 1000 secondary students across two deployments. In the first deployment, we observed large variations in the completeness and complexity of student models, and large variations in their iterative refinement processes. We also observed large differences in student modeling practices across different classrooms and teachers, and we were able to predict a student's teacher based on the observed modeling practices with a high degree of accuracy without significant tuning of the predictive model. In our second deployment, we saw improvements in the completeness and complexity of students models, suggesting benefits from improvements in modeling tool design.

These results highlight the value of this approach for extending our understanding of student engagement with scientific modeling, as well as the potential value of analytics for identifying critical differences in classroom implementation. These results shed light on potential improvements in tools and curricula. Before discussing our approach and results further, we first present the education and learning sciences theories underpinning this work and describe our research context and the EcoSurvey tool in more detail.

3.2 Related Work

3.2.1 Scientific models that support complete explanations of phenomena

Scientific models are tools for explanation and prediction. A complete scientific explanation should explain observed relationships between variables and describe the mechanisms that support cause and effect inferences about them ([58]). Thus, to support student explanations, a scientific model of a phenomenon should include important components (variables), their interactions (relationships), and define the mechanisms involved. This approach is similar to the Structure-Behavior-Function model (SBF, [25]). However, the focus on interactions in our approach allows for more abstract relationships that do not rely on one particular behavior or function.

When modeling an ecosystem, these correspond to the organisms in the ecosystem (animals, plants, insects, fungi, etc), how these organisms interact with each other and the environment (predator, prey, producer, decomposer, etc), and the involved processes (abiotic, biotic, etc). Professional biologists use this information to measure the biodiversity of an ecosystem in terms of species richness, evenness, and divergence ([19, 26, 50]).

In this work, we characterize variation in students models by examining the number of organisms present, the variety of types of organisms present, the number of interactions between organisms that students have identified, and the diversity of these interaction types. We also look at how these features are distributed within a model. These measures are used to understand the complexity of a student model. This approach is similar to prior research understanding student models, particularly work from [32]. Our approach to analyzing student models primarily differs from the focus on components and interactions rather than using the SBF framework. This change allows us to automatically characterize the different pieces of models, which can be used to support real-time analysis and feedback in the future.

Interestingly, understanding the complexity of an ecosystem has been shown to support students to develop empathy and other affective stances towards nature [36]. Student understanding the flow of matter and energy through ecosystems has also been shown to vary strongly across cultural boundaries [1], providing further motivation for supporting equitable opportunities to learn scientific modeling.

3.2.2 Strong student scientific modeling practices

Constructing scientific models is part of the inquiry tradition in science education, where students learn scientific concepts through hands-on doing [39]. Understanding what students are doing at a fine-grained level can provide teachers with useful insights into learning processes, as well as provide teachers with feedback as to where and when students need additional assistance. Towards this end, several scholars have developed frameworks characterizing effective student modeling practices [76, 3]. [76] identify a series of seven practices: (1) identifying the anchoring phenomena to be modeled, (2) constructing a model, (3) testing the model, (4) evaluating a model, (5) comparing the model against other ideas, (6) revising the model, and (7) using the model to predict or explain phenomena. [3] identify a similar set of practices as important to support student learning during modeling, namely (1) observation (paralleling the anchoring phenomena), (2) model construction, (3) model use, (4) model evaluation, and (5) model revision. Their research suggests that supporting students to engage in these practices can lead to positive learning outcomes [76].

Here, we focus on a subset of these practices - constructing, evaluating, revising, and using models - incorporating them into our analysis framework [10]. We focus on these four practices as they are directly supported through the EcoSurvey interface and can be readily observed and tracked in the usage log. In addition to these four practices, we examine the degree to which students engaged in iterative design of their models. Iteration occurs when students cycle between the other four modeling practices, where the four practices correspond directly to individual actions in the EcoSurvey interface, such as adding an organism or relationship (construction), editing an organism or relationship (revision), or generating a graph of the entire ecosystem to support explanations (using). Iteration is an important modeling practices that is used to both expand the scope of a model and to improve its accuracy [28, 3]. Learning analytic techniques are used to identify the degree to which students used these practices and to examine variations in student modeling practices. While these usage log analysis methods are an excellent passive way to collect data on student practices [65], it is important to note that these methods do not capture information about how students are reasoning with their models. Exploring student reasoning with models and how they generate explanation using models is beyond the scope of this study, and would require deep exploration of students cognitive processes using think-alouds, cognitive interviews or other learning and cognitive sciences research methods (e.g. [75]).

3.2.3 Learning Analytics of Student Activity

Understanding student activity is an active area of Learning Analytics research. There is a large body of work focused on detecting students skill acquisition using digital tools. One example of this is a study focused on young students ability to make numeric and fractional estimates in a number line game. In this study, they found significant gains in student accuracy over time.

One important area of research around understanding student activity focuses on detecting variance between students as a predictor of future activity. One example is [63], understanding student use of online classroom resources using features drawn from student clickstreams. These features focus on generalizing beyond the course content, focusing on simple frequency measures (e.g. number of clicks per day) and abstractions of how the content accessed relates to the course schedule, determining whether the content being accessed is being previewed or reviewed. The team used these metrics to categorize whether students engagement with online resources increased, decreased, or stayed the same, creating three different categories of students. Overall, they found that students who specifically increased their engagement with the tool had a better success rate in the class.

While these aggregate features are useful in understanding activity, another approach to understanding this activity and variance is the use of sequence modeling [91]. This approach focuses on a fine-grained distinction of different activities and analyzes patterns in how activities lead into one another. This approach parallels that used by [11], where they used sequential pattern mining to study student course enrollment patterns.

These innovations inspire my approach to analyzing student engagement with modeling practices. By adapting the approaches of aggregate activity and sequence analysis to scientific modeling in the classroom, I develop new insights into how students participate in this crucial activity.

3.2.4 Scaffolds

Scaffolding, or the use of external supports to help a person accomplish a task is a common approach in the learning sciences. [88] discussed such supports when describing the zone of proximal development for apprentices in trade fields. Scaffolds are natural extensions of this idea, tools that support a learner in this process, providing guidance or removing levels of complexity from the problem. These scaffolds can then be removed from the application, setting, or activity when the learner is ready to move on.

This scaffolding approach has been adopted in software-supported learning scenarios (e.g. [71, 16]). One particular line of work has focused on digital applications that support scientific inquiry. [72] describe three phases of scaffolding scientific inquiry:

> 1) Characterizing the cognitive tasks, social interactions, tools, and artifacts that constitute the scientific practices in which learners are engaged.
>
> 2) Characterizing the aspects of these practices in which learners encounter obstacles.

> 3) Characterizing scaffolding guidelines that specify ways that tools can alter the task to address the obstacles by helping make tasks more tractable and productive for learners.

Our work naturally builds on these guidelines. By focusing the design of EcoSurvey around the practices outlined in learning science literature on modeling in the classroom, we naturally connect to the real practices of the task. Our analytics are designed to automatically measure how successful students are at engaging with these practices and the points at which students run into difficulty. We have used this feedback in design, and our results show that students are using the revised modeling tool to create more complete models of their local ecosystem.

3.3 Context

EcoSurvey was developed as part of a larger collaborative design-based research project called the Inquiry Hub, which is focused on supporting teachers in developing student-centered approaches to curriculum and teaching [77], [67]. Inquiry Hub Biology is a digital high school biology curriculum developed in partnership with a large urban school district in the midwestern United States. Within the ecosystems unit of this curriculum, students are asked to choose a tree to plant on their school grounds or other designated site that will improve their local ecosystem's biodiversity and resilience. Classes use EcoSurvey to create a collective model of their local ecosystem. They use these models to provide evidence and construct arguments to support their choice about the type of tree they choose to plant. The recommended type of tree is then planted on the site, in collaboration with the local Parks and Recreation Department, based on the students' arguments and evidence. Thus, the models students create using EcoSurvey support them to construct arguments with real world consequences. To illustrate the use of EcoSurvey within this context, we follow the experience of Maria, a fictional student in Ms. Smith's 3rd period class.

3.3.1 Data Collection and Creating the Model

Ms. Smith instructs students to map the ecosystem within a selected site on their school grounds or in the local area, taking pictures and making field notes on the organisms and interactions between organisms that they observe. Maria's group makes observations along the creek that runs next to the school. They find a lady beetle, a honey locust tree, some mushrooms, a gray squirrel, and a few other organisms. Using their smartphones, they take pictures of these organisms and upload them to EcoSurvey, creating a "card" for each organism while out in the field. Each card automatically captures information about the date, time, and location of the observation being recorded. Cards also include a "relations" field to capture interactions between organisms and information about the organism's role in the ecosystem. Students begin entering this information as they observe it in the field, and then continue to augment this information back in the classroom through additional research. In figure 3.1, we see Maria's lady beetle card under construction. While in the field, she created the card, uploaded a picture, and added details about interactions they saw. At the same time, her team members are also creating cards for other organisms they are observing.

3.3.2 Evaluating the Model

As students create cards, their organisms are added to a shared class "survey". The survey view shows all of the organism cards and their detailed information, ordered by how recently they were edited. Maria can see that her classmates have created many cards, including a Blue Jay card (figure 3.2).

Ms. Smith organizes the student groups into pairs and asks each group to review the other's cards for correctness and completeness. Maria's group is paired with Group 2, who completed several cards. Andre, a member of Group 2, asks Maria to first review the blue jay card he created. Maria uses the search feature of the survey view to quickly find the blue jay among the cards. She notices that this card is missing many details, including interactions with other organisms.

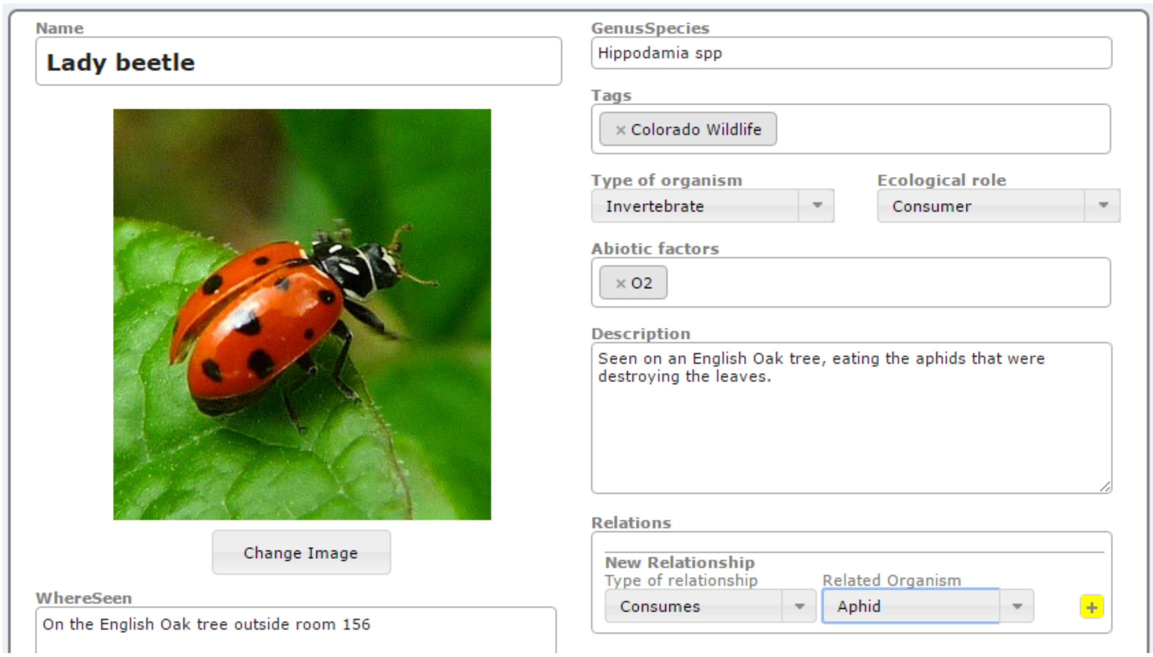


Figure 3.1: The edit view for Maria's Lady Beetle card.

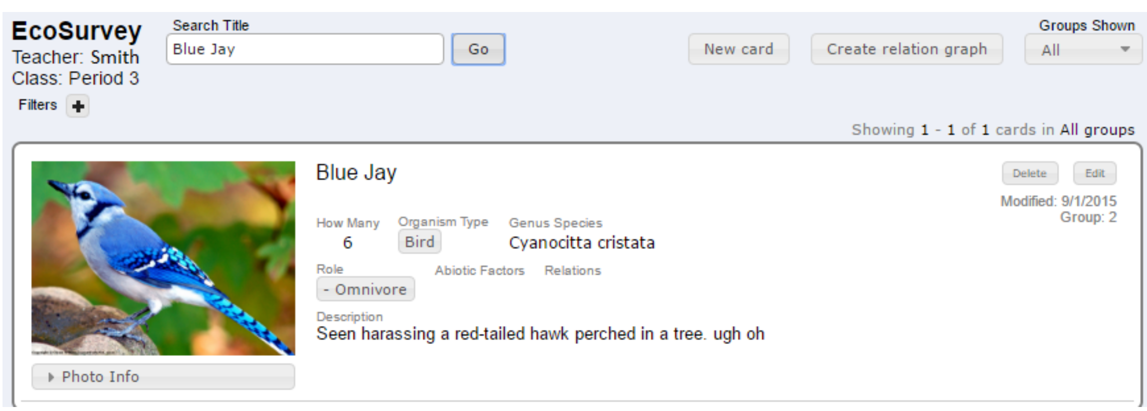


Figure 3.2: The main view of Ms. Smith's class survey.

3.3.3 Revising the Model

Maria recommends that Group 2 do further research into how the blue jay contributes to the local ecosystem. She also takes the chance to update her group's honey locust card. She discovered that blue jays nest in honey locust trees during her earlier research. She didn't realize that their school ecosystem included blue jays until she reviewed the work of her classmates, as her group did not see one. Once Maria has completed editing her group's cards, she continues her review of Group 2's cards. She uses the group select function to view only the cards created by members of Group 2.

Group 2 notices that two people in Maria's group created duplicate lady beetle cards. Maria decides to add her lady beetle information to the other card, since it is more detailed, and uses the delete function to remove her lady beetle card from the model.

3.3.4 Iterating the Model

In reviewing Group 2 cards, Maria sees a card for geese, but notices that the group did not add a predatory relationship to grass, even though she observed geese eat the grass on the soccer field. She uses the search functionality and discovers that no one in class created a card to document grass as an observed organism. Maria adds a new card for grass and includes a predatory-prey relationship with geese. By cycling back through earlier modeling practices (creating new cards), Maria is iteratively improving the class model to be more complete and accurate.

3.3.5 Using the Model

Once the class has created a robust model of their local ecosystem, students use this model to construct arguments for choosing a particular tree to plant. Maria presses the "create relation graph" button, which generates the graph representation of the model and exports it to a digital graphing tool(figure 3.3). Maria and her team study the resulting diagram that enables them to visualize the relationships (links) between all the organisms (nodes) they have cataloged. It is clear from looking at her graph that the English Oak trees are an important keystone species in their site, involved in a large number of relationships with a wide variety of organisms. The geospatial locations in the observational data indicate that there are only two English Oak trees located in their site; Maria and her group recommend planting an additional tree of this type.

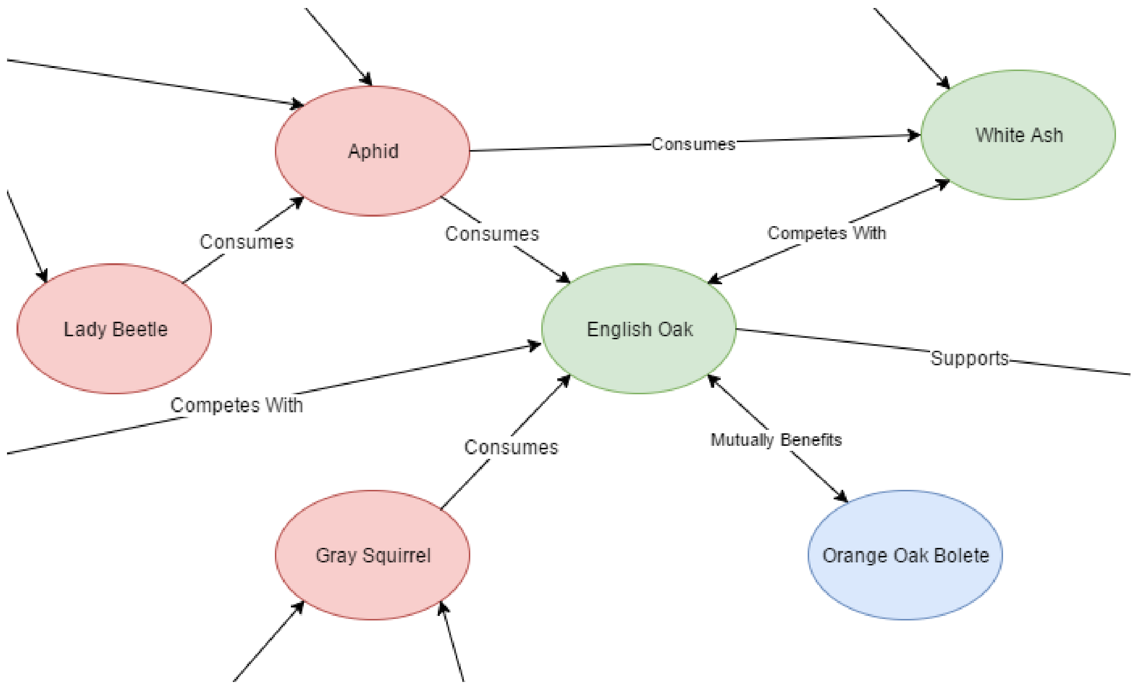


Figure 3.3: A section of Maria's final graph.

Maria's scenario illustrates how EcoSurvey supports students to engage in the practices of creating, evaluating, revising, iterating, and using models. To use a learning analytics approach to study modeling practices, we must map specific actions, or sequences of actions, taken in the EcoSurvey interface to specific modeling practices. Table 1 describes the mapping between modeling practices and specific EcoSurvey interface actions that we use in our analyses. As students interact with EcoSurvey, the system captures and logs each of the actions shown in table 3.1. Each log entry includes the time, user, survey, and action type.

3.3.6 EcoSurvey Design Iteration

Consistent with a design-based research approach, we are iteratively improving the design of the Ecosurvey tool and the supporting curriculum after each field deployment. Classroom observa-

Table 3.1: EcoSurvey Actions

Modeling Practice	Description	EcoSurvey Actions
Create Model	Create a new entry in the model	New Card
Evaluate Model	Explore the organisms and interactions in the current model	Group Select, Search
Revise Model	Edit or delete organisms and interactions included in the current model	Edit, Delete
Use Model	Export a representation of the model for use (e.g. constructing an argument)	Generate Graph, Download
Iterate	Cycle between creation, revision, and use practices	New Card, Edit, Delete, Generate Graph, Download

tions, feedback from users, and analysis of the usage patterns from the first version drove several important changes.

The foremost change is the redesign of the survey view, incorporating the graph representation of the model into the students main workflow as seen in figure 3.4. This view presents the model as a collection of components (organisms) and interactions (relationships). This development grew from results from our first deployment that students failed to engage with relationships for many organisms in their models, and many classrooms showed limited engagement with exporting their models to the graph view. Therefore, we designed the graph layout to emphasize the relationships between organisms, naturally promoting the task of adding relationships to disconnected cards.

The second change we made was to the types of relationships that could be added as seen in figure 3.5. In the first version, the relationship field was open and would accept any response. This led to a wide variety of responses, many of which did not accurately reflect possible relationships (e.g. brown trout is not a relationship type). While we normalized the relationships for analysis (as discussed below), this still left a large number of unknown relationships. By using a closed list of relationship options based on language from the science standards, version two of EcoSurvey scaffolds student model development and scientific understanding by driving them to consider how their intuitive representation of the relationship maps to the terms used by scholars in the field.

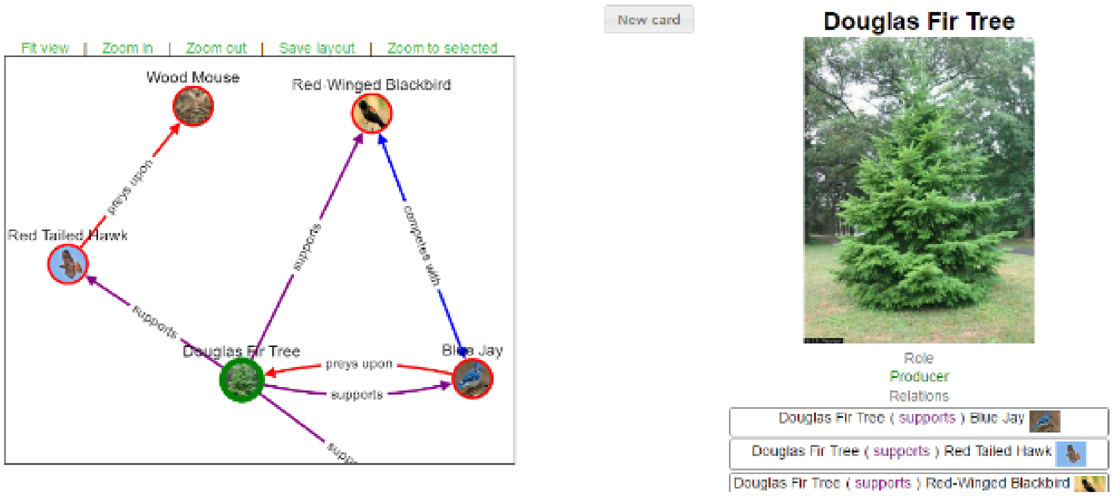


Figure 3.4: The graph view in version two

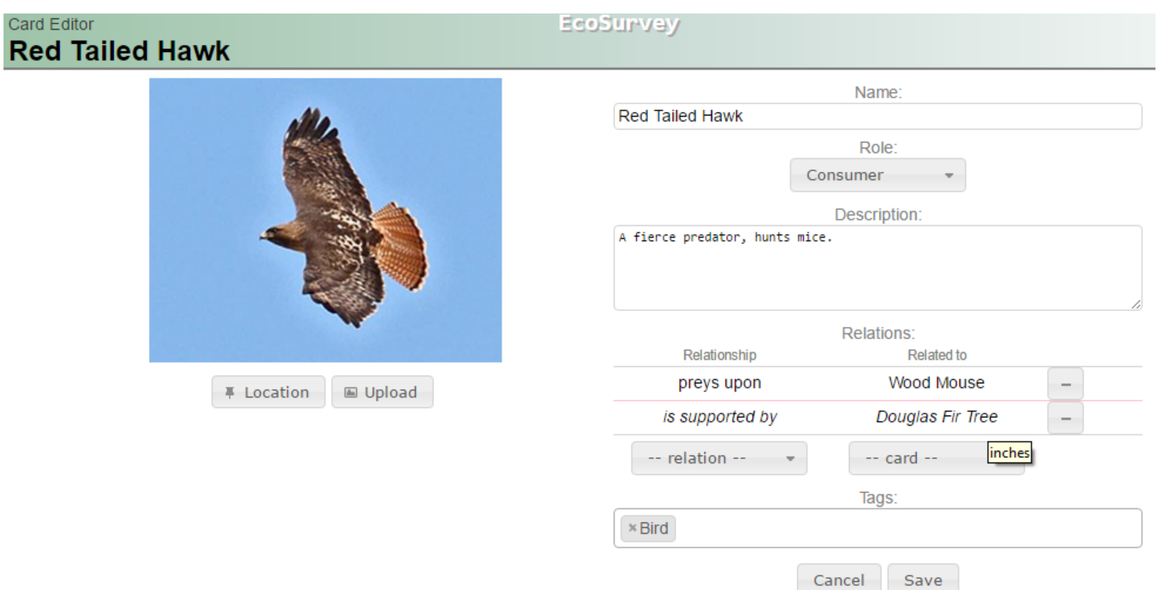


Figure 3.5: The card edit view in version two

3.4 Methods

Here, we describe data used in our analyses as well as the specific analytic techniques used to answer each of our three research questions.

This study was reviewed and approved by the University of Colorado Boulder Institutional Review Board, and all activities were conducted according to their rules and guidelines. Teachers gave written informed consent, and all student data were collected anonymously in the course of normal classroom activities. The University of Colorado Boulder Institutional Review Board waived the need for written informed consent to be obtained from the students' parents/legal guardians. All teachers names are pseudonyms.

3.4.1 Study Data

The work presented here builds on two deployments, one of each version of EcoSurvey. Both deployments took place in the same district and with the same professional development routines. However, the data we have analyzed for each deployment used a separate cadre of teachers, which allows us to avoid effects due to previous experience with using EcoSurvey in the classroom.

For our first deployment, EcoSurvey usage log data was collected across 10 high school classrooms during Fall 2015. A total of 9 models were created, which included 586 organism cards and 545 interactions. Our second deployment in Fall 2016 featured final models from students in 35 classes across 11 teachers. These models included 4,136 organism cards and 4,701 interactions.

The deployment of the first version of EcoSurvey also incorporated activity logging. From the sample, we recorded actions for 204 students, generating 3160 action logs, while 58 students did not record any activity. All classrooms in both samples followed a 3:1 device deployment where three students used one laptop together; thus it is not surprising that there are students with no recorded activity.

3.4.2 Model Complexity Analysis

Our first research question examines variation within student models, focusing specifically on the richness of students' models in terms of the number of organisms and their relationships. We analyze the relative number of organisms and interactions within each class survey. We also look at the balance of interactions per organism by evaluating both the average number of interactions per organism and variance in the distribution of interactions. Examining variance allows us to distinguish different patterns in the assignment of interactions to organisms. Some classes may create models where most organisms have a similar number of interactions, while other classes may create models where only a few organisms have been assigned many interactions.

We also analyze the distribution of relationship types using evenness. This measure considers how each type of relationship is represented within the survey. We calculated evenness using the shannon index, the same formula for species evenness in the study of ecosystems [78]. The shannon index gives an evenness score from zero to one. A survey with an equal number of relationships of each type would have a perfect evenness score of one. Conversely, a survey with many predator-prey relationships and few others would have a low program type evenness score. The shannon index is calculated using the following formula:

$$J' = \frac{H'}{H'_{MAX}} \tag{3.1a}$$

where

$$H' = -\sum_{i=1}^{N} P_i ln(P_i) \tag{3.1b}$$

where p_i = proportion of relationships of type i in the surveys and

$$H'_{MAX} = -\sum_{i=1}^{N} ln(S_i) \tag{3.1c}$$

where S_i = total number of relationships in the surveys

3.4.3 Influence of Tool Design on Models

Our second research question seeks to understand how design changes in digital modeling tools can have an impact on students' models. These measures are important for the iterative process of our design-based research approach, providing evidence of what impact the design changes have on students' models.

To evaluate the impacts of design, we aim to compare directly across deployments of different versions of EcoSurvey. We run the same statistical comparisons for each version and compare across conditions. In cases where direct comparison of means and variance is possible, we use a standard Student's t-test [82] to determine significance.

3.4.4 Variation in Modeling Practices

Our last research question examines variation in student modeling practices, focusing on action variety, frequency, and iteration. Action variety refers to the range of actions a student performed. For example, some students may have only created and edited cards, while others may have used the full range of EcoSurvey actions. Frequency refers to the total number of actions completed by an individual student and the number of usage sessions they engaged in. Sessions are defined by a series of actions from a single user without a large break in activity (greater than two hours). Defining a session using a two hour gap allows for any student activity within a long class period to occur within one session; several of our classrooms employ 1.5 hours block periods.

To characterize iteration practices, we look for evidence of design cycles within the log information. Design cycles can be recognized when students engage in multiple sequences of construct-revise-use practices. This focus on a sequence of practices is consistent with prior work (e.g. [76]), which characterized modeling practices as a series of steps. By extension, a design cycle consists of returning to a previous modeling step after moving on in the sequence (e.g. creating a new card after editing a different card). We counted the number of cycles as a measure of iteration.

Combined, these three metrics - action variety, frequency, and iteration - yield an eight feature vector for each student consisting of total number of EcoSurvey actions, total number of create actions, total number of evaluate actions, total number of revise actions, total number of use actions, total number of EcoSurvey action types taken, number of sessions, and number of iterations. We combined the feature vectors for students with the same teacher, and performed a Kruskal-Wallis H test ([42]) for each feature to determine differences between teachers. A Kruskal-Wallis H test is a non-parametric adaptation of an ANOVA to compare samples of different sizes, as we have in our groups. We further explored these differences using Tukey's HSD test [84] to test the significance of pairwise differences between teachers.

3.4.5 Predictive Value of Modeling Practices

This understanding of modeling practice allows us to characterize the variation in student activity in EcoSurvey. To expand on this characterization, we examine the degree to which we can use sequences of student modeling actions to predict that student's teacher. We plan to use this prediction in a support system for students and teachers embedded within the tool.

For this prediction task, we use the previously described features of variety, frequency, and iteration as well as automatically extracted sequence patterns. In our work, a sequence pattern consists of a series of EcoSurvey actions (e.g. "New Card", "Edit", "Generate Graph") embedded within a student's complete action log. To extract sequence patterns, we used the Colibri Core [86] software package. This software package, originally designed for natural language processing tasks, treats every action as a token and determines the frequency of consecutive token sequences (n-grams) from student usage logs. These token sequences can include wildcard actions (skip-grams). For instance, the software will extract the sequence "New Card", "Edit", "Generate Graph" as either an n-gram or as the skip-gram "New Card", {*}, "Generate Graph". This skip-gram will capture similar sequence patterns, where one action occurs between New Card and Generate Graph actions. This yielded 2,893 unique sequence patterns, that occurred at least three times, across all student usage logs. Once we extracted these sequence patterns, we used them as a new series of features to augment each student's existing feature vector.

To understand which features that characterize a student's modeling actions are most predictive of his or her teacher, we input subsets of each student's feature vector into four Naive Bayes classifiers using Weka [22]. The first classifier used the eight features related to variety, frequency, and iteration of actions. The second classifier used the full set of sequence pattern extracted by Colibri Core for each student. The third classifier implemented a best-first search [21], which automatically reduced the full set of sequence patterns to the eighteen most predictive features. The last classifier combines the eight variety, frequency, and iteration features with the eighteen most predictive sequence patterns. Each test was run using 10-fold cross validation.

3.5 Results

Our results are divided into sections based on the type of analysis performed. In the first two sections, results are further broken up by deployment version, allowing us to present each set of results independently and then discuss how they relate to our second question about the impact of design on student models.

3.5.1 Model Complexity Analysis

3.5.1.1 Version 1

As shown in table 3.2, there are substantial variations in the models created by students in different classrooms. We see that Anderson's students documented many more organisms (155) and interactions (264) than all other classes. Though Anderson had both of her classes work together to create one survey, the total number of students contributing to this model is comparable to the number of students contributing in other classrooms. We also see that students in Baker's three classes each documented significantly fewer organisms and interactions. One class only documented 19 organisms (less than one per student) and did not document any interactions. Chavez's classes exhibit wide variation, particularly in the numbers of interactions documented by each class.

The number of interactions per organism, a broad measure of model complexity, further

Table 3.2: Version 1 Final Models

Survey	Organisms	Interactions	Int. Per Org.	Int. Var.
Anderson 4/7	155	264	1.7	4.35
Baker 1	47	7	0.149	0.297
Baker 2	25	5	0.2	0.24
Baker 4	19	0	0	0
Chavez 1	88	70	0.795	0.663
Chavez 2	45	27	0.6	1.31
Chavez 6	60	57	0.95	3.78
Chavez 7	81	82	1.012	5.72
Chavez 8	66	33	0.5	0.826
Averages	65.11111111	60.55555556	0.6562222222	1.909555556

illustrates apparent classroom differences, with Anderson's class creating more complex models than Baker's and Chavez's classes. To better understand classroom differences, we examine variance in the number of interactions per organism. In Anderson's class, we see a high variance in comparison to the interactions per organism metric, which indicates that there are a small number of organisms with lots of interactions and many organisms with few interactions.

Chavez' P1 and P7 classes provide a particularly interesting case to examine this variation. On reviewing table 3.2, we see that the variance in the number of interactions assigned to each organism is significantly lower in P1 than in P7, while the actual number of organisms and interactions are comparable. Further analyses reveal that students in Chavez's P1 did not assign any interactions for 39% of their organisms, while students in P7 did not assign interactions to 74% of their organisms. A similar analysis revealed that 42% of the organisms documented in Anderson's model did not include interactions. In most classes, the majority of organisms have no documented interactions. It appears that students engaged significantly more with describing organisms, and spent far less time consistently documenting interactions.

3.5.1.2 Version 2

The models generated during the second deployment of EcoSurvey demonstrated both parallels and contrasts. First and foremost, the average number of organisms and interactions are both

Table 3.3: Version 2 Final Models

Survey	Organisms	Interactions	Interactions Per Organism	Interaction Variance
MIN - Jaques 1	6	1	0.166666667	0.138889
MAX - Lin 1	70	189	4.5	78.67857
Averages	118.1714286	134.3142857	1.136605416	6.539478711

higher in our second deployment group (as seen in table 3.3), with the number of organisms showing a trend towards significance (P = 0.09). Furthermore, the ratio of relationships per organism tends to be significantly higher, demonstrating significantly more complex models.

The variance in relationships per organism once again also leads to some interesting results. The high levels of variance across classes highlights continued imbalance in the distribution of assigned relationships; students are once again focusing on key cards when creating relationships. However, the wide range of percentages of cards without relationships (17% to 88%) demonstrates that classes are engaging with relationships in different patterns. Nevertheless, the average percentage of orphaned cards (50%) is lower than the first deployment.

3.5.2 Relationship Analysis

3.5.2.1 Version 1

Students did successfully engage with adding relationships to their models in the first version of EcoSurvey. However, analysis (as seen in figure 3.6) did suggest several trends of use that did motivated changes to the relationship system in EcoSurveys design. First and foremost, there were a large number of unknown relationships that could not be normalized. This trend represents the ambiguity in students models, which leads to an inability to construct explanations and arguments using the model. Second, we see an imbalance of relationship types (evenness = 0.749), with a particular emphasis towards predator-prey connections (55.8%). This indicates that students were not engaging deeply with exploring other types of relationships, particularly competitive relationships (5.7%) that are important for maintaining organism balance. Finally, this analysis emphasized that important bidirectional connections between organisms, such as symbiotic relationships of mutual beneficence, were not properly incorporated into EcoSurvey.

3.5.2.2 Version 2

Version two showed remarkable improvement in the evenness of relationship types (figure 3.7, evenness = 0.803), even with the added complexity of a new mutually benefits relationship type. In particular, we see a remarkable decrease in the relative abundance of predator-prey relationships (down 26.5%) and a substantial increase in the use of all other valid relationship types. In addition, though version two incorporated the ability to denote unknown relationships, this feature saw very little use (1.1%).

3.5.3 Practices Analysis

3.5.3.1 Analysis of Teacher Differences

There are significant differences between the student action sequences of our three teachers on all eight metrics related to variety, frequency, and iteration ($p < .001$). Our Tukey's HSD test for each feature shows that the three groups are each distinct to a significant degree in Create, Revision, and Iteration frequency (figure 3.8a, $p < .05$), as well as Overall Actions, Session Count, and Action Variety (figure 3.8b, $p < .05$). We also see Anderson's students performed significantly more Evaluate and Use actions than the other two teachers' students (figure 3.8a, $p < .05$), though the differences between Baker's and Chavez's students are not significant. Anderson's class also used EcoSurvey twice as much, as measured by session counts. Overall, Anderson's students engaged in more modeling practices than both of the other two groups, and Chavez's students engaged in more modeling practices than Baker's.

There were also differences in the modeling practices that students employed. Students in Baker's classes rarely engaged in three of the five modeling practices we are studying: revisions, iteration, or use. Chavez's class engaged with four of the five practices, but appeared to rarely use their models.

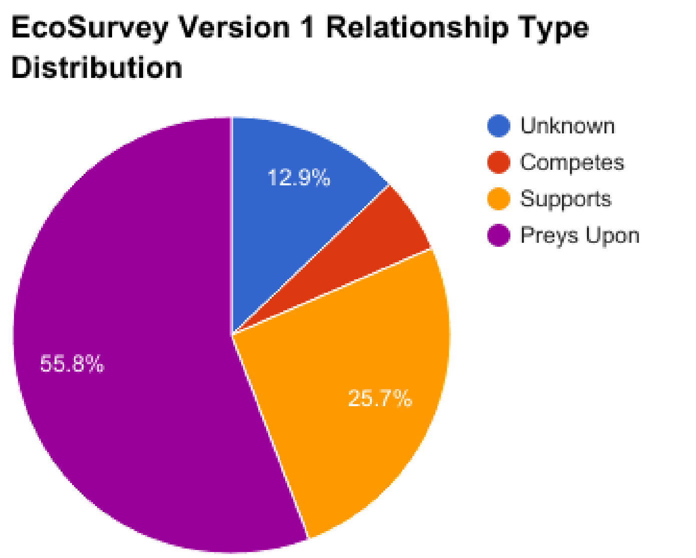


Figure 3.6: The relationship type distribution for Version 1

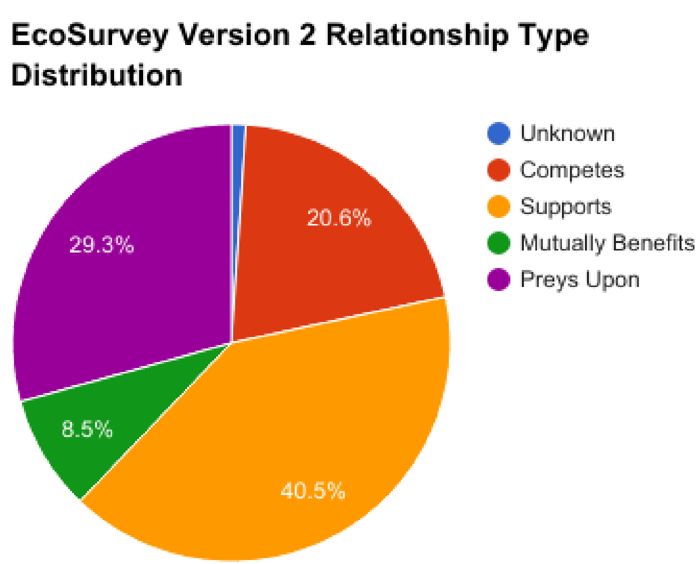
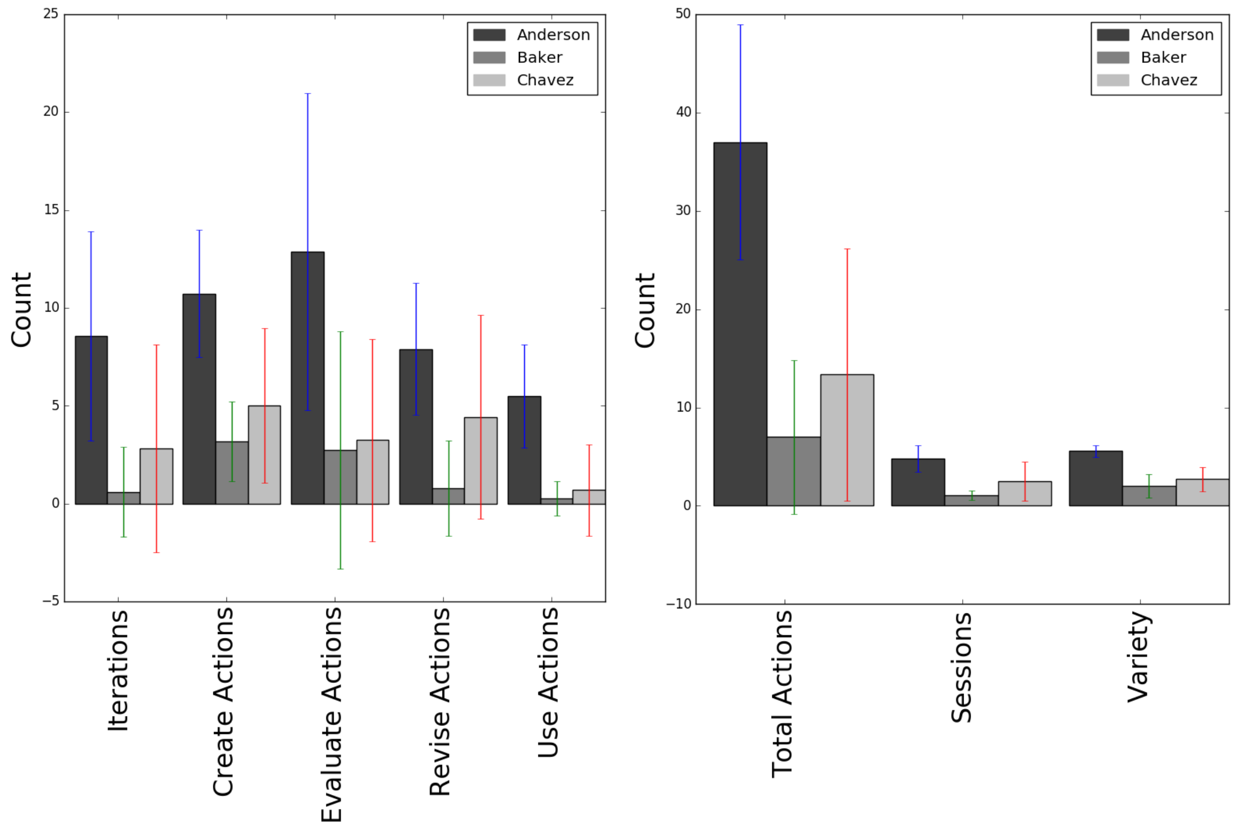


Figure 3.7: The relationship type distribution for Version 2



(a) The average number of actions by modeling practice type.

(b) The average number of actions, types of actions, and action sessions.

Figure 3.8: Student modeling practices for each teacher's students.

3.5.3.2 Predictive Value of Practices

As shown in table 3.4, student action sequences can predict their teacher with varying degrees of reliability depending upon the features used. Our baseline assumes that each student is in one of Chavez's classes; almost 52% of the students in this study were in one of his classes. All of the feature sets we studied improved performance over the baseline. Classifying based on all 2,893 sequence patterns improved our classification accuracy by almost 12%, whereas classifying solely based on our variety, frequency, and iteration features improved performance by over 15%. We also trained a model on the best sequence patterns, that is, the 18 most predictive patterns identified by Weka's Attribute Selection tool [21]; this yielded a nearly 25% improvement in performance. The best performing model was one that combined the most predictive sequence patterns with our variety, frequency, and iteration features. This combination resulted in a 30% improvement over baseline, correctly predicting a student's teacher 80% of the time.

The most useful features for classification accuracy are the 18 ”best” sequence patterns (table 3.5). A closer examination reveals that these sequence patterns correspond to our five modeling practices in interesting ways. These patterns prioritize model revision, evaluation, and iteration as distinguishing features, which correspond to the differences in classroom modeling practices discussed under research question 2.

To better understand the types of errors that our best performing model makes, we generated a confusion matrix (table 3.6). We see that 75% of the errors are due to the misclassification of 30 of Chavez's students as Baker's students. One possible reason for this misclassification is that some

Table 3.4: Predictive accuracy of each action sequence feature set

Feature Set	# Attributes	Naive Bayes Acc
Baseline	0	51.96%
All Sequence Patterns	2,893	63.73%
Variety, Frequency, and Iteration Features	4	67.65%
Best Sequence Patterns	18	75.00%
Combined Features	22	80.39%

Table 3.5: The most predictive action sequences.

New card, New card, {*}[1], New card, {*}, {*}, New card	New card, {*}, Group Select, {*}, New card	Group Select
Group Select, {*}, Group Select	Group Select, {*}, {*}, {*}, Group Select	Group Select, {*}, New card, {*}, New card
Group Select, Search	Search, {*}, {*}, {*}, Edit	Edit
Edit, Edit	Edit, {*}, Edit	Edit {*} {*} Edit
Edit, Search	Edit, Generate Graph, Download	Edit, Generate Graph, Download, Edit
Generate Graph	Download	Generate Graph, Download

students in Chavez's classes performed very few modeling actions overall, similarly to the majority of students in Baker's classes.

3.6 Discussion

Overall, these results demonstrate the capabilities of our analytic techniques to help us understand scientific modeling in the classroom. We have been able to discover the variance in students' models, the impact of design features on those models, and the variance in student engagement with modeling practices. These results inform work in learning analytics, modeling tool design, and the design of curricula and professional development for the Next Generation Science Standards.

While the second deployment showed limited evidence of improved student contribution, there is still large variance in the number of organisms and interactions documented at the classroom and teacher level. These differences could be due to a variety of factors, such as the time allocated to modeling during class, the teacher's dispositions and knowledge about scientific modeling, or the teacher's capability to support student use of EcoSurvey. These possibilities can be addressed

Table 3.6: Combined features confusion table.

		Classified As		
		Anderson	Baker	Chavez
Correct Class	Anderson	29	0	1
	Baker	1	64	3
	Chavez	5	30	71

through curriculum and professional development design around supporting student modeling, as well as through the interface design and the inclusion of teacher supports within digital modeling tools like EcoSurvey.

Our analysis of student models also revealed a disturbing similarity across all classrooms and teachers: all the models contained significant percentages of organisms that did not have a single defined interaction with another organism. Thus, these student models are missing critical elements of a complete and sound ecosystem model. It is unlikely that these models can support students to develop comprehensive explanations and predictions as called out in the *Framework* ([58]). There are multiple possible explanations for these behaviors, including weaknesses in the Inquiry Hub curriculum, the associated teacher professional development, or the design of the EcoSurvey tool.

In developing the second version of EcoSurvey, we made key design changes that we hypothesized would improve student models. As a first step, we made major changes in designing EcoSurvey version two to make it easier for students to establish relationships from multiple parts of the interface, to visualize established relationships through an integrated graph view, and to see which organisms are not connected to others in the model. In the second version, we did see gains in the level of completeness and complexity of students models, as well as a more even distribution of relations mapped in the system. However, these changes have only slightly reduced the isolated organism phenomenon. This result suggests that further mechanisms will be necessary to address these issues. Our planned approach is to incorporate adaptive feedback mechanisms within EcoSurvey, providing scaffolds for struggling students.

The large variance we observed in student modeling practices provides evidence of significant teacher-level differences. Clearly, these teachers are implementing EcoSurvey and the corresponding lessons differently in their classrooms, with wildly varying results. When teachers devoted more time to modeling, as measured by sessions, their students' engaged in a richer variety of modeling practices. Prior research suggests that there is a linkage between student engagement in modeling practices and future learning outcomes ([76, 3]). Thus, it appears that students in several of our participating classrooms lacked critical opportunities to learn ([52, 55]), that could ultimately impact their academic performance. In future work, we plan to examine the relationships between student engagement in modeling practices and their learning outcomes as measured by end-of-course school district assessments.

Our predictive analysis provided further evidence of significant teacher-level differences. The feature selection algorithm honed in on the presence or absence of three modeling practices - evaluation, revision, and iteration - as the features that best predicted a student's teacher. This suggests that future professional development and curriculum design should focus on these specific practices, ensuring that all students get an opportunity to participate in these parts of the modeling process. The most accurate classifier also benefited from additional features characterizing action variety, frequency (number of actions), and iteration. These features further highlight differences in student engagement, with some students missing the opportunity to explore, develop, and use their models over time.

In EcoSurvey version 2, we expanded features designed to support evaluation, revision, and iteration practices. By facilitating students to use (visualize) their models more frequently, we hope that this will prompt them to notice shortcomings and engage in modeling practices that were previously underutilized. A parallel clickstream analysis of our redesigned interface is a necessary next step in our future research.

While this study yielded many results that have informed our partnership design work, there are several limitations that are important to note. First, we cannot attribute our observed variation in models and modeling practices to student-level differences, due to the shared and collaborative nature of the deployment. All our participating classrooms asked students to work in groups and each group shared a single laptop computer; we are actually observing the collaborative modeling practices of small groups rather than individual students. Second, our practices analysis is only available for students in our first deployment cycle. This limited set shows potential for capturing differences in modeling engagement, but further data collection is required to explore the generalizability of these findings.

While our technique is designed to generalize across tools, our investigations thus far have only explored student use of EcoSurvey, limiting our ability to generalize our findings. Nevertheless, a core aspect of our analytic approach explicitly linked specific user interface actions in the EcoSurvey tool to individual modeling practices identified through prior research: creating, evaluating, revising, using, and iterating ([76, 3, 28]). This approach enabled us to work with theoretically and empirically sound features identified through prior classroom research. And, this approach enabled us to interpret the action sequences identified as salient by our algorithms in a theoretically-informed way, enabling us to link our findings back to instructional concerns, such as curriculum design and professional development. This method of linking interface actions to identified modeling practices can support generalizing this analytic approach to other tools that support scientific modeling, such as Model-It! ([31]), Dragoon ([87]), or activities within the Wallcology unit ([48]).

3.7 Conclusion

In this study, we demonstrated the utility of learning analytic methods for characterizing variation in students' scientific models and their modeling practices. We also showed that an individual student's modelling action sequences can be used to predict his or her teacher. Our results support Windschitl et al's findings documenting large variations in how teachers implement modeling in their classrooms [90]. While we did not conduct formalized classroom observations, our analysis revealed profound, quantifiable differences in the models that students constructed across different classrooms and significant differences in their classroom learning experiences as depicted in the range of modeling practices that they engaged in. This result confirms and expands upon the conclusions of [32] that modeling is handled differently across classrooms, but provides evidence that the variance is not only attributable to the teacher. The variance in model complexity within each teacher shows that student and class level variance can sometimes have a higher impact than teacher level variance.

One important aspect we plan to address in future work is the impact of modeling activities on student learning. Our team has been developing assessments to embed three dimensional assessments [9] within the ecosystems curriculum. Within this body of questions, we have designed prompts to elicit student understandings of modeling as a science and engineering practice as well as a cross-cutting concept, allowing us to measure student development of these skills while using EcoSurvey and the accompanying curriculum. In addition, we have developed protocols to evaluate students' final reports related to the unit-level challenge of choosing a tree to plant on their school grounds. By analyzing how students incorporate their models of the local ecosystem into their final choice, we can measure the impact of EcoSurvey on students' explanations of ecosystem phenomena [58].

We are also incorporating these findings into the next iteration of design and deployment of our modeling tool. Our biggest improvement is to provide these analytics in real-time feedback systems within EcoSurvey. We plan to work with pre-service and active teachers to design interfaces that support the needs of students in successfully developing complete models of their ecosystem, as well as interfaces to support teachers in understanding the activity and contributions of students towards their models.

3.8 Author Contributions

All the authors contributed to the design and development of EcoSurvey tools and analytics. David Quigley took the lead on writing, but all the authors took part in writing and reviewing content.

3.9 Acknowledgements

This work builds on [70]. This material is based in part upon work supported by the National Science Foundation under Grant Numbers 1555550 and 1147590. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. In addition, the authors would like to thank the Gordon and Betty Moore Foundation for their support.

Chapter 4

Understanding the impact of a scientific modeling tool on student engagement and learning

This chapter has been written to target submission to the Journal of Learning Analytics for their Human-Centered Analytics special issue this Fall. This report presents the data from the third year of EcoSurvey deployment, exploring the limits of my model and modeling practices characterization (RQ1). These results also expand our iterative deployment analysis (RQ2), finding ways in which students' modeling activities are influenced by tool design. This study also presents the first results on connections between student understanding of scientific modeling and their engagement with modeling practices (RQ3).

4.1 Introduction

Contemporary understandings of science have shifted away from considering each scientific discipline to be a siloed collection of facts and have instead begun to accept a more integrated perspective. This approach is particularly prevalent in science education, and is summarized in the *Framework for K-12 Education* [58] as three-dimensional (3D) science. The Framework describes these dimensions as 1) disciplinary core ideas, 2) science and engineering practices, and 3) cross-cutting concepts. However, while the perspectives of science educators, particularly researchers in the field, have been accepting this approach, the field has struggled with realizing this vision. This struggle is particularly evident when it comes to determining how well a learner has understood these multi-dimensional features, especially the implementation of science and engineering practices naturally embedded within other aspects of learning.

In addition to the expanding complexity of approaches for the learning of science, the tools we use to support learning have been growing. The advent of digital learning tools, including learning management systems [12] and intelligent tutors [18] have been helping learners to better track their knowledge and provide them with optimized supports and trajectories for learning new information. However, these tools often use a prescribed approach to gauge understanding of a core concept and working through a structured process, rather than providing an open space for exploration and development.

Each of these areas of research has made significant progress in supporting students' learning. Still, there is a pressing need to integrate digital supports in open exploration to support 3D science education. By understanding the ways in which students use open-ended digital tools to implement science and engineering practices, we can better gauge how to support students.

Our research examines student use of EcoSurvey [70], a tool used primarily in high school biology classrooms to support the science and engineering practice of modeling. Students use EcoSurvey to construct models of the components and interactions in the local ecosystems. This tool has been developed as part of a larger curriculum development project aimed at supporting 3D science teaching & learning [77]. Our team has undergone an iterative deployment process over three successive school years, allowing us to compare usage metrics and final models across versions of the tool [69]. We have also recently integrated real-time feedback into the tool, which provides targeted scaffolding and "next step" recommendations to students based on their previous modeling activity.

This longitudinal deployment of EcoSurvey has informed three research questions:

1) How do variations in tool designs impact students' final models and engagement with modeling practices?

2) What effect does the feedback mechanism have on student modeling practices?

3) What do students understand about scientific modeling after use of digital modeling tools?

Overall, we have found that while there was no significant difference in the size of students models, they had significantly higher engagement with several modeling practices when the tool incorporated new features. In addition, while there was limited use of the feedback mechanism, this feature did often have a positive impact on student activity when used. Finally, we conducted short student interviews in order to document students understanding and experiences of scientific modeling. Through these interviews, and connecting them with students modeling activity, we have determined that students understanding of scientific modeling can be predicted to some degree by their levels of engagement with modeling.

4.2 Background

4.2.1 Measuring Equity in Learning Experiences

One of the important approaches to defining improvement across learning sciences is to reduce the variance attributable to the sources inherently present in the education system [4]. By reducing the variance attributable to differences in teacher, school, or population demographics, we can provide a more equitable educational experience. Our objective in this work is to reduce the variance in students use of modeling practices, with a particular focus at reducing effects found at the teacher level. This reduction in variance will promote equity in the learning experience by giving all students opportunities to learn [28].

In order to draw conclusions about the effectiveness of an intervention, it is important to understand uptake. We approach uptake as how a learner responds to feedback, correction, or recommendations, as demonstrated by their actions following the moment of intervention [49]. This perspective is realized in EcoSurvey by looking for students to immediately engage with the modeling practice recommended by the system.

4.2.2 Student Scientific Modeling

Scientific models form a core component of engaging with scientific practice [8]. Furthermore, it is increasingly recognized that scientific modeling is an important skill to embed in science education [6, 17], and "Developing and Using Models" is a core science engineering practice in the *Framework for K-12 Science Education* [58]. Scientific modeling has been increasingly studied in the K-12 classroom (e.g. [76, 3]). One of the emerging theories around developing scientific modeling skills is to view modeling activities as engagement with a series of practices [76, 3]. These practices focus on the process of modeling, from ideation and conjecture formation to the final application of the model to help address a problem or question. In our work designing and studying tools that support students' engagement in modeling, we focus on creating model elements, reviewing the contents of a model, editing components and interactions, and attempting to use the model as a whole. These practices are primarily contained within a single modeling medium, such as EcoSurvey, rather than occurring in discussion or other settings. In addition, we emphasize the importance of iteration as both a modeling practice [3] and as a practice applicable to larger scientific inquiry activities [2].

Students' understanding of scientific modeling has been studied primarily through interviews and observations [76, 5]. These approaches are important to the field and allow us to gain a deep measure of the impact of intervention on a student's understanding. We utilize these same approaches to evaluate how students understand scientific models after using EcoSurvey. However, our work also emphasizes the inability of these measures to scale beyond the research setting. It is important for us to use scalable measures of student activity and understanding, and to find ways we can support teachers in assessing their students' status for intervention in real time. This need informs our approach of using new analytic techniques grounded in learning sciences frameworks to document students' activity while using digital tools.

4.2.3 Digital Classroom Tools

Many projects have built digital tools to engage students in scientific inquiry tasks, and research has shown that these applications can support and even improve student learning [35, 89, 56, 54]. The vast majority of literature in this area is focused on simulation; either simulating a model of a particular scientific phenomenon [35, 89] or simulating an immersive environment with discoverable scientific phenomena models embedded in the world [56, 54]. These activities are important to building students' understanding of models and modeling, but they do not provide support for the generation of new models. It is also important for students to build their own models, as a means of taking ownership of both the content and outcome. New model development is an important part of modern scientific learning expectations, but more importantly it is a necessary skill for participating in modern science. This is the central approach taken in the design and implementation of EcoSurvey and its accompanying curriculum.

A common thrust of digital learning tools across disciplines is to use feedback to guide students towards fruitful activity. This approach can be implemented in a number of ways, including the intelligent tutor approach to selecting the next example or problem [18], but intelligent tutors requires a prescribed semi-ordered series of activities rather than supporting open-ended student use. Our approach displays usage information and provides a customized recommendation of a next step. This approach builds on the idea of "informating" [92], providing students with information they can use in reflection as well as scaffolding the decision of what to do next without forcing a particular usage pattern. By providing a recommendation rather than automatically sending a student to do the "next step", we leave students with the agency in their model creation process.

4.3 Research Context

This research examines the iterative deployment of EcoSurvey, a digital tool for creating models of the components and interactions in a local ecosystem [70]. Our deployments take placeEcoSurvey exists within the framework of an entire year of high-school biology curriculum created by the Inquiry Hub team [77] and deployed in a large urban school district in the western US. Within this curriculum, students are faced with the engineering design challenge of choosing a tree to plant on their local school grounds. To support their work on this task, students use EcoSurvey to create a model of the existing ecosystem in their area, which allows them to determine which organisms may be impacted by planting a certain kind of tree. This tool is implemented as an online application, allowing access across devices including laptops and chromebooks increasingly found in classrooms along with tablets and smartphones.

We have been using a design-based implementation research [14] strategy to improve EcoSurvey and accompanying curricular materials. EcoSurvey was conceived and developed during design cycle for our ecosystems unit and first deployed in Spring 2015. Since then, our team has made yearly changes in response to teacher and student feedback, and deployed these revisions in the following academic year.

4.3.1 EcoSurvey

EcoSurvey is built around the concept of a "web" or "graph" model, in which components (organisms) are connected by interactions (prey upon, support, mutually benefit, compete), allowing students to visualize the ecosystem as a whole and determine which trees have the best impact. Students work together, building their models in groups to contribute to a shared class understanding of the local ecosystem.

To construct these models, students begin by going out and taking observations in their local ecosystem. They take field notes and pictures of trees, animals, and any organisms they can find. They also include any information they gain on relationships between animals.

Once students have captured these initial data, they then bring the information back to the classroom and input these organisms and relationships as "cards" and "interactions". Students can then review and edit this content as needed, as shown in Figure 4.1.

Students continue to add to and revise their models by accessing additional information, such as their local parks department's reports on the ecosystem or national ecological relationship databases. Once students have constructed an initial model, they can extract these individual elements to create a visual model of the ecosystem as a whole.

Throughout this process, students iterate on their models; when reviewing an organism's relationships, they may notice a missing interaction such as a food source which can inspire the creation of a new card for that organism, or an attempt to use their models can demonstrate to a student that they haven't considered the impact of one of their proposed trees, prompting them to

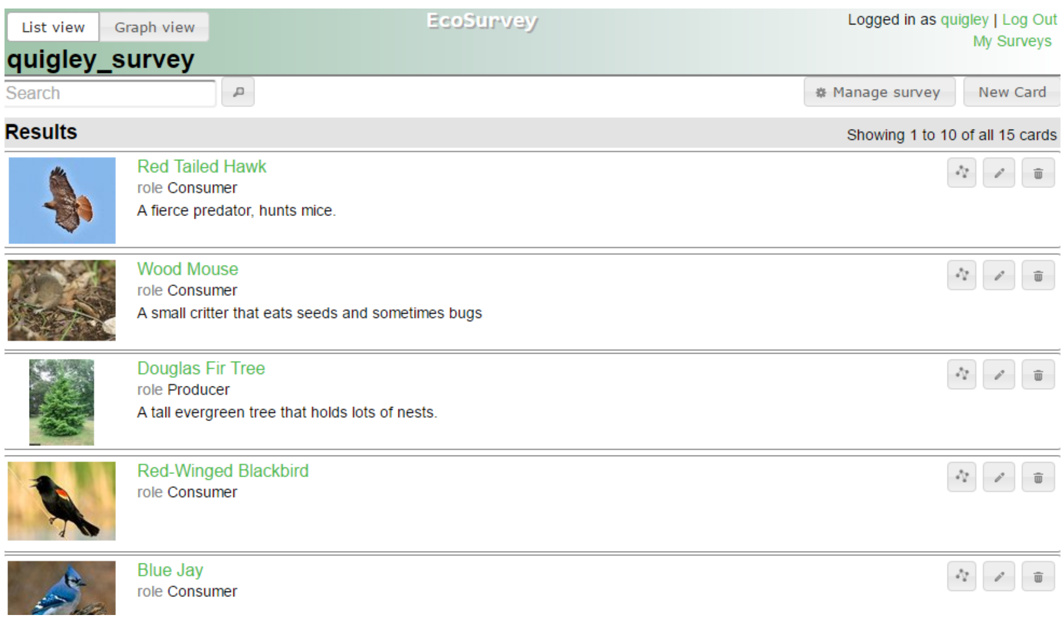


Figure 4.1: The list view of organisms within an ecosystem model.

go back and add, review, and edit that content.

In our prior work, we discussed significant teacher differences in classroom use of EcoSurvey [70]. These analyses led to a few important design decisions aimed at decreasing the variation across teachers and improving students' engagement with EcoSurvey.

4.3.2 EcoSurvey Design Changes

The first focus in our redesign efforts was to address student confusion within the modeling task and to improve student engagement with modeling practices. These approaches were designed to reduce the teacher-level variance in activity by scaffolding student activity.

4.3.2.1 Integrated "Use" Model View

As discussed in our prior work [69], one of our most important changes was to improve the "use" model view, as seen in Figure 4.2. In the first version of EcoSurvey, when students wanted to use their models, an export function allowed them to manipulate their models as graphs within a separate tool. However, our analysis found that uptake of this activity was very much differentiated by teacher, and the external and one-way nature of this process hindered iteration. To address these concerns, we integrated a graph visualization into the core functionality of EcoSurvey with the aim of seeing improved engagement with the "use" modeling practice. In addition, we hoped that this view would help students discover organism cards that have no relationships established ("orphaned" cards) in order to help reduce the percentage of these disconnected elements.

4.3.2.2 Adaptive Feedback system

In our analysis of year one deployment, we developed methods for predicting which student a teacher worked with by analyzing the student's individual sequence of activity with the tool. Since these teacher-level differences were so prominent both in the strength of final models and in the engagement with modeling practices, we treat this teacher prediction as a three-level proxy for measuring modeling success. Building on this idea, we implemented a real-time activity analysis

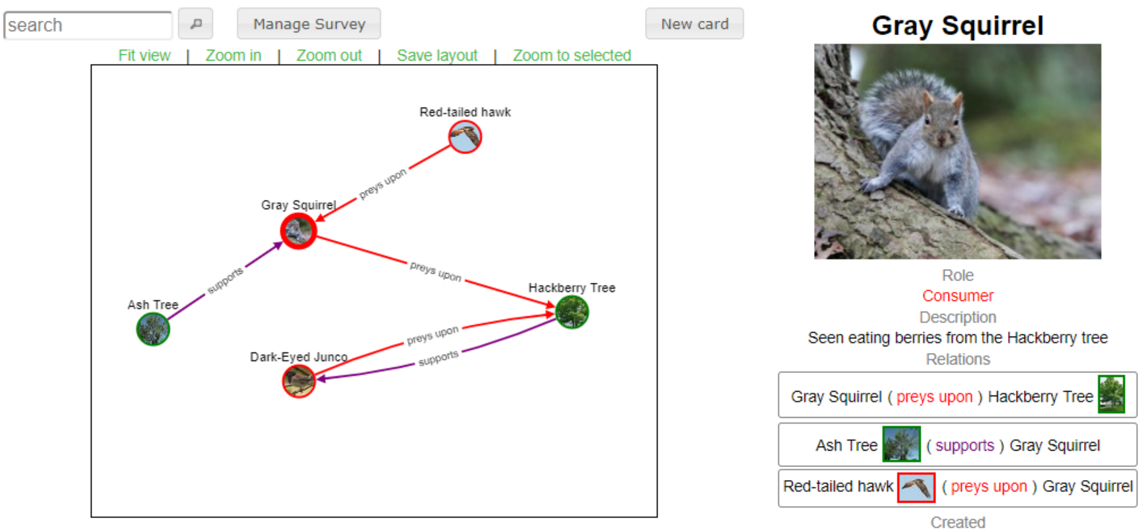


Figure 4.2: The integrated graph view in version 2.

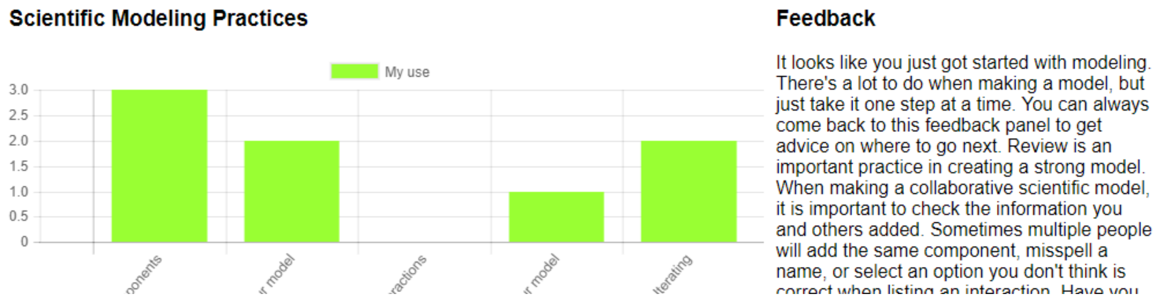


Figure 4.3: The feedback view in version 3.

system to generate feedback in the newest version of EcoSurvey. We used the same classification to categorize active students. This classification determined the level of scaffolded feedback to give to students. In addition, we analyzed the level of engagement with each particular scientific modeling practice to choose a ”next step” recommendation.

To view feedback, students went to a new panel, as seen in Figure 4.3. In this view, students were presented with a chart of their individual activity to date as broken down into practices and a small section of text providing the individualized feedback on how best to continue engaging with EcoSurvey. In Figure 4.3, we see a student who is currently classified as a low-strength modeler, so the feedback incorporates encouragement and additional scaffolds about the recommended next activity.

This feedback mechanism provided students with supports for metacognition in their activity by informing them of their activity up to this point. This feedback is intended to help students reflect on which modeling practices they had not yet engaged with, encouraging them to do tasks they might otherwise have left out. Teachers also cited a lack of insight into exactly what each student did and what support they needed within the tool; this feature could be used by teachers to visualize each student's engagement and provided a start for teacher intervention.

4.4 Methods

This article reports on student usage data from two separate years of deployment. In year 1, as previously presented [69], we collected data from 10 classrooms under 3 teachers. We had 262 students as unique EcoSurvey users in these classrooms. In year 3, we have data collected from 47 classroom groups under 15 teachers. Overall, we had 936 student accounts use EcoSurvey within the year 3 class surveys analyzed here.

We collected students' modeling activity by logging their clickstreams during use of EcoSurvey. In addition, as part of the larger Inquiry Hub project, students were given the opportunity to consent to individual data collection including brief interviews. We used this subset of consenting students, combined with teacher recommendation and student's additional assent at the time of interview, to select participants for a semi-structured interview about their understanding of scientific modeling.

4.4.1 Analysis of Student Models & Activity

We analyzed students' final models according to their size, complexity, and connectedness using the same metrics as presented in our prior work [69], including the number of components (organisms), interactions (relationships), and connectedness of components (orphaned cards). We used Tukey's HSD [84] to determine significance of differences between years.

Our iterative practices analysis process was built around the dual notions of increasing engagement with modeling practices and reducing the teacher-level variance in student modeling activity. To demonstrate changes in students' modeling activity, we used our deployments as different testing conditions and perform statistical comparisons. We used the Wilcoxon-Mann-Whitney test [59] as a group comparison to account for potential non-normality in the distribution of student activity, similar to other research in the area [64]. We used the Hedges G score [23] to determine the effect size of these differences.

It is impossible to perform direct comparison of the variance levels between deployments, due to the differences in the scale of deployment between the two years. However, within each year, the variance in how students' employ modeling practices can be broken apart, separating variance attributable to teachers from other sources. To measure this variance, we used a simple random effects model:

$$< feature >\sim teacher \tag{4.1}$$

This approach allows us to relate the variance due to teacher directly to the residual variance, which represents all other sources of differences between student outcomes. Subsequently, we report the proportion of variance due to teachers as a percentage of the total variance.

4.4.2 Impact of Feedback Systems on Student Activity

As discussed above, the EcoSurvey feedback system is designed to both scaffold the student modeling experience and recommend a specific modeling action that the student should participate in next. The scripted nature of these recommendations allows us to use the student's action sequence up to the moment of feedback to determine which recommended action the student received.

Our feedback impact metrics are designed around detecting and measuring a change in students' modeling activity after the moment of intervention by measuring uptake [49]. This approach focuses on the immediate activity taken by the student after the moment of feedback, looking for use of the recommended modeling practice. To quantify this approach and account for other actions that may be related to the suggested practice (e.g. reviewing your model to determine which organisms to edit), we used a geometric decay model analyzing each occurence of the recommended activity and its proximity to the moment of feedback. Our geometric decay model creates an impact score (S) by assigning a weight to each action from the moment of feedback to the end of the action sequence (n) based on its distance from the start (i) multiplied by a decay factor (d). This action weight is added to the impact score if it is of the recommended type ($a_r = 1$), otherwise it is ignored ($a_r = 0$). This approach can be summarized in the following formula:

$$S = \sum_{i=0}^{n} a_r * i * d \tag{4.2}$$

For this model, we used a decay of .9, allowing for nearby activities to be relatively similar in their impact score. We then binned these scores, assigning each feedback instance to high, low, or no impact categories. We determined .8 served as a proper threshold for high impact from recommendations, allowing for any use of the suggested activity within three actions to be significant as well as supporting a large burst of activity. Scores below .8 but above 0 were assigned low impact, indicating that there may have been some impact from the feedback but use of the targeted action may also have been driven by other factors. Impact scores of 0 indicate that a student never took the recommended action, which is classified as no impact. In addition, we completed a series of random effects models to determine the amount of variance in modeling practices engagement that can be attributed to using the feedback.

4.4.3 Student Learning of Scientific Modeling

To gain insight into students' understanding of scientific modeling, we conducted interviews with students at the beginning and end of their use of EcoSurvey. This approach allowed us to focus on individual students within this collaborative task and provided insights into each student's ideas and experiences around modeling without disrupting classroom flow.

The interviews were designed as a semi-structured process, with a series of checkpoints. To conduct these interviews, we met students during their normal class period. We selected students for interviews in conjunction with the teacher, requesting students with varied understandings of the material. Each student was brought to a private meeting room or space, asked for additional permission to conduct the interview, and reintroduced to their rights as participants. Once this step was completed, we interviewed them about: 1) what they had been doing in class so far this year, 2) their understanding of "scientific modeling", 3) their experiences with scientific modeling in biology, 4) their expectations for constructing an ecosystem model, and 5) their experience with using EcoSurvey (post interview only).

We developed a two-part coding scheme for the student interviews. The first dimension focused on the depth and generalizability of students' descriptions of scientific modeling (from 0 to 4), with a higher score awarded for descriptions that focused on modeling as a science and

Table 4.1: The statistics for model features from all 3 deployment years

Model Feature	Year 1 avg. (SD)	Year 2 avg. (SD)	Year 3 avg. (SD)
# Organisms	65.1 (38.5)	121.6 (86.4)	100.8 (94.1)
# Interactions	60.6 (77.1)	138.2 (161.1)	37.1 (46.3)
Orphaned Cards	50.0%	50.0%	50.0%

engineering practice that can be used across problems and disciplines. This coding captured their most advanced modeling description. The second dimension simply determined how far into the interview, and therefore at what degree of prompting, did students first give the description of modeling captured in dimension one.

Once we developed this score of each student's understanding, we connected these results to their activity streams within EcoSurvey. We used this link to determine the correlation between student's engagement with the practices to their understanding of scientific modeling as a science and engineering practice.

4.5 Results

4.5.1 Analysis of Student Models & Activity

In Table 4.1, we see the average and standard deviations for the number of (a) organisms and (b) relations in students' models during each EcoSurvey deployment year. Here, we see that the average number of organisms has fluctuated, though not to a significant degree (Tukey's HSD: year 1 - year 2 $p = 0.212$; year 1 - year 3 $p = 0.513$; year 2 - year 3 $p = 0.550$), and the same trend is seen for interactions (Tukey's HSD: year 1 - year 2 $p = 0.383$; year 1 - year 3 $p = 0.860$; year 2 - year 3 $p = 0.359$). Additionally, the table shows that the complexity of models remained consistent across years. This consistency is reinforced by the percentage of disconnected (orphaned) cards.

In Table 4.2, we report the average and standard deviations for the number of times students engaged with each modeling practice. In addition, we report the resulting P value for the Wilcoxon-Mann-Whitney test of differences between the two populations, and the Hedges' g value of effect size.

Table 4.2: The statistics for modeling practices engagement for year 1 and year 3

Practice	Y1 Avg (SD)	Y3 Avg (SD)	Wilcoxon P	Hedges' g
Iterating (model)	2.93 (1.66)	14.70 (23.44)	$< .001$	0.554
Creating	2.97 (5.67)	4.99 (6.12)	$< .001$	0.334
Reviewing	5.25 (4.40)	9.46 (16.692)	0.012	0.277
Editing	4.66 (7.86)	4.19 (7.24)	0.983	0.064
Using	3.918269 (5.818350)	2.787393 (5.012145)	$< .001$	0.219

The results in Table 4.2 demonstrate a significant medium-sized improvement in engagement with iteration, or the cycling back and forth between modeling practices (e.g. creating model elements, reviewing information, then going back and creating once again). In addition, we saw significant small-sized improvements in engagement with creating and reviewing model elements. Surprisingly, we also found a significant small-sized decrease in student use of their models.

In addition, the standard deviations for iteration and review have both increased drastically. Since it is impossible to have fewer than 0 examples of any practice, these changes show the dramatic increase in the size of the high-end tail of the distribution, meaning some students are engaging with these practices very heavily.

Tables 4.3 and 4.4 show the results of our random-effects model, teasing apart the variance due to teacher from other residual factors. Our random effects analysis indicates that in year 1, teacher differences are responsible for up to 61% of the variance in student engagement with modeling. Conversely, in year 3, teacher variance is responsible for 12 to 30% of the variance in modeling practices engagement, with a decrease shown for every practice in year 3 relative to year 1. This reduction in teacher effects is most significant in the case of review and iteration.

Table 4.3: The variance attributable to teachers and other factors for each modeling practice.

Practice	Year 1 T. Var.	Year 1 R. Var.	Year 3 T. Var.	Year 3 R. Var.
Iteration	2.232	1.425	115.3	353.6
Create	10.39	25.65	8.993	28.813
Review	9.261	13.631	28.59	203.63
Edit	21.9	48.32	17.45	39.62
Use	10.01	26.97	5.342	22.341

Table 4.4: The variance attributable to teachers and other factors for each modeling practice.

Practice	Year 1 T. Proportion	Year 3 T. Proportion
Iteration	61.0%	24.6%
Create	28.8%	23.8%
Review	40.5%	12.3%
Edit	31.2%	30.6%
Use	27.1%	19.3%

4.5.2 Impact of Feedback Systems on Student Activity

Examining student use of EcoSurvey during year 3, we saw 578 uses of the feedback page from 191 of the 936 users (20.4%). Interestingly, while there is some teacher-level variance in the scope of use, 13 of the 15 teachers in our study had one or more student use the feature in some capacity.

In Table 4.5, we show the distribution of impact due to feedback on student activity. It is first important to account for the feedback uses where students had already engaged successfully with modeling, so the system did not generate a recommended next action. These uses accounted for 27.5% of all feedback checks. After this first check, we found 34.7% of feedback users had high impact based on the impact score above 0.8 (showed high impact), 12.3% had an impact score below .8 but above 0 (showed low impact), and 25.4% had an impact score of 0 (showed no impact).

Another important measure of feedback impact is the change in feedback over occurrences. Overall, there were 120 instances of students making progress, accessing the feedback more than once and seeing a different, more advanced recommendation. In addition, 53 users accessed the feedback and received no recommendation, 32 of whom went from previously receiving feedback to

Table 4.5: The impact score distribution for feedback usage.

Feedback Impact Score (i)	# occurrences	% occurrences
N/A (no recommendation)	159	27.5
High Impact (i > .8)	201	34.7
Low Impact (.8 > i > 0)	71	12.3
No Impact (i = 0)	147	25.4

Table 4.6: The variance attributable to feedback usage compared to other factors.

Practice	Teacher Var.	Feedback Var.	Residual Var.	Teacher %	Feedback %
Iteration	87.43	85.47	394.6	0.154061674	0.1506079295
Create	8.2355	0.3005	29.4369	0.2168783527	0.007913538339
Review	20.77	36.29	229.89	0.07238194807	0.1264680258
Edit	13.681	3.643	36.262	0.2553092226	0.06798417497
Use	1.957	7.247	17.958	0.07204918636	0.266806568

the no recommendation group over their course of use. It is also interesting to note that 110 feedback visits were completed after a student had already entered the no recommendation condition.

Finally, when including feedback in our year 3 models, we see that this feature accounts for an additional segment of the variance as shown in Table 4.6. This inclusion further reduces the variance found at the teacher level.

4.5.3 Student Learning of Scientific Modeling

Figure 4.4 shows the distribution of how accurate and generalized students' responses were when discussing scientific modeling in their post interviews. These data show an interesting bimodal distribution, which gives an indication that there is some significant difference in students' development of their understanding. This difference is further highlighted by the correlation between engagement with modeling practices and the final depth of understanding score as seen in Table 4.7. These results show that a student's engagement with modeling practices, especially model use and a large number of overall actions, is predictive of a student's understanding of scientific modeling.

Table 4.7: The correlation between the number of modeling actions of each type and a student's final depth of understanding score.

Practice	Correlation	P-Value
Iteration	0.426	0.078
Create	0.233	0.353
Review	0.432	0.073
Edit	0.363	0.139
Use	0.501	0.034*
Total Actions	0.488	0.040*

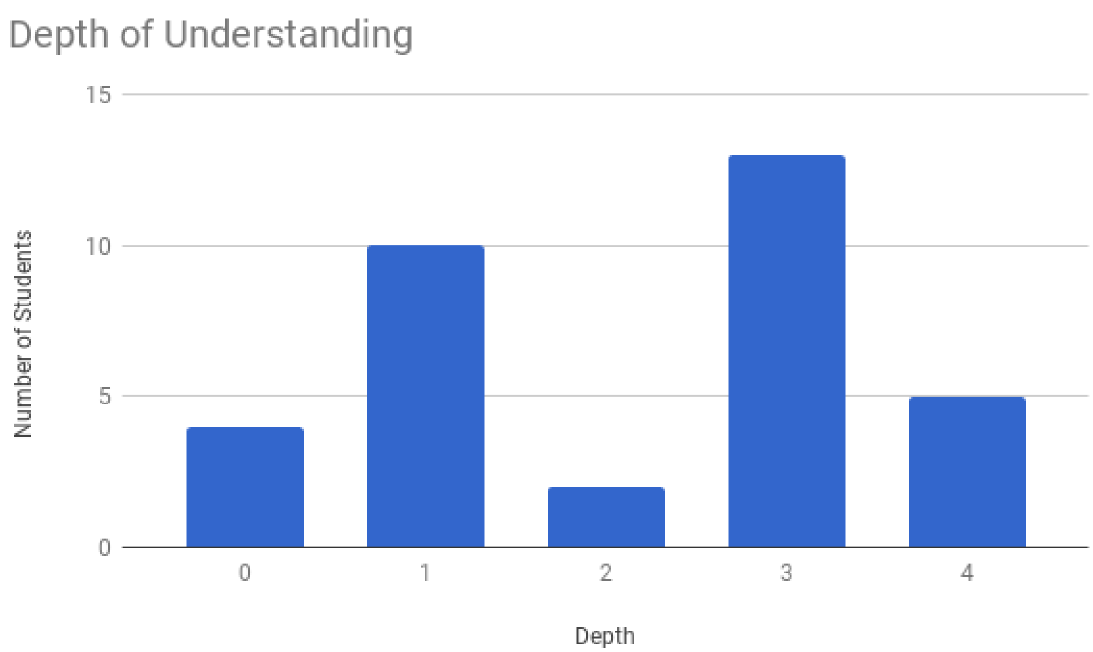


Figure 4.4: The distribution of student explanation scores in post-interviews.

One additional important trend we noticed in students' post interviews was the dichotomy between students' perceptions of modeling and their latent understanding. For example, only 32% of students discussed EcoSurvey or ecosystem models when asked "What kinds of models have you been making in science class this year?" during the interview, while the rest either discussed models generated prior to the study period or said that they had not modeled at all.

4.6 Discussion

4.6.1 (RQ1) Iterative Analysis of Final Models and Student Activity

Unfortunately, our analysis of the final models created by students showed that there were no significant changes to the size, complexity, or connectedness of these models between designs. Instead, the high variance in these model features demonstrated that there is further progress to be made in supporting the broad spectrum of students in creating robust models.

However, while we did not see a difference in students' models, our results demonstrate moderate success in our attempts to reduce teacher variance. One of the biggest improvements in our newer version of EcoSurvey was the positive impact on students' iteration, cycling back to previously conducted activities. As discussed above, iteration is a key component in successful scientific modeling, and improving students' engagement with this practice is a significant step towards equitable modeling experiences.

One of our biggest concerns is the drop in the "use" activity, visualizing their model as a whole. In our first year analysis, we determined engagement with model use varied significantly by teacher. Interestingly, we found a reduction in the between-teacher variance in model usage, and an even larger variance accounted for by use of the feedback tool. It seems only a certain subset of students within most classrooms ever used the model. One possible reason for this division is that, when they worked in groups, students divided up tasks and either trusted certain group members to review the use stage or they underwent a shared experience, with multiple students looking at one screen to discuss the complete model. In any case, it is important to rethink our techniques for better supporting students in the use of these ecosystem models to address their core design challenge.

It is important to consider the implications of an improvement in student activity with no corresponding improvements in their final models. Our interpretation of this result is that the measures of "good" student models are incomplete. While our metrics are designed around generic measures of models outside of a particular discipline, it is still important to remember students are using these models "to test a design, or aspects of a design, and to compare the effectiveness of different design solutions," ([58], pg 58) in this particular case to choose a tree to plant. Therefore, the "goodness" of the model is intrinsically tied to the accurate representation of the problem space, which is not captured by these generic features.

4.6.2 (RQ2) Impact of Feedback Systems on Student Activity

We saw a lack of utilization of our analytic feedback mechanism, with only around 20% of users ever visiting the page. That said, we noticed significant differences between feedback users and non-feedback users, with these differences not attributable to teacher differences. Our aggregate approach demonstrates that using the feedback system can account for a student exhibiting more robust modeling practices. Similarly, our geometric decay analysis suggests that the system showed a significant impact in 34.7% of use cases. In addition, the changes in successive feedback showed that students had increased their modeling practices engagement over 20% of the time. In addition, of the 159 access actions that received no feedback, 110 of those were repeat visits after previously having already received no feedback. These visits indicate that some students have interest in the graph of modeling activity, rather than just the personalized feedback.

4.6.3 (RQ3) Student Learning of Scientific Modeling

Our most startling finding is the latent nature of students' understanding of scientific modeling. While students said many things during the interview that demonstrated a strong understanding of scientific modeling, these responses were often not in response to general questions about scientific modeling. Rather, students reached their most sophisticated explanation when asked about specific modeling activities from the curriculum. In some instances, students did not even recognize their activities as ”modeling”, as seen in this excerpt from one interview:

> [Interviewer]: Have you all been making models in the past few weeks?
>
> [Student 30]: We actually haven't, we've been mainly seen looking at tree profiles and how to create them.

Yet, later in the interview:

> [Interviewer]: What would you say is the purpose of EcoSurvey?
>
> [Student 30]: To learn how different species, whether a tree or animals interact with each other and to learn the relationships they have.

Here, we see a student using generalized modeling language (”interact”, ”relationships”) in conjunction with domain specific modeling language (”species”) when discussing the use of EcoSurvey. However, when simply asked to reflect on recent modeling activities (which primarily focused on using EcoSurvey), the student did not consider the classroom activity to be ”modeling”.

In addition, it is important for our work to recognize the significant correlation between modeling activity and student understanding. While this correlation does not imply that using EcoSurvey caused a deeper understanding, it is nevertheless useful to know that these features found in the logs from our digital modeling tools can be predictive of students' understanding of modeling. This finding provides further evidence for the importance of using these generalized features to generate scaffolds for learning.

4.6.4 Limitations

Working within a design-based intervention research program has enabled the study of our design at scale and over time, but it is important to take note of the nesting (students within classroom groups within teachers) and other confounding effects of our results. While we have conducted mixed model tests, the interaction of factors between changes to EcoSurvey, other changes to the curriculum design, and implementation in the classroom are difficult to untangle. However, our approach of iterative design and deployment within a single curriculum minimizes these differences, and while our group of implementation teachers has grown and changed, we have kept our analysis within the same district.

4.7 Conclusions & Future Work

This research makes important contributions towards understanding students' scientific modeling at scale. We have demonstrated a variety of capabilities and limits of our generalized approach to analysis and used these features to understand the impact of changes in the design of a digital modeling tool. This work is building towards a more equitable learning experience through examining scaffolds for the modeling process and deploying automated feedback based on comparison with other modelers. These results demonstrate that there is some benefit to be gained from our approaches, but more work is needed.

Our future work includes an expanded focus on supporting the teacher as a mediator in this modeling process. As a first step, we made each students' feedback visible to their teacher, but a significant amount of work is required to make that insight easy to use and actionable for teachers while still supporting their unique insight into the needs of each student. In addition, while these results give some insight into how we can measure students' modeling activity, it is important to consider how these efforts can expand to document and incorporate a students understanding of modeling as a science and engineering practice. We also plan to leverage some interesting developments in the realm of 3D science assessment, creating tasks and scenarios that require students to incorporate modeling into their domain-specific solutions.

Chapter 5

Conclusions

Overall, my work has led to some key conclusions.

1) Student models can be characterized using generalized metrics, which can inform digital modeling environments. However, these features do not capture the entirety of model differences, and domain-specific features must be considered to understand the quality of a model.

2) Modeling practices can be successfully mapped to activity within a modeling tool. This mapping can provide important insight into a students understanding of the science and engineering practice of modeling.

3) Scaffolds and feedback built within a modeling tool can have an impact on both students models and their engagement with modeling practices, and reduce the influence of teacher differences on student success. However, designing tools and supports as optional features can lead to a self-selection bias in how successful these improvements are at reducing variance across the board.

5.1 Revisiting the Research Questions

5.1.1 RQ1) How can we automatically characterize students models and their engagement with modeling practices at scale?

Over the three studies, I have demonstrated a variety of metrics for understanding both models and engagement with modeling practices.

In understanding student models, I began with a simple metric of the number of components and interactions in each model. This metric provides a baseline of contribution and effort. In study

1, I mentioned the orphan phenomena - a significant percentage of model components were not connected to any others. This is one metric for model goodness; in real ecosystems, most (if not all) organisms are interacting with one another in some way. Additionally, I used an ecological approach of evenness to understanding the distribution of interaction types in the students models. By looking for a better representation from predatory, mutually beneficial, supporting, and competing relationships, we can determine if students were considering the breadth of ways in which organisms can interact.

In understanding student engagement with modeling practices, I developed a scheme of mapping user activity in a clickstream to their relevant modeling practices, which allowed for automatic, real-time characterization of the features in a students modeling behavior. I use these features to explore differences among populations of modelers. Through this analysis, I discovered significant differences in the way students behaved based on their teacher. These results, in turn, motivated design changes and gave me an opportunity to reuse these metrics to compare across multiple deployments of EcoSurvey.

5.1.2 RQ2) What methods can we use to promote successful scientific modeling?

My understanding of how we can promote successful scientific modeling comes from the iterative design of EcoSurvey and its features. In particular, I developed an integrated approach to using the models, implemented a closed list for assigning types of relationships, and incorporated a user dashboard featuring a real-time visualization of previous activity and personalized feedback.

The impact of these changes on students final models is mixed. I found that students use of the relationship types showed greater evenness in years with predetermined relationship options and an integrated model visualization, creating models that more accurately represent the ecosystem. However, I found no significant impact on the number of components or interactions in final models, nor an impact on the percentage of cards with no relationships in each model, across the three deployments.

Conversely, modeling tool design had more significant impacts on students engagement with modeling practices. We found that our design changes significantly reduced the teacher-level variance in reviewing, using, and iterating with their models, meaning that students did not face as many teacher differences in their exposure to these key modeling practices when using later versions of EcoSurvey. Furthermore, we found that use of the feedback feature is associated with increased engagement with modeling practices.

5.1.3 RQ3) How do students modeling practices relate to their understanding of scientific modeling?

The student interviews demonstrate that many students have a significant depth of understanding of scientific modeling. In addition, the correlation analysis of interviewee activity within EcoSurvey revealed that lower modeling activity can be associated with a less sophisticated understanding of modeling. While this does not necessarily imply a causal relationship, we can nevertheless use automated modeling activity detection and analysis to estimate a students understanding, which can then inform the scaffolding provided to the student during the modeling process.

5.2 Future Work

This dissertation opens the door for a whole area of new research. First and foremost on my mind is expanding these methods and analyses to other modeling tools and scenarios. There are a variety of projects that support scientific inquiry [54, 56], scientific reading comprehension [45], and modeling [35, 89] which provide an opportunity to generalize my findings across tools. It would be particularly interesting to explore modeling from the perspective of students using existing models (e.g. [35, 89]) rather than students process in generating new models.

Another important area for research is the continued exploration of design features within EcoSurvey. My analyses have found mixed success for the impact of several design decisions, but there is an open opportunity to continue to apply new techniques that can improve students modeling. I am most driven to explore ways to improve engagement with model use and uptake with feedback in future iterations of EcoSurvey. In addition, there are opportunities to expand information visualization research both with what feedback students and teachers can see and how that information is displayed.

Finally, there are opportunities to improve the methods used to predict student activity. As implemented in our most recent version, the predictive model uses aggregations of students modeling activity and the presence of key sequence features identified from our first years deployment as predictive of teacher differences. However, this approach does not yet adapt to the expanded patterns found in our larger deployments, nor does it distinguish between good modelers that have just begun their process and modelers who have struggled for a significant period of time. There is also room for improvement in the machine learning models used for prediction; our early work relied on the relatively strong performance of a simple nave bayes model, but more advanced and adaptive modeling techniques could be used to address some shortcomings in the current implementation.

Bibliography

[1] Megan Bang, Douglas L Medin, and Scott Atran. Cultural mosaics and mental models of nature. Proceedings of the National Academy of Sciences of the United States of America, 104(35):13868–13874, 2007.

[2] Srinjita Bhaduri, Katie Van Horne, J Ristvey, R Russell, and Tamara Sumner. From toys to tools: UAVs in middle-school engineering education. In 2018 ASEE Annual Conference & Exposition, In Press.

[3] C Bryce, V B Baliga, K de Nesnera, D Fiack, K Goetz, L M Tarjan, C Wade, V Yovovich, S Baumgart, D Bard, D Ash, I M Parker, and G S Gilbert. Exploring Models in the Biology Classroom. The American Biology Teacher, 8(1):35–42, 2016.

[4] Anthony S Bryk, Louis M Gomez, Alicia Grunow, and Paul G LeMahieu. Learning to Improve: How America's Schools Can Get Better at Getting Better. ERIC, 2015.

[5] Brandy LE Buckingham and Brian J Reiser. What is a model? experienced students beliefs about the nature and purpose of scientific models across modeling contexts. In annual conference of the National Association for Research in Science Teaching (NARST), Philadelphia, PA, 2010.

[6] Meng-Fei Cheng and David E Brown. The role of scientific modeling criteria in advancing students' explanatory ideas of magnetism. Journal of Research in Science Teaching, 52(8):1053–1081, 2015.

[7] Ming Ming Chiu, Bonnie Wing-Yin Chow, and Sung Wook Joh. How to assign students into sections to raise learning. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference, pages 95–104. ACM, 2017.

[8] Richard K. Coll and Denis Lajium. Modeling and the Future of Science Learning, pages 3–21. Springer Netherlands, Dordrecht, 2011.

[9] National Research Council et al. Developing assessments for the next generation science standards. National Academies Press, 2014.

[10] Mutlu Cukurova, Katerina Avramides, Daniel Spikol, Rose Luckin, and Manolis Mavrikis. An analysis framework for collaborative problem solving in practice-based learning activities: A mixed-method approach. In Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, LAK '16, pages 84–88, New York, NY, USA, 2016. ACM.

[11] Mathieu d'Aquin and Nicolas Jay. Interpreting data mining results with linked data for learning analytics: motivation, case study and direction. In Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, 2013.

[12] Ryann K Ellis. Field guide to learning management systems. ASTD Learning Circuits, page 7, 2009.

[13] Mingyu Feng, Andrew E Krumm, Alex J Bowers, and Timothy Podkul. Elaborating data intensive research methods through researcher-practitioner partnerships. Proceedings of the Sixth International Conference on Learning Analytics {&} Knowledge - LAK '16, pages 540–541, 2016.

[14] Barry J Fishman, William R Penuel, Anna-Ruth Allen, Britte Haugan Cheng, and NORA Sabelli. Design-based implementation research: An emerging model for transforming the relationship of research and practice. National society for the study of education, 112(2):136–156, 2013.

[15] EricB B Fretz, Hsin-Kai Wu, BaoHui Zhang, Elizabeth A Davis, JosephS S Krajcik, and Elliot Soloway. An investigation of software scaffolds supporting modeling practices. Research in Science Education, 32(4):567–589, 2002.

[16] Ujwal Gadiraju and Stefan Dietze. Improving learning through achievement priming in crowdsourced information finding microtasks. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference, pages 105–114. ACM, 2017.

[17] John K Gilbert. Models and modelling: Routes to more authentic science education. International Journal of Science and Mathematics Education, 2(2):115–130, 2004.

[18] Arthur C Graesser, Mark W Conley, and Andrew Olney. Intelligent tutoring systems. 2012.

[19] Lance H Gunderson. Ecological Resilience–In Theory and Application. Annual Review of Ecology and Systematics, 31:425–439, 2000.

[20] Isabelle Guyon and André Elisseeff. An Introduction to Variable and Feature Selection. Journal of Machine Learning Research, 3(3):1157–1182, 2003.

[21] M.A. Hall and Geoffrey Holmes. Benchmarking attribute selection techniques for data mining. IEEE Transactions on Knowledge and Data Engineering, 15(6):1437–1447, 2003.

[22] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H Witten. The WEKA data mining software. ACM SIGKDD Explorations, 11(1):10–18, 2009.

[23] Larry V Hedges. Estimation of effect size from a series of independent experiments. Psychological bulletin, 92(2):490, 1982.

[24] Martin Hlosta, Zdenek Zdrahal, and Jaroslav Zendulka. Ouroboros: early identification of at-risk students without models based on legacy data. 2017.

[25] Cindy E Hmelo-Silver, Surabhi Marathe, and Lei Liu. Fish swim, rocks sit, and lungs breathe: Expert-novice understanding of complex systems. The Journal of the Learning Sciences, 16(3):307–331, 2007.

[26] C S Holling. Resilience and Stability of Ecological Systems. Annual Review of Ecology and Systematics, 4:1–23, 1973.

[27] Kenneth Holstein, Bruce M McLaren, and Vincent Aleven. Intelligent tutors as teachers' aides: exploring teacher needs for real-time analytics in blended classrooms. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference, pages 257–266. ACM, 2017.

[28] Jack B Homer. Why We Iterate: Scientific Modeling in Theory and Practice. System Dynamics Review, 12(1):1–19, 1996.

[29] Eric Horvitz, Andy Jacobs, and David Hovel. Attention-sensitive alerting. In Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence, pages 305–313. Morgan Kaufmann Publishers Inc., 1999.

[30] Michal Huptych, Michal Bohuslavek, Martin Hlosta, and Zdenek Zdrahal. Measures for recommendations based on past students' activity. 2017.

[31] S Jackson, Joseph Krajcik, and Elliot Soloway. Model-it: A design retrospective. Innovations in science and mathematics education: Advanced designs for technologies of learning, pages 77–115, 2000.

[32] Rebecca C. Jordan, Steven Gray, Amanda E. Sorensen, Samantha Pasewark, Suparna Sinha, and Cindy E. Hmelo-Silver. Modeling with a Conceptual Representation: Is It Necessary? Does It Work? Frontiers in ICT, 4(April):1–8, 2017.

[33] Thomas J Kane, Daniel F McCaffrey, Trey Miller, and Douglas O Staiger. Have we identified effective teachers? Validating measures of effective teaching using random assignment. Bill and Melinda Gates Foundation, 2013.

[34] Tanja Käser, Nicole R Hallinen, and Daniel L Schwartz. Modeling exploration strategies to predict student performance within a learning environment and beyond. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference, pages 31–40. ACM, 2017.

[35] CJ Keller, ND Finkelstein, KK Perkins, and SJ Pollock. Assessing the effectiveness of a computer simulation in introductory undergraduate environments. In AIP Conference Proceedings, volume 883, pages 121–124. AIP, 2007.

[36] Stephen R Kellert and Edward O Wilson. The biophilia hypothesis. Island Press, 1995.

[37] Diane Jass Ketelhut, Brian C Nelson, Jody Clarke, and Chris Dede. A multi-user virtual environment for building and assessing higher order inquiry skills in science. British Journal of Educational Technology, 41(1):56–68, 2010.

[38] Elizabeth Koh, Antonette Shibani, Jennifer Pei-ling Tan, and Helen Hong. A Pedagogical Framework for Learning Analytics in Collaborative Inquiry Tasks : An Example from a Teamwork Competency Awareness Program. Proceedings of LAK16 6th International Conference on Analytics and Knowledge 2016, pages 74–83, 2016.

[39] Janet L. Kolodner, Paul J. Camp, David Crismond, Barbara Fasse, Jackie Gray, Jennifer Holbrook, Sadhana Puntambekar, and Mike Ryan. 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, 2003.

[40] Joseph Krajcik, Phyllis Blumenfeld, Ronald Marx, Kristin Bass, Jennifer Fredricks, and Elliot Soloway. Inquiry in Project-Based Science Classrooms: Initial Attempts by Middle School Students. Journal of the Learning Sciences, 7(3):313–350, 1998.

[41] Andrew E Krumm, Rachel Beattie, Sola Takahashi, Cynthia D'Angelo, Mingyu Feng, and Britte Cheng. Practical measurement and productive persistence: Strategies for using digital learning system data to drive improvement. Journal of Learning Analytics, 3:116–138, 2016.

[42] William H Kruskal and W Allen Wallis. Use of Ranks in One-Criterion Variance Analysis. Journal of the American Statistical Association, 47(260):583–621, 1952.

[43] Heather Leary, Samuel Severance, William R Penuel, David Quigley, Tamara Sumner, and Holly Devaul. Designing a deeply digital science curriculum: Supporting teacher learning and implementation with organizing technologies. Journal of Science Teacher Education, 27(1):61–77, 2016.

[44] Richard Lehrer and Leona Schauble. Scientific thinking and science literacy, 2007.

[45] Megan K. Littrell-Baez and Donna Caccamise. A Cognitive Perspective on Chemistry Instruction: Building Students Chemistry Knowledge through Advancing Fundamental Literacy and Metacognitive Skills, chapter 3, pages 31–42.

[46] Derek Lomas, Dixie Ching, Eliane Stampfer, Melanie Sandoval, and Ken Koedinger. " battleship numberline": A digital game for improving estimation accuracy on fraction number lines. Society for Research on Educational Effectiveness, 2011.

[47] Bote Lorenzo, L Miguel, and Eduardo Gómez Sánchez. Predicting the decrease of engagement indicators in a mooc. 2017.

[48] Michelle Lui and Tom Moher. Learning to model ecosystems with interaction food webs in middle school classrooms. Philadelphia, PA: International Society of the Learning Sciences., 2017.

[49] Roy Lyster and Leila Ranta. Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. Studies in second language acquisition, 19(1):37–66, 1997.

[50] Norman W. H. Mason, David Mouillot, William G. Lee, J. Bastow Wilson, and Heikki Setälä. Functional richness, functional evenness and functional divergence: The primary components of functional diversity. Oikos, 111(1):112–118, 2005.

[51] K E Maull, M G Saldivar, and Tamara Sumner. Understanding digital library adoption: a use diffusion approach. Proceedings of the 11th annual international ACM/IEEE joint conference on digital libraries, pages 259–268, 2011.

[52] L. M. McDonnell. Opportunity to learn as a research concept and a policy instrument. Educational Evaluation and Policy Analysis, 17(3):305–322, 1995.

[53] Katherine L McNeill, Diane Silva Pimentel, and Eric G Strauss. The impact of high school science teachers' beliefs, curricular enactments and experience on student learning during an inquiry-based urban ecology curriculum. International Journal of Science Education, 35(15):2608–2644, 2011.

[54] S Metcalf, C Dede, T Grotzer, and A Kamarainen. Ecomuve: Design of virtual environments to address science learning goals. In American Educational Research Association (AERA) Conference, Denver, CO, 2010.

[55] Yun Mo, Kusum Singh, and Mido Chang. Opportunity to learn and student engagement: A HLM study on eighth grade science achievement. Educational Research for Policy and Practice, 12(1):3–19, 2013.

[56] Tom Moher, Brian Uphoff, Darshan Bhatt, Brenda López Silva, and Peter Malcolm. Wallcology: Designing interaction affordances for learner engagement in authentic science inquiry. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '08, pages 163–172, New York, NY, USA, 2008. ACM.

[57] Emerson Murphy-Hill and Gail C Murphy. Recommendation delivery. In Recommendation Systems in Software Engineering, pages 223–242. Springer, 2014.

[58] National Research Council. A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. The National Academies Press, Washington, DC, 2012.

[59] Markus Neuhäuser. Wilcoxon–Mann–Whitney Test, pages 1656–1658. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[60] Jakob Nielsen. Iterative user-interface design. Computer, 26(11):32–41, 1993.

[61] Donald A Norman. Affordance, conventions, and design. interactions, 6(3):38–43, 1999.

[62] Zacharoula Papamitsiou, Eirini Karapistoli, and Anastasios A Economides. Applying Classification Techniques on Temporal Trace Data for Shaping Student Behavior Models. Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, pages 299–303, 2016.

[63] Jihyun Park, Kameryn Denaro, Fernando Rodriguez, Padhraic Smyth, and Mark Warschauer. Detecting changes in student behavior from clickstream data. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference, pages 21–30. ACM, 2017.

[64] Miranda C. Parker, Kantwon Rogers, Barbara J. Ericson, and Mark Guzdial. Students and teachers use an online ap cs principles ebook differently: Teacher behavior consistent with expert learners. In Proceedings of the 2017 ACM Conference on International Computing Education Research, ICER '17, pages 101–109, New York, NY, USA, 2017. ACM.

[65] Radek Pelánek, Jirí Rihák, and Jan Papoušek. Impact of Data Collection on Interpretation and Evaluation of Student Models. Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, pages 40–47, 2016.

[66] William R Penuel, Jeremy Roschelle, and Nicole Shechtman. Designing formative assessment software with teachers: An analysis of the co-design process. Research and Practice in Technology Enhanced Learning, 2(01):51–74, 2007.

[67] William R Penuel, Katie Van Horne, Samuel Severance, David Quigley, and Tamara Sumner. Students responses to curricular activities as indicator of coherence in project-based science. Transforming Learning, Empowering Learners, 2016.

[68] Dilhan Perera, Judy Kay, Irena Koprinska, Kalina Yacef, and Osmar R Zaïane. Clustering and sequential pattern mining of online collaborative learning data. IEEE Transactions on Knowledge and Data Engineering, 21(6):759–772, 2009.

[69] David Quigley, Conor McNamara, Jonathan Ostwald, and Tamara Sumner. Using learning analytics to understand scientific modeling in the classroom. Frontiers in ICT, 4:24, 2017.

[70] David Quigley, Jonathan Ostwald, and Tamara Sumner. Scientific Modeling: Using learning analytics to examine student practices and classroom variation. In The 7th International Conference on Learning Analytics & Knowledge, pages 1 – 10, Vancouver, 2017. ACM.

[71] Chris Quintana, Brian J Reiser, Elizabeth a. Davis, Joseph Krajcik, Eric Fretz, Ravit Golan Duncan, Eleni Kyza, Daniel Edelson, and Elliot Soloway. A scaffolding design framework for software to support science inquiry. Journal of the Learning Sciences, 13(3):337–386, 2004.

[72] Chris Quintana, Brian J Reiser, Elizabeth A Davis, Joseph Krajcik, Eric Fretz, Ravit Golan Duncan, Eleni Kyza, Daniel Edelson, and Elliot Soloway. A scaffolding design framework for software to support science inquiry. The journal of the learning sciences, 13(3):337–386, 2004.

[73] Brett Roads and Michael Mozer. Improving categorization training with structure-sensitive scheduling. Journal of Vision, 16(12):402–402, 2016.

[74] William Sandoval. Conjecture mapping: An approach to systematic educational design research. Journal of the learning sciences, 23(1):18–36, 2014.

[75] Leona Schauble, Robert Glaser, Richard a. Duschl, Sharon Schulze, and Jenny John. Students' Understanding of the Objectives and Procedures of Experimentation in the Science Classroom. Journal of the Learning Sciences, 4(2):131–166, 1995.

[76] Christina V. Schwarz, Brian J. Reiser, Elizabeth A. Davis, Lisa Kenyon, Andres Achér, David Fortus, Yael Shwartz, Barbara Hug, and Joe Krajcik. Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. Journal of Research in Science Teaching, 46(6):632–654, 2009.

[77] Samuel Severance, William R Penuel, Tamara Sumner, and Heather Leary. Organizing for Teacher Agency in Curricular Co-Design. Journal of the Learning Sciences, 25(4):531–564, 2016.

[78] Claude E Shannon and Warren Weaver. The mathematical theory of communication. University of Illinois press, 1998.

[79] Chuan-Fong Shih and Alladi Venkatesh. Beyond adoption: Development and application of a use-diffusion model. Journal of Marketing, 68(1):59–72, 2004.

[80] Elliot Soloway, Amanda Z Pryor, Joseph S Krajcik, Shari Jackson, Steven J Stratford, Michele Wisnudel, and John T Klein. Scienceware's model-it: Technology to support authentic science inquiry. T.H.E. Journal, 25(3):54–56, 1997.

[81] Mary Kay Stein and Linda Kucan. Instructional explanations in the disciplines. Instructional Explanations in the Disciplines, pages 1–237, 2010.

[82] Student. The probable error of a mean. Biometrika, 6(1):1, 1908.

[83] Tamara Sumner and CCS Team. Customizing science instruction with educational digital libraries. In Proceedings of the 10th annual joint conference on Digital libraries, pages 353–356. ACM, 2010.

[84] John W. Tukey. Comparing individual means in the analysis of variance. Biometrics, 5(2):99–114, 1949.

[85] UNEP. What is Biodiversity ? United Nations Environment Program, pages 3–9, 2010.

[86] Maarten van Gompel and Antal van den Bosch. Efficient n-gram , Skipgram and Flexgram Modelling with Colibri Core. Journal of Open Research Software, 4(e30), 2016.

[87] K. VanLehn, J. Wetzel, S. Grover, and B. v. d. Sande. Learning how to construct models of dynamic systems: An initial evaluation of the dragoon intelligent tutoring system. IEEE Transactions on Learning Technologies, 10(2):154–167, April 2017.

[88] Lev Semenovich Vygotsky. Mind in society: The development of higher psychological processes. Harvard university press, 1980.

[89] Uri Wilensky. Netlogo. evanston, il: Center for connected learning and computer-based modeling, northwestern university, 1999.

[90] Mark Windschitl, Jessica Thompson, and Melissa Braaten. Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. Science Education, 92(5):941–967, 2008.

[91] Zhengzheng Xing, Jian Pei, and Eamonn Keogh. A brief survey on sequence classification. ACM SIGKDD Explorations Newsletter, 12(1):40, 2010.

[92] Shoshana Zuboff. In the age of the smart machine: the future of power and work. New York: Basic, 1988.