NORTHWESTERN UNIVERSITY

Measuring Model-Based Learning of Complex Systems with Multiple Data Streams

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Learning Sciences

By

Kit Martin

EVANSTON, ILLINOIS

December 2021

Abstract

The dissertation builds on my current research to demonstrate the connection between affect and learning through machine learning and qualitative analysis of interactions where players use a complex systems game. The project is threefold: First, I developed a thinking and learning intervention, the agent-based modeling simulation Ant Adaptation. I showed that the intervention can shift people's schemas from a process schema to an emergent schema during 10-minute museum interactions. Second, to track that conceptual change common in agent-based modeling interventions, I developed a novel form of concept mapping, constructivist dialogue mapping (CDM), which is particularly useful as a learning analytic. Through CDM, I analyzed participants' spoken elaborations in small subsets to study how people develop their understanding of a system or museum exhibit over time. Third, through video analysis, I developed a method of affect detection to identify how participants are engaged across 17 facial action units. I map those affective-states to moments of learning using machine learning methods. Because the data source is video, the method has outsized potential for scale to predict unseen data. In short, in my work I have applied advanced methods to the design and evaluation of educational interventions.

Acknowledgements

I would like to thank the Center for Connected Learning and Computer Based Modeling (CCL), the Tangible Integration Design Lab and Learning Lab (TIDAL), and The Technological Innovations for Inclusive Learning and Teaching (tiilt) labs for their feedback throughout the project. I would like to thank the supportive Northwestern community. I would like to thank Uri Wilensky, my adviser, for keeping me focused on the main objectives of the project and not getting too far in the weeds. I would like to thank Marcelo Worsley for teaching me the methods of multimodal learning sciences research, and providing an awesome group of fellow scholars. I would like to thank Mike Horn for all his insight in museum learning, and approaches to learning and design. I would like to thank Corrie Moreau, and her Ant Lab for invaluable feedback on Ant Adaptation, and being a kindred myrmecological collaborator. Finally, I would like to thank the National Science Foundation (NSF STEM+C: Award # 1842375) and the Multidisciplinary Program in Education Sciences (IES: Award # R305B090009) for supporting this work.

Table of Contents

List of Figures

Chapter 1: Ant Adaptation: An Introduction to Measuring Model-Based Learning of Complex Systems with Multiple Data Streams

Tell me, and I will forget. Show me, and I may remember. [Immerse] me, and I will understand.
–Confucius (circa 450BC)

We live in a physical world whose properties we have come to know well through long familiarity. We sense an involvement with this physical world which gives us the ability to predict its properties well. For example, we can predict where objects will fall, how well-known shapes look from other angles, and how much force is required to push objects against friction. We lack corresponding familiarity with the forces on charged particles, forces in non-uniform fields, the effects of nonprojective geometric transformations, and high-inertia, low friction motion. A display connected to a digital computer gives us a chance to gain familiarity with concepts not realizable in the physical world. It is a looking glass into a mathematical wonderland.
–Sutherland (1960), creator of the first mixed-reality headset, in *The Ultimate Display*

1 Introduction

As the world is interconnected like never before, in recent years there has been a greater focus on teaching and understanding complexity (Bar-yam, 1999; Bryson & Crosby, 2005; Hmelo-Silver & Pfeffer, 2004; Morin, 2008; Wilensky & Rand, 2015; Wilensky & Resnick, 1999). The global interconnectedness means that for core problems, from climate change (League et al., 2019) to income inequality (Atkinson, Muro, and Whiton, 2019; Escobar, 2011) to traffic (Barceló, 2010), or pandemic disease (Jenkins et al., 2020), no one is really in charge (Bryson & Crosby, 2005). To effectively address these issues, we need an informed public along with educational environments that teach how to think about complex systems (Blikstein & Wilensky, 2009; Sengupta & Wilensky, 2009; Wilensky & Resnick, 1999). Students understanding the complex world we live in will improve the chances of overcoming these pressing problems. With this remarkable interconnectedness, it is clear we are all now planetary citizens; but it is clear our educational systems have not yet prepared us for this situation (Morin, 2008).

In this project, I use two methods that have been successful in supporting learning about complexity: constructionism and agent-based modeling. Constructionism, a mnemonic term (Papert, 1986) coined to describe a species of constructivist thought (Piaget, 1983), focuses on the benefits of learning from the external construction of an artifact alongside the internal construction of a mental model. This pedagogical theory argues that learning is the emergence of new knowledge structures, or mental models (Papert, 1980; Papert & Harel, 1991a) aided through building public entities (Bamberger, 2001; DiSessa et al., 1991; Holbert & Wilensky, 2019; Nemirovsky & Tierney, 2001; Papert, 1980; Papert & Harel, 1991a). One area of research that has been emphasized in this field involves the use of agent-based models to create constructionist learning experiences for complexity education (Goldstone & Wilensky, 2008; Klopfer et al., 2009; Wilensky, 1999b; Wilensky & Rand, 2015; Yoon, 2018). In particular, this work has focused on understanding complex systems of science phenomena, for example: bees (Guo & Wilensky, 2014), material science (Blikstein & Wilensky, 2009), ecology modeling (Wilensky & Reisman, 2006), electricity (Sengupta & Wilensky, 2009), evolution (Wagh & Wilensky, 2018), and ants (K. Martin, Horn, et al., 2019). With growing work in understanding social systems, like income inequality (Guo & Wilensky, 2018), that may have outsized leverage (Wilensky & Papert, 2010).

To meet the growing need to improve the teaching and understanding of complexity, I have developed Ant Adaptation (K. Martin & Wilensky, 2019), a complex systems-learning environment developed to foster communication and positive engagement around core concepts of complexity, including feedback and self-organizing systems. In the environment, users play a game where they set their own goals while rearing two ant colonies.

This project has two major components. The components weave into the entire project, but generally, the first component is explored in the 3rd chapter, and the second component is conveyed and explored in 4th and 5th chapters. In the first component, I explore and describe the design of the constructionist environment Ant Adaptation, an agent-based modeling game that integrates complex system learning with gameplay. In this part of the project backgrounded in chapter 2, and expanded on with empirical findings in chapter 3, I discuss the design of Ant Adaptation and its impact on visitors. I discuss the creation of the environment for deployment in the informal environment of a museum, based on prior complexity science education literature. Here, the goal of Ant Adaptation was to motivate museumgoers to investigate complex systems, which helps users create mental models of complex systems. In this first component I demonstrate that Ant Adaptation can improve the development of complex systems intuitions. However, while this first iteration of Ant Adaptation allowed users to develop intuitions about decentralized control and development, the limitations of the findings indicated the need for better understanding of how participants construct knowledge and engage with the learning environment.

As a result, in the second component of this project detailed in chapters 4 and 5, I use the product of the first iteration, Ant Adaptation, and implement two new ways to research how users come to understand complex systems: constructivist dialogue mapping (CDM) (Martin, Horn & Wilensky, 2020) and affective engagement that builds on the work of computationally augmented ethnography (K. Martin, Wang, et al., 2019). As I discuss in this second component of the thesis, the two approaches measure and assess the learning (using CDM) and engagement (using computationally augmented ethnography) of users during complexity science education. New methods to study participant learning and engagement are needed because constructionist learning environments encourage open-ended exploration. This shift is a restructuration (Wilensky & Papert, 2010) of education resulting from the impact of the advent of powerful computation.

1.1 Restructurations

A restructuration is a new way of doing something that affords additional abilities. Wilensky and Papert discuss a restructuration —namely, Hindu-Arabic numerals. In that transition the Hindu-Arabic numerals supplanted the romans numerals, first in accounting, but later through computation because they were easier to perform operations with (for example, it is much easier to divide 66 by 11 than it is to divide LXVI by XI).[1] This non-trivial process eventually led to a

[1] While addition and subtraction is relatively straightforward in Roman numerals, and multiplication isn't even so bad, division is the real problem. For example, in the Roman Empire to perform the calculation LXVI divided by XI

transition to the use of Hindu-Arabic numerals.

The transition happens along 5 dimensions. In a restructuration, Wilensky and Papert (2010) delineate 5 properties researchers should attend to: powerful ideas, cognitive properties, affective properties, social properties, and diversity properties.

- Powerful Ideas: A restructuration must afford what could be done before, but for it to spread, it also must be able to do more. In the case of Hindu-Arabic numerals this is the power of multiplication and division massively expanded. When we open education for students to explore computational microworlds on their own recognizance, this is a powerful example of what could be done before—understanding physics and biology for instance—but with a huge new set of ideas in complexity, emergence and self-organization. This notion undergirds the complex systems exploration afforded by the design of Ant Adaptation, as it lowers the floor for exploration of these multiagent, self-organizing systems.
- Cognitive properties: A successful restructuration is more easily learned, while still providing the augmenting power of the old tools. In the current case, this is the idea of understanding ecology, while adding the ability to explore the role of emergence and complexity in how an ant colony's ecology sustains.
- Affective properties: A new structure can be more or less engaging. Computation media offer one example of more engaging properties. One question is how to measure this engagement, a field the second component of this dissertation takes a deep dive into.
- Social properties: As a structuration is more shareable, it spreads through a group more readily. In the case of Hindu-Arabic numerals, their use in business made them highly sharable. In the case of Ant Adaptation, I have tried to design it to encourage folks to want to draw friends and family into its use. The properties of this sharing are key to the restructuration's spread.
- Diversity Properties: With the power of design and iteration, a restructuration spreads, and as it does so, it comes to ever more match people's style of use.

We can see these properties operation as learning environments grapple with the incorporation of more affectively and cognitively engaging curricula that students want to share

---

(66 divided by 11) was an arduous process of trial an error that really came down to repeated subtraction. But because the numbers are not both even or divisible by 5 or 10, a scribe calculating sums late at night had no shortcut. As Mark Chu-Carroll said: "Division is the biggest problem in roman numerals. There is no good trick that works in general. It really comes down to repeated subtraction. The only thing you can do to simplify is variations on finding a common factor of both numbers that's easy to factor out. For example, if both numbers are even, you can divide each of them by two before starting the repeated subtraction. It's also easy to recognize when both numbers are multiples of 5 or 10, and to do the division by 5 or 10 on both numbers. But beyond that, you take a guess, do the multiplication, subtract, repeat." (Chu-Carroll, 2006)

that attend to the unique accommodations and diversity of students while lowering the floor to understand powerful ideas. Though the evolution is happening, microworld exploration in a mathematical wonderland of computational environments has not yet hit the mainstream. When learning is no longer one-size-fits-all, it becomes more difficult to develop uniform, standardized assessments. While we could say it's only a matter of time, there is one elephant in the room holding back the spread of open-ended individually designed learning: assessment. The uniqueness of the learning trajectories affording by the exploration of these mathematical wonderlands, has made it harder for education researchers to develop strong assessments for constructionist learning environments. The lack of evaluation slows down their implementation and adoption in educational environments. This is unfortunate because this type of thinking—complex systems and the outcomes of self-organizing systems—is exactly what society needs to deal with core problems, such as climate change, in a complex world. The goal of the second component of this project is to demonstrate how two methods I developed give greater insight into the cognitive and affective properties of student learning of complex systems through the design of an agent-based modeling game that teaches complexity. The approach quantifies their affective and cognitive engagement with Ant Adaptation. The approach is applicable to studying learning environments more generally, and especially useful when studying open-ended exploration of mathematical wonderlands and microworlds.

The remainder of this introduction has five sections. In the first section, I define design-based research and situate the iterative development of Ant Adaptation within this research methodology. In the second, I provide an overview of my thesis with chapter summaries. In the third section, I describe long term goals and outcomes of the project. In the fourth, I list my research questions, and in the fifth, I summarize the findings of my project.

2 Design Research

How do you design a restructuration? You attend to the five properties of a restructuration and iteratively design. Design research is a methodology largely associated with the learning sciences community. Since being formally named in 1992 (Brown, 1992; A. Collins, 1992), design research has evolved to become an accepted paradigm of educational research, with Gutiérrez and Jurow (2016, p. 593) going so far as to say, "The field of the learning sciences developed to overcome… limitations through emphasizing interdisciplinarity and focusing on real-world problems." This process led to design research, a signature method of the learning sciences. The method has allowed for the processes of learning to be designed, documented, refined, and integrated as part of the dynamic activity context in which learning unfolds (Barab & Squire, 2004; Cobb et al., 2003).

Design research (DR) is intended to produce innovative learning environments—which would then provide knowledge about how such environments work in the intended setting or context—and elucidate some fundamental knowledge about learning or teaching (Cobb et al., 2003; Design-Based Research Collective, 2003; Edelson, 2002). However, there have been some well-organized critiques of DR (Kelly, 2004; Phillips & Dolle, 2006). The critiques focus on the one hand with Kelly stating that design research lacks clear standards or methodological rigor (Kelly, 2004), while on the other hand, Phillips and Dolle argue that DR cannot *simultaneously* evaluate designs while improving theory.

Replying to these critiques, Sandoval (2014) clarified some of the principles of design research and how it applies to learning science research. Sandoval proposes conjecture mapping, a methodology to address the two critiques. Conjecture mapping begins by assuming that the design of learning environments stems from theoretical activity, and that learning involves the embodied hypothesis of how learning happens through the designs we put into this world (Cobb et al., 2003; W. A. Sandoval, 2004). As a result, it would be beneficial to the field of Learning Sciences to be as initially explicit as possible about the ideas we are putting forth into the world.

For Sandoval, conjecture mapping is an attempt to provide a means for specifying such design relationships: to make them concrete. Conjecture mapping begins with a) high level conjecture about how some activities support certain outcomes, such as learning, and b) embodiments (usually the tools tasks, persistent structures, and discursive practices). These embodiments produce mediating processes, which are the observable interactions of people in the area: people interacting with the tools, the structures, and discourse practices. They are also the mediating artifacts that are created through the embodiments (i.e., NetLogo models (Wilensky, 1999b)), 3D printed designs, or handcrafted straw houses made of popsicle sticks). These mediating processes should lead—through our theoretical conjectures—to outcomes, including things that matter to us, such as the learning, engagement, interest, or motivation of participants.

Conjectures, according to Sandoval, connote often highly provisional ideas about how to design a learning environment at the start of a design research project. Design research typically aims to create novel contexts for learning, which our previous reading—our theory—suggests could be productive but are not common or well understood (Design-Based Research Collective, 2003). The decision to use a design research methodology allows the project to adapt to new technologies and accommodate refined learning goals.

2.1 The Design Research of Ant Adaptation

Ant Adaptation has been a multiyear design-based research project to create a restructurated unit: a complex systems learning environment that is more affectively and cognitively salient. The decision to implement new ways of researching how users understand complex systems resulted from the process of iterating designs to make a better tool and deepen fundamental knowledge about learning.

As shown in Figure 1, when developing Ant Adaptation I started with a high-level conjecture: dynamic interaction with a complex system model would lead to overcoming the twin challenges of complexity learning: levels confusion (Wilensky & Resnick, 1999) and not understanding the parts of the system (Chi et al., 2012). In practice, that meant gathering peers around a touch table or digital screen where their discussion about a game mediates learning complex systems, and by doing so be able to overcome these challenges. I put forward an embodiment that has taken on several forms. I designed Ant Adaptation, as an agent-based modeling game with a built-in task structure. For people to ideate around the space, they needed peers to talk to, and I needed to have a way of understanding them. To do this, I talked to them about ants in a pre-interview, and, after providing a short introduction, I let them play in an open-ended way. This would give participants enough guidance to preliminarily scaffold the interactions that then falls away as they took over, giving them agency to speak aloud to each other and compete and contest different notions about the idea—namely, the embodiment, Ant Adaptation.

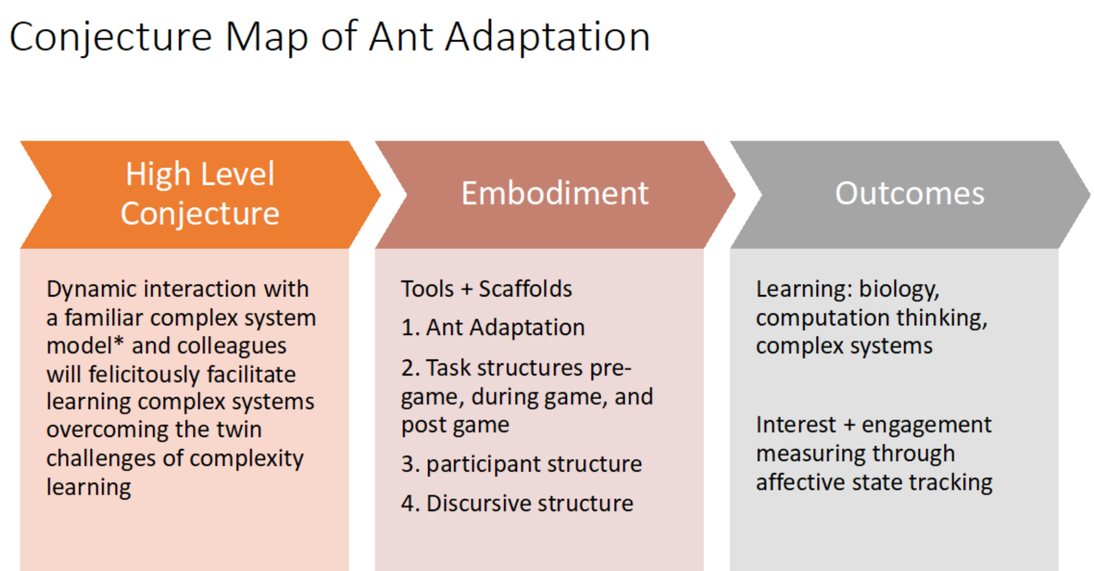

Figure 1: Conjecture Map of Ant Adaptation: the high level led to embodiments, which were designed to lead to outcomes

As seen in Figure 1, under embodiments, while I tested this system, the participant structure was changing due to social restrictions from the pandemic. Though I began by recruiting participants at the midwestern museum, I had to adapt to the onset of in-person social restrictions due to the COVID-19 pandemic, and subsequently I had participants play the game in an online, remote leaning environment where they were still able to freely contest and work out their ideas. The discursive structure was open-ended: there were no wrong answers, and each participant would play with a complex system in order to get to know the system, like a visitor to Paris would become familiar with French by being immersed in the language (Papert, 1980).

As seen in Figure 1**,** the mediating processes were just-in-time interactions with a dynamic digital artifact, through a digital tool, and conversations and competitions with other members of the group. Through these processes, the participants created artifacts: parameter sets about how the aggression level and size of their ant affected their macro-successes, or the distance to food sources. In other words, these sets reflected the participants' own understanding of how an ant colony works and what values matter in the success of the colony. I projected that this approach would have specific outcomes for participant learning in education in biology and ecology, through the instantiated mode of learning about Ant Adaptation; and Computational Thinking (CT), as they considered loops inside the simulation, algorithmic thinking, and self-organization concepts such as hill climbing algorithms. This all led to, and would scaffold, thinking about complex systems, feedback, and the success of the colonies as a result the success of the individual ants. In the last iteration, I have begun to look at the varied motivation of participants. To do this, I used engagement measures, such as affect tracking, word choice, and cognitive engagement in a group.

Table 1: Design and theoretical conjecture lead to the proposition.

Design Conjecture:	**Theoretical Conjecture:**	**Proposition:**
The design conjecture of the system is if learners engage with ant adaptation, and other players, through competitive inquiry of the system, then participants will engage through play, and learn about complex systems.	The theoretical conjecture of Ant Adaptation is if play occurs, we will observe more smiles, engagement and deep learning of complex systems.	The design conjecture can be stated as a proposition for Ant Adaptation. This proposition is as follows: a collaborative inquiry with an ABM of ants will lead to a clearer understanding of simple rules that lead to super colony, macro-outcomes. Following from this proposition, students will come to a greater understanding of these numerous overarching problems in humanity, such as climate change and social and economic inequality.

Sandoval (2014) brings up an important consideration of design-based research: commonly available tests are not appropriate measures of ambitious outcomes. We see these ambitious outcomes when we design to improve affective and cognitive engagement and measure them through novel approaches. Desired outcomes may not be very well conceived at the start of the design research project, and early cycle design research may be needed to clarify how these outcomes might be measured. This is the process I have taken in this project. In this project, I developed both the measures and embodiments in multiple iterations. I first looked at the shift of schemas while developing constructivist dialogue mapping to track the deep learning that I was initially interested in but did not have the tools to study. Further, I knew that participant play, and engagement mattered, though I did not have the tools to measure moment-to-moment engagements like joy, sadness, and confusion. In this way the development of the design-based research project that surrounds the design of Ant Adaptation has been fecund ground for researching these measures which I discuss in chapters 4 and 5.

3 Overview

I conducted a multiyear design-based research project with participants interacting with Ant Adaptation in multiple environments. I collected pre-post tests, interviews, video, audio, field notes, and screen captures to measure learning and engagement with the game. In this dissertation, I will first explore the theoretical underpinnings of this work in the literature review. In the literature review, I will start with the theory of constructionism and complexity and there use in teaching complexity, then introduce museum studies, focusing on concept elaboration around prolonged interactions with exhibits. I will then review the affective computing literature to focus on the means of measuring engagement through affect, and the limits of those measures. I will follow this with the introduction and design of Ant Adaptation, which builds on the constructionist learning environment design and museum studies literatures. In the description of the design of Ant Adaptation, I will highlight its key learning objectives and design rationale. Following that I will discuss methods of measuring engagement and learning that I developed including constructivist dialogue mapping, affective state tracking with sensors, and the triangulation of these measures. After the literature review, I present three more chapters on iterations of Ant Adaptation. Finally, I will present my conclusions.

Chapter 2, the literature review, is broken into four sections: in section 1 of the literature review, I will describe the history of constructionism, and how it has led to restructuration theory (Wilensky & Papert, 2010) through processes that better align with students' learning. In section 2, I review the complexity science. Then I review their promise and the challenges people face learning them. Finally in section 2.3.4, I will talk about constructionist learning environments that teach complexity—focusing on computer aided micro worlds of STEM subjects.

In section 3, I will introduce museum studies, specifically, active prolonged engagement, concept elaboration through talk as means to measure learning, symbolic interactionism to examine interactions around and between exhibits, and how I will design a tabletop game for interaction in a museum. This work also motivates my use of tracking elaboration in talk with constructivist dialogue mapping.

In section 4, I will look at affective computing. Affective computing shows many physiological measures that can be tied to human emotion. I will review the history of research into emotion and some of these measures. I will focus on two modes of measurement in affective computing (Picard, 2000), that is skin conductance and facial tracking using facial action units that each can be used to measure moments of high stimulation during an interaction. Then I will discuss what research says about the role of affect and intelligence and learning, measuring mood, use of these tools in research, manual coding of faces. In section 4.5 I review the methods limitations (Barrett, 2019; Crawford, 2021). I conclude by asking about whether emotion is the same as affect and the role of affect in human choice.

In section 5, I review the literature that supports the design of the microworld used in this project, Ant Adaptation. In 5.1 I cover learning about complexity with models and games. Then I discuss prior uses of ant-based modeling.

In section 6, I discuss to measurements collected to examine open-ended learning environments. This part review Piagetian clinical interviews and describes how engagement has been measured.

After the literature review, the dissertation has three chapters and a conclusion. Researching Ant Adaptation over the last four years through a design-based framework has taken on three iterations. I present each iteration in its own chapter. In chapter 3, I cover the design principles of this design journey, and my findings from testing at the Field Museum. In the first iteration of this part of the design-based research, I asked 114 people at the museum to play Ant Adaptation and discuss it for between 5 and 30 minutes. This does not account for the other 50 people who tested it during the design. The main findings shown in chapter 3 (Martin, Horn and Wilensky, 2019) are 1) participants could shift their understanding of an emergence in less than 10 minutes of interaction. The second iteration was kicked off when I realized the measures I was using to study learning were insufficient for the contexts I was investigating.

In chapter 4, responding to this limit, I develop the method of constructive dialogue mapping (CDM) to meticulously track what people were discussing and how they were discussing it as they played the game. During this phase, I developed CDM by recapitulating Piagetian (Piaget, 1926, 1929, 1952b, 1983) theory and differentiating it from other means of concept mapping such as dialogue mapping, narrative mapping, and the broader category concept mapping. I put forward the idea that CDM is more apt for the researcher to understand the ontological entities employed by the participant, than it is as a means by which the participant can understand the change themselves. The chapter shows CDM's utility in researching learning as concept elaboration through interaction with technology and a facilitator. This demonstrates two key parts of constructivist theory: 1) elaboration of discussion improves understanding and by tracking the paths of elaborations and its sometimes-dual conceptions we can observe conservation of concepts in transcripts and 2) knowledge fluidly grows through action taken in a complex system. The dynamic interaction with the model enabled users to learn through mediation with the computer system, each other, and a facilitator. From these findings, I suggest the method could be scaled to other learning environments and tracked through CDM. At the end of this phase I realized the work had focused too much on cognition. This is the cognitive view of learning (Gutiérrez & Jurow, 2016). Leaning on Gardner (Gardner, 1987), Gutiérrez and Jurow describe the view as follows: "In the Western conception, learning has historically been treated as an individualist, cognitive endeavor. The result of this has been a de-emphasis on the social world and how its organization affects the nature of problems and the way people deal with them" (2016, p. 593). And I had fallen into the trap of designing merely the cognitive measures. As a result, embracing computational ethnography and close measures, I developed a means to synchronously study how people affectively engage and learn, as they worked together in groups to play Ant Adaptation. This resulted in the 2019 paper Computational Ethnography, in which Bain, Wang, Worsley and I look at early measures to study how people feel as they learn with Ant Adaptation. This paper was the grist that led to the third phase of the project. In the last iteration, covered in chapter 5, I wanted to pull together the different strands to study each of these pieces.

In chapter 5, I re-implemented Ant Adaptation, holding most of the interviews remotely (due to restrictions on social interaction imposed by the coronavirus pandemic). I used video analysis to track how people were engaging as they played in the system to better leverage the regulation people employ as they learn and to study the pathways of affective states that best lead to their engagement with each other. Remote learning was theorized as an exploded museum, as in an exhibit beyond the walls of the hall. In this way, folks were invited to play like they were at a museum, as opposed to other informal learning arrangements. In this iteration, I also deployed constructivist dialogue mapping, interweaving and triangulating these measures to understand more about the person as an affective agent, as they come to understand the systemic and structural nature of the agent-based model, and hopefully become more historically aware agents themselves that are more equipped as change agents in their society.

In chapter 6, the discussion and conclusion, I explore how I investigated my core questions: How do pheromones work? What would you like to be different about the game? Why did you like the game? In what ways can a game be scientific? Do you know how ants collect food? How do evolutionary adaptations impact the ants, and her colonies success? How do ants know what to do? And how are you feeling while talking about this? These are questions to explore complex systems, self-organization, the design of a game to enhance that exploration, and the feelings of the user. At the center is a user centered approach to designing a learning environment. In each iteration of this project, I have sought to improve the design, and measure users cognitive and affective engagement with the lessons embedded in the digital tool and the discursive practices.

In the future work section of the conclusion, I situate the work going forward in social design research, and admit *I have not yet reached that goal*. This is a project of continuous iteration situated in the cannon of design research to develop a complex systems education game. As a result, the discrete chapters are merely markers along the way to the ultimate purpose of the design that is not yet fully realized. In the discussion I discuss how the restructurating has led to better complex systems learning, as previewed below in the summary of findings (see the end of this chapter). In that section, I show that short integrations lead to a better understanding of self-organization, they smile a lot while doing it, and the methods developed demonstrate this complexity learning more clearly.

A new frontier for design research is social design experimentation, introduced by Gutierrez and Jurow (2016). Social design seeks a process pathway to fundamental social transformation. As I seek to educate students about complex systems for them to become social agents that can affect the big changes mention at the beginning of this introduction (climate change, health crisis or income inequality), my design research is moving in this new direction. Social design research tries to transform work, rather than working inside the existing institutions. It seeks equity, climate change action, justice for the wronged, new forms of society, and better gender relations. In short, social design experimentation seeks to create more intentional historical agents that can manipulate the structures and systems that result in injustice by creating significant reorganizations of systems of activity in which participants become designers of their own futures— in my words, architects of their own destinies. We see this work influence new curriculum standards, such as California’s recent initiative in math education reform for equity

(California Department of Education, 2021). This reorganization of activity is seen as the best way to cause fundamental change in situations perceived as unjust by non-dominant communities. This reorganization will require that we support and create a critical mass of people who see themselves as historical actors capable as social and educated change agents. And while embedded contextual understandings of inequalities and injustice are crucial, I seek also to disseminate an understanding of the structures and systems within which we exist.

4 Long Term Goals and Outcomes

- The long-term goals and outcomes of this dissertation are as follows:
- To provide a social insect-based curriculum for use in informal learning environments, to help students understand complexity.
- An evaluation of that curriculum in terms of how well that approach works to help students understand complexity.
- Develop design principles for how to structure open-ended learning environments.

5 Research Questions

These three goals inform the following set of research questions. I answer these questions through qualitative and quantitative methods including clinical interviews, pre- and post-tests, video analysis, automatic facial coding, modeling and constructivist dialogue mapping. With these methods, in this dissertation I answer the following six questions:

- What do students learn about complex systems by engaging with a social insect-based curriculum?
- What process of iterative design can result in a system like Ant Adaptation?
- Are ants, as social insects, a good path into teaching about complex social systems? Do these thinking tools help them think about core global problems?
- Does the technique of CDM add new abilities to gain new insights into how learners, in group conversation, can advance their learning?
- What new can we learn from physiological measures of affective states, while people engage with a museum exhibit?
- Is there a relationship between high stimulation and learning as measured by CDM and affective computing signals? Within moments participants elaborated, is there a relationship between positive affect and learning?

My first and third question relate to, how students make sense of complex systems while engaged in Ant Adaptation. I will study these questions through qualitative methods. I code the interviews for moments of complex thinking. I will pay particular attention to the following to elevate the learning and design questions: did participants learn complex systems through their interaction? Did individual learning happen about complex systems? Do these systems overcome a deterministic centralized mindset? During these investigations I will also keep an eye on questions 5 and 6: do the multimodal learning analytics serve to figure out where and how the learning is happening? Does constructivist dialogue mapping help? Answering questions one and two, handled in chapters 3 and 4, involved looking at the type of learning the design of Ant Adaptation—and more generally social insects as examples of complex systems—affords. Complex system models have been used to teach several STEM fields, including material science (Blikstein and Wilensky, 2009), electricity (Sengupta and Wilensky, 2009), and evolution (Wagh and Wilensky, 2017). Following in this tradition, Ant Adaptation attempts to teach complexity science through an agent-based modeling game, Ant Adaption (K. Martin, Horn, et al., 2019). As a result, I investigate the learning affordances of the environment using qualitative coding of the interviews. I study a game in use in the context of a museum. In this environment, I focus on group learning, as players collaborate to advance in the game. Further, I have developed a technique, that attempts to ferret out the learning in the group, constructionist dialogue mapping. To answer my third question, I will research the affordances of the use of CDM in comparison to more traditional qualitative methods. The environment is designed as an open-ended environment, where players engage with ants and each other, with the target of getting insight into general issues of complex systems.

Answering question 2, requires to attending to the process of designing Ant Adaptation. As a result of the enmeshed nature of the social technical system (Hutchins, 1995) I explore this design throughout the dissertation to answer the question of what process of iterative design can result in a system like Ant Adaptation? In chapter 3, I explore the design of Ant Adaptation, which mixes the foundational agent-based learning environment work of environments such as GasLab (Wilensky, 1997) or NetLogo Investigations in Electromagnetism (Sengupta and Wilensky, 2009), but aiming to lower the interaction time to fit into the short interactions typical of informal learning. As I explore in Chapter 3, 4 and 5, when players compete around the application to contest ideas, a rich set of ideas emerges through the play. There are several design elements of the game that facilitate this interaction as I preview below in the Summary of Findings and expand on in the concluding chapter. But more than just the design of the game, creating the environment required attending to the discursive and social structure around the computational representation. Attending to a robust conversation, is an ongoing process.

Answering questions four, is about the efficacy of the tool of CDM. Answering question four and five in chapter 5 involves looking at how much students update their understanding during moments of high affect, as compared to other times while using Ant Adaptation. Theory shows that moments of high stimulation (D'Mello & Graesser, 2012a; McGaugh, 2003, 2004) are associated with learning. Affective states pathways, specifically the engagement-confusion-delight-engagement pathway, has been hypothesized to facilitate advanced problem solving if the confusion state does not overwhelm the learning (D'Mello & Graesser, 2012). Andres *and colleagues* (Andres et al., 2019) found while researching informal learning with *Betty's Brain* that the only emotional pathway associated with learning gains was sustained delight. Horn *et al.* (Horn et al., 2016) found post-test gains associated signifyingly with all of the affect words they considered such as "*wow*" and "*cool*" after a museum exhibit. Each of these uses of affective state tracking, point to greater integration of affect detection to improve design, interactivity and analytics in learning technology. Question four will take these findings one step deeper and investigate the process of learning in moments of stimulation and investigate the role of positive affect in that process of complex system's thinking. The proposed mechanism is as follows: moments of high stimulation are related to things people remember more. I want to know what leads up to those moments. Which would mean investigating if good design impacts positive affect or those learning moments. What is prompting positive affect? Is it related by the design or the way the social interactions play out? It is a system of mediation, but how is it connected? This work aims to underscore the connection between learning and affect.

6 Summary of Findings

During the study I have several key findings for the six research questions. Though I expound on these answers in the concluding chapter, here I provide a summary.

6.1 Top level findings

- Through playing Ant Adaptation, learners advanced in their conceptions of complexity during the period of the activity. The intervention was short duration and could be deployed in settings such as a museum.
- CDM is an effective method of studying knowledge as it fluidly grows in individuals and in groups. Therefore, it can be used in open-ended constructionist learning environments.
- Predictive models based on facial expressions are a good way analyze unseen video for moments of interest, such as moments participants elaborated as indicated by CDM.
- We can use random forest and XGBoost tree models to triangulate these findings to better understand learning environments

Next, I provide an answer to each question.

6.2 Question 1

For question one, *What do students learn about complex systems by engaging with a social insect-based curriculum?* I found students come to understand properties of complex systems by engaging with the curriculum. First, through play, Thomas learned that entities like ants have a mechanism such as laying trails to attract other ants to flowers in a cycle by recursively following the chemicals or wandering. As shown in chapter 4, Rebecca first identifies a phenomenon she does not understand, namely, purple tracks that she later understands are pheromone trails. Then through observation, she makes a conclusion that that phenomenon attracts other ants in a self-organizing foraging system. From this prediction she expresses its function, but maintains some confusion. This confusion is the affective state that accompanies her process of accommodation, that pushes her to account for greater and greater parts of the complex phenomenon. Thomas employed complex systems thinking to reach the goal he set for himself of maximizing population. He set this goal in communication with his teammate in the open-ended constructionist learning environment afforded by our design of Ant Adaptation. This group goal setting and exploration is key to the learning engendered from learning with this insect-based curriculum. One of the primary motivations of the learning environment was to teach that the simple rules of agents could lead to complex social patterns that sustain a population of ants. Thomas's descriptions indicates he came to understand the impact of ants' simple rules on complex social behaviors. Thomas also realized that sometimes this process can lead ants astray as shown by his use of vinegar to redirect them out of deleterious local optima. While he tries to reverse the deleterious effects of local optima, he did not learn the term. Students learn how self-organization happens. As Mar showed in chapter 5, Mar learned three functions while playing the game: 1) how ants group up, 2) collect food, and 3) use pheromones. Learning these functions allowed her to better understand how ants know what to do. Later in chapter 5, Bob and Mel used the system to understand two parts of an insect social system, 1) Mel established the need for equilibrium between colonies to not tip into instability, but

2) Bob found the need for an insect society to deploy various adaptations depending on a mercurial shift of an ecosystem. In this way, Bob argued for the use of adaptation in an evolving ecosystem. In chapter 3, users shifted their schemas from process to emergent schemas by using Ant Adaptation. For example, three participants in two groups changed their schema from direct to emergent. For example, Ed changed from a direct schema to an emergent schema after playing. Some people did not shift their schemas. Both Stacy, a seven-year-old girl in Group 20, and Pri, a sixteen-year-old girl in Group 27 held their original direct schema. Finally, when playing Ant Adaption, the learning is unidirectional: people do not shift from emergent schema to process schema by interacting with the complex systems about ants, it's a one directional learning pathway.

6.3 Question 2

For question two, *What process of iterative design can result in a system like ant adaptation?* I find that providing consequential decisions in the game creates the tensions that players explore and come to understand complex systems. For instance, in chapter 3 when Ed says "I don't want to be too aggressive" as the two teams stare at the touch screen, it shows a burgeoning realization that adaptations affect his ant colony's life cycle. This leads through the interaction to the younger brother, Thomas, teaching his older brother, Ed, that the direction of pheromones leading successfully toward food increases the ants' population through the process of feedback. This is an example of the discursive structures leading to better learning. Thus, players of the design discover the principles of complex systems.

Users appreciate in the design that you figure it out without guidance. While scaffolding at the start enables beginning the exploration, players want to be architects of their own knowledge. As Thomas says in chapter 3, this is why he liked the game: "Yeah, you had to figure it out and the–you have to have some flowers, see, and then you put the chemicals and lead it to there, then they'll bring it back, and like, if you want to get rid of the chemicals you use the vinegar." He says this as if he is teaching his siblings, telling them what they should think of the situation. Here he is taking on the role as a peer-teacher. In other words, the game is enabling the social properties of the restructuration as kids want to tell other kids about what they are finding, which is the beginning of a replication cycle.

Another key feature I found during testing NetLogo on a touch screen is that close timing of players' actions with changes, such as adding pheromones to the model, changing how ants move, and learning about pheromone trails may have accelerated understanding the complex system. In digital game design, it is argued that when actions and results happen within one frame, or approximately 10 milliseconds, connections between cause and effect happen better. We see this happening when players can touch the screen to interact with the complex system. This design principle seems important when designing future complexity learning games to hasten learning. In short, the design causes proximity to user actions and effects.

The design of Ant Adaptation facilitates identifying and describing learning like in earlier microworld work by surfacing the discussion between players. This talk reveals how players reconstruct provisional theories considering dialogue between theory and evidence (Wilensky & Reisman, 2006). The decision built into the main action of Ant Adaptation—whether to peacefully collect food to increase population by employing feedback cycles, or go to war to eliminate their opponent—sets up a crucial engagement where the uncertainties make the testing immediate and productive (D'Mello & Graesser, 2012a). As I present in chapter 3, the design also hastens the learning, compacting it into play sessions under 10 minutes. Therefore, as a design principle, we should design better ways to facilitate complexity learning where people engage part of a complex system, attempt their best theories in real time, and receive dynamic feedback from the computer and each other. These learning moments may happen most when players notice breaks where to get out of their confusion (D'Mello & Graesser, 2014a), learners must engage in effortful—intense, purposeful, psychological effort – and problem-solving activities (D'Mello and Graesser, 2012). The game in the museum had the following four perceivable effects: First, construction of their own colonies, in competition with an opponent, afforded comparisons, which allowed for dynamic theory validation and imitation. Specifically, Thomas placing ants close to the nest and drawing ants to them by laying chemical trails showed he understood proximity and the connection of flowers to the nest aided population growth. This is an example of learning to make micro- to macro-level connections through an agent-based model. Second, the other team copied his strategy. Sharing strategies allowed players to update their operating theory. Third, taken together, these scaffolds facilitate players' exploration and learning about the complex system. Fourth, within less than a quarter of an hour of play, the game facilitated one player to switch from a direct schema to an emergent schema during a conversation with his brother. Lastly, I find in Chapter 5 that designing for affective engagement creates a happy hypothesizing interaction where users smile as they contest these immediate and productive uncertainties. As we see in chapters 4 and 5, we see this process through affect tracking and follow this elaboration through CDM. In short, if we turn over knowledge authoring to the user of the game, provide enough scaffolding at the start, and allow for players to address uncertainties, they will learn how the system works. If that system is about powerful ideas such as self-organization and reasoning between individual agents and macro-scale outcomes like population growth, they will come to overcome the twin challenges of complexity learning, and understand those through immersion in the microworld.

6.4 Question 3

For question three, *Are social insects a good path into teaching about complex social systems?* I found in chapter 5: In terms of Bob and Mel's knowledge building we see the following six aspects: First, they took a key moment during the interaction to re-order their knowledge, demonstrating a shift from what they brought into the interaction, and what interacting with the game engendered. Second, the restructuring led to them explicating how ants know what to do and the role of flowers and ants interacting leading to population growth. Third, they explicated the rules of how ants know what to do through trail formation, and the connection between create cost and food intake. Fourth, learning about the complex system allowed users to describe their different goals in the game: Bob wanting to optimize parameters in a changing ecosystem, thus role-playing evolution in this system. Whereas Mel set herself the goal of discovering how to make the two colonies equal, expressing an equity lens. She set about running experiments to try to figure out how to achieve her goal. This different goal setting played a tacit, but crucial step, in their interactions. Fifth Bob and Mel learn how ants can control their own colony, grow the population and integrate into a complex ecology. Sixth this new information allows them to set their own objectives in the learning activity. This process was aided by their prior familiarity with ants and grew through interacting with the ant game. They didn't have to ask about ants, because they were familiar to them, and as a result, went about the task of understanding the complex system without thinking if they knew enough.

In terms of Mar's knowledge building, we see the following three aspects: First, Mar elaborated functions while she worked with Ant Adaptation. The three functions together describe an understanding of how an ant colony can collectively feed itself and grow the population, without the need for a central controller. Second, she explicated the rules of how ants group up, how ants pickup and carry food, and how the rules of how ants walk drives their ability to find new food, and exploit food patches they have already found using pheromone trails. This function expanded on her prior knowledge about ants by giving her a way to explain what she knew coming in, that ants walk, carry things, and group up. In short, I have evidence that Mar happily learned how ants can control their own colony, grow the population, and integrate into a complex ecology. Third, Notably, Mar did not share much about her goal setting, this might be because she played by herself, and so had less reason to explicate her goals. In short, by learning from the complex system of an ant social system, users learned deeply about the role adaptation, equity and ant behavior serve in the system. They were able to explore these interests in the sandbox mode the Ant Adaptation affords. Unfortunately, there is no evidence this helps them understand other core problems.

6.5 Question 4

For question four, *Does the technique of CDM add new abilities to gain new insights into how learners, advance their learning?* I found overall that CDM demonstrates learning as concept elaboration over time through the proxy of changes in speech. As shown in chapter 4, tracking elaboration in discussion improves researchers understanding of the learning. We see with Rebecca, we can observe conservation of concepts in transcripts more clearly. Using CDM researchers can capture visitor's moment-to-moment sense-making. In Rebecca's case, Moment-to-moment expressions stick together as they account for what Rebecca sees. She connects the declarative knowledge we probe with but when they no longer helped her understand the ants, she dropped the ideas. Knowledge fluidly grows through action taken in a complex system. The dynamic interaction with the model enabled users to learn through mediation with the computer system, each other, and a facilitator. As track by CDM, this dialectic interaction could be scaled to other learning environments. Effortful problem-solving activity is the process of science (Conant, 1947), and that is the process constructivist dialogue mapping tracks. The CDM approach introduced to capture the type of learning common in this project was able to capture changes in a player's understanding of agents, functions and properties (entities' mechanisms), while they learned complex ant systems through playful interaction with Ant Adaptation. CDM captured the changes as utterances occurring during a short interaction. By analyzing changes in talk pre- and post- play, we found that players learn about feedback, and employed that learning at multiple levels to maximize an ant population. As we will see in chapter 5, CDM affords tracking how people come to understand the functions—such as ants grouping up, gathering food, or laying pheromone trails—that drive agent-based models as they discuss the behavior of ants in the agent-based model. Finally, as we see in chapter 5, CDM affords identifying moments where players stop and then restate and reformat their knowledge, as we see 16 minutes into Bob and Mel's interaction.

6.6 Question 5

For question five, *What new can we learn from physiological measures of affective states, while people engage with a museum exhibit?* I found three aspects. First, in Mel and Bob's case in chapter 5, I found joint users display widely disparate readings of affective state intensity. Second, we can describe who does most of the talking. In multiple user cases, the ability to notice when users engage, and disengage is important. From the silence we can ask what Bob was doing during that time. Time series analysis makes that possible. Second, in all cases, we can describe the frequency of affective expressions. We can say quantitatively that Bob in the first half of the interview did not engage as much, became more ebullient and loquacious towards the latter half. This approach to studying affective engagement affords this insight.

6.7 Question 6

For question six, *Is there a relationship between high stimulation and learning as measured by CDM and affective computing signals?* Within learning moments, is there a relationship between positive affect and learning? In chapter 5, in Mar's case, the main emotional pathway around moments of learning was the delight pathways. In all cases, the primary affective expression was smiling. Moments of learning are accompanied by a lot of smiling. There is a positive correlation between learning identified by CDM and positive affective states. During Mar's interaction around learning moments, delight is the predominate emotional pathway, accounting for 29,490 of the pathways around learning moments. The next two biggest emotional pathways are contempt and contempt delight pathways, account for 2,713 occurrences. As Mar's sip of coffee highlighted, researchers need to pay attention: not all the moments expression analysis bring to light are salient.

To address this I performed predictive modeling, using facial expressions to predict moments participants elaborated. In chapter 5, through random forest modeling, I achieve a high the degree of predictive power using expressions to identify moments participants elaborated. For the first case, Mar's, getting 96% accuracy, 97% precision and 81% recall. For the second case, Bob and Mel's, I get 90% accuracy, 96%precision, and 73% recall. Additionally, using an XGBoost model I find support as well. Using XGBoost to identify moments of elaboration for the first case, Mar's, XGBoost identifies 97% of the 0s and 76% of the 1s, with a total accuracy of

93%, precision of 76%, and recall of 89%. For the second case, Bob and Mel's, XGBoost correctly identifies 84% of the 0s and 70% of the 1s, with a total accuracy of 79%, precision of 70%, and recall of 70%. With XGBoost top predictors of moments participants elaborated are smiling (AU06 and AU12). The findings show, using a train test split, and the extensive data provided by affect detection, we can develop algorithms to predict moments participants elaborated identified through CDM. If we can identify real time learning moments, these findings would allow us to offer recommendations and respond to the learners within the critical 10 millisecond response range discussed in chapter 3. Furthermore, with further affective analysis, with these methods researchers could identify moments where students get stuck and provide real time feedback in an instructive learning agenda. This opens up the possibility of doing the analysis at scale, as we develop more accurate models. Lastly, the process of synchronizing qualitative methods, and learning analytics provides an example of combining measures from multiple data streams to provide real-time learning analytics, and refocus on outcomes of learning such as learning engagement. This is a powerful means of understanding the impact of our learning designs. My process of dedicated design of a restructurated unit to teach complexity along with the innovation of learning methods is only just getting started, but the results are already promising.

Chapter 2: Literature Review

A large body of research has informed the study I am proposing for this dissertation including research on constructionism, restructuration, complexity, teaching complexity science, museum studies, designing Ant Adaptation for open-ended museum education, and the challenges of measuring learning in open-ended learning environments. In place of a lengthy review on the full body of work done studying informal learning and evaluation, I choose to emphasize the literature on the use of affective measures (Picard, 2000) to study people's responses, the design of digital interventions to restructure education (Wilensky, 2020; Wilensky & Papert, 2010), and on the use of carefully tracking talk in order to study learning around museum exhibits settings (Leinhardt et al., 2003; Leinhardt & Crowley, 1998; Vom Lehn et al., 2001). I begin this section by introducing constructionism, a pedagogy. Then I recapitulate the literature on representations and learning in a constructionist paradigm (Wilensky and Papert, 2010), as it guides the discussion of the design of digital learning environments to teach about complexity in educational settings. Then I will review complex systems and provide examples. I will discuss museum studies to situate the work in the context of the study. Then, I will review affective computing and justify my selection of measurement modalities. I will introduce the theory behind the measurement modality constructivist dialogue mapping (CDM) and discuss knowledge in pieces (diSessa, 1993, 2018) and the node-mode framework (Sherin et al., 2012a), emphasizing their connection to CDM.

1 Constructionism

Constructionist thinking influences learning sciences and educational research, particularly when addressing learning technologies and reform of mathematics and science education. Papert coined the name constructionism, as a mnemonic (Papert, 1986), to describe a species of constructivist thought (Piaget, 1983). Four aspects are crucial to the design of Constructionist learning environments: learning through the activities of designing, personalizing, sharing, and reflecting (Brennan, 2015). These aspects are key, because Constructionist environments focus on the benefits of learning from the external construction of an artifact beside the internal construction of a mental model, or framework that is then shared to a wider community. For example, Logo, the name derived from the Greek for *word* or *thought,* was the first constructionist programming language. Wally Feurzeig and Seymour Papert invented Logo in 1967 (Logo, 2015). Using tools like Logo, constructionists organized learning environments where learning was free from time constraints and highly collaborative. Papert coined the term mathetics to describe the art of learning, and argued “My mathetic point is simply that spending relaxed time with a problem leads to getting to know it, and through this, to improving one’s ability to deal with other problems like it” (Papert, 1996). This relaxed time was meant to be spent in a community of other learners, creating a Constructionist learning environment.

Constructionism burst into the public eye after Papert published his seminal work *Mindstorms: Children, Computers and Powerful Ideas* (1980). Papert used Logo to operationalize many of the big ideas described in *Mindstorms,* using the idea of the turtle. The turtle is a single software agent that can represent many different organisms, such as a turtle in sheep’s clothes. Constructionism has inspired many powerful tools for education including NetLogo (Wilensky, 1999), a multi-agent programming environment, and Scratch (Resnick et al., 2009), a blocks-based programming environment. These environments have been used to make powerful mathematics and scientific exploration tools that afford learners the ability to learn as a French student in France, immersed in the subject (Papert, 1980).

1.1 Restructurations

One recent outgrowth of constructionism is Wilensky and Papert's theory of restructuration (Wilensky, 2020; Wilensky & Papert, 2010). A restructuration (Wilensky & Papert, 2010) is a new way of doing something that affords additional abilities. For example, Wilensky and Papert discuss the advent of Hindu-Arabic numerals. In that transition the Hindu-Arabic numerals supplanted the romans numerals, first in accounting, but later through computation because they were easier to perform operations with. Wilensky and Papert position the process as follows: "Hindu-Arabic numerals were not invented with an educational intent. But they could have been, and that allows us to show the need for a new branch of the learning sciences with the mission of understanding, facilitating and even designing shifts similar to the shift from Roman to Hindu-Arabic numerals" (2010, p. 2). We could call the new branch of learning sciences 'the science of designing new better ways of thinking embedded in our tools.' But Wilensky and Papert present a better name— namely structuration. They define the term as "the encoding of the knowledge in a domain as a function of the representational infrastructure used to express the knowledge" (2010, p. 2).

In a restructuration, Wilensky and Papert argue there are 5 properties researchers should attend to: power properties, cognitive properties, affective properties, social properties, and diversity properties.

Powerful Ideas: A restructuration must afford what could be done before, but for it to spread, it also must be able to do more. In the case of Hindu-Arabic numerals this is the power of multiplication and division massively expanded. When we open education for students to explore computational microworlds on their own recognizance, this is a powerful example of what could be done before — understanding physics and biology for instance — but with a huge new set of ideas in complexity, emergence and self-organization. This notion undergirds the complex systems as it lowers the floor for exploration of these multiagent, self-organizing systems.

Cognitive properties: a successful restructuration is more easily learned, while still providing the augmenting power of the old tools. In the current case, this is the idea of understanding ecology, while adding the ability to explore the role of emergence and complexity in how ant colonies ecology sustains.

Affective properties: a new structure can be more or less engaging. Computation media offer one example of more engaging properties. One question is how to measure this engagement, a field the second component of this dissertation takes a deep dive into.

Social properties: As a structuration is more shareable, it spreads through a group more readily. In the case of Hindu-Arabic numerals, their use in business made them highly sharable. In the case of Ant Adaptation, I have tried to design it to encourage folks to want to draw friends and family into its use. The properties of this sharing are key to the restructuration's spread.

Diversity Properties: with the power of design and iteration, a restructuration spreads, and as it does so, it matches people's styles of use ever more closely. For instance, the first hammer — likely a rock —had few design elements to match individual's needs. It was heavy, had no handle, and had few options designed for different styles of use. But now — 10,000 years later or so — you can find a hammer for a plethora of use cases, in many sizes, colors and weights. The structuration came to match people's styles of hammering.

In short, Wilensky and Papert argue novel restructurations have powerful impact on learners. They describe the move from Roman numerals to Arabic numbers, operations such as multiplication and division became significantly easier because of the new representation's affordances. Importantly, this power comes as the community wide literacy with the tool, reduces the difficulty of performing certain sums. This results from the fact that the various new representations are the focus of both internal reflection and external action that foster shared meaning, positively mediating groups' sense making about actions taken in the world. These restructurations can move science education past rote memorization of the traditionally identified core elements of disciplines to study (Clark et al., 2009). They argue well-designed digital games and simulations help learners build accurate intuitive understandings of the concepts and processes embedded in the games "due to the situated and enacted nature of good game play" (p. 3). These types of games provide excellent objects to use to think about scientific concepts and processes (Holbert & Wilensky, 2019). It is a version of these deep restructurations that I study in this dissertation, as they engage learners and shift their thinking in interaction with each other, and Ant Adaptation.

This restructured education will better align with students of the future. Papert (1993) argued that the advent of the restructuring of digital worlds that children can explore in computers will create less patient, accepting students. "Children who grow up with the opportunity to explore the jungles and the cities and the deep oceans and ancient myths and outer space will be even less likely than the players of video games to sit quietly through anything even vaguely resembling the elementary-school curriculum as we have known it up to now" (p. 9). As a result, Ant Adaptation restructurates the learning of complex systems to make it more intuitive, play-based, and occurring in a group.

To describe the system Wilensky and Papert introduce three examples of restructurations: the turtle, agent-based models, and the tick model.

1.2 Agent Based Models

Imagine an ant. She opens her eyes and looks around, she sees many of her nestmates nearby, and she is hungry. So, she starts to wonder around, and look for food. When she finds some, she has a bite, and then brings the rest back to the nest to feed to other workers along the way and babies in the nest. Here we have told a story about an ant collecting food.

Now imagine a lot of ants who each follow these same rules to feed their nest and provision their young. By creating these “agents” in a computer, that follow these rules, we can model this situation. NetLogo (Wilensky, 1999b) is a modeling and simulation platform for this kind of modelling called agent-based modeling (ABM).

Agent-based modeling (ABM) is a powerful structuration that has emerged from complex systems theory (Epstein & Axtell, 1996; Grimm & Railsback, 2005; Wilensky & Resnick, 1999). Agent-based modeling use simple computational rules as the fundamental modeling elements, in contrast to more traditional mathematical modeling. When modeling with equations we observe a phenomenon and try to fashion an equation that fits the observed data. In the Lotka-Voltera equations, used to model the change in predator and prey population levels over time (Hastings, 1997), the core elements of the model are variables that refer to population level parameters. In the Lotka-Volterra equations, the core elements are H, the population level of the hare prey; L, the population level of the lynx predators; and K, the interaction constant that describes the average predation. To understand the state of the system at a future time *t*, you solve the equations at time *t*. In contrast, as shown in Figure 1, in the agent-based modeling game, the core elements are computational objects or “agents” that represent individual agents, such as ants. Each of these agents has variables that describe its state including, variables such as age, energy level, hunger, or location. The behavior of the agents is determined by the computational rules that tell each agent what to do at each "tick" of a clock. As the modeler, we frame the rules from the agent's perspective. For example, if the agent is an ant, the rules might say: each tick, move a step in the direction you are headed; if you find food, pick it up, if you have food, return it to the nest, if you are returning to the nest, leave a trail other ants will follow, if not turn randomly, etc. To determine the state of the system at future time *t*, you run the system for *t* clock ticks. This is a set of simple rules, but it also is a complex system.

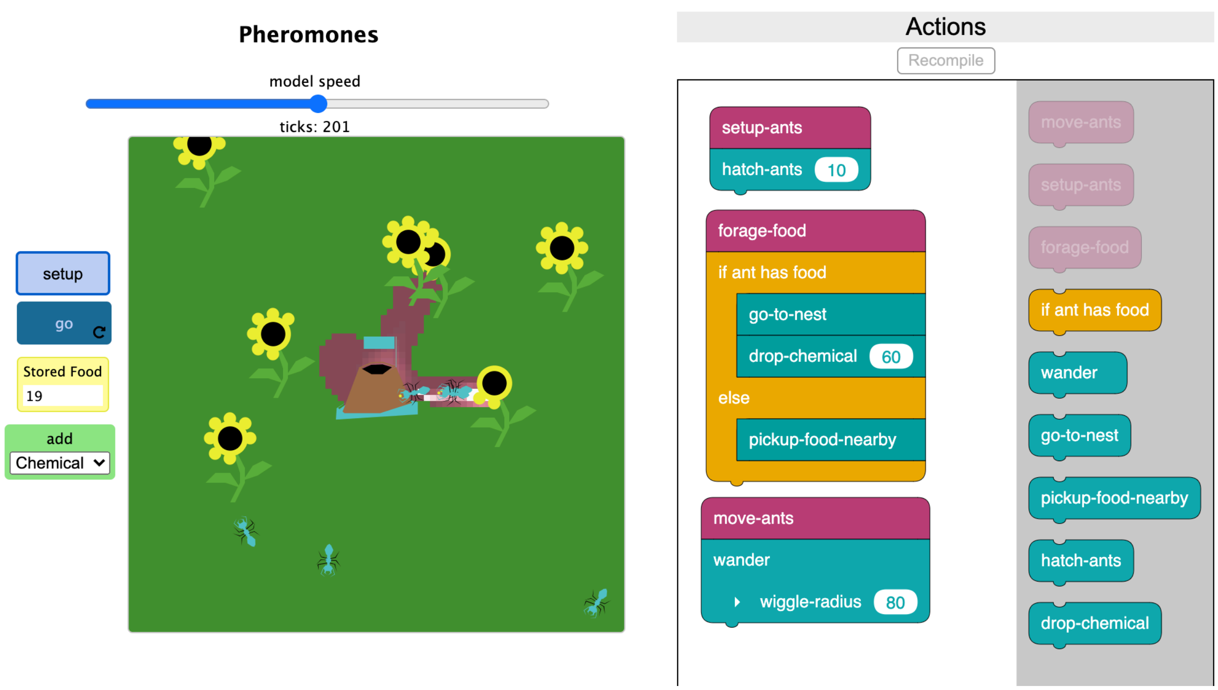

Figure 2: Ants collecting food in Ant Adaptation, following the simple rules of forage-food and move-ants. try it yourself: https://antomology.netlify.app/

2 Complex Systems

Since Weiner's definition of cybernetics (Wiener, 1948) and von Neumann's early work on cellular automata (Von Neumann, 1951), researchers have studied systems in which simple parts organize themselves into complex wholes. Self-organization has been defined as making meaning out of chaos (Atlan, 1986) the integration of elements perceived as disorder into more encompassing, often larger, organizations. These are important structures for understanding science and society. We see these systems in many structures, including minds, social insects, crystals, flocking, climate change, income inequality, evolution, and many other domains. These complex systems also touch some of the deepest issues in science and philosophy, including randomness vs determinacy and order vs chaos (Wilensky & Resnick, 1999).

2.1 Decentralization in Scientific Models

For 300 years Newtonian physics dominated the world of science and even more so people's perception of science (Davies, 1989; Morin, 1992). Newton offered an image of the universe as a machine, a clockwork mechanism. Newton’s universe is ruled by linear relationships of cause and effect. A gear turns another gear, and this creates linked causal chains. In the Newtonian world mutual interaction is not emphasized. When we think of interaction and the Newtonian universe, we think of one object acting on another object. One object is in control. For example, Jupiter controls its moons. The other is acted upon. In that view, the moons do not control Jupiter. Most of the focus goes to two laws of motion with a focus on how force influences the motion of objects (Resnick, 1994). Less attention goes to the intersubjective action, underlined in Newton's third law, that focuses on the reaction that accompanies every action. In the 20th and 21st century, the Newtonian worldview has been challenged on many fronts.

One of the more serious challenges comes from the growing interest in this kind of thinking called complex systems. In many fields, scientists have shifted metaphors about the nature of science, viewing the clockwork mechanisms as antiquated and instead employing a metaphor of complex ecosystems (Bar-yam, 1999; Capra, 2016; Davies, 1989; Laughlin, 2005; Morin, 2008; Pagels, 1989; Wolfram, 2002). Davies describes a move away from the classical scientific worldview toward complexity: "The new paradigm emphasizes the collective, cooperative, and organizational aspects of nature; its perspective is synthetic and holistic rather than analytic and reductionist" (1989, p. 2). Rather than looking at the world through one agent or an individual acting on another in a neat causal chain, researchers view the world in terms of decentralized interactions and feedback loops and looking at how complex emergent patterns emerge from interactions among these simple components.

Because thinking about complexity is difficult, this encourages reductive cause and effect thinking. poses a problem. Trying to think about complexity poses a problem. When we need to talk about structures that emerge, we oftentimes throw up our hands saying it is too difficult to study. In machine learning, for example, when we talk about node structures in a neural network, we often discuss the black box and focus on what it outputs, because it is difficult, if not impossible to inspect. This focus on emergence can sometimes feel almost pseudo religious or mystical. Throughout his career Morin proposed means to study these social interactions trying to abolish the separation between studying humans and nature (Montuori, 2004; Morin, 2008), arguing children need to come to understand complexity (Morin, 2001). This approach does not want to eliminate causal thinking, nor does it argue that the Newtonian worldview of causal change is wrong. Rather, the models are simply inappropriate for trying to make sense of certain types of phenomena (Bonabeau et al., 1999; Gause, 1934; Resnick, 1994; Wilensky & Rand, 2015).

Therefore, we need new models, to operate at a different level from the Newtonian physics, focusing on the behaviors of groups of individuals, not the actions of individual agents. This echoes

Wolfram's *New Kind of Science (2002)*. The book contains a systematic, empirical study of computational systems such as cellular automata. Wolfram defines these systems as *simple programs* and argues that the methods and scientific philosophy for the study of simple programs are relevant to other fields of science. In the book, Wolfram takes a programmatic approach. He takes a sequence of simple programs and runs them to see how they behaved. To his surprise, what he found, despite the simplicity of the programs, was that the behavior was far from simple. He argues this is a profound moment in the history of science.[2] The work to explore these processes and understand how that reshapes our epistemology are exciting directions in the future of complexity research. Below I present a few of those explorations to illuminate the work of complexity.

2.2 The Mind: A Complex System

These ideas that we need to look at, groups and interactions, have also influenced our thinking about how the mind works. This shift in thinking affects the methods of this dissertation. I am looking at decentralized units of thought inside individuals and how they share between groups using the method of constructivist dialogue mapping and affective state tracking. The change in thinking about how minds work underpins a fundamental change of worldview that has occurred in the last half century. Descartes argued, *Je pense, donc je suis (Descartes, 1637)*. His book marked a major cleavage: he saw two worlds: one of objective, scientific knowledge, and the other

[2] "It took me more than a decade to come to terms with this result, and to realize just how fundamental and far-reaching its consequences are. In retrospect there is no reason the result could not have been found centuries ago, but increasingly I have come to view it as one of the more important single discoveries in the whole history of theoretical science. For in addition to opening up vast domains of exploration, it implies a radical rethinking of how processes in nature and elsewhere work." (Wolfram, 2002, p. 2)

subjective constituted by internally known, reflexive knowledge (Morin, 2008). The latter is world made of soul, spirt, feeling and literature, the former the world of science, technology, and math. This division sets up an idea of the self as the part that does the thinking. Descartes so clearly demonstrated that while I can doubt, I cannot doubt that I am the one doubting. Consequently, I am the one thinking (Morin, 2008).

Though nothing seems more obvious to most people than the singular nature of mind and a unitary understanding of the self, we experience self-consciousness. In the book *Descartes' Error* (1994), Antonio Damasio aims to restore our appreciation of the role of the rest of body, and the shared balance of powers from which we emerge as conscious persons. As Morin argued in On Complexity "We human beings know the world through the messages our senses transmit to our brains" (Morin, 2008, p. 62). We feel like we have a single unified presence in this world. However, as Donald McIntosh pointed out: "The idea of self-control is paradoxical unless it is assumed that the psyche contains more than one energy system, and that these energy systems have some degree of independence from each other" (1969). If there was a single unitary self, there would be no problem of self-control. In 1968, Heinz Von Foester indicated the paradox of self-organization: "Though self-organization obviously signifies autonomy… a self-organizing system is a system which must work to construct and reconstruct its autonomy, and this requires energy." The philosopher of mind Daniel Dennett (1993, 2017) argued each of us imagines our own mind as *I*, not a collection of *we*. The idea of a unified centralized mind has fallen by the wayside in the past century, with accelerated motion over the last three decades. Resnick describes in his book (1994) that people began to study the mind from a multi-agent, complex system perspective. Freud's construct of the subconscious, which obviously could be used to explain issues of self-control, challenges the idea of a single executive in charge of the mind. Freud argued the unconscious is an equal participant with the conscious in the workings of the mind. Freud is not saying the subconscious is a repository of forgotten ideas that comes up merely in dreams, but is a lively agent actively organizing thoughts. He further subdivided the mind into the ego, the super ego, and the id, which pull the ego in different directions. Objects relations theory proposes a collection of inner agents within the mind that primarily seek relationships. With that view, the self emerges from interactions among internalized objects.

Marvin Minsky and the field of artificial intelligence in the 1980s studied the mind from a perspective very different from that in psychoanalysis, but it too moved towards a studying the mind as a complex system 1980s (Minsky, 1988). Minsky, a pioneering computer scientist, developed an elaborated theory of how the mind works, describing it as composed of many computational agents that comprise a society of mind. In the middle of the 1980s, and again in the 2000s, there was a renewed enthusiasm for neural networks. In Minsky's book, *Society of Mind* (1988) an image of society of mental agents working together to do things that no mental agent could do on its own is put forward. Minsky's later work argued that intuitions, emotions, and feelings are not distinct, but instead different ways of thinking. Through examining these different forms of mental activity, Minsky says, we can explain why thoughts sometimes takes the form of carefully reasoned analysis and at other times pivot on emotion (Minsky, 2007).

Philosopher and cognitive scientist Daniel Dennett has argued for a model with multiple drafts of consciousness. The idea builds off the Uexküll and Sebeok work on *Umwelten*, usually translated as the self-centered world view, of the 1960s. Dennett's work argues that there is no single stream of consciousness in the mind. Instead, consciousness is a manifest image (we) construct. To understand Dennett consider the mind like multiple news feeds on Twitter or Facebook. Multiple narratives are created and edited in different parts, the dominant ones are trending and lead to consciousness at a particular time. The idea of a single stream of consciousness, Dennett argues (Dennett, 1993), implies a single functional unit, a homunculus, that controls it all, where it all comes together, like on one news feed. He says, such a centralized narrative does not exist. Furthermore, relying on the ideas of a special center in the brain is the most tenacious, bad idea, bedeviling our attempts to think about consciousness.

The field of quantum mechanics offers a lens that facilitates the shift from this bad idea. Quantum mechanics is a fundamental theory of physics which provides a description of physical properties of nature at the scale of atoms and subatomic particles. Quantum mechanics arose to describe nature smaller than the macroscopic (ordinary human) scale. Quantum mechanics has three key differences from classical mechanics: 1) that quantities of a bound system, like energy and momentum are restricted to a discrete value; 2) objects can have characteristics of both a particle and a wave (wave-particle duality); and 3) the uncertainty principle that there are limits to how accurately a value can be predicted prior to measurement. The third difference is the principle of *indeterminacy.* This idea has become a key metaphor in the study of the mind as a complex system.

The advent of the field of quantum mechanics brought with it new metaphors for thinking about the mind. For example, Lambert-Mogiliansky and Busemeyer (2012), colleagues of Nobel laureates in economic sciences, Kahneman and Tversky, formalized this situation in their paper *Quantum type indeterminacy in dynamic decision-making: Self-control through identity management*. Indeterminacy is like an ambiguous picture. For instance, the duck rabbit, in one way you look at it, it is clearly a duck, but by another, it is rabbit. The tricky part is it cannot be both at the same time.

Indeterminacy has a role in thinking about societies of mind and societies of people. The need to identify the levels clearly, mapping where minds end, and individuals form societies becomes a key task. Both societies of mind and societies of people recursively reproduce through a process of indeterminacy. Heisenberg (1942) and Bohr (1991) recognized the similarities between human societies and quantum mechanics. The fundamental similarity is that in both human societies and quantum mechanics: “the object of investigation cannot – always – be separated from the process of investigation (Lambert-Mogiliansky and Busemeyer, 2012 p. 98). For Lambert-Mogilansky and Busemeyer, this makes it fully justifiable to study human behavioral phenomena through the mathematical formalism of quantum mechanics. Such investigations hold out the potential to discuss the effects of multiple fields’ effect on multiple selves within the individual that each are vying for their own utilities. In other words, we can discuss the limits imposed in multi-voiced situations, and theoretically, the interplay of multi-self-individuals’ interactions in one field, and potentially multiple fields. In other words, It’s a recursive process. The multi-self produces the experience of the self. Collections of these selves produce society. Societies, and the context in which the individual lives, produces the multi-selves.

2.3 The Ants: A Complex System

Morin argued that complex systems can be difficult to understand because of paradox: they are recursively created (Morin, 2008). He argued in a recursive process the effects and products are required by the process that creates them. Ant colonies are also recursive processes. For example, in ants, like other social systems, the colony as an organizing and organized superorganism, feeds back to produce individual ants through social learning (Franklin, 2014) and colony specific ways of acting (Hölldobler & Wilson, 2009). These individual ants through their interactions, produce society, that then produces future individuals that reproduce it. This is a recursive process. Through organizing this production, in interaction with the environment, the colony organizes itself, it maintains its tunnels, and clears the detritus of social life. If necessary the colony repairs itself. And if the food foraging goes well, and ant eaters do not intervene, the colony develops as it develops its production. As Morin highlighted the role of regeneration, "There is no recipe for equilibrium. The only way to fight against degeneration is permanent regeneration, in other words, the aptitude of the whole of the organization to regenerate, and to organize itself by facing all disintegration" (2008)

Early work on agent-based modeling was inspired by the behavior of social insects (Langton, 1997; Resnick & Wilensky, 1992, 1993). Ant behavior has inspired games. SimAnt (McCormack & Wright, 1991) is based on Hölldobler and Wilson's (1990) *The Ants*. The collective, self-organizing, recursive behavior of ants has been simulated using agent-based models many times. Starting in the early 90s. StarLogo was used to model the collective behavior of social insects (Resnick & Wilensky, 1992, 1993). Then, Wilensky (1997b) modeled food source preferences resulting from pheromones. Meanwhile, Bonabeau investigated the role of agent-based model in pattern formation (Bonabeau, 1997). The team more broadly studied swarm intelligence, the study of self-organizing systems (Bonabeau et al., 1999).

On the biologically side, Pratt (2005) modeled collective nest selection of *Temnothorax albipennis* also using an agent-based model and Sumpter and Pratt's joint work explored the importance of group decision making with quorums (2009). Their work showed that when choosing a destination together, cooperation reduces the probability that an individual will suffer predation. Robinson, Ratnieks, and Holcombe (Robinson et al., 2008) used an agent-based model to explore attractive and repellent pheromones in pharaoh ants.

The biology inspired work in computer systems. Frameworks, such as Anthill, have been used to support the design, implementation, and evaluation of technical systems, such as peer-to-peer networks (Babaoglu et al., 2002). Their work drew on examples of complex adaptive systems to justify engineering and user applications because complex adaptive systems exhibit resilience, adaptation, and self-organization that are seen as valuable in social applications. Work on ant optimization models helped develop the routing of information packets that supports the internet ("Ant Colony Optimization Algorithms," 2021).

2.4 Complex Systems for Education: promise and challenges

A central claim of this dissertation is that Ant Adaptation (Martin & Wilensky, 2019) helps people understand these complex systems. Understanding science and society in the twenty-first century requires students to understand conceptual perspectives about complex systems that arose in the field of complexity (Bar-yam, 1999; Goldstone & Wilensky, 2008; Jacobson & Wilensky, 2006; Morin, 2008; Wilensky & Jacobson, 2014). That being said, a review by Perkins and Grotzer (2000a) found students tend to explain complex phenomena with simple causal explanations, gears turn other gears. There are well documented examples that people resist conceiving of the world in self-organizing terms (Resnick, 1994; Wilensky & Resnick, 1999). It's not just a lack of knowledge, or information, but there is a cognitive bias, that we need to complicate with new models and explanations.

Learning about complex systems can be difficult (Chi et al., 2012; Jacobson et al., 2011; Wilensky & Resnick, 1999). Researchers have documented the difficulties learners have in understanding complex systems. For example, in interviews conducted in the 1980s and 1990s, Wilensky and Resnick describe people's resistance to ideas of explanations of self-organizing systems (Resnick, 1994; Resnick & Wilensky, 1992; Wilensky & Resnick, 1999)

Before proposing designs to ameliorate the difficulties, here I outline two threads: first, the promise scholars' see in teaching complexity, and second, I describe the scholar's views on the difficulties humans have in learning complexity.

2.4.1.1 Promise of Complex Systems Learning

Our world is full of complex interactions that self-organize. For instance, in the introduction to *An Introduction to Agent-Based Modeling* Wilensky and Rand (Wilensky & Rand, 2015) ask us to consider an ant. The ant wakes in the colony, goes outside, wanders randomly until it finds food. When she finds food, she leaves a trail as she returns the food to the colony, bringing energy back to the colony. When the next ant comes out, she follows the trail to food. As ant after ant repeats the loop of following simple rules the colony gains food. Additionally, a trail network emerges that the ants move along. Each ant, following simple rules leads to a complex, emergent pattern that feeds the colony without any central control. Resnick introduces his 1994 book, *Termites, Turtles and Traffic Jams,* by discussing the choreography of flocking birds. While many people assume bird flocks are led, he describes the coordination emerging. He asks, how did the movements become so orderly and synchronized? There is no leader-bird. Rather the flock is an example of what people call a self-organizing system. Each bird follows simple rules, reacting to movements of the birds nearby it—namely, distance and proximity functions. Orderly bird flocks emerge from these simple, local interactions. Bird flocks and ant colonies are not the only thing that follow these rules, minds, along with market economies, traffic jams and immune systems follow patterns not set by some centralized authority of a queen or president, but local interactions among several centralized components. In ant colonies, as ants forage for their food, the trail patterns are determined not by the organizational machinations of some Machiavellian queen plotting the colonies every move, but the local interactions among thousands of self-organizing workers. Difficulties in Learning Complex Systems

Starting with interviews that Wilensky and Resnick conducted in the 1980s and 1990s, they describe people's resistance to ideas of explanations of self-organizing systems (Resnick & Wilensky, 1993, 1998; Wilensky & Rand, 2015; Wilensky & Resnick, 1995, 1999). During the interviews, Wilensky and Resnick met with people from diverse backgrounds and asked them to explain emergent patterns. The results suggested a cognitive pattern called the deterministic centralized mindset (DC mindset). The pattern emerges from two empirical findings. First, people do not see the role of randomness in creating structures they observe in the world. Instead, the participants describe randomness as a destructive force to patterns, rather than a force that creates patterns. Second, subjects describe patterns that arise from the actions of a single centralized controller or leader, like a conductor in an orchestra. When asked how birds got into formation, interviewees typically responded the leader bird in front guides his lieutenants, or a mother bird in front tells her children what to do, or perhaps the biggest bird pushes back the air and the next strongest follows. Further, when asked to explain traffic jams, interviewees hypothesized that there was perhaps a broken bridge, radar trap, or even an accident that was causing the problem ahead. All these ways of understanding complex systems reflect a DC mindset where something is in control. Ideas revolve around leadership: a leader that orchestrates, agents moving into formation because of orders, or some specific centralized cause driving the effect. Further interviewees saw these as deterministic following on bias noted earlier to casual explanations of a complex systems. Counter to these intuitive understandings most traffic jams arise from random entry and exits of cars onto the freeway. These merges result in a statistical distribution of cars and speeds. For example, Patterns of traffic on the 5 in Los Angeles emerge from the interactions among drivers in individual cars making local decisions. Furthermore, birds do not stay at the apex of a V formation; rather, the formation changes over time, when the current leader drops back and another bird flies in the front. The formation emerges.

Second, in further analysis Wilensky and Resnick identified key components of this DC mindset that pose obstacles to thinking — namely the need to think in levels. Emergent phenomena can be described at two levels: at the level of the individual and at the level of the system (such as the colony of ants). Many people fail to distinguish between these levels, and instead slip between levels to give properties from one level, such as the individual, to the entire colony. This ability makes people think the stability is in the flock, whereas birds move in the formations, as individuals. People persistently account for the organization they observe by assuming a leader is in charge, even when the system is self-organizing.

This description persists, even when the system is self-organizing. Interviewees have a bias to describe complex phenomena in terms of leaders (Resnick, 1994). For example, Resnick writes of Marvin Minsky walking into his office at MIT and watching slime molds on a monitor and saying immediately, this is not self-organizing. There must be localized food that is drawing these slime molds together. Even when Resnick explained, Minsky persisted for some time. He also describes two high schoolers creation of ant cemeteries, and their resistance to the use of self-organization in modelling ant behavior.

These biases are not just for inexperienced learners. Wilensky and his colleagues (Jacobson & Wilensky, 2006; Sengupta & Wilensky, 2009) have shown that even expert researchers find emergent levels difficult to understand. In other words, it is not just a misconception of the scientifically naive. Instead, it seems to affect the thinking of nearly everyone.

Perkins and Graezer (2000b) argue that in contrast to everyday lived experiences of people, many scientific models involve more than causal explanations of ordinary events. These models introduce invisible entities like rule systems such as Ohm's law that govern the global behavior of systems, electrons, and large-scale patterns of action that are "emergent" from small-scale interactions, as with the gas laws. Wilensky and Resnick (1999) describe the difficulties people have in “thinking in levels,” exhibiting “levels confusion” and difficulties with distributed control and stochastic processes. Other researchers have noted that not only do learners have a hard time thinking across multiple levels such as disease of the whole body resulting from microscopic pathogens, but they tend not to think about phenomena such as the flow of ink dropped into water as the processes of collectives of agents interacting (Chi et al., 2012). If the glass of water changing color is explained by the individual parts of ink interacting with H2O, then the process becomes more intuitive. Coming from a non-constructionist background, Chi et al., (2012) argued that most people, are not familiar with ink particles. These two points of view have created two strands in the literature (Sengupta & Wilensky, 2009). Next, I will describe each of these perspectives.

2.4.2 Emergent vs Process Schemas

For Chi et al. (2012) processes, can be categorized in two ways – sequential or emergent. Sequential processes can be subdivided into a sequence of events, like an assembly line where the metal is rolled, pressed, stamped, and smoothed before being turned into cans and filled with tomato sauce. This is a process with multiple agents: the machines, their operators, and a manager. It makes sense to speak of sequential processes resulting from a single agent’s actions. For example, we could say the manager increased the efficiency of the assembly process by increasing the speed of the conveyor belt. Even though all the agents participated in the process, we can focus on the goal-setting manager's actions to account for the change. As a result, Chi et al. (2012) say one can give special controlling status to the agent that caused the change in a sequential process. This means if someone thinks of a process, like ant colonies, as sequential, people oftentimes assume interactions at the agent level are done to reach the goal of the higher level, whether that be producing cans of tomato soup at the managers, or the queen ant giving orders. The causal mechanism that leads to the result comes from the summing of the assembly line's outcomes directed by the manager. Many people confuse this control when thinking about emergent processes.

Emergent processes, like ants searching for food, marching in orderly paths, or getting stuck in a doorway are slightly different. These processes result from each ant taking actions and some of these actions are random. The results emerge from the repetition of the action, but no agent is in control. These processes are encountered in school standards such as osmosis and diffusion, electrical current, and buoyancy.

Chi and her colleagues argued that misconceptions about emergent phenomena in general result from an incommensurability (Chi, 2005; Slotta & Chi, 2006) or incompatibility (Chi et al., 1994) of amateur and expert schemas. Chi et al. (2012) argued that all processes can be categorized into two types: sequential and emergent schemas. This stands in contrast to the constructionist perspective led by Wilensky and colleagues, who argue that the difference between amateurs and "experts" may not be so distant for their explanations of emergent phenomena. Moreover, researchers—even experts—sometimes find emergent levels counterintuitive and difficult to understand, which I will cover next.

2.4.3 Difficulties Learning Complexity, Regardless of Experience

The second strand argues amateurs and experts can both find it hard to reason about these complex emergent systems. As Resnick (Resnick, 1994) would argue, this “centralized mindset” is one of the major impediments to teaching complexity. Wilensky and colleagues argue that the difference between amateurs and “experts” may not be so distant for their explanations of emergent phenomena. Moreover, researchers—even experts—sometimes find emergent levels counterintuitive and difficult to understand. When a person sees a pattern in the world, they often assume centralized control, even when it does not exist (Wilensky and Resnick, 1999). The assumption leads people to tend to look only for sequential chains of causes and effects even when systemic, emergent patterns are at play (Perkins & Grotzer, 2000b). Jacobson (2001) examined expert–novice comparisons of complex systems thinking by examining interview transcripts of undergraduate students and complex systems experts. He found that students favored simple causality, central control, and predictability. In contrast, complex systems experts’ explanations demonstrated decentralized thinking, multiple causes, and the use of stochastic and equilibration processes. This data indicates a bias in reasoning about complex situations towards simple causality, but that with experience, some people can learn to view situations from a decentralized mindset.

Thus, in the two strands, we see a disagreement over whether people can learn a to think in complex system terms. Regardless, it is crucial for people to understand complexity to understand important issues. According to current research, species are emergent entities of the processes of evolution over time (Perkins & Grotzer, 2000b). Emergent processes also help explain serious issues for humanity like climate change, nuclear arms proliferation, and income inequality.

Previous research suggests ants are an exemplary way for people to come to understand complex systems (K. Martin, Horn, et al., 2019; Wilensky, 1997b). In ant colonies, we can see the emergence of ants searching for food. They wander somewhat randomly until they come across a morsel. They pick it up and bring it back to the colony, leaving a pheromone trail as they do. Then other colony mates move toward the food, filing in a straight line along the scent trail made by each ant leaving pheromone as it returns. When the flow of ants returning from the food fades, the ants discover and construct new trails through their random walks and feedback. The pheromone trail is a self-organizing process that organizes ants without a central agent controlling their actions. This process requires randomness. The organization is hard for some people to accept because of a "deterministic mindset" in line with Einstein's erroneous statement that "God doesn't play dice" (as in Wilensky, 1997c). In fact, randomness is central to many phenomena. As Taylor (2003) writes, "disorder does not simply destroy order, structure, and organization, bus is also a condition of their formation and reformation." As people come to understand the randomness of how ants find food and lay trails, they form an intuitive understanding of complex environments.

2.5 Constructionist Learning Environments to Teach Complexity

Learning about complex systems can be difficult (Wilensky and Resnick, 1999; Chi, Roscoe, Roy and Chase, 2012; Jacobson, 2011). Computer mediated learning may help with complexity learning. We designed a museum game to teach complexity in short interactions in a museum. Inserting the complex system content into video games may be a useful design because 97% of children in the United States play video games at some point in their lives (Lenhart et al., 2008). One type of game, constructionist video games, seems particularly promising. By encouraging open-ended engagement and exploration, games can support learning across a wide variety of topics and contexts, providing a powerful way for learners to construct new knowledge and understanding (Kafai, 2006; Kafai & Resnick, 1996; Papert & Harel, 1991a). Constructionist learning games try to strike a balance between open-ended play and targeted treatment of learning content through providing learners objects to think with (Holbert & Wilensky, 2019; Weintrop et al., 2012). Constructionist video games employ traditional game structures infused with constructionist ideals to create a game experience that both encourages exploration and delivers targeted content. These games mediate open-ended learning.

These difficulties motivate the design of Ant Adaptation to facilitate learning about complex systems with social insects. In this proposal, we seek to study the use of an agent-based model, Ant Adaptation, to improve learning about complexity.

These difficulties motivate the design of Ant Adaptation to facilitate learning about complex systems with social insects. In this dissertation, I study the design and use of an agent-based model, Ant Adaptation, to improve learning about complexity.

2.5.1 Learning from Insect Systems

In the earlier work, I showed that some participants can shift schemas through the use of complex systems models (K. Martin, Horn, et al., 2019), lending weight to Wilensky and colleagues strand. In the first two chapters, (K. Martin, Horn, et al., 2019; K. Martin et al., 2020), I presented visitors with the opportunity to create an interactive digital ant ecosystem by exploring a microworld (Edwards, 1995; Papert, 1980). Doing so allowed learners to better understand the world of ants, a complex system, by immersion. The work done promises to overcome the twin-challenges of complex systems education, including: level slippage (Wilensky & Resnick, 1999) and not accounting for the individual particles in a system (Chi et al., 2012). This work builds on past work with agent-based models of social insects microworlds that have been effectively used for classroom teaching of complex system ideas. Examples of past work includes *Ant Food Grab,* a yearlong block-based programming curriculum with ants (K. Martin et al., 2018; Sengupta et al., 2015), and *Beesmart*, a curriculum about the hive-finding behavior of honeybees (Guo & Wilensky, 2014). These examples highlight that observing and/or interacting with insect microworlds allow students to construct their own understandings of complex systems through exploring and adapting models in a classroom setting. Like other complex systems interventions, the students build their understanding of complex systems in natural systems. The literature has depicted learning through diverse simulations of natural systems, including: particle simulation of water transitioning from ice to steam in mixed reality (DeLiema et al., 2019), bees (Danish et al., 2011), material science (Blikstein & Wilensky, 2009), multi-level modeling (Wilensky & Reisman, 2006), electricity (Sengupta & Wilensky, 2009), and evolution (Wagh et al., 2017).

2.5.2 Constructionist Microworlds

Constructionist environments have often been long term explorations of motivating problems. In the museum setting, however, we needed to deploy a system to explore, but that provides an experience in as little as two minutes, which is the average playtime in museums. This design constraint motivated some modifications in my use of constructionism to teach complexity. With *Ant Adaptation*, we examined the type of learning that occurs between a user, and a complex system model. In previous work, we provided an example of a learning environment that scaffolds individual meaning-making through touch, discussion and play for museum patrons (K. Martin, Horn, et al., 2019). In the current work, we will investigate the roles of engagement and concept elaboration in learning among groups of museum patrons.

Building on Wilensky and Resnick's interviews in the 1980s and 1990s, in the book

Turtles, Termites and Traffic Jams, Resnick (1994) outlines a three good points that motivate this thesis. First, he develops new tools—at the time using StarLogo to investigate ants, termites, traffic jams and many other self-organizing phenomena. Second, he puts forward an argument that the best way to promote thinking in a decentralized way is having learners construct and play with such systems. Third, to make that possible he argues as did Feurzeig and Papert (Feurzeig et al., 1970; Turkle & Papert, 1992) for programming languages with low floors and high ceilings, such as Logo and NetLogo (Wilensky, 1999b), where people can simply program the actions of thousands of computational objects to follow simple rules to create elegant, emergent patterns. The three parts together are a powerful restructuration (Wilensky and Papert, 2010) of learning in this topic which I employ in this project.

Next, I review literature of museums studies to provide context to the project, as well as support the use of constructivist dialogue mapping to analyze constructionist learning environments. CDM builds on the museum studies' perspective of elaboration as learning (Leinhardt & Crowley, 1998) during active prolonged engagement (Humphrey & Gutwill, 2005) as groups interact around exhibits (Vom Lehn et al., 2001). This literature review motivates the design and analysis of Ant Adaptation.

3 Museum Studies

Next, I review literature on museum studies that look at participant engagement with museum exhibits. The issues and methods described here inform both the study design and analysis. Below I will review several key aspects of museum learning. First, at the Exploratorium, building off Oppenheimer's embrace of constructivist learning, Active Prolonged Engagement (APE) facilitates museum studies to think about changing what it means to be engaged in museums.

Second, the definition of learning as conversational elaboration has become a key aspect of measuring learning in museums. Third, Symbolic Interaction will be reviewed as an important lens to research museum exhibits. Finally, some alternative museum designs will be reviewed, and the design rationale for Ant Adaptation based on prior exhibits will be presented. First, I will introduce active prolonged engagement.

3.1 Active Prolonged Engagement

The exhibit design work that I engaged in for *Ant Adaptation* was influenced by the design principles of “active prolonged engagement” (Humphrey & Gutwill, 2005). Thomas Humphrey and colleagues argue that hands-on interaction with museum exhibits is not the same as active, prolonged engagement (APE). Consequently, informal learning designers need to research the art of fostering active prolonged engagement. APE was built on the hands-on revolution stemming from the impact of constructivism on informal education. Thomas Humphrey, a researcher at the Exploratorium in San Francisco, outlines how from the Exploratorium's early days, as part of the hands-on revolution of the 1970s, exhibits got labeled with signs of how to engage with them. Examples of signs to visitors included, “What to notice?”, “What to do next?” and “What is in fact going on?”. When the experiences follow this formula, the Exploratorium calls it a planned discovery (PD). Building on the planned discovery hands on exhibit, Humphrey outlined the APE framework to aggressively promote self-discovery by the following five steps. One, minimize instruction and explanation. Two, encourage your visitors to initiate play. Three, encourage experimentation. Four, encourage observation. And, finally, encourage speculation.

This all led to Murphy’s team putting forward their first principle:

Principle 1: Museums should shift authority from the museum to their visitors.

This would look like the user trying all the gears in the activity and sequencing their own action.

This was part of the pedagogical revolution of constructivism. As they said, "invest in them the authority to participate actively in their own learning."

From this background, I can infer the archetypical museum's activities. The museum's activity should have five properties. One, encourage questioning that generates exploratory activity. Second, facilitate critical and uncritical observation. Third, deepen investigation along branching pathways. Fourth, provide ample collaboration with other visitors. Fifth, searching for and reflecting on causal explanations for exhibit phenomena helps visitors learn.[3] I designed for each of these properties in creating Ant Adaptation.

Interestingly, Humphrey wondered if APE required focused hypothesis testing, or could it be more free form exploration at exhibits? This is a question I also am investigating. This research had a few findings. Humphrey found two approaches to visitor interaction at APE exhibits, that I also observed in tests of Ant Adaptation (Humphrey & Gutwill, 2005).[4] These were in opposition to off task or just doodling behavior. They noted *investigative*: analytically well-formed lines of thought when investigating the system. Then there was *exploratory*: these were visceral chains of action and attempts to arrive at an aesthetic or Gestalt result. Visitors demonstrated both types of behavior while playing Ant Adaptation.

The team went on to find that visitors asked more questions at APE exhibits. However, the average questions per minute were the same at the APE as compared to normal exhibits. This means that they engage longer but asked the same number of questions per minute.

The kinds of questions they asked also changed, becoming more action and exploration-

---

[3] Humphrey *et al.* promoted these activities through a variety of strategies. First, at the Exploratorium they worked to promote social interaction between the visitors. They also provided opportunities for creation at the exhibits and offered multiple activities. Moreover, they worked to pose challenges visitors could solve with the exhibit. Finally, they emphasized behavioral over content goals.

[4] They also found two more types of activity observation, activity and construction.

oriented, such as: can you turn it faster? What if we connect this to this? How does that work? Or why did that happen? Additionally, they answered the questions using the exhibit or asking each other as opposed to referencing the signs or talking to some museum staff. They also engaged longer, three and a half minutes, as opposed to one minute and six seconds, which led to more questions asked. Finally, they left most often due to external factors such as their family pulling them away, instead of completing everything in the exhibit, like in PD. These findings led to their final four principles:

Principle 2: multi-station design helps ameliorate multi-user interference in the activity,

Principle 3: keep groups together to reduce external draw away from exhibits,

Principle 4: provide individual activities sometimes, and

Principle 5: reduce visitor interference with each other at all costs. This has changed over the years, as now museum researchers find some conflict can be productive, as I will outline in the design of the exhibit.

3.2 Learning as Elaboration

Learning Conversations in Museums (Leinhardt et al., 2003) proposes means to study how learning actually occurs in museums. Their work in museum learning motivates their attempt to solve a core issue: what constitutes learning in museum research, and specifically what terms or actions do we track to measure it. From these problems, the authors propose three outcomes, the last of which is important for CDM. They argue that their approach will provide a novel, stable, and disseminatable methodology to conceptualize, collect and analyze conversations as a process and as an outcome of learning in the museum context. The qualitative method we use in this study, constructive dialogue mapping, has the same aim and could solve core operations issues with the original framework, which we will demonstrate in the following 3 chapters.

The core issue surfaced by Leinhardt and Crowley, how to broaden the definition of learning beyond idiosyncratic and implicit definitions of learning, created a healthy dialogue on the subject (Leinhardt & Crowley, 1998, p. 3). This dialogue pivoted the definition of learning to be dynamic for each study rather than a static definition, as they see static definitions as taking insufficient account of the social contexts of learning. They make clear they are not arguing for a major debate over what constitutes learning, but instead they argue for accountability and clarity in whatever definition of learning is used in a particular study. In this paper, I present an operationalizable definition of learning in the museum context based on the CDM method of analysis. After introducing the problems and outlining them, Leinhardt and Crowley (1998) describe a pragmatic approach to study learning in museums: define learning as Conversational Elaboration. They chose how visitors elaborate conversations for their pragmatic approach and operational definition of learning.

They focus on conversational elaboration because it is both a naturally occurring part of the museum experience while also being a product of the experience. By elaborations they mean a particular kind of talk that occurs within a group both during and surrounding a museum visit.[5] Conversation is important because it reflects the "inter-twining of social with cultural processes" (White, 1995, p.1). Sociocultural theories of voicing (Rogoff, 2003; J. Wertsch, 1991) emphasize that inter-twining of voices (in the Bakhtinian sense) as the primary activity through which knowledge is constructed and appropriated across people.[6] This process is important as

[5] In the George Herbet Meade sense of within an individual, both in and around a museum exhibit. This learning definition could be useful to track learning generally. For example, if we tracked how a young man engaged with extremist literature through the internet or books, and change his conversation, then sought like-minded individuals and changed his speech, we could track how a young man learned to become a militant extremist through conversation elaboration).

[6] Bakhtin (Wertsch,1991) argues we do not exist in an open world, but one of Saussurian *parole*, in other words in a dialogically contextually constrained world. The major constraints of addressivity between voices {??}. These voices can be other people, artifacts, and even the multi-voiced nature of the self. This notion aligns with Bourdieu's

conversation is a natural process and consequence of an enjoyable, shared experience, such as visiting a museum as a family.

The approach particularly looks at four processes that groups will undertake. If there is learning, after an interaction with an exhibit, a Coherent Conversation Group (CCG) will:

- Refer to more items,
- Include greater detail about those items,
- Synthesize elements to elements from their prior knowledge, and
- Increase the level of analysis of the phenomena that they share.

This approach moves away from focusing on the amount of talk or types of talk while building a strong foundation between amount, type, and the process of learning. The big takeaway here is that a part of the field has argued that it is essential to attend to visitor talk as both a process of learning and a learning outcome. This motivates CDM to be used as a methodological tool for this comparative study. We expand on the work of Crowley and Leinhardt while offering new tools for researchers and practitioners working in museums.

Looking ahead, how does this connect to my research? In the end, the methods of collection and analysis Crowley and Leinhardt (1998) suggest are interesting, but they are not sufficient. With text analysis, and affective sensors, I can measure the learning and engagement more fully than Leinhardt and Crowley did. I expand on the method of CDM in the second Chapter. Then, in the third chapter, I will use CDM to study elaboration in an informal learning experience to expand

---

idea of self-censoring in mediated association with power to explain a lot of human actions. As Holland posits (Holland & Lachicotte, 1998), figured worlds are materially and perceptibly expressed activities defined and confined jointly, by meanings, structure of privilege and influence. These figures "interanimate" (Wertsch, 1991, p. 54) each other and larger societal and trans-societal forces. Through this interanimation, a sense of the individual is fashioned. But to be clear this individual is not unitary; it is a composite of selves. In other words, technically, we do not author a Self, because if contextually indicated, and objects can be mapped to multiple symbolic meanings, every individual has more than one context, and so logically, more than one contextually bound self. So instead of fashion self, let us say they rehearse performances of an individual. For as McIntosh pointed out: "The idea of self-control is paradoxical unless it is assumed that the psyche contains more than one energy system, and that these energy systems have some degree of independence from each other." (Donald McIntosh Foundations of Human Society. As in Thaler page 103 of Misbehaving) an exploration of dynamic learning and use affective sensors to study engagement.

3.3 Exhibiting Interaction: conduct and collaboration in museums and Galleries with Symbolic Interaction

In this study, we not only have participants use one model of ant colonies, like we did in previous iterations (Martin et al., 2019; 2020), I have participants use a scaffolded series of models to learn from this informal design. When thinking about sequencing informal learning experiences, I considered the work in symbolic interaction to prepare the design. Dirk vom Lehn, Christian Heath and Jon Hindmarsh (Vom Lehn et al., 2001) explore how people alone, and together examine exhibits and galleries. Their work focuses on ways people interact around and throughout exhibits. The work is grounded in symbolic interaction that ties together the work of Mead (Mead, 1932a, 1932b, 1934) and Blumer (1969), who place the object and how it is socially organized at the center and tie the object to behavior in public spaces. It considers how people animate exhibits for others and how this structures others' experiences. The work echoes Wertsch and Bakhtin, but this literature is not leaned on here. Instead, they focus on how meaning and experience arise, through action with objects. The study explores how people in museums and galleries continually coordinate their conduct with each other.[7] While several people interacting on a Zoom meeting is not the same scale as a crowd in a museum, the ideas still bear consideration.

Visitor studies, a field of study developed in the 1990s, investigates social interaction with and around exhibits, to understand how people behave in an experience (Dierking & Falk, 1992). It is crucial to understand museums provide an opportunity to interweave the social constitution

[7] This intersects with Bourdieu's work, (Bourdieu, 1984) which examines how people learn to understand art through etiquette. Bourdieu's work also connects to social display rules that constrain what affective states people display (Picard, 2000) (Picard, 2000).

of the object, and the material environment with the study of social interaction. This leads us to study people with exhibits, both with their companions and with strangers, to study how they interact. The work argues that knowledge, like conversational elaboration, emerges moment to moment. This indicates researchers should study moment to moment actions, such as conversational elaboration, and bio-markers indicated through affective computing signals, to study learning in informal environments.

To undertake the study, Dirk vom Lehn et al., did a standard field notes and video study, collecting as people interact with the museum. Traditionally, people looked at people interacting with an exhibit, measured how long they stayed there (i.e., hold time), how and who was attracted to the exhibit, and did it encounter by encounter. The problem with this approach is that people construe meaning with each other, and even with strangers. So, the list of people who show up is important, as people are using the setting to continuously construct an understanding. For example, a man examining an exhibit of prints, each print is seen in light of the last one he looked at (Vom Lehn et al., 2001). Consequently, this ongoing constitution of meaning seems at odds with the idea that particular exhibits have stopping power (Shettel, 1976). For *Ant Adaptation*, this is important, because users co-construct the meaning of the simulation while they use it. For instance, there are overlaps between groups playing, groups talking to each other; throughout the game they are setting the goal of the interaction with the exhibit by watching other users and talking. In the third iteration of the experience, we provide a scaffolded sequence of Ant Adaption models, which further heightens the need to consider the sequencing of experience of participants.

This is the work Dirk vom Lehn, et al., (2001) studied. They studied points of action as a more dynamic and flexible notion of how participants view and constitute their experience at exhibits. They found people negotiate with their companions the access to museums. Companions influence three aspects: first, they influence people's draw to look at parts of the exhibit. Second, they influence what they see. And third, the conclusions they draw from their interactions. Through interaction with each other, visitors negotiate access to and participate in the exhibit. People shape each other's interaction: what is seen, how it is seen, what is said, discussed, and the experiences that people raise in and through interactions with each other and the objects. In other words, experience in a museum is an emergent phenomenon constructed between users as they share their feelings and watch each other’s interactions. I discuss more about how affect shapes human choice in the affective computing section.

Symbolic interaction analysis is like visual pheromone trails for people. For example, in ants, there are pheromone trails. When an ant finds something of interest, like some good food, she will leave a trail as she walks back to the nest, letting everyone know where to find good food. People show each other where the good parts are in a museum. Ants leave pheromone trails when they are carrying food they found, this encourages other ants to find the same food. The process allows the self-organization of foraging. Likewise, people glean information about what might be interesting, or novel about an exhibit, and how an exhibit might be used or interpreted by watching other individuals at the exhibit. If you are being critical, you might say that the author is saying people will follow groupthink or herd mentality, but it seems more like stygmergy where ants socially construe the meaning of the environment to come to an understanding of it to enable foraging and colony care.

I am reviewing this literature because symbolic interaction — in particular, the Dirk vom Lehn way of analyzing people in museums — is connected to my work in two ways. First, based on my experience from implementing Ant Adaptation in the Field Museum last time, I noticed people come by and watch what other people are doing and then they implement the interactions they observe. Because I am not using the exhibit in a purely clinical trial, but instead in an open space. In other words, there is some impact on their interactions on how they saw other people do it. Previously, I saw some groups copying the last group they saw in terms of how they interact with each other and the tabletop game. For example, a kid comes up and then stays for a little longer and joins the next group sharing his previous experience. Second, the museum exhibit process is impacted by family structure: their kids are coming up, and the family structure is mediating the interaction with the exhibit, and the users are, in some ways, controlling each other's interaction with the exhibit just as Lehn describes. For instance, a father told his son, "Please talk about this game more as a scientific thing. You're making this too much into a game." In this way, the family is modulating the interaction with the exhibits. In other words, because I am trying to bring Ant Adaptation into a wild-type study, the groups, the exhibit, and the space I am doing it in, all impact this analysis. As a result, symbolic interaction is important in understanding the exhibit's role in a museum experience.

3.4 Design with Alternative Learning Technologies for Informal Learning

Different exhibit designs elicit different interactions. Alternative design technologies include: augmented reality (Yoon et al., 2012), immersive technologies (Snibbe & Raffle, 2009), tangible user interfaces and touch interfaces Horn et al. (In press), and gigapixel imaging to bring science to the public (Louw & Crowley, 2013).

In the first two chapters of this dissertation project, I used a touch interface to engage museum visitors. In choosing this technology I needed to decide between touch and tangible. There is a draw to tangible and touch interfaces. Prior work showed they both are compelling; however, tangible interaction attracts children and families more. But interactive and tangible were both equally effective at supporting visitors' interaction with exhibits (Horn et al., In Press). Large multi-touch displays are commonplace in museums. Tangible interactions make the use of tangible objects or the whole body in interaction. Based on existing knowledge, visitors can use tangible interactions, leveraging our real-world experience.

Prior research on building interactives in museums informed my design. Work on agent-based models in informal learning (Olson & Horn, 2011) shows that interactive tabletops can be productive tools for cooperation and learning, but the research also suggest tabletops are fast paced and children tend to change configurations quickly (Fleck et al., 2009; Hornecker et al., 2008). Current research on multi-touch tables for museums suggests several key design elements (Davis et al., 2015; Horn et al., 2016) such as enjoyment, comparability, and productive conflict to promote their benefits. Enjoyment, expressed through affect words such as “wow,” “cool,” and “hah,” is significantly correlated with learning measures considered by Horn et al. (2016) and so focusing on situations that promote positive affect may be beneficial. The tree of life game is a tabletop game where users can investigate different taxonomic lineages. In the game, facilitating comparisons-aided learning; players who drew comparisons between lineages learned more easily and were more likely to use terms of interest in open-ended questions on post-tests. Block et al. (2015) found that groups of two spend more time at an exhibit and engage more with scientific content than groups of three or more. Finally, conflict can be productive. Davis et al. (2015) and Falcão and Price (2011) argue that interference between users on and across a multi-touch interface can be productive for learning when it triggers argumentation and collective knowledge construction. This finding is notable because though it contradicts one of the active prolonged engagement principles to reduce conflict, the finding aligns with Dirk vom Lehn’s description of how people shape each other’s interactions around museums.

In the third Chapter, I used an interface to engage participant remotely. In choosing this technology, I needed to restructure the activity from the touch interaction in the first and second iteration for remote learning using a mouse. These adaptations were technically straightforward but could impact the use of Ant Adaptation as result of allowing all users to reach the entire screen, as well as giving one user, that holds the mouse, outsized authority.

Because positive affect and engagement is associated with better learning in museums Horn et al. (2016), I am proposing to study positive affect more closely. I will review this literature next to motivate the use of affective measures of mood with the refactored interaction with Ant Adaptation implemented in the third Chapter.

4 Affective Computing

Affective computing allows for characterization of users' affective states, which can be used for analysis, as in the third chapter of this dissertation, or to allow the computer to interact based on a user's affect. Technologies are deployed both to sense affect and impact users' affect. For example, Mattel released a short-lived project, *Hello Barbie.* Produced in collaboration with *ToyTalk*, a San Francisco-based company, the project worked toward the fulfillment of an ancient dream: toys that can talk to us (Vlahos, 2015). In the past, the idea of talking toys had been a bit of a party trick; toys used recordings to fool the user. But that is all changing. In the project, Mattel released an affect detecting, speaking Barbie. The project was built on the past decade of breakthroughs in artificial intelligence and speech recognition that gave devices around us — such as smartphones, computers, and toys — the ability to engage in conversation to generate intelligent responses to users. As the technology improves, it may become the primary way people engage computers (Vlahos, 2015). The success of the *Echo* and *Alexa* augur this future. Though it did not last long, *Hello Barbie* was discontinued only months after release with 51% of reviews on Amazon at 1 Star. The main complaint was that the technology was not ready: the battery would go flat, or the Barbie would not understand the child. Nonetheless, the promise, and perils of the technology persist. The technology was specifically designed to appeal to children by creating a friend that remembers and gets to know the user. While the implementation is not there just yet, affective agents and detection are getting better.

These systems are not just theoretical, there have been practical means to use affect detection. For example, designers have used it to alleviate user frustration increasing the probability of users liking software agents (Bickmore & Schulman, 2007) as *Hello Barbie*. Moreover, affect detection also assists in measuring voters' candidate preference based on affective responses to election debates (McDuff et al., 2013). Furthermore, as outlined in the section below on affect and learning, affective states pathways, especially the engagement-confusion-delight-engagement pathway, have been hypothesized to facilitate advanced problem solving if the confusion state does not overwhelm the learning (D'Mello & Graesser, 2012a). Andres and colleagues (2019) found while researching informal learning with *Betty's Brain* that the only emotional pathway associated with learning gains was sustained delight. Horn et al. (2016) found post-test gains associated with positive affect words like "*wow*" and "*cool*" after a museum exhibit. Each of these uses points to greater integration of affect detection to improve design, interactivity, and analytics in learning technology.

4.1 History of Affective Research

Since ancient times, emotion has been researched (Healey, 2014). In ancient China, the Yellow Emperor, the source for traditional Chinese medicine for more than two millennia, argued emotion resided in the body, and excesses of emotion could damage a person's life-energy. In ancient

Greece the physician Hippocrates theorized that the body was composed of four humors. These humors drove a person's physiological and behavioral patterns, and imbalances could lead to deleterious outcomes. Imbalances would lead to choleric, melancholic, phlegmatic, or sanguine outcomes in the blood, respectively. Aristotle also had a physiological view of emotion. He viewed them as "passions", comparable to physical states like hunger, thirst, or desire. These ideas still influence our thinking about emotion (Healey, 2014).

Further, moving into modern times, Hans Eysenck (1947) cited humors as the inspiration for his dimensions of personality such as neuroticism. Additionally, William James (1893), the first theorist of modern times to put forward a theory of emotion, viewed the physical response as primary in the feeling of emotion. James believed the stimulus triggered the activity or the response in the autonomic nervous system (ANS), which caused an emotional response in the brain. For instance, he would argue people feel sad because they frown or cry. Similarly, Carl Lange proposed a theory that was very similar to James, and so this theory has become known as the James-Lang theory of emotion. Continuing his earlier work, James described the physiological activities that accompany emotion. Simultaneously, Charles Darwin also started documenting the observable physiological responses in both animals and people (1872). He focused on fear reactions and used it to classify and differentiate species based on affective states. These descriptions of physiological states and patterns are the theoretical underpinning for using physiological signs and patterns to recognize emotion and affective computing, which I propose to use to examine how people respond during the implementing of a simulation in a museum.

Walter Cannon, a Harvard professor, significantly criticized the James-Lange theory along five lines, including that the autonomic nervous system or “visceral change” (1927, p. 112) was too slow and nonspecific to be unique to each emotion. Therefore, Cannon argued emotion must primarily be a cognitive event (Cannon, 1927).

Building on mounting evidence for a purely physiological or cognitive source for emotion, Schachter (1964) spelled out the implications of a cognitive-physiological formulation of emotion. Schachter (1964) proposed a compromise between James-Lange and the Cannon theories, i.e., his *two-factor* theory of emotion. In his studies, he injected epinephrine into subjects, but was not able to detect emotion after the introduction of this chemical. However, when he added a situational context, such as scaring the person, he could get an enhanced response after injecting epinephrine. Therefore, he concludes feelings such as fear, anger, or happiness have a physiological part. In other words, physiology was part of emotional experience, but emotion is made up of the combination between physiological changes and cognitive interpretations. Though this two-factor theory was important, Schachter informed but did not entirely explain complex interactions between physiological and cognitive influences of emotion. Physiological responses to emotion are an interaction of cognitive responses and physiological responses that create complex interactions we can call emotions. Notably one major criticism was that epinephrine is too coarse of a measure to study this complex (Zajonc, 1994).

4.2 Affect and Intelligence

In her seminal work, Rosalind Picard, an MIT researcher, studies the promise, challenges, and potential reward of augmenting artificial intelligence with a core component of human intelligence — namely, emotion (Picard, 2000). The ability to recognize emotion is among the key aspects of emotional intelligence, which is a facet of human intelligence. Furthermore, emotion plays a key aspect in perception (Picard, 2000). There are many examples that point to an intervening role for emotions in perception. Studies show how mood influences participants' perception of ambiguous stimuli. As shown in Figure 1, this results from the fact that brains have two perception pathways, quick and dirty stimulus processing through the limbic system, and a slower, more accurate system through the cortex (Picard, 2000). The interaction of the two systems drives intelligent systems.

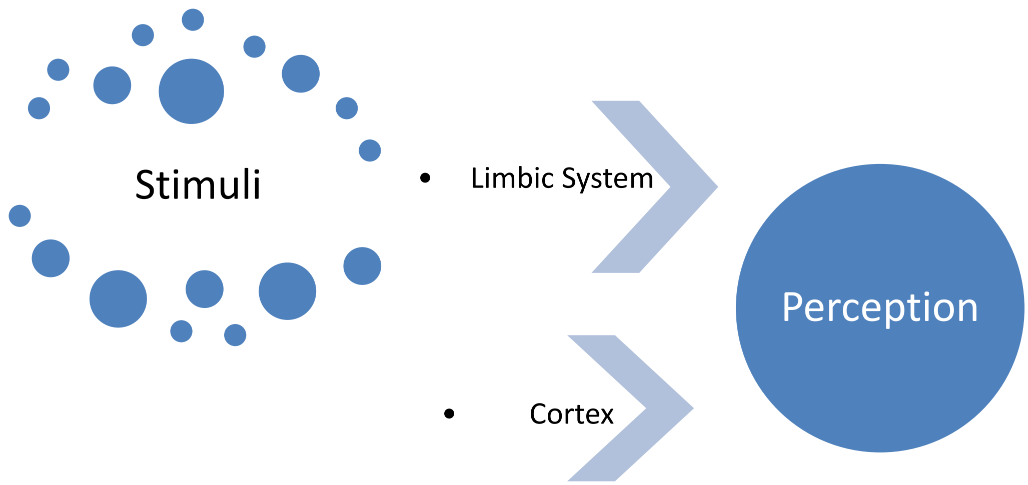

Figure 1: Brain has two pathways for perception, Limbic and Cortex

This dual role inserts passions into rationality. Picard is not disturbed by this paradox, instead, she fully embraces it: "Even reasonable behavior is neurobiologically directed by these so-called passions" (Picard, 2000, p. 8). As a result, we need to understand affect's role in that equation in order to understand noise in human decision making (Kahneman et al., 2021), or what is more, build artificial intelligence (Picard, 2000), or study learning (Calvo & D'Mello, 2011; D'Mello & Graesser, 2012a; Singh et al., 2002).[8] This suggests my research questions that explore the role of affect in learning and engagement in chapter 5.

How do researchers track that? To understand the connection, they need to track physical responses. The connection between limbic and cortex thought modulation, such as facial

[8] See *Descartes Error,* Damasio, 1994, Thaler and Ganser's *Misbehaving* (2015) for additional powerful arguments of the role of emotion in rational decision making.

expressions, posture, and voice inflection, are physical means by which an emotional state is expressed. Humans primarily communicate emotion through these embodied forms (Picard, 2000). Ekman (Ekman, 1992; Ekman et al., 1993) argues “basic” emotions have particular musculature, movement pathways and social display rules associated with them. His theory led him to innovate a facial action coding system. While recent research (Barrett, 2019) has put the one-to-one mapping between emotions and affect in doubt, how humans recognize emotion in others is still primarily through physiological communication. As a result, researchers can measure mood by measuring physical responses. Mood, or affective state, are associated with processes of learning regulation, which I will turn to next.

4.3 Affect and Learning

Increasingly, the affective processes associated with learning are of most interest for researchers who are trying to understand how students regulate their learning processes (Calvo & D’Mello, 2011; D’Mello & Graesser, 2012a; Singh et al., 2002). There are affect-sensitive or affect-aware educational systems that dynamically tailor their instructional activities based on student affect (D’Mello & Graesser, 2014b). Reviewing 24 studies having 1,740 participants, D’Mello and Graesser’s take-home message was that happiness, engagement/flow, confusion, boredom, curiosity, and frustration were the major affective states shown by learners (2014). These states are good candidates for modeling affect in learning. Furthermore, happiness was the only basic emotion to feature in education environments frequently. They found this not entirely surprising because there is inadequate theoretical justification to expect notable levels of some of the basic emotions, such as disgust and fear, during short one-on-one learning sessions with computers. Several researchers have studied how students cycle between emotions during learning (Baker et al., 2007; McQuiggan et al., 2010; Ocumpaugh et al., 2017). In their influential 2012 work, D'Mello and Graesser argued that flow states, punctuated by confusion, may contribute and assist in complex problem solving (D'Mello & Graesser, 2012a). Specifically, they hypothesized that there are two main pathways that students follow when transitioning between affective states: one that promotes learning (the engagement-confusion-delight-engagement pathway) and one that reduces learning (the engagement-confusion-frustration-boredom pathway). In should be noted that while educational research has explored the role of student's affective states in learning, they may not have captured the frequency or meaning of transitions between states (Andres et al., 2019). As a result, Andres' colleagues expand the number of pathways studied. Their results suggest that increases in the frequency of two affective state patterns, engagement-confusion-delight-engagement and sustained delight, relate to an increase in post-test scores. Their results should be interpreted carefully because they used non-sensor-based detection methods of affect detection. Nonetheless, the results are promising as they indicate that delight, and the affective states occurring around it are related to learning measured on post-tests. Curiously, the only pathway associated with learning gains was sustained delight (Andres et al., 2019). This finding is interesting as it connects to the museum research that positive affect words are associated with post-test gains (Horn et al., 2016). Even though they are relatively infrequent, the two hypothesized pathways from D'Mello and Graesser (2012) appear to influence knowledge measures. Therefore, the evidence supports the hypothesis that moments of high stimulation, especially of delight, would correlate with gains on post-tests, because high stimulation is implicated in memory formation (McCaugh, 2004). In short, affect plays a role in learning.

4.4 Measuring Mood

To study these physiological displays, research needs to choose bio-markers. Creating a system to measure physiological signals is the first step to measure and automatically recognize physiological patterns associated with emotion. This step will be crucial to measure affective signals during the use of *Ant Adaptation*. Consequently, I will review what bodily signals to measure. In the widest view, all bodily signals could be viewed as physiological signals, including brain signals, voice patterns, facial expressions, or body chemistry. While the descriptions of physiological responses that Darwin, James, and Lange described make sense to humans, computers need quantifiable metrics to detect features for their analyses. As a result, affective computing researchers use digital recording devices and electronic sensors to calculate such features as heart rate acceleration and skin conductivity metrics to classify emotions (Ekman et al., 1983; Levenson, 1992). However, all the nuance of emotion may not be captured with these sensors.

Two of the most promising technologies for examining human response using quantifiable metrics are facial recognition (Cohn & De la Torre, 2014) and skin conductance (Healey, 2014). I will review each of these techniques below.

Emotions are degenerate, meaning they have more than one bio-marker to determine them (Barrett, 2019). As a result, researchers should look to more than one-time variant signals. There are several bio-markers that vary with time (Picard, 2000), including the following:

- Voice intonation
- Facial expression
- Posture
- Heart rate
- Diastolic and systolic blood pressure
- Pulse
- Pupillary dilation

- Respiration
- Skin Conductance and color
- Temperature

Physiological measures have long been associated with emotion (Healey, 2014). In this study, I will collect one of these bio-markers: facial expression (Cohn & De la Torre, 2014). I also layout a second measure, skin conductance (Healey, 2014), to pave the way for future research. These two signals are important because they show high utility and signal fidelity when participants are active, such as when playing in a museum. Below, I will review the two techniques, first covering skin conductance and then facial expression coding, in order to provide context.

4.4.1 Skin Conductance

An area of interest is skin conductance. Galvanic skin response, GSR or electrodermal activity is commonly used to measure affect. It indirectly measures the amount of sweat in a person's sweat glands because skin is an insulator and its conductivity primarily changes in response to ionic sweat-filling sweat glands (Healey, 2014). Sweat gland activity is a measure of the sympathetic nervous system activation. GSR is a robust and non-invasive way to measure the sympathetic activity (Cacioppo et al., 2000). GSR was first used by Carl Jung to measure ‘negative complexes’ in word association tests (Jung et al., 1969). This measure is used in lie detectors that measure the amount of stress a lying person manifests (Marston, 1938). Skin conductance has been shown to vary linearly with the emotional aspect of arousal (Lang, 1995). Skin conductance has been used to differentiate between states such as fear and anger (Ax, 1953).

4.4.1.1 Placement and Use of Probes

Researchers usually place skin conductance on the most sensitive sweat glands. This includes the soles of the feet and the palms of the hands (Boucsein, 2012). In the lab, often lower segment of the middle and index finger of the dominant hand can also be used. Low conductivity gel is used to ensure good contact. Low voltage injected into the skin and the resulting change in the voltage is measured. Ambulatory study adoptions can use conductors on the wrist, arm, or leg. Picard's E4 Empatica wristband is a useful option. Alternatively, researchers could use jewelry or clothing (Healey, 2011; Picard & Healey, 1997). These measures have slightly lower signal validity, but do not get knocked off in motion like some of the other placement options would, making them better for studying children at play.

4.4.1.2 Signal Interpretation

There are several ways of interpreting signals. Features extracted from skin conductance include mean conductivity level, variance, slope, max and minimum levels, features of originating responses such as the tightness of the signal to a response such as sound (*See* Damasio, 1994), the amplitude latency, and the rise time. There are several confounding factors of skin conductance. First, it is nonlinear and time-variant. People are not a closed linear system; they change throughout the day. Research must account for this complexity. Second, there's a baseline drift. I am not building on the same system over time. Third, conduction's change due to the electrode contacts change. So, as the sensor falls off the skin or moves around, the voltage will change in the contact. This shift will be correct for using a noise filter provided with EDA Explorer (S. Taylor et al., 2015). Fourth, it is difficult to extract affective signals in the presence of motion. Because humans are not linear or time-invariant systems, they are complex systems. Fifth, there is currently no method for accurately predicting an individual's physiological response to emotion. It is even more difficult to attempt to subtract it off from the emotional signal (Healey, 2014). There is no source separation method for emotional versus non-emotional responses.

Healey (2014) also covers heart rate extensively. However, *there* are additional factors regarding heart rate besides emotion that make it a poor fit for our current study. For instance, factors including age, posture, level of physical conditions, breathing frequency and circadian cycle can confound the signal. Besides these factors, assumptions arise. HRV assumes a relatively unchanging subject, making the method much more useful for use in a supine hospital patient than an active subject such as children in a museum. As a result, the signal is not useful for our current study.

4.4.2 Facial Expression

Automated face analysis for affective computing is maturing (Cohn & De la Torre, 2014). The face conveys information about a person's age, background, sex, and identity (Bruce & Young, 1998; Darwin, 1872; Ekman & Rosenberg, 2005). Facial expression regulates face-to-face interaction, indicates interpersonal repulsion and attraction, reciprocity, and communicates subjective feelings between members of various cultures (Bråten, 2006; Fridlund, 1994; Tronick, 1989) among other details. As a result, behavior scientists have focused on the face (Cohn & De la Terre, 2014). Kanade (1974, 1977)) began computer scientists' interest in the face as a potential biomarker. This interest led to work with computer vision to automatically analyze facial expressions (Ekman et al., 1993; Parke & Waters, 1996). Infrastructurally, the facial action coding system (FAC) has enabled this work (Ekman et al., 2002; Ekman & Friesen, 1978).

Recent research with facial expressions is struggling with several problems (Cohn and De la Torre, 2014), including the following: (1) Challenges with detection in naturalistic settings, (2) low base rates, (3) partial face occlusion, (4) pose variation, (5) rigid head motion, and (6) lip movement associated with speech. As researchers have overcome these challenges, automated face analysis (AFA) is beginning to realize the goals of advancing human understanding (Ekmam et al. 1992; Cohn De la Torre, 2014).

4.4.2.1 Manually Coding Faces

When researchers code faces, there are three common approaches: (1) message, (2) sign, and (3) dimensional measurements. First, message-based measurement is where human coders make inferences about human emotion or affective state. This is based on Darwin's work (1872) where he coded 30 emotions. Ekman and colleagues refined these (Ekman & Friesen, 1978) to what they called basic emotions. An appealing assumption of this method is facial analysis provides a direct readout of emotion (Buck, 1984). But this appeal is problematic. Emotions are contextually dependent. There is a reason to be dubious of one-to-one mappings between expression and emotion (Cacioppo & Tassinary, 1990). Second, sign-based measurement starts with appearance. It is a descriptive sign-based system that you follow up with experimentally or observationally to discover, or assign, the relationship between signs and the emotion. The facial coding system (FAC) is the most popular (Cohn et al., 2007) (Cohn, Amador, Ekman 2007). Third, dimensional coding is a way to use human coders. While the first and second options for coding emphasize the differences between emotions such as joy and confusion, an alternative dimensional measurement emphasizes the similarities (Schlosberg, 1952, 1954). The approach posits that facial expressions can conform to a circular surface, where for instance pleasant-unpleasant (valence) and attention-rejection, are principal axes.

An attendant issue with each of these measurement methods is the code reliability (Cohn & De la Torre, 2014). Reliability is how repeatable, consistent and free from various errors a measurement is (P. Martin et al., 2007). This affects all measurement approaches. What researchers need is first *agreement* and *consistency*. Agreement means each coder assigns the same score to a measurement. Consistency means the degree to which ratings from different sources are proportional when expressed as deviations from their mean scores.

Combining these three methods, and upholding reliability, researchers can begin to use automated face analysis (AFA). AFA seeks to detect the movement of one or more of these measurement types, the process requires several steps, including: (1) face detection, (2) tracking, (3) feature extraction, (4) recognition, and (5) learning. In studying learning it is common to use supervised methods where label data of a limited set of observations trains a classifier for work on a larger corpus of data. There are several difficulties with AFA, including registration, in which non-frontal angles on cameras and head motion can cause serious problems. I will address this issue by mounting a camera in the center of the tabletop display to capture faces as they look down to play the game. Additionally, action units, the motions I detect of a face, modify the appearance of the face. These changes interfere with computer vision's detection. Also, since facial actions are subtle, I need to be aware of these micro-actions. Furthermore, the non-standardized face shape of people can undermine generalizations of facial detection across people. Moreover, the temporal dynamics of facial motions need to be attended to. Additionally, variable classification can be overfit, where the training data, the people in our study, do not generalize to a larger population.

4.4.2.2 Automated Face Detection: Steps

The difficulties and challenges of automating facial detection have created a plethora of algorithms and applications. These steps and the work to avoid these difficulties has led to the exponential growth of solutions (De le Torre & Cohn, 2011). Regardless of the application, AFA begins with face detection. For frontal detection of faces, Viola and Jones' (2004) algorithm is the most used. There are two basic ways of going about this detection. The first one is sparse, where observers just look for minimal features such as eye contour. The second way is dense, which involves eye contours and other permanent facial features. This allows us to get good yaw, pitch, and roll for 3D face inference and detection (Cohn an De la Torre, 2014).

AFA has several parts (Cohn an De la Torre, 2014). First, we have face detection, where we use a camera or a set of cameras to detect the face through contour analysis. Then we have registration, where we register the face of the primary parts we are going to use to follow it through the video. This is followed by feature extraction. For instance, we could use geometric, the appearance, or the motion for the feature extraction. We then go through a process of data reduction and selection to reduce the dimensionality of our data. This is important because facial recognition data is often high-dimensional and through reduction, we can get more generalizable data. Researchers put this data through supervised learning usually. Supervised learning has input label data on a small set that we then use to detect similar features in a larger set of data. Though sometimes we do use unsupervised learning in these cases (*See*, De la Torre and Cohn, 2011). In state modeling, it used to be that neural networks were the primary means by which we analyze the data, but research has been moving towards support vector machines (SVMs) which the best open-source facial detection library, *OpenFace*, uses. A crucial part of these projects are reliable databases, there are two that stand out: DISFA (Mavadati et al., 2013) and the Binghamton-Pittsburgh 4D database (Zhang et al., 2013).

From this pipeline there have been many exemplary applications (Cohn and De la Torre, 2014). Applications have included action unit detection, intensity monitoring, checking for physical pain, depression and physiological distress, deception detection, interpersonal coordination, and expression transfer between subjects in a test. Other applications have included marketing, distinguishing subtle expressions, drowsy driver detection, autism aids and, important for our project, instructional technology (Craig et al., 2007; Whitehill et al., 2011). As a result, facial detection provides an interesting affective computing signal to look at subtle differences of affective state, and a powerful collection of research that maps affective state to emotional state in participants. As such, I use AFA in Chapter 3 to track engagement.

4.5 Limits

Affective Computing approaches have powerful affordances for tracking learning and engagement, but it also unearths a challenge of a broader category: incorporating AI in education research. AI systems have been shown to be less accurate at identifying the faces of dark-skinned women, to grant women lower credit-card limits than their husbands, and to erroneously predict Black defendants will commit future crimes more than whites. When applying these systems to learn about education, many of the solutions require having a human in the loop. As a result, a key domain of my research is investigating the diversity challenges posed by AI, while leveraging the technologies' benefit for education. The connections between affect and learning, as well as the diversity challenges of applying AI to education will be my double-barreled research program going into my career.

4.5.1 Is Emotion the Same as the Signal

I think I should pause, however, and ask, can computers detect emotions? Lisa Feldman Barrett (2019), an esteemed researcher of human emotion at Northeastern University, works with the Harvard Medical School and MIT on the use of emotion detection in computing. Her work covers the excitement of the advent of emotion reading gadgets. She asks, "Is it possible for machines to read emotion in a body?" (2019). Companies are claiming to do it, founded on the idea that emotions can be read in bodies (Picard, 2000). She offers a warning and claims that those who say they are reading emotions successfully are exaggerating what they can accomplish. It is not that machines cannot interpret emotions, but that researchers misunderstand what emotions are and how they are generated. As a result, there is the chance to solve this fundamental problem by drawing on the unique advantages and breakthroughs of affective computing in bio-marker detection.

Affective computing, especially facial recognition, is based on modulations of sentic signals, such as the face or electrodermal conductance. These lead companies and researchers to say they detect happiness, when really, they are detecting smiles. There is a difference between our perception of facial movements versus the actual facial movements. The common stereotype is that a person will frown when sad, smile when happy, scowl when angry, etc... However, individuals are different. Some people may laugh when they are afraid, or cry when happy. In fact, only 30% scowl when they are angry (Barrett, 2019). This means scowling is a weak predictor of anger, but anger does not cause scowls. In other words, software can detect scowling, under ideal conditions, well. But that detection is different from detecting an emotion. The detection of emotions from looking at a face varies by culture, situations, and even among people in the same situation (Gendron et al., 2018). Grendron and colleagues argue while the scientific evidence suggests people sometimes smile when they are happy, frown when they are sad, and scowl when they are angry, the ways people communicate anger, disgust, fear, happiness, sadness, and surprise varies substantially across cultures, situations, and even across people within a single situation. Therefore, while it has been claimed configurations of facial movements can be universally recognized as emotional expressions because they provided information in situations critical to gene propagation for our hunting and gathering hominin ancestors, experiments have called this particular evolutionary hypothesis into doubt (Gendron et al., 2020).

One of the long-established ways to see how people predict emotions is having them read a story and then either pick a face that most aligns with the protagonist's feelings or to describe the feeling. This method may be flawed: the experimental method of reading a story and then assigning the emotion the characters have manufactures the evidence of universal emotions (Barrett, 2019). Barret suggests that if we want to analyze emotions in humans using computers, the evidence is not strong enough for real-world outcomes such as legal determinations. She describes how people move their faces in different ways with the same emotional state and that variation is the norm. So, you cannot simply measure a face to monitor human emotion. The faces we use for emotion detection are basically stereotypes, or "a science of emojis" (Barrett, 2019). We do not want to build AI analysis on stereotypes but on real, non-impoverished emotional episodes. A feeling like anger, or another emotion, is not static. Emotions are dynamic and use hundreds of bio-markers. Emotions are degenerate, meaning they have more than one bio-marker to determine them.

For instance, even if we can make determinations of facial expression, there are several complicating factors of emotion detection and their connection to displayed affect. Here are seven examples (Picard, 2000):

1. The intensity of emotion;
2. Type of emotion, i.e., there are many types of love (See *Symposium*);
3. How the state was induced, i.e., from a film or being a genuine experience, such as a conflict;
4. Social display rulers: whether a person was encouraged to express or suppress the emotion. In short, the context of the detection matters;
5. Hormone, diet, and medications;
6. Mood-state cues memories associated with that mood: positive moods tend to cue positive memories. Ledoux's work shows emotion can hijack the cognitive centers of the brain; and
7. Person-independent emotion detection is difficult because people display emotions differently, "given that a particular emotion is felt, a variety of factors influence how the emotion is displayed" (Picard, 2000, p. 32).

Consequently, the use of affect detection should be used based on Negro Pontes's suggestion which is in a person-dependent way, and I would add, a context dependent way. We can measure a person within a context to measure moments of high facial expression display or EDA levels. Currently, these may not be generalizable beyond the individual. As such, I measure relative changes of stimulation.

4.6 Role of Affect in Human Choice

Emotion is critical for forming memory, attention, and rational decision making (Picard, 2014). Not only does mood influence judgment of seemingly objective perceptions, but it also affects memory retrieval (Picard, 2000; McCaugh, 2004). Picard argues convincingly the following:

> Whether or not affective computing is an area in which you conduct research, you are using emotion when you choose to read this. You are involving your emotion system when you decide where to spend your time—when you act on what matters most to you. Affective computing researchers have a chance to elucidate how emotion works: how to build it, how to measure it, how to help people better communicate and understand it, how to use this knowledge to engineer smarter technology, and how to use it create experiences that improve lives.
>
> (Picard, 2014, p. 19)

Once a research team determines how to measure affect, they can use it to study its role in human behavior. Affect impacts decision making. Researchers have found several interesting impacts of affect on behavior. Positive affect facilitates creative problem solving (Isen et al., 1987). After attending a movie, moviegoers' decisions about people changes depending on if they watched a scary, thrilling, or comedic film (Forgas & Moylan, 1987). Good typography induces a good mood while reading leads to higher performance on relative subjective duration and certain cognitive tasks (Larson & Picard, 2005). Picard argues that emotions' influence on cognition may happen primarily "through emotion's influence on memory" (2000. p. 40). This would be exciting, as it would tie to McCaugh's work (2004) that strong emotional states lead to strong memory encoding. Work on the power of negative emotions, some of the most cited papers in the social sciences literature, influences this work (Baumeister et al., 2001; Rozin & Royzman, 2001). Their work shows that negative emotions outweigh positive ones in affecting lives. Bad parenting, or an insult, impacts behavior much more than good events. Trauma is one of the most impactful events in most people's lives and there is not even a word in the English language for the positive equivalent (Tierney & Baumeister, 2019).

5 Ant Adaptation

The goal of designing Ant Adaptation is to create an agent-based modeling environment for learning complex systems. Social insects provide a compelling context to explore and learn about complex systems. Social insects, like minds or the climate, are self-organizing systems. This means an overall order arises from local interactions between parts of an initially disordered system. This self-organization offers affordances to learning about complex systems, including reasoning about local actions leading to global outcomes. A branch of complex systems theory, Swarm intelligence (SI) in artificial life, shows simple creatures repeating simple rules can display surprising amounts of efficiency and complexity (Beekman et al., 2008; Bonabeau et al., 1999). Social insects, like ants, are excellent examples of these emergent principles (Resnick & Wilensky, 1992; Wilensky, 1997b; Wilensky & Rand, 2015). Though individual ants are quite limited in their functions, ant colonies construct overall order—including bridges, farms, highways of food, and information through local interaction—from an initially disordered system. Ants are good exemplars of the principles of self-organization. In this work, I have designed a digitally embedded curriculum based on ants' self-organizing behavior, and in this section, present the research that informs the design.

5.1 Learning about Complexity with Models and Games

Learning about complex systems can be difficult (Wilensky and Resnick, 1999; Chi, Roscoe, Roy and Chase, 2012; Jacobson, 2011). Computer mediated learning may help with complexity learning. I designed a museum game to teach emergent schemas in short interactions in a museum, Ant Adaptation (K. Martin & Wilensky, 2019). Inserting the complex system content into video games may be a useful avenue of intervention because 97% of children play video games at some point in their lives (Lenhart et al., 2008). One type of game, constructionist video games, seems particularly promising. By encouraging open-ended engagement and exploration, games can support learning across a wide variety of topics and contexts by providing a powerful way for learners to construct new knowledge and understanding (Holbert & Wilensky, 2019; Kafai, 2006; Kafai & Resnick, 1996; Papert & Harel, 1991a). Constructionist learning games try to strike a balance between open-ended play and targeted treatment of learning content through providing learners objects to think with (Holbert & Wilensky, 2019; Weintrop et al., 2012). Constructionist video games employ traditional game structures infused with constructionist ideals to create a game experience that both encourages exploration and engages desired content (Egenfeldt-Nielsen, 2006). These can games mediate open-ended learning through play (Vygotsky, 1966). In this proposal, I propose the use of a game, *Ant Adaptation,* to improve learning about complexity and show children's learning of it in a museum. Taking the previous use of ants as a means to understand, research, and teach complex systems, I designed a complex system model to deliver complexity learning in the short, open interactions that are normal for museum's educational experiences without the mess of installing 20 million safari ants in the Midwest.

6 Measures

Previous museum research has used a very basic, flat coding scheme. These were done at a very superficial level. Constructivist dialogue mapping, introduced in detail in chapter 4, offers a rich connected view since conceptual understandings are elaborated in real time. Beyond high-level codes, these earlier scholars are not really engaging because they only talk about high-level ideas but do not talk about how concepts are elaborated. This is reasonable because they are practitioner focused, but if researchers want to study learning as it unfolds, they will need to track the ideas as they develop and elaborate in a much more rigorous way. Humphrey et al. point towards this when they say, "visitors seem to be constructing a conceptual understanding." (p. 20). This would support the use of CDMs inside of museums. It also fits into the work by Gaea Leinhardt and Kevin Crowley (1998) on learning as elaboration.

Because agent-based models of complex systems are an effective way to learn complex systems, and complex systems underpin the solutions to some of the world’s most pressing problems, educators need more effective ways to teach the subject and measure the learning while they do. But researchers and policymakers have a problem: how do we understand what students are taking away from open-ended learning about this hard-to-understand way of thinking? We need to evaluate the effectiveness of open-ended learning environments. However, as Jonassen argued, “perhaps the thorniest issue yet to be resolved regarding the implications of constructivism for learning is how to evaluate the learning that emerges from those environments” (Jonassen, 1991, p. 1). In general, the literature has documented the difficulties in evaluating learning in informal settings delivered in a constructivist way. Previous empirical studies (Segers, 1996) have shown constructivist environments do not always result in the expected outcomes. Possible explanations have been proposed to explain why these environments do not live up to their promise

(Delva et al., 2000). One problem with these studies is that they treat constructivism as a narrow pedagogical approach rather than a broad-based theory of how people learn (Renkl, 2009). For a review of several constructivist evaluation methods and their shortcomings see Rikers and colleagues (2008). This broad-based theory should include how people regulate learning, which is an affective question; ideas are not just thought they are felt. Even reasoned behavior, such as learning, is neurobiologically directed by the affective states of the learners (Picard, 2000).

Below, I address how evaluation in constructivism, affect's role in this regulation, have been used analyze learning and engagement in open-learning environments. In the chapters of this project, I develop two extensions: constructivist dialogue mapping (CDM) and user engagement tracking with affective computing measures. Then, I will triangulate these measures to identify moments where users are (a) elaborating their ideas and (b) engaged.

6.1 Studying Open-Ended Learning

Piaget's method of clinical interviews studied how people talk and how their speech references the mental models they are using (Piaget, 1926, 1929, 1952b). The Methods highlights the possibility that concepts are not stable. This matters when we study open-ended, constructionist (Papert, 1986; Papert & Harel, 1991) learning activity. The new methods are needed because of a weakness in evaluation methodologies of open-ended environments (Ochoa & Worsley, 2016). I argue in chapter 4 that a form of deep learning with complex systems models can be documented by using our verbal interactive tool: constructivist dialogue mapping. This tool was developed from constructivist and conceptual change theory (Martin, Horn & Wilensky, 2020).

6.1.1 Piagetian Clinical Interviews

How are concepts formed? In my constructivist view, concepts are conserved ontological entities a person builds through action in the world (Piaget, 1952a). So, how do we get at these concepts that people hold? Piaget invented a method, the clinical interview, which is an approach to documenting an open-ended conversation designed to illuminate the way a child thinks or explains a particular phenomenon. Even though Piaget widely employed the clinical interview to examine how children construct their knowledge, there is surprisingly little discussion of the method in his work (Posner & Gertzog, 1982). Piaget elaborated on his data collection method most in the introduction to *The Child's Conception of the World* (1929) and in the preface to *The Language and Thought of the Child* (1926). His method of analysis of the development of cognitive constructions involved observing children as they reason about unusual phenomena that he presented in designed settings. The method involved the following four steps:

- Design an activity.
- Let the child talk.
- Notice the way the thoughts unfold.
- Pose probing questions.
- Do not just notice the answers the child gives to questions posed, but also follow
- the child's line of thought.

Piaget argued that "If we follow up each of the child's answers, and then, allowing him to take the lead, induce him to talk more and more freely, we shall gradually establish for every department of intelligence a method of *clinical analysis* analogous to that which has been adopted by psychiatrists as a means of diagnosis" (*emphasis added,* Piaget, p. 276 as in Claparéde's preface to Piaget, 1926). This approach is a useful way to follow children's understanding of the world. It focuses on the knowledge, or mental models, children construct during an activity. This work fits nicely with Leinhert and Crowley's (1998) idea of learning as Conversational Elaboration. For their pragmatic approach and operational definition of learning, they chose how visitors elaborate conversations. They focus on conversational elaboration because it is a naturally occurring part of the museum experience while also being a product of the experience.

6.1.2 Knowledge in Pieces

Knowledge in pieces is an approach that attempts to describe how conceptual change takes place in an individual over time. Knowledge in pieces (diSessa, 2018, diSessa, 1993) addresses the following questions: what the elements of knowledge are, how do they arise, what level and kind of systematicity exists, how does that system evolve, and what can be said about cognitive process that underly the system and its operation. The study of this framework looks to study moments of learning that are consistent with constructivism. The theory proposes many fine-grained bits of knowledge (p-prims). P-prims are micro-generalizations that people abstract from experience. They are small knowledge structures that get enacted by being recognized and cued to an active state on the basis of the perceived configuration. P-prims are mid-level cognitive elements, neither the low-level sensory information, nor the high-level named concepts or categories. P-prims are activated based on a cuing priority. diSessa does not focus on the origins of p-prims, but instead focuses on their life histories, especially how they might "become embedded in more physics sophisticated thinking" (diSessa, 1993, p. 114). In this way of thinking, one might see the expressions documented by CDM as unstable explanations, perhaps not learned during the interaction, but instead expressions arising from prior experience. In that perspective, participants may not retain much of what they are talking about during these interactions. This is an empirical question: in what ways do the examples of CDMs developed during play inform, or reorder p-prim networks related to complex systems heuristics? Thankfully, the method of CDM allows us to study the stability overtime, and is designed to study the evolution of the ideas. To conduct the study requires a larger, longitudinal study.

6.1.3 Node-mode Framework

In the article, *Some Assembly Required: How scientific explanations are constructed during clinical interviews*, Sherin et al. (2012) argue that conceptual dynamics are messy, but not impossible to fathom. And this fathoming helps us move beyond questions of coherence. The paper attempts to research, and come up with new questions to investigate, common sense science knowledge. They propose a tractable framework the node-mode framework. This framework focused on temporary explanations and dynamic mental constructs. In prior research, it has been found that students exhibit misconceptions, and these misconceptions are resistant to instruction. Their resistance arises, because people enter science with such a substantial body of information about the natural world. As a result, learning science must involve transformation of existing knowledge. But there's still a lot of debate: the field has not reached consensus on fundamental attributes of students' prior knowledge of the natural world. Prior knowledge is a conceptual system. But there's a debate here over whether that conceptual system is relatively coherent theories—as Vosniadou would argue (Vosniadou & Brewer, 1992) or is common sense, scientific knowledge fragmented—as diSessa (1993, 2018) argues. Sherin et al. posit a third way to reconcile this question: suggesting we need a new way of thinking of this question.

In terms of coherence, Sherin et al. propose that sometimes students have coherent theory, and sometimes they don't. And this change depends on the context: Most substantially on the population under study and the structure of the questions. For example, in some contexts students answer as a consistent model, such as flat earth, or hollow earth. But under validation studies, validation seems to fail. And this is not because of some issue with the nature of knowledge but instead, results from the questions. As a result, we must ask how are ideas elicited from participants. The original study (was Brewer and Vosniadou 1992) had open style questions whereas the validation studied used multiple choice. As a result, Sherin concludes that the images we get of common-sense science knowledge depends on the way questions are posed. Answers we obtain will be specific to the population of the subject. The view we get of common-sense science knowledge will critically depend on the methods we employ. In other words, interviews must be tuned to a particular kind of concept layer. Each type of interview provides a complex window into common sense science knowledge of students, and we need to see through data to student knowledge that governs the response; we must understand and model student reasoning as it occurs in the interview we employ to study common sense science knowledge

Sherin et al. (2012) begins a program of inquiry into conceptual dynamics and provides several examples of how to use a framework: node-mode. Some students, such as Leslie from the paper, seem to be working out explanations as they answer problems.

In the paper, they provide examples about the seasons, arguing that just about everyone knows something about seasons but even with that familiarity, it is still very challenging for most people to explain seasons because the explanation is subtle, having to do with correctly and adequately accounting for the tilt of the earth in respect to the plane of its orbit.

6.7.1.1 Definitions of K knowledge in the Node-Mode Framework

To help us understand the node-mode framework Sherin et al. provides a language for describing knowledge. The framework has three basic entities, node, mode, and dynamic mental constructs. The node is a multi-type data structure. It can encompass propositional knowledge as well as lower-level knowledge such as diSessa's p-prims. And while some work has been done to exhaustively list and categorize the nodes, this has proven difficult (Lee et al., 2006). The second entity in the framework is the mode. A mode is the interconnected nodes within the conceptual ecology, that tend to be triggered under certain queueing principles in response to a particular class of cognitive structures. The last part are dynamic mental constructs (DMC). A dynamic mental construct is a product of this modal reasoning—namely, networked reasoning. For example, in one case, a student is asked to explain the seasons, the DMC includes the current state of explanations they have constructed up to that point. It's dynamic because it's temporary, highlighting the fact that a DMC might change rapidly. At any given point in the interview, the map about the nodes and DMC connected to one another to form an explanation, assuming that there is an explanation at that point. So, nodes and modes are considered to be part of long-term memory. They can be active or inactive. When nodes are gathered into modes they allow for modal reasoning, which leads to a temporary dynamic mental construct (DMC). In short, DMCs are temporary and made of nodes to form a mode, a DMC might change rapidly, moment to moment, but it can also show resistance and persistence, as the authors show in the article.

They provide the example of a mental model in textual reasoning to explain the kind of modal reasoning that comes out of this view. In textual reasoning, readers will have a working understanding of a text (Just & Carpenter, 1987) and the mental model forms of the text. This mental model will change and evolve as the text is read.

So, while modes and nodes, are prior knowledge that help understanding, a DMC is the working understanding in the moment, which might change rapidly, and is constrained from moment to moment by the mode, subject and context and is constructed for that moment, so after it will no longer exist. The DMC has several features a DMC can include partly formed ideas, and is a working mental state. It includes knowledge, not incorporated into a working model. And for experts, when the mode is highly constrained, such as a professional astronomer, DMC is, where there is some temporary constriction, which leads to stability in DMCs.

In this way, an interview can be described as questions that trigger a mode, consisting of nodes, particular to the context. An example of a DMC would be the sun, which is a source of energy, and days longer in the summer, shorter in the winter, which accounts for why the shift in seasons.

This DMC, the authors believe, contains a p-prim: more effort/input leads to greater results. Though there's no evidence of this p-prim from the data. In the remainder of the article, they give examples of different DMCs and say, "for many reasons we cannot be very certain about the particular knowledge that underlies Leslie’s utterance". This admission is the fundamental difference between constructive dialogue mapping and the node-mode framework I believe we cannot be certain of anything below the observable speech, which is the primary data. To infer about the mental constructs that lead to that knowledge is simply not possible from speech data. The fundamental argument for the node-mode model, is shared with CDM: we do not need to look upon these models as final models, it is possible and sometimes important to examine how participants get there. We can look at the knowledge that participants drew upon, the process that led her to assemble the knowledge, and the way she did. The point is that the more inclusive real time analysis of Leslie's reasoning is tractable and provides a more complete picture of Leslie's knowledge (Sherin et al., 2012, p. 177).

Sherin et al. point out some common DMCs include rapid shifts and that stable DMC s which are most akin to the type of mental model identification appearing in studies of common sense science. They introduce some important phenomena, such as mode skimming, where when asked a question, a student will survey their knowledge and science context to prove what they know. Sometimes there is mode constrained DMCs.

I would ask after reading this paper, how can a DMC be stable, if it is temporary and dynamically constructed. The answer being only when the context of both the interview, and the subject population is stable. DMCs also can shift. So, there are both stable DMCs, which are resistant to challenge, but also the DMCs which are abandoned, which do not resist a challenge. A good example of this was in Angele's interview. In this view we have shifting modes. Consequently, we have modal reasoning, like trying to put puzzle pieces together, instead of like discarding the entire puzzle and starting anew. An important question that the paper does answer is why a DMCs in certain situations are so stable, and they say for two reasons. The prior knowledge overlaps of the participants and the consistency in the interviews. This means the persistence of some of the same constraints gave the dynamics across students' preference particular DMCs providing the observed stability.

In the end, their approach wants to see through the features to see the student knowledge that in part generates the data that we perceive. This seems like an overly idealistic goal since we only have speech, and we must infer what's below. In this way, the speech is the proxy for these dynamic mental constructs. The manuscript is largely an exercise in attempting to convince the reader that we can possibly see through the observable features to do this work.

They close the chapter with an exciting consideration: The node-mode framework and DMCs in a social cultural theories of science learning. The node-mode framework ultimately is characterizing individual knowledge, but that does not mean they don't think of group knowledge. In this way, they see the interviews as an exemplar dynamic interaction that must themselves be understood to properly analyze our data. That the approach is amenable to studying larger groups. It also is amenable to studying interactions between groups and representational infrastructure. This is an affordance that constructivist dialogue mapping is continuing to explore. In Sherin et al.'s approach DMCs rely on the students' prior knowledge, but also things like the affordance of a drawing which is an example of the representational architecture. They assume that the behavior they see in an interview emerges in interaction with the interviewing context. They suggest this may be useful in teaching because they say teaching when they're trying to elicit particular conceptions, is kind of like an interview. They suggest teachers are likely to see mode skimming and the way they pose questions should impact students' response.

There are two main lingering questions for the framework:

- How are the modes made?
- Can a DMC turn into a node or mode?

6.1.4 Constructivist Dialogue Mapping Compared to Node-Mode

There is a tension though: how long do we think knowledge structures pointed to by constructivist dialogue mapping persist? On the one hand, there are short explanations that people come up with based on their prior experience, and on the other hand long term, stable cuing explanations, resulting from networks of p-prims. Sherin et al. (2012) argued for the field to move away from questions of persistence, and instead argues that understanding the intermediate versions of reason will help us better understand the origins. It is possible and sometimes important to examine how participants' reasoning gets there. We can look at the knowledge that a participant drew upon, the process that led her to assemble the knowledge, and the way she did. The point is that the more inclusive, real time analysis learner's reasoning is tractable and provides a more complete picture of knowledge. Constructivist dialogue mapping attempts to present these intermedia steps that form during learning interactions. In node-mode terms, an individual node in a CDM is node in node-mode. The entire constructivist dialogue map is a dynamic mental construct. The mode would be the context cued by the setting. We would expect the dynamic mental constructs to have stability in as much as the setting does, that is the subject, the context of learning, and the interaction. In other words, in a persistent social group, playing the game in a persistent group, we would expect from node-mode to see persistence in the dynamic mental constructs and observe slowly changing CDMs. But Sherin et al. also noted rapidly changing dynamic mental constructs, and this is because they are temporary, and develop to explain the current moment. The remaining question is when does a dynamic mental construct turn into part of long-term memory, forming nodes and modes? Constructivist theory (Ackermann, 2001; Piaget, 1952b, 1983) suggests this happens through action in the world, a world that for most of us shares many shared cuing characteristics which can lead to stable CDMs from stable modes.

The main connection between constructivist dialogue mapping and Node-Mode is we both paid attention to the fact that there are these dynamic constructions during interviews. When Sherin et al. wrote the paper, people hadn’t really paid attention to this dynamic in the moment explanation. And I am paying attention to that same thing.

So how long do CDMs persist? One could ask that about any teaching at all, how long does any of it persist? I would say that if any knowledge is touched at all by the interaction, then it adds to plausibility there is learning in the interaction; there might be some effect on it. The question I begin to investigate in chapter 5 through affect-detection, but could get at the spirit of the question is that there could be these critical constructions, a student gets a *click*, and a student gets a mental bookmark, because it’s pushed a satisfying button, I have an intuition these moments are special kinds of moments in these interviews. Identifying these moments of click is a key idea of investigating the idea of persistence of knowledge going forward, that could help us evaluate moments of learning more broadly in education. Studying these moments could potentially be done through affective-state tracking, an approach I employ in chapter 5.

6.2 Measuring Engagement

Measures of engagement in museums have been rather pedestrian, like numbers of minutes engaging, but we have new technologies that more directly measure engagement. We can collect biometric or facial expression data. These are a more direct measures of physiological and psychological engagement (K. Martin, Wang, et al., 2019). Facial/biological measures of affect can help us understand the relationship between positive affect and learning (D'Mello & Graesser, 2014b; Picard, 2000). In this study, as I lay out in detail in chapter 5, I will track facial expressions using videos of faces (Cohn and De la Terre, 2014).

Chapter 3: Prevalence of Direct and Emergent Schema after Play[9]

1 Introduction

Learning about complex systems can be difficult (Chi et al., 2012; Jacobson et al., 2011; Wilensky & Resnick, 1999). Computer mediated learning may help with complexity learning. I designed a museum game to teach complexity in short interactions in a museum. Inserting the complex system content into video games may be a useful avenue of intervention because 97% of children in the U.S. play video games at some point in their lives (Lenhart et al., 2008). One type of game, constructionist

Fig. 1. Luck and skill weave the action of an emergent system of ants into the players' attempt to beat their opponents as they swipe and touch interacting with an ant microworld.

[9] This chapter was previously published in Informatics in Education.

video games, seems particularly promising. By encouraging open-ended engagement and exploration, games can support learning across a wide variety of topics and contexts, providing a powerful way for learners to construct new knowledge and understanding (Holbert & Wilensky, 2019; Kafai, 2006). Constructionist learning games try to strike a balance between open-ended play and targeted treatment of learning content through providing learners objects to think with (Holbert and Wilensky, 2019; Weintrop, Holbertand Wilensky, 2014). Constructionist video games employ traditional game structures infused with constructionist ideals to create a game experience that both encourages exploration and engages desired content (Egenfeldt-Nielsen, 2006).. These games mediate open-ended learning. In this chapter I propose the use of a game, *Ant Adaptation,* to improve learning about complexity and show children's learning of it in a museum.

1.1 Social Insects in Complexity Education

The exhibit presents visitors with the opportunity to create an interactive digital ant ecosystem by exploring a microworld(Edwards, 1995; Papert, 1980) Doing so allows learners to better understand the world of ants. Previous work with agent-based models of social insects and microworlds has been effectively used for classroom teaching (Wilensky and Reisman, 2006), but less so in informal museum settings. Examples of my classroom use include *Ant Food Grab,* a yearlong block-based programing curriculum with ants (K. Martin et al., 2018; Sengupta et al., 2015), and other world with social insects include *Beesmart*, a curriculum about the hive-finding behavior of honeybees (Guo & Wilensky, 2014, 2014). These examples highlight that observing and/orinteracting with insect microworlds allows students to construct their own understandings of science through exploring and adapting models in a classroom setting. Like other complex systems interventions, the students build their understanding of complex systems in natural systems. The literature has depicted learning through diverse simulations of natural systems: bees (Danish et al., 2011), material science (Blikstein & Wilensky, 2009) (Blikstein and Wilensky, 2009), multi-level modeling (Wilensky & Reisman, 2006), electricity (Sengupta & Wilensky, 2009) and evolution (Wagh & Wilensky, 2018). However, there is much less work that leverages the pedagogical value in these earlier microworld based interventions in museums (Horn et al., 2014; Strohecker, 1995a, 1995b) a dearth the present chapter partially addresses. Informal education is different from formal education in a few key ways, including:

(1) Informal education does not provide teachers.
(2) In classrooms, teachers may take days or weeks, to cover a topic whereas in museums, a topic can be covered in as little as two to ten minutes.
(3) In open-ended learning there is no coercion, as the participant can walk awaywhenever.

In fact, *holding time*, that is the number of minutes a participant stays at an exhibit, is a key measure in museum evaluation.

As described in chapter 2, although microworlds can be useful for learners to explore complex systems, developing robust understandings of complex systems can be challenging. Wilenskyand Resnick (1999) describe the difficulties people have in "thinking in levels," exhibiting "levels confusion" and difficulties with distributed control and stochastic processes. Not only do learners have a hard time thinking across multiple levels such as disease of the whole body resulting from microscopic pathogens, but they tend not to think about phenomena such as the flow of ink dropped into water as the processes of collectives of agents interacting (Chi *et al.*, 2012). If the glass of water changing color is explained by the individual parts of ink interacting with H^2O, then the process becomes more intuitive. Most people, however, are not familiar with ink particles. Therefore, building off earlier work in classrooms (K. Martin et al., 2018) and theory about reasoning about population dynamics(Wilensky & Papert, 2010), I hypothesize that thinking about ants isa powerful way to restructure both of these challenges. The game, *Ant Adaptation* (K. Martin & Wilensky, 2019), scaffolds thinking in levels (Wilensky and Resnick, 1999) such as between the colony and the individual ant, and seeing success and failure of the colony as a result of ants interacting through processes. To research this claim, this chapter explores how visitors made sense of the self-organization of ant colonies, and how this changed their schemas.

1.2 Constructionism

Constructionist thinking influences the learning sciences and educational research, particularly when addressing learning technologies and reform of mathematics and science education. Papert coined the name Constructionism, as mnemonic (Papert, 1986), to describe a species of constructivist thought. It focuses on the benefits of learning from the external construction of an artifact beside the internal construction of a mental model, or framework. Logo, the name derived from the Greek for *word* or *thought,* was the first constructionist programming language. Wally Feurzeig and Seymour Papert invented Logo in 1967 (Logo, 2015). Using tools like Logo, constructionists organized learning environments where learning was free from time constraints. Papert coined the term mathetics to describe the art of learning, and argued “My mathetic point is simply that spending relaxed time with a problem leads to getting to know it, and through this, to improving one’s ability to deal with other problems like it”(Papert, 1996, p. 12).

Constructionism burst into the public eye after Papert published his seminal work *Mindstorms: Children, Computers and Powerful Ideas* (1980). Papert used Logo to operationalize many of the big ideas described in *Mindstorms* using the idea of the turtle. The turtle is a single software agent that can represent many different organisms, such as a turtle in sheep's clothes. This paradigm allowed for programming from an agent-based perspective. While, the first versions of Logo had only one turtle, Kala and Blaho's *Imagine Logo* (Blaho et al., 1993) included object-oriented features and multiple turtles. Since this time, the notion of constructionism has inspired many powerful tools for education including NetLogo (Wilensky, 1999b), a multi-agent programming environment, and Scratch (Resnick et al., 2009), a blocks based programming environment. These environments have been used to make powerful mathematics and scientific exploration tools that afford learners the ability to act in sequence or simultaneously on multiple representations of a phenomena. Such a creation can produce contexts in which group-level understanding is constructed and contested. These novel restructurations have powerful impact on learners (Wilensky and Papert, 2010). For example, they describe when accountants moved from Roman numerals to Arabic numbers, operations such as multiplication and division became significantly easier because of the new representation's affordances. The various new representations are the focus of both internal reflection and external action that foster shared meaning, positively mediating groups' sense making.

Papert (1993) argued that the advent of the restructuring of digital worlds that children can explore will create less patient, accepting students. "Children who grow up with the opportunity to explore the jungles and the cities and the deep oceans and ancient myths and outer space will be even less likely than the players of video games to sit quietly through anything even vaguely resembling the elementary-school curriculum as we have known it up to now" (p. 9). Though constructionist environments have often been long term explorations of motivating problems, in the museum setting we needed to deploy a system to explore, but that provides an experience in as little as two minutes. This design constraint motivated my use of the constructionism to teach complexity. With this model, I examined the type of learning that occurs between a user, and a complex system model because I, like Papert, agree that in the coming educational environment, learners will be less patient as they seek their own meaning. In this chapter, I provide an example of a learning environment that scaffolds individual meaning-making through touch, discussion, and play for museum patrons.

1.3 Ant-based Modeling

Early work on social insects motivated work on agent-based modeling. Early workon agent-based modeling was inspired by the behavior of social insects (Langton, 1997; Resnick & Wilensky, 1992, 1993). Ant behavior has inspired games. SimAnt (McCormack & Wright, 1991) is based on Hölldobler and Wilson's (1990) *The Ants*. The collective behavior of ants has been simulated using agent-based models many times. StarLogo was used to model the collective behaviorof social insects (Resnick and Wilensky, 1992, 1993). Wilensky (1997a; 1990) modeled food source preferences resulting from pheromones as well as the formation of ant trails (Wilensky, 1997b). Bonabeau investigated the role of agent-based modelsin pattern formation (1997) and more broadly, looked at swarm intelligence (Bonabeau et al., 1999). Pratt (Pratt, 2005) modeled collective nest selection of *Temnothorax albipennis* also using an agent-based model. Pratt's work explored the importance of group decision making with quorums (2009). The work showed that when choosing a destination together, cooperation reduces the probability that an individual will suffer predation. Robinson, Ratnieks, and Holcombe (2008) used an agent-based model to explore attractive and repellent pheromones in pharaoh ants. Likewise, frameworks, such as Anthill, have been used to support the design, implementation, and evaluation of technical systems, such as peer-to-peer networks (Babaoglu et al., 2002). Their work drew on examples of complex adaptive systems to justify engineering and user applications because complex adaptive systems exhibit resilience, adaptation, and self-organization that are seen as valuable in social applications. While these earlier models provided insights and enjoyment, none of them delivered their lessons in the short interaction times typical of museums. These earlier works helped experts understand ant colonies and applied lessons learned to understand other self-organizing systems. Taking this previous use of ants to understand, research, and teach complex systems, I designed a complex system model to deliver complexity learning in the short, open interactions normal for museum's educational experiences without the mess of installing 20 million safari ants in the Midwest.

1.4 Making Sense of Complexity

Chi *et al.* (2012) argued that all processes can be categorized into two types: sequential and emergent. All processes share seven characteristics:

(1) They can be described at the agent or at the pattern level[1]. In ant colonies the pattern level is the whole colony, whereas the individual is the behaviors of each ant. She provided the caveat that the agent itself is a collection of patterns atthe micro-organismal level. Though in some systems, I can leave this aside in studying sense making in complexity, the aside becomes important in ants. The colony is a superorganism, that is composed of several ants. The ants can bedecomposed into microsystems like their microbiomes, and macrosystems like colonies, or collections of colonies. Mentally moving up and down these levels lets us think about causation.

(2) At both agent and pattern level, agents can be clustered into aggregates of subgroups, such as chasing

wolves or blocking wolves on a hunt, or workers that clean up a colony's waste, or forage for food.

(3) Behaviors can be distinguished into levels and behaviors can differ in different levels.
(4) At the agent level, the behavior of agents can be conceived of as interactions between agents and not simply individual agents.
(5) Conditions that elicit behavior at each level differ.
(6) Processes can be visible or invisible. For instance, in our model, I hide the actions that occur inside each colony.
(7) Information about each level can be infinite, but misunderstanding them is not the same thing as lacking this information.

Processes, all of which share these seven characteristics, can be categorized in two ways – sequential or emergent. Sequential processes can be subdivided into a sequence of events, like an assembly line where the metal is rolled, pressed, stamped, and smoothed before being turned into cans and filled with tomatoes sauce. This is a process with multiple agents: the machines, their operators, and a manager. It makes sense to speak of sequential processes resulting from a single agent's actions. For example, I could say the manager increased efficiency of the assembly process by increasing the speed of the conveyor belt. Even though all the agents participated in the process, I can focus on thegoal setting of manger's actions to account for the change. As a result, Chi *et al.* (2012) say one can give special controlling status to the agent that caused the change. And this control means that interactions at the agent level are done *to* reach the goal of producing cans of tomato soup. The causal mechanism that leads to the result comes from the summing of the assembly line's outcomes directed by the manager. Many people confuse this control when thinking about emergent processes.

Emergent processes, like ants searching for food, marching in orderly paths, or getting stuck in a doorway are slightly different. These processes result from each ant taking actions, where the result emerges from the repetition of the action, but no agent is in control. These processes are encountered in school standards such as osmosis and diffusion, electrical current and buoyancy. These processes also appear in serious issues for humanity like climate change and nuclear arms proliferation. In ant colonies I can see emergence (Wilensky, 1997b, 1997a) in ants searching for food, then filing in a straight line along a scent trail toward food, made by each ant leaving pheromone as it returns. Then when the flow of ants returning from the food fades, the ants discover and construct new trails, and through feedback, select to file along that pheromone scent's gradient. The pheromone trail is a self-organizing process, that organizes ants without a central agent controlling their actions.

Users could understand this process as either a sequential process, where a queen tells them what to do, or an emergent process. In an emergent process, all the agents follow simple rules of taking a random walk until they find food or a strong pheromone trail. If they find food, they return it to the nest leaving a pheromone trail. If they find a pheromone trail, employing a hill-climbing mechanism, they follow it toward the strongest scent; the strongest scent is always toward where the last ant just came from carrying food. An emergent process can be identified by four characteristics (Chi *et al.*, 2012):

1. Cannot attribute action to one agent, such as a queen.
2. All ants have equal status.
3. No actions are goal directed.
4. Pattern is a collective outcome.

As a result, Chi *et al.* argues that people misconceive emergent processes as sequential processes. "Misconceptions reflect the use of attributes of an alternative Direct Schema to explain non-sequential processes that ought to be explained by attributesof an Emergent Schema" (Chi, *et al.*, 2012, p. 12). They assume the emergent schemais missing for most students. I set out to test people's beliefs in direct or emergent schemas and to test whether a short interaction with an agent-based model might shift their schemas.

Chi posits that the direct and emergent schemas are incommensurate, that there are not transitions, and that emergent schema have to replace direct schema. Wilensky*et al.* (Levy

& Wilensky, 2009; Sengupta & Wilensky, 2009) rejects the claim of incommensurability, and argues that many features of direct schemas, can be used to understand emergent phenomena, and there is a continuous back and forth between these schema. Some of the data I describe below supports that shifting in as little as 10 minutes.

1.5 Research Question

Working with agent-based models of insect colonies may improve students' understanding about macro level, emergent patterns, such as population graphs that result from all the agents' actions. In this chapter, I explore how users make sense of the self-organization of cooperation between ant colonies in competition with each other while they use an agent-based model built in NetLogo (Wilensky, 1999b). In this way, I buildoff the arguments of Wilensky and Papert (2010) to test Chi *et al.*'s assumptions and assertions about emergent schema. Wilensky and Papert argued that people can better reason about populations when they can observe the individual agents in them:

> *Students can reason about and visualize individual animals in an ecology far better than they can population levels. They can draw on theirown body and sensory experience to assess and/or design sensible rules for the behavior of individuals. They can therefore make much greater sense and meaning from the agent-based representations.* (Wilensky & Papert, 2010, p. 8)

I researched how users shifted their schemas when using the model of ant colony life. I asked the following question:

1. How does an experience designed to facilitate change in users' direct schema affectusers in a short interaction in a museum?

To test this, I specifically asked before and after the use of an agent-based model how ants know what to do, how they collect food, and how they deal with traffic. Each question was asked to elicit an understanding of people's emergent or direct schema. If they answered that it is the queen ant that tells them what to do, this would show direct schema of a single agent. If they answered that the collective action of all the ants following and leaving scent leads the ants to do what they do, it would show an emergent schema. If they said animal instinct guides them, I inferred whether they meant that some global force controlled the individuals (direct), or whether they meant that the agents' individual actions led to global patterns (emergent).

We, inter-rater reliability codes and I, coded the data for users pre- and post-gameplay for direct and emergent schema.

I had four hypotheses:

H_1: *Users had a direct schema before and after* (Direct → Direct).

H_2: *Users had a direct schema before but changed to an emergent schema after* (Direct → Emergent).

H_3: *Users had an emergent schema before and after* (Emergent → Emergent).

H_4: *There was one more possibility, that people changed to a direct schema from an emergent schema prior, but based on previous theory (Chi et al., 2012), this seemed highly unlikely.* (Emergent → Direct).

2 Design: Agent-based Modeling Game for a Museum

Next, I will describe the model/game and discuss the design decisions I took because of implementing in the museum, both of which I describe next.

With *Ant Adaptation*, I aim to realize the promise of agent-based modeling originally illustrated by systems such as Gas Lab (Wilensky, 1999a) or NetLogo InvestigationsIn Electromagnetism (Sengupta & Wilensky, 2009), but in a rich tangible interaction form factor for walk-up-and-play use in an informal learning space. As shown in Fig.2a, *Ant Adaptation* simulates two ant colonies side by side. It tracks a user's touch placed on the digital displays surface. The sensing area of the screen contains five widgets for each team. As demonstrated by the corresponding indexes in Fig. 2a: (1)At the top, there is a counter of the ants' population labelled "Black Ants" on the left, and "Red Ants" on the right. (2) The three widgets at the bottom left and right are sliders players can use to adjust their ant's size, aggressiveness, and the maximum amount of energy, or basically how long ants can walk without eating. Adjusting any of these sliders will change the Create Cost. These sliders can be adjusted at any time duringthe game to experiment with different settings. (3) In the middle, *Ant Adaptation* pro-

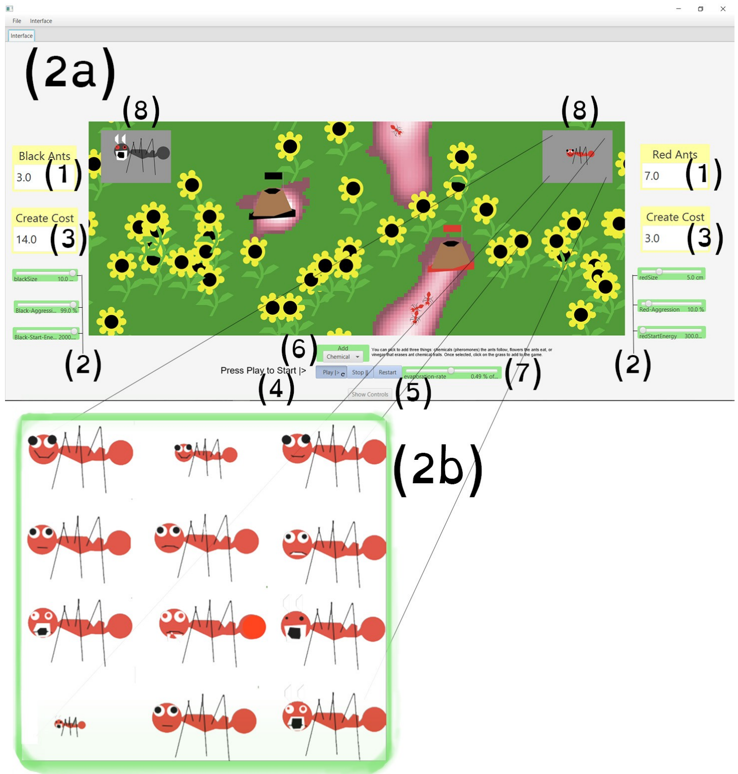

Fig. 2. 2a: Above, Ant Adaptation has two ant colonies that players interact with through touch and control widgets that count the ant population, display the cost to create another ant, and set what the user adds to the model. 2b: We see nine ant displays for varying aggression and size. There are 100 total possibilities to display providing immediate feedback to players as they adapt their ants.

vides a widget "Create Cost." The Create Cost widget shows the summed value thatit currently costs to produce one more ant. This is an example of what Chi *et al.* callan opaque summing mechanism that NetLogo designers employ (2012), where "the collective mechanism is computed by the NetLogo system itself, thus [left] opaque tothe students" (p. 21).

Specifically, the colony produces a new ant when the stored food is greater than the Create Cost. The cost is calculated by adjusting three sliders. A colony's Create Cost is equal to the current size of ants divided in half, plus the current aggressiveness divided by 15, plus the current maximum energy ants can pick up divided by 1,000. Consequently, if players set their ants to be size 10, with 100 aggressiveness, and starting energy of 2,000, each new ant would have a Create Cost of 14 food stored in the colony. In other words, because 10 / 2 = 5, 100 / 15 = 6.6, and 2,000 / 1,000 = 2, if we round up to the nearest integer, the Create Cost equals 14. Because the outcome of the calculation is the current cost for the colony to birth one more ant, this summing mechanism was the topic of many players' strategizing as they tried to maximize their populations.

To supplement NetLogo (Wilensky, 1999), I have developed software designedfor touch interaction with the model. The software allows users to adjust the slider widgets and swipe on the screen to interact with the model with their fingers. The users share five widgets in the center of the screen. As shown in Fig. 2a, (4) Play and Stop, which control the model's time. (5) Restart, which sets the model back to initial conditions. (6) Add, which allows users to control what is added to the model when they touch the screen. They can choose to add chemical, flowers, or vinegar. Chemical is a pink pheromone that attracts ants toward the highest concentration of it displayed in whitish-pink shades. Flowers are these ants' main food source. Ants collect and eat the flowers to feed themselves and bring food back to the colony for collective rearing of young. Vinegar erases ants' trails allowing the playerto mask pheromones, disrupt communications, and clear the ground by applying vinegar to the chemical trails. (7) A slider to control the evaporation rate of pheromones. (8) Lastly, there are two representations of the players' ants in the top right and left of the play space. These show the user how large and aggressive their ants are when born. As shown in Fig. 2b, the display changes according to the mixture of aggressivenessand size the player chooses for their team. This provides the player immediate feedback for changing slider parameters, giving them a better sense of cause and effect inthe model. This is important because adjusting the sliders only affects new ants born, instead of extant ants. So the effect of the interaction is longer than the 30 to 60 Hz, 16.6 to 33.3 millisecond periods, people associate with cause and effect within games (Gregory, 2018).

In this chapter, I present findings of how users shifted their schemas with our platform design by reporting on a deployment in a major natural history museum – the first deployment of NetLogo Touch.

To provide context for the analysis below, I describe the action of the game with and without user interaction. Even without user interaction, in the game I created, *Ant Adaptation*, ants go out to collect food and return to the nest. As they return to the nest, ants lay down a pink pheromone that attracts others nearby. Other ants walk toward the strongest chemical smell, which in most cases is where the first ant just arrived. When ants find a flower, their food source, they return, lay down more pheromones, and thus reinforce the pink trail. This creates an emergent feedback loop that routes more and more ants to successful sites of forage. As the ants exhaust a food source, they must find new locations and thus repeat a cycle. When two or more ants of opposing colonies encounter each other, they fight or scare each other away also leaving chemicals that attract more ants. For the winner, this works to protect the food source from competing colonies. The ant queen reproduces when the ants in her colony collect enough food, in other words when collection surpasses the current Create Cost. Flowers periodically growing up around the map, add food to the game.

The player interacts with this complex system by adding pheromone trails that the ants follow, as well as adding sources of food (i.e., flowers) to the system, thus changing the amount of food in the game. By interacting with the system, users form a functional understanding of the ants and their mechanisms of action (i.e., agents and their rules) in the model. This design scaffolds experimentation. Players must simultaneously

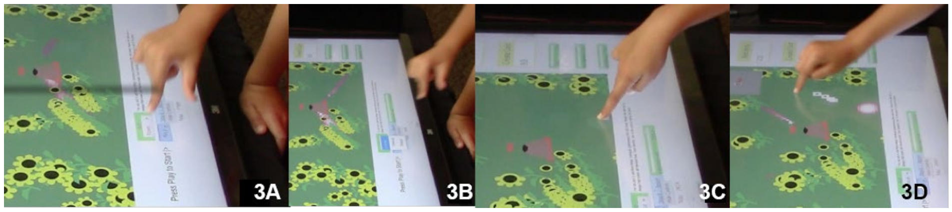

Fig. 3. John selects to add chemical (3A3B), experiments with the touch user interface (3C), and then lays down his first pheromone trail (3D).

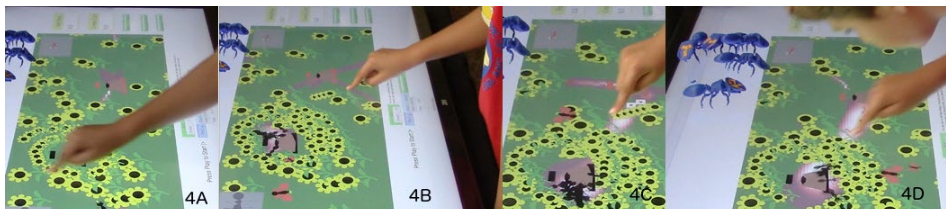

Fig. 4. Player draws flowers (4A), draws pheromones to lead ants (4B), tries to connect his two colonies (4C), and drives his ants through flowers (4D). Leaning forward as he becomesso engaged.

make choices. As shown in Fig. 3, players can touch the screen to add pheromones the ants will follow. As shown in Fig. 4A, at the flick of a switch, players can add more flowers anywhere they like in the game, acting like an ant or a seed, but with a bird'seye view. Lastly, they can choose to apply vinegar (Fig. 4B), which erases trails. Erasing trails was used by some game players (like Thomas, discussed later) to get ants out of a feedback loop that had them stuck in local optima, and was leading them nowhere. For players to achieve their goals in the competitive environment, they are required to understand the emergent consequences of simple ant behavior.

Players can decide how big and aggressive ants are. When the size of ants increases, they become slightly faster and stronger in a fight. Each level of increase adds up. At the highest levels they are 13 times stronger. When players make their antsmore aggressive, it increases the radius in which ants detect opposing ants and thusthe probability that they will attack. Increasing either the size or the aggressivenessalso increases how much food is required to raise an ant, so the largest ant requires13 times as much food to feed to adulthood. This gamification impacts how muchfood ants must collect to make a new baby ant. Increases in either of these parameters reduces the expected population of the colony, by increasing the Create Cost, thoughit increases their likelihood of fighting and winning through emergent interactions of parameters (size and aggressiveness) and agent actions (collecting food, leaving trails, and fighting).

This sets up the main action of the game as a series of strategic choices–to decide whether to pacifically collect food, thereby increasing the population, or to go on the warpath where big, aggressive ants conquer their opponents. Either methodof play could lead to high populations or the elimination of the opponent through better-controlled food resources. For example, after learning about the consequences of strategic choices through gameplay, players strategize by increasing ants' size, aggressiveness, or both. This might lead them to win the game by annihilating the other group's ant colony. However, bigger and/or more aggressive ants consume more foodto reproduce and potentially reduce the colony's population size. Thus, a player might strategize by adding more flowers and pheromone tracks around the colony to helpthe larger ants survive. This learning and strategy cycle interweaves the learning into the gameplay.

The game has four affordances that support two learning objectives. In *Ant Adaptation*, playing with parameters allows players to, (1) construct their colony in competition with an opponent; (2) share strategies through comparison; (3) discuss whatis happening through observer scaffolding such as parents' intervention or interaction between players, including slapping hands; and (4) learn about the emergent impactsof colony behavior arising from individual ant behavior in a complex system game.This approach allows visitors to learn (1) the impacts of adaptation on ant colony life and (2) how attractants such as pheromones work in ants' organization to increase the population. These together have also the potential to scaffold learners in switching from a direct to an emergent schema. I built the game to be engaging while building off theliterature of designing digital interactives for museums, which I will cover next.

2.1 Designing Digital Multi-Touch Tables for Museums

Prior research on building interactives in museums informed our design. Current researchon multi-touch tables for museums suggests several key design elements (Davis et al., 2015; Horn et al., 2016) such as enjoyment, comparability, and productive conflict. Enjoyment, expressed through affect words such as "whoa," "wow," "cool," and "hah," is significantly correlated with learning measures considered by Horn *et al.* (2016). Facilitating comparisons-aided learning in the case of a tree of life game where players who drew comparisons between lineages learned more easily and were more likely to use terms of interest in open-ended questions on post-tests. Block *et al.* (2015) and Horn*et al.* (Horn et al., 2014) found that groups of two spend more time at an exhibit and engage more with scientific content than groups of three or more.

Finally, conflict can be productive. Davis *et al.* (2015) and Falcão and Price (2011) argue that interference between users on and across a multi-touch interface can be productive for learning when it triggers argumentation and collective knowledge construction. From the review of this literature, I hypothesized that designers should encourage discussion and comparison in a competitive game mode where the biology and complexity science weave into the experience. To design this experience, I created interactive, agent-based, complex system tabletop games for museum settings that expedite learning in these short interaction times. As a result, (1) I implemented turns into the play, as Block *et al.* (2015) found that groups who take turns spent longer times and engaged more with biological content in the tree of life exhibit. Taking turns both increased the use of learning terms and comparisons with the biological content. (2) Our design included two teams, which allowed players to explore the game's possibilities and compare between strategies.

These two design elements facilitated comparison and discussion between teams. Exploration involves players moving their bodies and hands across a digital tabletop to engage in a game. This process engages the group at play more than mousing at a keyboard. I hypothesized based on Wilensky and Papert (2010) that a body in motion, talking out ideas would create a rich discussion and a problem-solving mindset around the game interface. The game includes hampering the competing colony through players' dexterity, or at times, physically blocking other players' hands to develop another colony's strength following on earlier work on productive conflict (Pontual Falcão & Price, 2011). This competition is mediated through the luck of the stochastic system. Tradeoffs of adaptations and complex systems thinking are woven into the game, which allows users to explore and learn about complex systems and ants by making strategic choices both in the digital microworld and while standing in the museum. The design encourages talk and comparison to maximize learning about ant behavior, a complex system, in a short interaction at the museum. The argument is that highly engaged play with a compelling, complex, biological systems model teaches users about a complex system through open-ended play.

As shown in Fig. 5, in the deployment I set out a table, a large poster motivating the game, and the multi-touch display in a hallway of the museum where users could walk up and use the game. The hallway was a medium traffic zone chosen for better acoustics for recording interview audio.

As shown in Fig. 6, the poster scaffolds the interaction by proposing a scenario, that players should imagine they are playing as natural selection. Then it suggests some exciting adaptations ants have evolved. Then it suggests a few things to try, such as pressing Restart or Play. Finally, it poses several questions players should think about while playing.

Fig. 5. Setup of the game in the museum. Poster was set out to scaffold participants' interactions. A video camera recorded play sessions.

Fig. 6. Poster Displayed Scaffolding the Experience.

3 Methods

We, other interviewers from the lab and I, tested the game in a major natural history museum outside of a large, popular exhibit. The game was used over a six-day period, and we ran a supervised treatment, with 114 visitors playing in 38 groups. Additionally, I ran the game unsupervised for two 30-minute segments on each of the six days. I collected video, audio, and field notes of all participants' play.

We developed the data in this chapter by watching the videos during these periodsof use and analyzed the transcripts. In the chapter, I present coding of the pre-and post-gameplay interviews for the number of participants that held direct and emergent schemas before and after the intervention in groups that contained at least one participant less than 18 years old. Finally, I show through analyzing the video how one user, Thomas, taught his older brother, Ed, an emergent schema. Interestingly, he also was solidifying an emergent schema through his own interaction with the game and his family at the same time.

I analyzed each of the pre- and post-gameplay interviews for groups with at least one player between ages 6 and 18 years old. I analyzed 27 people (23.7% of the total participants) in nine groups (23.7% of total groups). I examined the participants' answers on three questions (Table 1):

(1) How do ants collect food?
(2) How do ants know what to do? and
(3) How do ants deal with a situation like traffic, where they all keep bumping into each other?

For each group there could be an answer for each individual, and so I coded each individual in the group separately. Often, however, one interlocutor answered most of the questions. An inter-rater reliability coder and I coded 20% of the data and had greater than 90% interrater reliability. Disagreements were discussed as a group, and we reached a consensus.

In our coding, emergent schema was defined by four characteristics (Chi, *et al.*, 2012):

(1) The person cannot attribute action to one agent, such as a queen ant;
(2) All the ants have equal status;
(3) No actions are goal directed; and
(4) The pattern is a collective outcome.

We used these characteristics to code users' statements.

We had four hypotheses about how participants would change after playing *Ant Adaptation* and discussing:

H1: Direct → Direct

H2: Direct → Emergent

H3: Emergent → Emergent

H4: Emergent → Direct

Table 1

Coding book for pre- and post-gameplay interviews Code Book

	Typical Answer Before	Typical answer After
Q1: Have you noticed anything about ants?	"They can carry, like 50 timestheir own body weight?"	"They make trails to follow."
Q2: How do the ants collect food?	"They use their snappy thingies.I can't remember what they are called." (Unclear)	"They would go and pick up a piece. And one by one keep working and working. They went off one by one to grab a piece one by one. Then theywould bring it back to their ant. And then keep working and working." (Emergent)
Q3: How do ants know what to do?	"Well. They got big heads. They are really smart. They use their heads. They use their antenna to know what to do." (Direct)	"They just keep following the chemicals?" (Emergent)

4 Findings

I tested the game in a major natural history museum outside of a large, popular exhibit. The game was used over a six-day period by 114 museum visitors (87% White, 4% Black, 5% Asian, 3% Latinx). Of the players, 60 were male (51.57%) and 54 (48.43%) were female. This contrasts with the museum-wide attendance demographics of 70% White (difference of +16.61% points), 5% Black (difference of -0.54% points) and 14% Latinx (difference of -11.32% points). Fig. 7 shows players ranging in age from 2 to 55, with the age distribution skewed to lower ages. The average length of time people played was 387 seconds, as opposed to a museum-wide average interaction time with digital interactives of 105 seconds (as reported by internal museum evaluations).

The results indicate that the design held participants more than the average interactive experience in the context of the museum. Unlike Block *et al.* (2015) and Horn *et al.* (2015), I found groups of three or more spent more time at our exhibit than groups of two. Worryingly, the design seemed to appeal to certain demographics. When I bifurcated playtime by race and gender, however, I found that non-white users who engaged with the game were some of the most engaged users.

As shown in Fig. 8, fourteen of the seventeen non-white users engaged with the game for longer than four minutes. Though most players in our sample were white museum patrons (97), the non-white patrons engaged with the game longer than four minutes, on average. Because players could stop playing whenever they wanted, unlike a standardized test, this longer engagement is an interesting proxy measure for interest. More study is required to understand the implications of the design on the audience engagement.

In the results below, I show that many participants of the 27 participants in9 groups held direct and emergent schemas both before and after. A few shifted their schemas. Then I discuss the process two users had when they changed their schema through interaction with each other and *Ant Adaptation*.

I found that among the participants at the museum who played *Ant Adaptation*, before the game, five participants held an emergent schema and seven held a direct schema. After the game, 11 people held an emergent schema. Unfortunately, not all participants had a clearly defined trace apparent in their verbalizations. In other words,

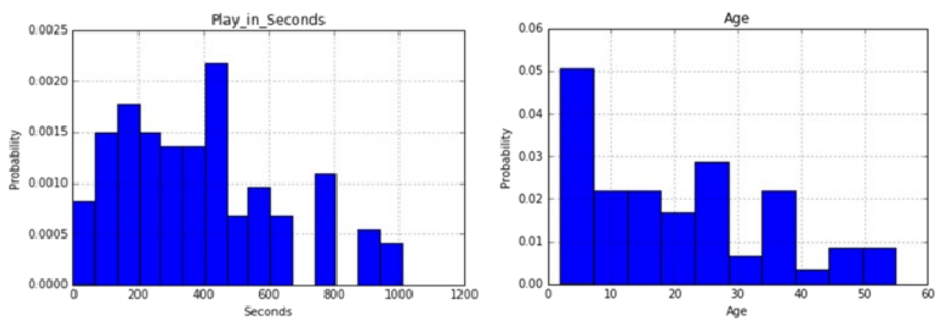

Fig. 7. Histograms of play time show that most players engaged for 400 seconds. The average age was 20 with a sizable number of players under 10 and some as old as 55.

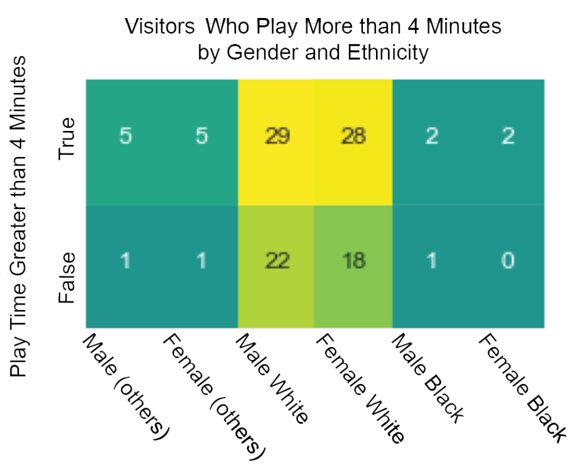

Fig. 8. Heat map of visitors' interaction time above and below twice the museum average play time (4 minutes) by ethnicity and gender. The majority of players, in the middle,were white. Notably, of the 17 non-white players, 14 played

for twice the average engagement time in museum interactives. Visitors were gender balanced.

not all participants verbalized a codable utterance both before and after play. As such, only seven of the participants had a discernable learning trace on this item. In other words, seven participants made a statement that we could code as a schema both before and after play.

As shown in Fig. 9, I found that three people switched their schemas after playing *Ant Adaptation*, and none switched from an emergent schema to a direct schema. To be counted in Fig. 9, users had to have a learning trace, with a demonstrated schema both before and after play. This limited the sample due to language ambiguities or users not responding to the protocol completely.

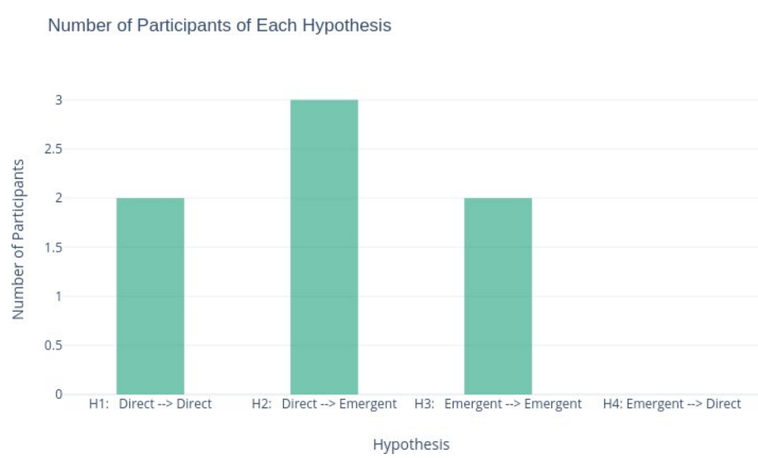

Fig. 9. Four participants held their schema constant. Three participants switched schema from direct to emergent.

4.1 H1: Users had a direct schema before and after

As shown in Fig. 9, two groups with children under 18 years of age held direct schemas both before and after playing *Ant Adaptation*. Both Stacy, a seven-year-old girl in Group 20, and Pri, a sixteen-year-old girl in Group 27 held their original direct schema. When asked what she knew about ants, Stacy said, “They move around a lot. They follow paths. They are always in the cracks of cement.” When asked how ants collect food,she said, “They collect it. They go and collect it." When I asked how ants deal with traffic she said, "Set paths." For Stacy, paths seem to tell the ants what to do. She didnot talk about how ants' actions construct paths. Thus, she sees it as a one directional effect on ants. Afterward, her direct schema was reinforced by the game. When asked about how ants know what to do, she said, "Pink stuff kind of leads them where they want to be. But if you put too much in one space it becomes a problem, because they really don't go anywhere else." Here she describes the paths as organizing, and at times determining ants' final location. She had no awareness of ants' own actions mutually constructing paths and directions. Pri likewise held onto her direct schema. Beforehand, when asked how ants know what to do, she said, "I think [they] navigate earth's magnetic field," indicating that a pervasive global force directs the ants. Afterward, she argued that they follow paths to "where the ants are supposed to go," indicating a controlling nature of pheromone trails.

4.2 H_2: Users had a direct schema before but changed to an emergent schema after

As shown in Fig. 9, three participants in two groups changed their schema from directto emergent. For example Ed, a thirteen-year-old white child in Group 10, while playing with his younger brother Thomas, changed from a direct schema to an emergent schema after playing.[10] Before the game, he had some idea about ants. When asked what he knew about ants he said, "They can carry 50 times their body weight", indicatingthat he has heard

[10] An additional group also changed from direct to emergent schema, but their interaction was half in Chinese, and so while I translated it, due to the nuances of language I feel presenting it here would go beyond the current study. I will present it in future work, as it may contribute a cross cultural dimension to emergent schema.

about ants in some cursory way. When asked how they collect food he said, "They, they scavenge for they--they're scavengers," which indicates he knows something about different animal feeding habits. His brother, Thomas, age nine, then interjected, "Like, they can like go hunt for food. They can like, um, try like, get to some, like, maybe some food on the ground like in the city or like in a park, or they can just eat a leaf." The interviewer immediately asked them how the ants know what to do. They replied with a direct schema:

Ed:	*Um, animal instincts.*
Interviewer:	*Animal instincts.*
Ed:	*Or the queen tells them to, if there's a queen ant. I don't know.*
Interviewer:	*So what's the diff--how does the queen tell them to?*
Ed:	*It's something with their antennas [sic].*

Ed thinks either a global force of animal instinct or the queen provides ants directions. His younger brother Thomas seems to agree with him but does not verbalize completely. After playing *Ant Adaptation*, Ed shifts his schema. And Thomas became clearer about how he was thinking about how ants are directed.

After playing, Thomas held an emergent schema. When the interviewer asked the group how to play the game, Thomas argues that feedback plays a big role:

Thomas:	*If you just do one path that leads to the sunflowers ants will just get the energy and just keep going backand forth and back and forth and that's how we got 21 [ants].*
Interviewer:	*And how did they get--how do they go back and forth and back and forth? What are they doing?*
Thomas:	*Like they bring the energy back, then keep going. It'slike a cycle.*
Interviewer:	*So how do they know to go back to that same place instead of somewhere else?*
Thomas:	*Because they just keep following the chemicals.*

Here Thomas articulates (1) if there is a path, ants will follow the path, (2) antsbring back energy that creates a cycle of fetching food and following a path, and (3)that this cycle leads to their highest population of 21 ants. Taken together, these three points indicate that Thomas understands the role of ants in following paths cyclically leads to higher population. This is a connection between the micro-level agents' actions and the macro-level pattern of population change. His older brother Ed agrees with Thomas. After Thomas' explanation, when the interviewer asked Ed what the pink chemical does, he said, "It attracts the ants to a certain location." And Thomas affirms him, "Yeah, it does that." Because this coding is crucial to our analysis there is one caveat to it. The interviewer did not ask them, as the protocol dictates, "How do ants know what to do?" From watching the video, and his self-report, it seems he was about to ask but hesitated, because, it seemed to him in the moment that he would get the same answer of the question. The participants had already established that they think ants follow chemicals. So instead, he asked Ed, "What does the pink chemical do?" to check if Ed agreed with Thomas. Therefore, I argue that in this context, "How do you think the ants know what to do?" and "What does the pink chemical do?" are the same question in different formats. Thus, I think the participants hold an emergent schema. In other words, these two visitors do not think ants follow the queen but the chemical, which agents lay down to form cycles of food collection.

4.3 H_3: Users had an emergent schema before and after

As shown in Fig. 9, two users, a 37-year-old mother and a 19-year-old participant in a second group, both accompanying younger participants, held an emergent schema both before and after the intervention. These were beyond the scope of this chapter, which is examining school age children's change.

4.4 H_4: Users changed from emergent schemas to direct schemas

As mentioned, there was one more possibility: people changed to a direct schema from an emergent schema, but based on previous theory (Chi *et al.,* 2012), this seemed highly unlikely. Indeed, as shown in Fig. 9, I found this never happened in the data.

It should be kept in mind, that of the 27 people in the 9 groups who played thegame, our coding only found seven people with a learning trace, who we could clearly code both before and after play. In future, I want to improve our coding schema,and protocol to understand the processes and frequency more deeply by which people change their mind about process schema. After completing this coding for frequency,our question became how did Ed shift his schema? What process of change duringthe game affected him? Below I present and analyze the experience in more detail.

4.5 Ed Shifts His Schema

Ed, Thomas, Sam, and Sally ages 13, 10, 7 and 6 played the game together. The group was all white. The experience started when an interviewer asked if they wanted to playa game about ants. When they agreed, along with their parents, interviewers told them aboutthe study. Then after consenting, standing by the multi-touch display, Thomas watching the display, and Ed standing off to one side to represent the group, the interviewer asked them a series of questions from a semi-structured protocol. When they were asked,"Have you ever noticed anything about ants," the oldest, Ed, showed some prior knowledge. When I checked how much they knew about ants self-organizing behavior they showed some conflicting idea of the process:

Interviewer: *How do ants collect food?*

Ed: *They, they scavenge for they--they're scavengers.*

Interviewer: *What do you mean?*

Thomas: *Like, they can like go hunt for food. They can like, um, try like, get to some, like, maybe some food on the ground like in the city or like in a park, or they can just eat a leaf.*

To clarify how ants, know what to do, following the protocol, the interviewer next clarified how ants know to scavenge. With this question I was identifying whether thevisitors had a direct schema, or an emergent schema. They offered both:

Interviewer: *Okay. That's fair. And then, how do ants know what to do? So you said they pick up leaves, or they scav enge, but how do they know to do that?*

Ed: *Um, animal instincts.*

Interviewer: *Animal instincts.*

Ed: *Or the queen tells them to, if there's a queen ant. Idon't know.*

Interviewer: *So what's the diff--how does the queen tell them to?*

Ed: *It's something with their antennas [sic].*

Interviewer: *Okay.*

Ed: *I think.*

Here Ed offers the idea that queen ants control all the other ants somehow through antennae, or the ants individually have an "animal instinct" that organizes behavior globally. The first idea is a direct schema. The latter possibility could either be an emergent schema, in which each ant follows simple rules that emerge in a pattern; but it seems more like a direct schema, wherein a global force animal instinct is giving orders.

To clarify this dual understanding, next in the protocol I asked how ants would deal with traffic.

Interviewer: *Fair. Okay, so let's imagine we're in a colony and there's all those ants. How do you think they deal with something like traffic? Are they all bumping into each other? How do they fix that?*

Thomas: *They make paths.*

Interviewer: *They what?*

Thomas: *Make path--more paths.*

Thomas indicates that the ants organize by constructing paths to get around the emergent issue of traffic. While not very elaborate, this seems to indicate he may already have an emergent schema in that he organizes his thinking through self-directed activity. Alternatively, I could interpret this to mean that ants are ordered, by the queen or animal instinct, to accomplish the goal of constructing more paths.

Thus far, Ed is of two minds, either ants are organized by their internal instincts, or the queen directly tells them what to do. Through the intervention, on the one hand, Ed shifts his understanding, which seems to annoy him. On the other hand, Thomas builds on his understanding of self-directed behavior, which he shares with his family.

In summary, before the game, when he was asked how ants know what to do, Ed saidhe thought it was animal instinct or that the queen tells them what to do. He also suggested that there is something in their antennae that makes them do what the queen tells them to. To us, this is the direct schema because he attributes ant behaviors to a higher power, whether instinct or the queen ant.

4.6 The Instructions to Play

The interviewer then explained the game. First, the interviewer organized the four players into two teams, then he showed them how to set the size and aggressiveness of their ants by moving sliders all the way to the right. Then he instructed the players in using theselector to choose to add chemicals or flowers to the game by touching the screen, and hereminded them to take turns. He demonstrated ants following his finger as the chemical is placed, and showing the flowers appear as he drags his finger across the screen. Then,he said that with the chemical “you can kind of control them." Then he showed how increasing each slider increases the "Create Cost," or how much food the ants would need to collect for the colony to reproduce another ant in the colony. Afterward, players chose how big and how aggressive they wanted their ants by talking to each other and setting their sliders.

4.7 The Gameplay

Ed says, "I don't want to be too aggressive" as the two teams stare at the touch screen. Then Sam touches the sliders on his side to increase the size of their ants, and the little sister increases their ants' size. Ed and Thomas adjust their controls and say, "We willbe [size] seven. No, six." Then they ask each other "Ready?" Sam controls the screen pressing Restart and then Play.

Ed asks, "How do I create?" tapping on the "Create Cost" widget on his side of the display then placing his finger on the screen he says "Boop," as if to imitate the interfacefeedback he is expecting when playing a game and pressing a button. He expects to be able to intervene in the game and add an ant directly. Thomas then says, "Make more flowers." It is ambiguous whether Thomas is responding to Ed, because as he says it, he is placing flowers on the table. Meanwhile, Sam is pressing on his Create Cost widget mimicking Ed's action moments before.

They proceed tapping the screen. Then Ed asks, "How do I do this?" Thomas shows him how to add flowers. And Ed imitates him, adds his own flowers, and says, "Here have some flowers," as if he is talking to the ants.

They then lean back a bit and view the ants as they collect the food, they have put nextto the colony. Ed says, "They are gaining food, I think." Thomas adds some more flowers close to the colony and says, "We have 15 ants." Then he turns a bit toward Thomas and says, "Just keep doing what you are doing.” When a flying ant emerges, Thomas asks, “Is that a queen?” Then pointing at the population counter on the heads-up display of their ants, he shows he understands the population level. “We have 21 black ants.” Following up he says, “Hey look at our aggression,” while he points at the heads-up display widget controlling aggression. Then Thomas and Ed stop playing. Ed says, “Ahh we’re doing pretty good, I’m just gonna let it do its thing,” indicating he understands the ants are organizing their own behavior.

Then Ed suggests an intervention. “Oh wait, we need chemical. Excuse me,” he says and pushes Thomas’ hand out of the way while he was adding flowers right next to the ant colony, and begins to just place chemical everywhere first, but then says, “We need to draw,” and starts trying to draw the ants to the flowers. “We need to get all these flowers. All of these.” Thomas then inquires, “We have 16 ants. What is going on?” His head leans forward as he examines the contest. Then Ed starts to take control of the ants, forcefully drawing a pheromone trail he says, “Go this way.” It sounds like he is trying to order them. But instead he has added the chemical in such a way that the ants are going in a circle, getting stuck. When they don’t do as he wants he says, “Vinegar will kill everything. Kill everything.” Thomas switches to the vinegar and begins to erase the ants’ pheromone trails. He methodically removes the pheromone from between the flowers and then from around the colony. Ed then says, “OOOOOOOOooohhhh I see what you are doing. Smart.” And then they start drawing the ants toward them by laying down new pheromone trails.

Thomas shows he knows that misplaced pheromones lead the ants astray. When Thomas adds new pheromones, he always adds them starting from the colony so that ants will walk towards the most intense part of the trail. When Ed adds pheromones to the screen not originating from the colony, Thomas swats his hand away, demonstrating he has a theory of the optimal placement and direction of pheromone trails.

Thomas teaches Ed that the direction of pheromones leading successfully toward food increases the ants' population. As he draws the pheromone trail to the flowers closest to the nest he says, "See they are adding more energy," indicating that the ants gathering food increases the colony's energy. Ed says he will just watch, arguing, "You know what you are doing, so I will just watch," and leans back a bit setting his handson the edge of the table the screen is sitting on. Thomas then crosses his arms and says, "All right let's just see it go on." Meanwhile, Sam has also learned through watching these two to use pheromone trails to direct his ants, drawing a long pheromone trailfrom Thomas and Ed's colony to his own, potentially confusing Thomas and Ed's ants. Thomas realizes this and says, "Hey, hey, hey!" demonstrating annoyance with the first aggressive move of the game. Through scaffolding, Thomas taught both his older and younger brother how to draw pheromone trails so that the ants interact with them, and how to use them to have the ants initiate a feedback loop of grabbing food and returningit to increase the population. When Sam tries to add another trail, Thomas hits his hand. "And that's when you ruined our plans," Sam responds. When they stop messing with it Thomas bemoans the falling population.

Thomas: *We have four. We have three--what? We have five, yo. That's even worse than what we had. We used tohave 20 [ants].*

Ed: *It's like--we had like 21 [ants].*

Ed corrects Thomas, indicating they both think their population has fallen. Then to recover, they see their ants are fighting the red ants. Thomas says, "Vinegar. They are notgetting enough [food]". Ed echoes the idea and says, "Clear out everything." He swipes his finger over the fight to try to clear the fight. Then even Sam reaches over and tries tohelp clear the fight out.

All the sudden while they are trying to break up the fight, Thomas and Ed's colonydies. "Wait what?" Thomas says. "Oh wait, where--where--where'd our thing go?" Ed responds in reference to their colony. Sam celebrates, "We're won it--we're--we won!" Sally joins in, "Yay!" Ed seeks confirmation, "Wait, we died?" Sally says, "It was smartto get some food," but Sam is a little less gracious about Thomas and Ed's play. "Cheaters," he says under his breath. Ed responds, "Cheaters? How do--I don't--I never knew how to play the game in the first place," referencing again that he did not feel completelyin control of the scenario.

Unprompted, Thomas offers why he liked the game. "Yeah, you had to figure it out and the--you have to have some flowers, see, and then you put the chemicals and lead itto there, then they'll bring it back, and like, if you want to get rid of the chemicals you use the vinegar." He says this as if he is teaching his siblings, telling them what they should think of the situation. Additionally, it shows he now understands that drawing trails to flowers starts a process where the ants will bring food back to the nest. He then mentions that if the process goes wrong, you can use vinegar to break them out. Hethen goes on and links this process to a repeating cycle that leads to a population level change. "So, um, you put some sunflowers down, then you get the chemicals and lead it to the sunflowers and if--if there's too much then the ants aren't getting the sunflowers and you--then they'll just like, then you use the vinegar and erase it. But if--if you just do one path that leads to the sunflowers, it'll just get the energy and just keep goingback and forth and back and forth and that's how we got 21 [ants]."

Thomas's breakthrough was reflected in Ed's change of how he thought. He moved from his equivocal answer that the queen controls the ants to a more nuanced view.

When Ed was asked whether this game shows an ant colony or a computer program, he answers, "ants," reasoning that because the ants would not listen to him, they mustbe ants and not computer programs. His answer implies an assumption that programs should be controllable. When asked whether he thinks this game is scientific, he answered, “Definitely,” because ants in the game would not follow his direction.

Though the interviewer did not ask the question, “How do ants know what to do?” from his answer, it appears that Ed started changing the schema from direct schema to the emergent one. He used to think that the queen might direct ant behaviors based on his answer to the pre-gameplay interview. And in the game the players attempted toplay the agent role of the queen ant directing others. I infer this may be because as a gamer, he has assumptions he is in the game as a leading agent with the higher hierarchy than other agents in the video game. Instead, it appeared to him that no other agents followed his orders as the queen. In other words, he learned ants were self-organizing, and that annoyed him. He had an expectation of the video game that as the player hewas in charge and through the intervention came to understand ants as autonomous, non-goal-oriented agents in the system.

5 Discussion

The example of Ed shifting his schema could be alternatively interpreted. He could have ceded the idea that he knew how the game worked. In that analysis, he did not shift his schema; instead, Thomas stated an emergent schema and Ed simply parroted his younger brother to avoid the embarrassment he seemed to be manifesting in his inability to understand the game. Regardless, the presence in our data of people who employ an emergentschema both before and after, and the potential that some of them changed their mind through play indicates interventions like this one may be able to effect rapid change on people’s emergent schemas. Given that experts have employed emergent schemas to understand crucial situations like climate change, global oil markets, and nuclear proliferation, I find the potential to increase people's use of emergent schemas through this restructuring, or restructuration (Wilensky and Papert, 2010), of informal learning important. So finding emergent schemas in a single data collection, and some evidence that a short museum intervention can impact participants' thinking is a hopeful development in complexity education. I can design practical, small-scale interventions that move the needle on the dissemination of this difficult and important mental schema to understand multi-party, multi-level interactions, such as currencies' value.

One limitation of this work is that I have no data on whether these ways of thinking about a complicated situation transferred beyond this scope or impacted longer-term thinking. Another limit is that I could not track the parts that affected Thomas and Ed's mental construct development. This second limit I address in the next chapter.

One design feature I found while testing NetLogo Touch was that the close timing of players' actions with changes, such as adding pheromones to the model, change how ants move, and may have accelerated learning about pheromone trails. In digital game design, Gregory (2018) argued that when actions and results happen within one frame, or approximately 10 milliseconds, connections between cause and effect happen better. This seems important when designing future complexity learning games. When effecting change, people need to map between the actions they take and the results. For instance, here I present them with tangible controls, to control the model, and see how the agents respond to those actions. This is not unique to *Ant Adaptation*. Every agent-based model uses a control interface to allow the user to directly control the conditions of the simulation and through observation form an understanding of (1) How the agents in the model react, and (2) The patterns or outcomes of the model, i.e., the emergent phenomena of interest. Usinga mouse and keyboard to change sliders, or in this case a touch-screen to affect this controlled observation, allows the user to come to understand an emergent pattern, a process, through repeated observation. In other words, though touch-based manipulation is different in the userinterface, it is not fundamentally different in the fact that the user learns by employing a user interface to control the agents and observe their simple rules. The difference here isthe effect of touching the model, and how the agents react to user touches is much faster,usually within milliseconds. It seems this closer connection between cause and effect may hasten users' learning with the model. Further study will be needed to explore this, but it seems an exciting avenue for future learning game design research. But if it bears out, I will likely make the sliders directly affect extant ants on screen's size and aggressiveness,instead of only changing the attributes of the ants born afterward.

The approach developed here can allow us to identify and describe learning by examining how players reconstruct provisional theories considering dialogue between theory and evidence (Wilensky and Reisman, 2006). Through this approach to teaching natural history in short interactions, designers can bring theory-building to players, who can, like Newton, Einstein, or Darwin, organize abundant data as part of the theory they are building (diSessa & Cobb, 2004). The decision built into the main action of *Ant Adaptation*–whether to peacefully collect food to increase population by employing feedback cycles,or go to war to eliminate their opponent–sets up a crucial engagement where the uncertainties make the testing immediate and productive (D'Mello & Graesser, 2012a).

6 Conclusion

One young man seems to have shift his schema from a direct to an emergent schema through interacting with his younger brother around a game in a museum. We built the interaction particularly for museums.

Our design had its intended draw to users. The design of the experience drew users in, with people playing for up to a quarter of an hour with average playtimes over twice the normal interaction times with exhibits in the natural history museum. Considering the engagement with this design, I conclude it reasonable to extend the design principles of *Ant Adaptation* and create complex systems arcades for natural science learning in informal settings, which I have now done with a 3D version built in Unity. The game in the museum had the following four perceivable effects:

(1) Construction of their own colonies, in competition with an opponent, afforded comparisons, which allowed for dynamic theory validation and imitation. For example, Thomas placing ants close to the nest and drawing ants to them by laying chemical trails showed he understood proximity and connection to flowers to the nest aided population growth. This is an example of learning to make micro-to macro-level connections through an agent-based model.
(2) The other team copied his strategy. Sharing strategies allowed players to update their operating theory.
(3) Taken together, these scaffolds facilitate players' exploration and learning about the complex system.
(4) Within less than a quarter of an hour of play, the game facilitated one player to switch from a direct schema to an emergent schema during a conversation with his brother.

People have argued that learning about complex systems is hard (Chi et al., 2012), and for an individual it is. When people engage part of a complex system, attempt their best theories in real time, and receive dynamic feedback from the computer and each other, we can design better ways to facilitate complexity learning. These learning moments may happen most when they notice breaks where to get out of their confusion, learners must engage in effortful – intense, purposeful, psychological effort – and problem-solving activities (D'Mello & Graesser, 2012a). That effortful solving activity is the process of science, and that is the process players in *Ant Adaptation* took.

While Ed vocalized most of the learning, Thomas was showing him how, and their younger siblings tested their approaches and copied them. During this research, I noted that the social aspect of the multi-touch interactive allowed discussion to guide the theory exploration. As a result, in future work, I hope to better share the joint sense making mediated through technology and each other.

As the contribution of this work, I claim that an interactive game – built around a complex, multi-agent model of ant behavior placed in a public space in a large museum – not only attracted many interested players, but shifted at least one of these players' schemas from a direct schema to an emergent schema about how ants know what to do. In other words, a computational environment, thought to be extremely difficult to understand, can elicit learning about complexity in a few players in an informal setting that lasts only for short periods of time, in less than 15 minutes. This learning happens through group discussion around a game on a touch screen through self-directed learning.

The theory-building exercise at the heart of the game was engaged with in the process of this motivated play. The game afforded that type of play to happen. This work, in short, expedites the activities found in earlier agent-based modeling and theory-building exercises that have been used in schools and with swarm intelligence researchers. The design allowed us to work in a novel context—namely, the fast interaction times typical in informal learning environments.

7 Acknowledgements

I would like to thank Robert Grider who wrote NetLogo Multitouch for this project. Also, Corrie Moreau and the Field Museum's Ant Lab that gave amazing feedback and support throughout the design of the exhibit as well as their welcome to and assistanceat the Field Museum. I also would like to thank the Center for Connected Learning, TIDAL lab, and TIILT lab for their feedback throughout the project, and the supportive Northwestern Community. Additionally, Marcelo Worsley and Emily Wang provided copious insight. Thank you to Rui Han for her endless support in discussing the findings, coding, and support in the process. Finally, we would like to thank the IEF for their generous support of this work, and the Multidisciplinary Program in Education Sciences for funding the project (IES: Award # R305B090009).

Chapter 4: Constructivist Dialogue Mapping Analysis of Ant Adaptation[11]

1 Introduction

To study open learning activities in a museum, I created a method to encode group learning, through participant dialogue. The approach builds on Piaget's method of clinical interviews, which studied how people talk, and how their speech references the mental models they are using (Piaget, 1926, 1929, 1952b). I extended that approachto pictorially represent how people understand complexity. The method creates hierarchical concept

[11] This chapter was previously published in Informatics in Education

maps, called constructivist dialogue maps, of what people say, focusing on the temporality of when they advance new ideas. The method highlights the possibility that concepts are not stable. The approach, constructivist dialogue mapping (CDM), was innovated to study open-ended, constructionist (Papert, 1986; Papert & Harel, 1991a) learning activity in a museum (K. Martin, Horn, et al., 2019). The approach was developed because of a weakness in evaluation methodologies of open-ended environments(Berland et al., 2014; Ochoa & Worsley, 2016). The main objective of this chapter is to argue that a formofdeep learning with complex systems models can be documented by using CDM. I developed this tool from constructivist and conceptual change theory. In this chapter, I first introduce constructivism and my definition of concepts. Then, I differentiate CDM from similar methods, including concept mapping more broadly, as well as semantic mapping, dialogue, and narrative mapping particularly, and I provide pictorial examples of those methods. From that point, I narrate how museum studies focus on elaboration to track learning (Leinhardt et al., 2003; Leinhardt & Crowley, 1998). In the methods and results sections, I show how CDM can be used to follow how users elaborated their ideas. Then, I discuss next steps in theprocess of automating the approach in the discussion section. Finally, I argue a CDM depicts theory development during learning in groups, with a computer, and with a facilitator. To begin the discussion, next I turn to the theory that underpins each of these visualization methods: constructivism.

1.1 Constructivism

Piaget was a Swiss psychologist who focused on the development of the child's understanding. He presented a notion of knowledge that is constructed. For example, he found that young children, under 3 months old, could not remember where an object was. He would hold his watch on a chain in front of the child and slowly drag the watch out of sight. When his children were young, they had no "object permanence," and so, when the watch came back, they were surprised, but as they aged, he found they could track objects. He concludes that humans "distinguish between these changes of position and changes of state and thus contrast at every moment the thing as it is with the thing as it appears to our sight; again, this dual distinction leads to the permanence characteristic of the object concept" (Piaget, 1952a, p. 7). In other words, as the child develops, she becomes able to construct a notion of an object. The list of *all* objects a child has constructed in this way is her ontology. As a result, knowledge is constructed, and a child builds her ontology through action in the world.

To elaborate, consider if the child grew up in a gas cloud in some remote universe. At first, the twirling nebular shapes would appear to have no order, and she would have no ontology for them. Therefore, when a particular eddy returned to the developing child's perception, she would not attach to it any particular permanence because the learner would not have *conserved* its continuous change into an identified concept. As the learner lived there longer, however, Piaget's theory predicts she would combine the gas cloud eddies into predictable movements based on experience to create concepts of gas cloud objects through the process of assimilation and accommodation.[12]

---

[12] One interesting experiment that should be done on this theory is about whether it is that children construct the idea of permanence, or whether it is that they construct the idea of time. That is, at first they experienceall time simultaneously and so the object for them is technically in both places, all places, but through experience come to see the object first as here and then there, or in other words in linear time. Basically, do children, only through experience, construct linear time? This question would have implications for physics, because some of the issues with the unified theory of the universe can be resolved by dropping thenotion of time from the equations. According to the Julian Barber: "If you try to get your hands-on time, it's always slipping through your fingers people are sure time is there, but they can't get hold of it. My feeling is that they can't get hold of it because it isn't there at all" (Franks, 2011).

This process would be conservation: seeing an object despite its changing features, or ever-changing atomic arrangement as a continuation of its prior existence.[13] In this way, all matter is variant, but the child can construct an invariant version out of this directionless, senseless mass, to form a theoretical frame. With this frame, the child can operate on their constructed ontology to make predictions and act. This is like a biker, who sees a group of people walking, and can bike past them very closely because she can assume they will continue to move the way she has observed them up to now. This prediction of the pedestrians' operation is prone to error, but as a working hypothesis for riding a bike, it is close enough. Concepts are fluid but have a stability to them. In other words, in childhood, we are thus able to see enough rhythm and repetition around us to create meaningful concepts.[14]

1.2 Concepts

While there are several perspectives on what a concept is, I take a physicalist theoryof mind: a concept is a mental representation the brain employs to mark a class ofthings, like a nebular swirl or flock of pedestrians, in our world. Concepts are mental representations that allow researchers to draw appropriate inferences about the type of entitieswe encounter in our everyday lives (Murphy, 2004). Murphy argues concepts make sense of our

[13] When learning from an agent-based model, this could be described as learning an agent's operators, or rules.

[14] One interesting possibility for constant construction and reconstruction of knowledge is the possibility that there is no memory. In other words, like the original Macintosh, there is no onboard memory; instead, knowledge is always reconstructed by assimilating, or imposing one's order on the world in the context at hand. One's order then would have to be maintained, but memory would be unnecessary. This possibility becomes particularly interesting in light of the claim made by Ackermann (1996) that due to the 3D reconfiguration task required for perspective taking, that some views are harder to reconstruct than others (p. 6).

lived experience: "Concepts are the glue that holds our mental world together. When we walk into a room, try a new restaurant, go to the supermarket to buy groceries, meet a doctor, or read a story, we must rely on our concepts of the world to help us understand what is happening. We seldom eat the same tomato twice, and we often encounter novel objects, people, and situations. Fortunately, even novel things are usually like things we already know, often exemplifying a category that we are familiar with" (Murphy, 2004, p. 1). Thus, concepts let us see the changing world as consistent and predictable.

Children construct conceptions. Adults do too. These conceptions have several uses, such as that they can be placed in categories, that aid learners in making inferences. For example, 4 to 7-year-old dinosaur aficionados can generate many appropriate inferences about an unfamiliar dinosaur after they categorize it on the basis of surface and taxonomic features (Chi & Koeske, 1983). This categorization allows people to theorize about missing information in a theory (Chi, 2008).

For instance, in Fig. 1 a scientific illustration of Hadrosaurus shows a head. However, there is no fossil record of Hadrosaurus heads. The artists must extrapolate, deduce, or imagine based on other hadrosaurids. The concept artists construct through experience allows them to infer or deduce the missing information.

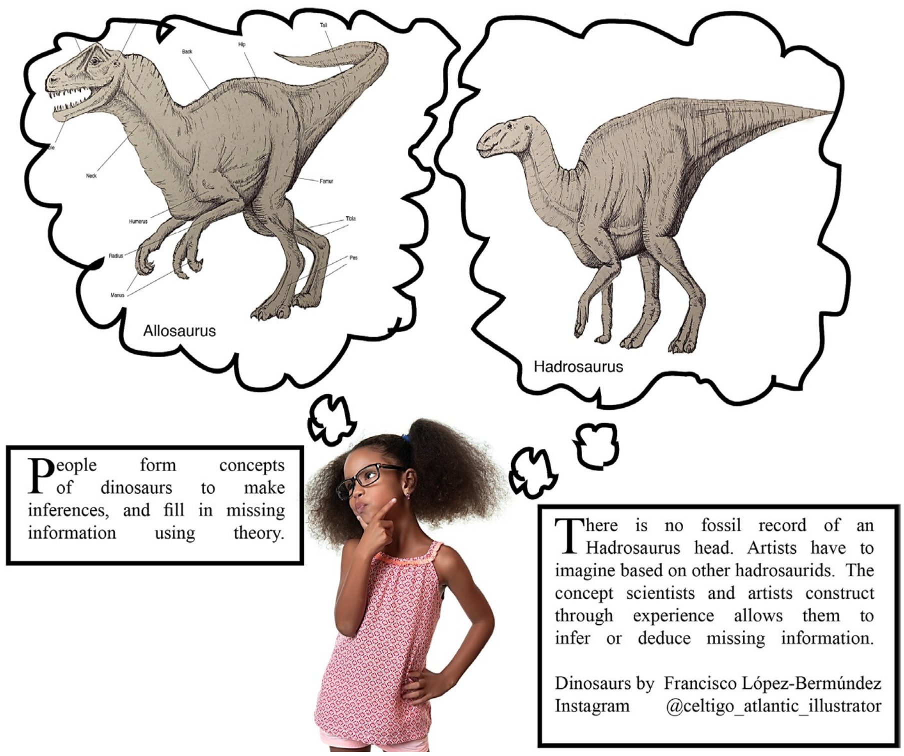

Fig. 1. People form concepts about organisms. They learn the parts of the organism, its functions, and its properties, such as color or size. Based on what they know about similar organisms they infer other unknown parts. (Right) Hadrosaurus interpretation. (Left) Allosaurus with taxonomic identification Hadrosaurus and Allosaurus drawn by Francisco López-Bermúdez.

Categorization is a crucial use of concepts in a physicalist view. Humans identify and assign concepts to the category to which it belongs through categorization (Chi, 2008). Once a concept is categorized it can inherit from where we assign the knowledge. For instance, if we know that *ants* are *insects* and *insects lay eggs,* we can infer that *ant lay eggs* even if we were never told that fact. When a student knows an ant is a kind of insect, she can infer that *ants* inherit the properties of *insects,* such as being six-legged. Categorization of concepts through assignment is powerful because a student can use knowledge of the category to make many inferences and attributions about a novel concept/phenomenon (Medin & Rips, 2005).

From Aristotle to the 1970s, philosophers argued that concepts have a definitional structure. Concepts defined this way have a list of features, that are both *necessary* and *sufficient* to determine membership into a class, i.e., a bachelor is an *unmarried man.* There are several challenges to this view. For the current project, the biggest challenges are that categories can be "fuzzy" and that there is no psychological evidence for humans using concepts as strict definitions (see Margolis & Laurence, 1999 for a complete accounting of the challenges).

Despite this challenge, however, categories remain very useful heuristics for interaction in the world, letting us fill in the missing head of hadrosaurids, and the pedestrian movement for bike riding, among many other daily tasks.

1.2.1 Piagetian Clinical Interviews

How are concepts formed? In the constructivist view, concepts are conserved ontological entities a person builds through action in the world (Piaget, 1952). So, how do we identify concepts that people hold? Piaget invented a method, the clinical interview—an approach to documenting an open-ended conversation designed to illuminate the way a child thinks or explains a particular phenomenon. Even though Piaget widely employed the clinical interview to examine how children construct their knowledge, there is surprisingly little discussion of the method in his work (Posner & Gertzog, 1982). Piaget elaborated most on his data collection method in the introduction to *The Child's Conception of the World* (1929), and also in the preface to *The Language and Thought ofthe Child* (1926). His method of analysis of the development of cognitive constructions involved observing children as they reason about unusual phenomena that he presented in designed settings. The method involved the following five steps:

(1) Design an activity.
(2) Let the child talk.
(3) Notice the way the thoughts unfold.
(4) Probe them with questions.
(5) Do not just notice the answers the child gives to questions posed, but also followthe child's line of thought.

Piaget argued that "If we follow up each of the child's answers, and then, allowing him to take the lead, induce him to talk more and more freely, we shall gradually establish for every department of intelligence a method of *clinical analysis* analogous to that which has been adopted by psychiatrists as a means of diagnosis." (*emphasis added,* Piaget, p. 276 as in Claparéde's preface to Piaget, 1926). This approach is a useful way to follow children's understanding of the world; it focuses on the knowledge, or mental models, children construct during an activity. This process is the process of theory development.

1.3 Theory Development

Theory development is the idea that students' mental models are built out of theories of how the world works, the sum of lessons learned from thinking that builds knowledge. DiSessa and Cobb (2004) argue that from Newton to Darwin and Einstein, theories embody generalizations to organize overly abundant data that is subsequently viewed aspart of a new theory. In this way, diSessa and Cobb (2004) posit theory as a lens, "teaching us how to see" (p. 4). How we see the world is the crucial part of these theories since the lenses constructed through experience take actual form. Just as "[T]he worldis not just sitting out there waiting to be to be uncovered, but gets progressively shaped and transformed through the child's, or the scientist's, personal experience" (Ackermann, 2001), constructivist thought highlights transformation and molding as the work of mental models. These models can form into more stable theories. This transformation happens through conceptual change which I will introduce below. Next I will introduce a fragmentary view of knowledge: knowledge in pieces.

1.3.1 Knowledge in Pieces

Knowledge in pieces (diSessa, 1993, 2018) addresses the following questions: what are the elements of knowledge, how do they arise, what level and kind of systematicity exists, how does that system evolve, and what can be said about cognitive process that underly the system and its operation. The study of this framework looks to study moments of learning that are consistent with constructivism. The theory proposes many fine-grained bits of knowledge (p-prims). P-prims are microgeneralizations that people abstract from experience. They are small knowledge structures that get enacted by being recognized and cued to an active state based on the perceived configuration. These are mid-level cognitive elements, neither the low-level sensory information, nor the high-level named concepts or categories. P-prims are activated based on a cuing priority. DiSessa does not focus on the origins of p-prims, but instead focuses on their life histories, especially how they might "become embedded in more physics sophisticated thinking" (diSessa, 1993, p. 114).

1.4 Conceptual Change

There are at least three circumstances under which a person can learn (Chi, 2008). First, they have no prior knowledge, so they *add missing* knowledge. Second, the person may have *incomplete* but correct knowledge, so the learner is *filling the gap.* Third, the student may need to amend prior knowledge. In this third case, we refer to the process of knowledge acquisition as *conceptual change* (Chi, 2008). Students undergo conceptual change when they perceive a mismatch between previously constructed knowledge and novel experience (Chi, 2008; Posner & Gertzog, 1982). When the learner experiences contradiction, she can reconstruct knowledge structures to account for the novel experience. If the context of learning is aligned with conical scientific knowledge, the resulting structure will more closely resemble scientific knowledge (Chi, 2008; Posner & Gertzog, 1982.

There is a connection between observation, contextual knowledge construction, and stable theories. These states convert one to another through conceptual change. When a learner modifies their knowledge, the knowledge-in-pieces framework (KiP) indicates construction happens incrementally over time (sometimes years) (diSessa, 2018). The learner has subtle knowledge infrastructures that yield conceptual understanding through knowledge structure synthesis. These knowledge structures only exist in context, not as stable structures (Hammer, 1996). Instead context activates specific cognitive building blocks known as phenomenological primitives (p-prims) (diSessa, 1993). These structures can fire differently depending on context, so what may appearas a misconception, may result from a re-ordering of p-prims from different contexts. KiP indicates that instruction should activate p-prims in the context in which the instructors intend knowledge to be used. Even though knowledge structures are notstable and they operate at different scales (Duit et al., 2008), some emerge and reinforce to become large, broad, stable knowledge structures stored in memory, like a bikers understanding of pedestrian movement. We call these *theories* (Darner, 2019). Theories are cohesive mental models that explain causal relationships between several different contexts or phenomena. This chapter attempts to track thismessy, context dependent path through a type of concept mapping called constructivist dialogue mapping (CDM).

I developed CDM to study learning in an informal environment based on constructivist theory. In constructivist theory, a learner's mental model drives his or her construction of understanding and internal cognitive structures. This process includes accommodation and assimilation (Piaget, 1952), maintaining a balance between stability and change, continuity and diversity, and closure and openness (Ackermann, 2001) when exploring the world. For Piaget, children are not just incomplete adults (1952). Their ideas function very well for their current context and as a result, their mind changes through experience. Even though knowledge is contextually bound, so are actions and habits. As Ackermann (2001) said, children's conceptual changes are like those of scientists: they happen through "action-in-the-world" (p. 3) to accommodate for experiences, and most likely through a host of internal cognitive infrastructures. "Knowledgeis not merely a commodity to be transmitted, encoded, retained, and re-applied, but a personal experience to be constructed" (Ackermann, 2001, p. 7). Thus, with CDM I hope to track these changes as elaborations of ontologies, as they happen, during knowledge construction to facilitate the study of learning. An example of this sort of learning environment can be found in constructionist microworlds (Edwards, 1995; Papert, 1980), or constructionist video games (Holbert & Wilensky, 2019), where the learning is interwoven into the gameplay.

1.5 Elaboration as Learning in Museum Studies

In *Museum Learning as Conversational Elaboration: A Proposal to Capture, Code, and Analyze Talk in Museums*, Leinhardt and Crowley propose means to study how learning actually occurs in museums (1998). Their work in museum learning motivates their attempt to solve a core issue: lack of theoretical coherence in the museum learning research. They suggest three problems: first, a need for a learning definition; second, the "univariate" issue where researchers attach a particular indicator, like terms, to a single factor, such as age; third, the museum diversity problem, where research does not span the different types of museums; and finally, the division between quantitative and qualitative research methodologies. From these problems, the authors propose three outcomes, the last of which is important for constructivist dialogue mapping. They argue that their approach will provide a novel, stable, and disseminable methodology to conceptualize, collect, and analyze conversations as a process and as an outcome of learning in the museum context.

After introducing the problems and outlining them, Leinhardt and Crowley (1998) describe a pragmatic approach to study learning in museums: define learning as **Conversational Elaboration**. For their pragmatic approach and operational definition of learning they chose how visitors elaborate conversations. They focus on conversational elaboration because it is a naturally occurring part of the museum experience while also being a product of the experience.

By elaborations they mean a particular kind of talk that occurs within a group. Conversation is important because it reflects the "inter-twining of social with cultural processes" (White, 1995, p. 1). Sociocultural theories of voicing (Rogoff, 1990; J. V. Wertsch, 1997) (emphasize that inter-twining of voices (in the Bakhtinian sense) is the primary activity through which knowledge is constructed and appropriated across people. In this chapter I take this process of studying learning as conversational elaboration, and then use concept mapping to study how these elaborations occur. Next, I introduce the different types of concept mapping.

2 Concept Mapping

A concept map is a diagram that depicts relationships between concepts. It is a graphical tool that instructional designers, engineers, technical writers, and others use to organize and structure knowledge. In their overview of qualitative methodology, Miles, Huberman and Saldana (2014) describe concept mapping as useful for qualitative methods. Joseph Novak invented concept maps to assist learners (1983). This theory was based in the philosophical and epistemological origins of Conant (1947), Kuhn (Kuhn, 1962), and Toulmin (Toulmin, 1972). Concept mapping is a process that focuses on a construct of interest or a topic. The process generates input from one or more participants. As shown in Fig. 2, from it the participant or a researcher produces an interpretable pictorial view (concept map) of their ideas, concepts, and how these are interrelated.

These maps are usually organized hierarchically, and can have the links between them labelled, or not. They can also be organized in other shapes, such as a radiating pattern. There have been varied uses of concept mapping in education over the years. Novak and coauthors' method was used to bring concept mapping into students' hands to better understand science content. This work used concept mapping to improve student learning. Trochim (Trochim, 1989) used concept mapping to guide action in planning and evaluation. Novak (1990) and Kinchin

(2014) reconceptualized concept mapping as an educational tool through a wider literature on curriculum development. Jackson and Trochim used concept mapping for analysis of open-ended survey responses, an alternative method to existing code-based and word-based text analysis techniques (2002). There have been many types of concept mapping, including mind maps, many forms of note-taking, such as Cornell notes, the presentation software Prezi, and repertory grids. Following in this tradition, constructivist dialogue mapping is a type of concept

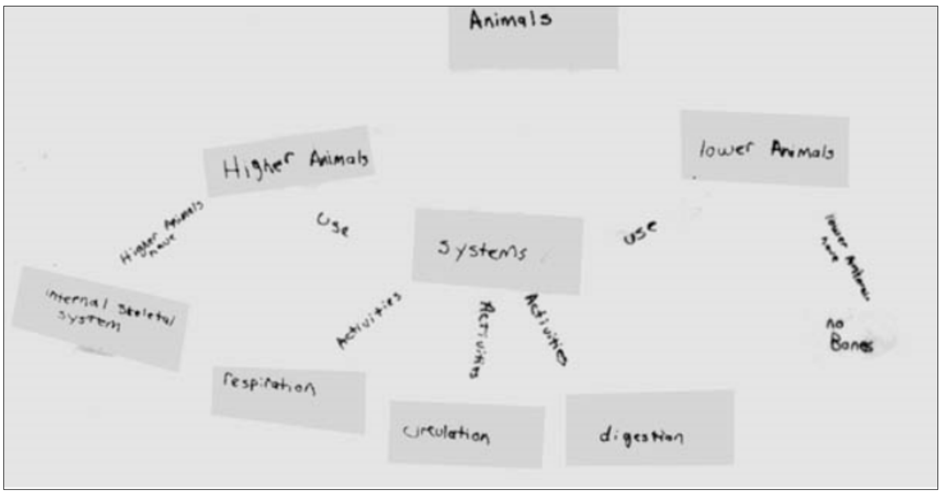

Fig. 2. From Novak, 1983, and example of concept mapping of the concept of animals, differentiating between animals with and without internal skeletal systems.

mapping meant to track the temporality of when people have ideas during an intervention or interview, and how those ideas elaborate over time. There are several other typesof concept mapping that are closely related, which I will introduce next, including semantic, dialogue and narrative maps.

2.1 Semantic Mapping

Semantic mapping is a method of the study regarding the conceptual knowledge, where concepts are represented as a semantic maps of nodes and their related properties in a network of nodes and links (Anderson, 1976; A. M. Collins & Loftus, 1975; A. M. Collins & Quillian, 1969; Linton et al., 1975). As shown in Fig. 3, Attributes of the network structure are assessed by the number of links between nodes, the strength of linkages, and the cohesiveness of the entire collection of concept nodes in semantic memory.

Chi, Hutchinson, and Robin used semantic mapping to represent children's conceptions of dinosaur knowledge (Chi et al., 1989). For Chi *et al.,* semantic mapping is the structure of the conceptual knowledge. It is the study of the properties of a representation in a static state. It concerns what a representation is composed of (nodes), how the nodes in the representation are clustered and related to each other, and the cohesiveness of the structure. However, it is not about representational comparison between an adult and a

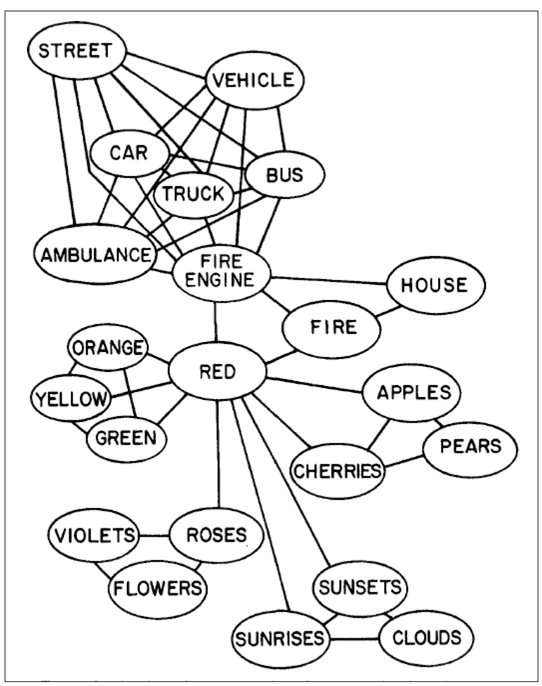

Fig. 3. Adapted from Collins and Loftus (1975), an example of a semantic map of stereotypical map of human memory.

child, who are believed to have different representations. Likewise, comparison between students of different ages may also be suspect. These hierarchies can be useful when learners employ them to perform categorical reasoning. Categorical reasoning can allow learners to infer that a particular kind of robin, can lay eggs, if they only know that birds lay eggs, because robins are a type of bird (Chi, 2008).

2.2 Dialogue and Narrative Mapping

Concept mapping is related to narrative mapping. Veterans in a social situation, like Alcoholics Anonymous, often use narrative maps to orient, inform and advise newcomers (Pollner & Stein, 1996). Pollner and Stein describe narrative maps as having several impacts, including:

(1) Affecting recruitment into a social sphere through improving the attractiveness of a situation.
(2) Contributing to socialization and social reproduction by transmitting values.
(3) Shaping action by changing probable activities in an area.

In this way, narrative mapping is part of a process where a social world is talked into being constituted. In other words, narrative mapping is both part of representing a social world, and part of the process of reproducing that social world. Narrative mapping is used to share cultural world resources from current members to new members through a systematic, pictorial representation.

Concept mapping is also related to dialogue mapping. Dialogue maps can provide a visual tool to promote shared understanding between novice (student) and expert

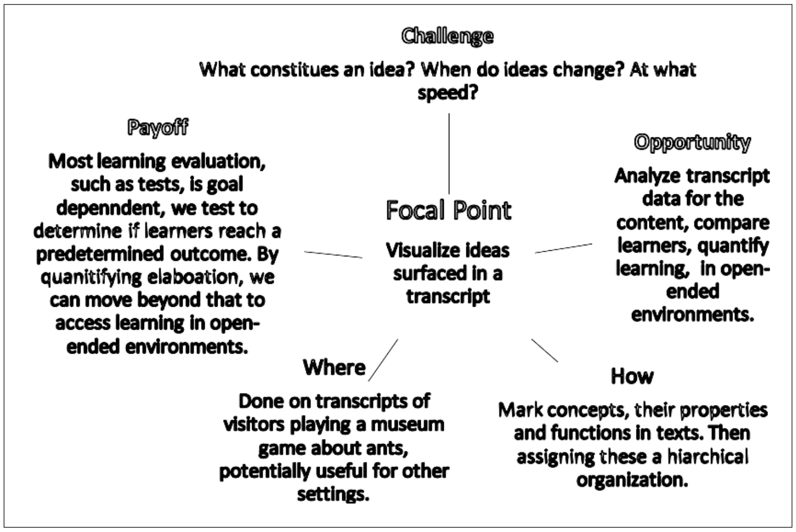

Fig. 4. Adapted from Pollner and Stein (1996, A Narrative Map about Constructivist Dialogue Mapping.

(teacher) to increase meaningful learning in biology education (Kinchin, 2003). Dialogue mapping can represent and scaffold students' argumentation (Okada & Buckingham Shum, 2008). The work to improve student argumentation with dialogue mapping has been assisted by *Compendium*, a dialogue mapping software (Okada, 2008).

2.3 Uniqueness and Differences of the CDM Coding Method

Constructivist dialogue mapping is a means of producing a pictorial representation of knowledge, but instead of trying to use the representation to improve student learning,it is used to track the instability of knowledge as students learn and associate those changes with learning interventions. I use it to represent players' knowledge states at different times as they develop. CDM is developed to show what learners believe exists (i.e., their ontologies) as they come to understand a complex system modeling game.This method is useful in representing learning in systems where participants elaborate their understandings in interaction with each other, teachers, and technology to form new theories to account for novel experiences. While this method is particularly useful in studying non-goal directed, constructionist learning – where it is difficult to write summative tests, because the learning is not goal directed – constructivist dialogue mapping can be useful for evaluating other learning activities, an idea I expand upon in chapter 6.

The dearth of evaluation methodology for constructionist, open-ended learning environments (Ochoa & Worsley, 2016) motivated my design of CDM. The lack of evaluation can be easily explained by the lack of a learning trace, whereas computer-aided learning systems, if designed for it or not, provide ample data of how the user interacts with the environment. Open-ended people-to-people interactions provide far less data. Because I want to know how participants structure the knowledge they learn in these environments, I developed CDM using constructivist theory—namely, following ideas as they appear, are contested, and fluidly change, based in a knowledge construction view.

3 Theory: Studying Learning in Informal Learning through Constructivist Dialogue Mapping

This theory led me to construct a methodological innovation that is useful for the interactions typical of constructionist learning environments. With CDM we track the construction of conserved concepts through the available proxy of change and elaboration of speech. I built on concept maps, methods as they emerged through transcript data focusing on the moment-to-moment manifestations of ideas. The method is differentiated from age-specific representations as semantic maps have been. I do not attempt to use the maps to improve student learning of the connections between material as much of theconcept mapping has. Instead, I track learners' moment-to-moment manifestations of conceptions as they construct them into less variable forms. At this level, these maps represent how a player's speech references the ontological entities he or she is forming or has formed. The maps capture these ideas of the understanding players demonstrate during play in a hierarchical map. I present concept maps of players elaborating their ideas about agents through observation and interaction that accounts for the changing nature of ideas through activity. Maps visually depict the ideas players share through what they say and how they interact with the game.

I argue that I can research what people conserve (that is, learn through accommodation demonstrated by what they say) by filling in a map with what people say and do during play. Constructivist dialogue mapping allows us to track how people's words and actions indicate learning through short play periods. The more researchers time stamp when new concepts arise and change through talk, the closer we can track the messy contestation that develops new ontological entities and modifies existing ones. However, these maps are only a proxy built from the externally observable action of speech. In this chapter, I use transcript data, but the observable data could include log files of their actions, or gestures that describe actions or entities. Importantly, we do not have direct connection to the internal cognitive infrastructure. In other words, the maps are a methodto pay close attention to a participant's words, a naturally occurring feature of museum play, while keeping an eye on what it says about her ontology.

The advantage of this approach is that I can present the smallest parts of what we observe. For instance, researchers can note when users first identify an unknown activity, such as "purple line" and then how they build out an understanding of the subprocesses of that task such as "attracts ants" or "fades away." As a result, I read the transcript word by word to see how users construct such operational knowledge on the representations theysee. A limit here is that I study what participants name, those higher-level cognitive structures, but things they might not retain. Theoretically, these are somewhere between p-prims and higher order named categories. CDM demonstrates learning as concept elaboration over time through the proxy of changes in speech. I will explain the implementation of CDM analysis in the methods section below. First, I introduce the research questions. Then, following Piaget's method, I explain the design of the unusual activity I use to understand how people come to understand complexity.

4 Research Questions

In this chapter, I use CDM to explore how users make sense of the self-organizationof cooperation between ant colonies in competition with each other while they use an agent-based model built in NetLogo (Wilensky, 1999b).

I researched how users build knowledge when using the model of ant colony life, Ant Adaption (Martin & Wilensky, 2019; Martin, Horn, & Wilensky, 2019). I was particularly interested in how visitors made sense of the behavior of individual ants in the simulationto and in turn reason about aggregate-level outcomes of groups of ants.

Specifically, I researched the following questions:

1. How can we capture visitors' moment-to-moment sense-making while they explore complex systems notions such as emergence?
2. Can we see evidence that new knowledge structures emerge through game play?
    a. How do these knowledge structures shift or remain stable over time emerging as theories explaining the context?

5 Design: Agent Based Modeling Game for a Museum

To show the environment I evaluated with CDM, I will describe the model/game and discuss the design decisions I took because of implementing in the museum.

5.1 The Game: Ant Adaptation, Agent-Based Modeling in Museums

As shown in chapter 3, Ant Adaptation (Martin & Wilensky, 2019) was used to study schema formation. In the following chapter, I studied the connection between affect and cognition. In this chapter, I use it to collect data and use CDM to analyze that data to demonstrate constructivist dialogue mapping. As I did in chapter 3, to provide context for the current analysis, I describe the game below.

Ant Adaptation is a fully functioning agent-based model. In the model, ants go out to collect food and return to the nest. As they return to the nest, ants lay down a pink pheromone that attracts others nearby. Other ants walk toward the strongest chemical smell, which in most cases is where the first ant just passed. When ants find a flower, their food source, they return, lay down a pheromone trail, and thus construct and reinforce pink trails. This creates an emergent feedback loopthat routes more and more ants to successful sites of forage. As the ants exhaust a food source, they must find new locations, and thus repeat a cycle. When two or more ants ofopposing colonies encounter each other, they fight or scare each other away, also leavingchemicals that attract more ants. For the winner, this works to protect the food source from competing colonies. The ant queen reproduces when the ants in her colony collect enough food. The player interacts with this complex system by adding pheromone trails that the ants follow, as well as adding sources of food to the system, thus changing the amount and distribution of food in the game. While interacting with the system, students form a functional understanding of the ants and their mechanisms of action (i.e., agents and their rules) in the model.

This design scaffolds experimentation. Players must simultaneously make choices. As shown in Fig. 5D, players can touch the screen to add pheromones the ants will follow. At the flick of a switch, they can add more flowers anywhere they like in the game. Lastly, they can choose to apply vinegar, which erases trails. Erasing trails was used by some game players (like Thomas discussed below) to get ants out of a feedback loop that was leading them nowhere (a local optima). For players to achieve their goals in the competitive environment, they need to understand the emergent consequences of simple ant behavior.

Players can decide how big and aggressive ants are. When the size of ants increases, they become slightly faster and stronger in a fight. Each increase in size/level adds up. At the highest levels, the ants are thirteen times stronger. When players make their ants more aggressive, it increases the radius in which ants detect opposing ants and thus the probability that they will attack. Increasing either the size or the aggressiveness also increases how much food is required to raise an ant, so the largest ant requires thirteen

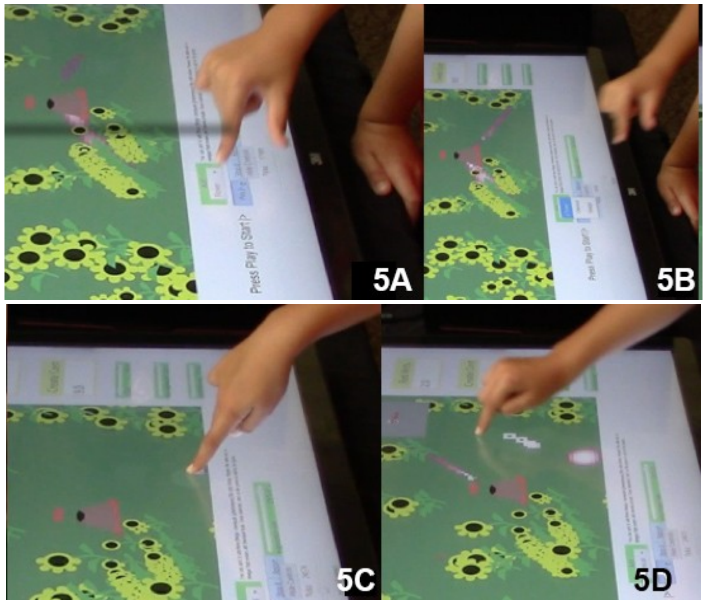

Fig. 5. John selects to add chemical (A) experiments with the touch user interface (B-C) and then lays down his first pheromone trail (D).

times as much food to feed to adulthood. This gamification impacts how much food antsmust collect to make a new baby ant. Increases in either of these parameters reducesthe expected population of the colony, though it increases their likelihood of fightingand winning through emergent interactions of parameters (size and aggressiveness) and agent actions (collecting food, leaving trails, and fighting).

This sets up the main action of the game as a series of strategic choices: to decide whether to passively collect food, thereby increasing the population, or, to go onthe warpath where big, aggressive ants conquer their opponents. Either method of play could lead to high populations or the elimination of the opponent through better control of food resources. After learning about the consequences of strategic choices through gameplay, players strategize by increasing ants' size, aggressiveness, or both. This mightlead them to win the game by annihilating the other group's ant colony. However, biggerand/or more aggressive ants consume more food to reproduce, and potentially reducethe colony’s population size. Thus, a player might strategize by adding more flowers andpheromone tracks around the colony to help the larger ants survive. This learning and strategy cycle interweaves the learning into the gameplay.

6 Method: Constructivist Dialogue Mapping

CDM presents learning as concept elaboration (Leinhardt & Crowley, 1998) through transcript analysis. I developed the method to record the agents, properties, and actions that users notice in learning environments. In this study, I conducted two treatments: the first with 6 participants in hour-long individual clinical interviews in a lab setting; and the second with 38 participants in approximately 4 to 20-minute group interviews in a natural history museum.

6.1 Treatment One

In the clinical interviews I conducted a semi-structured interview protocol. This was administered while people discussed their understanding of Ant Adaptation. In treatment one, there were 6 participants (50% White, 16.16% Black, 16.16% Asian, 16.16% Latinx) being 50% male and 50% female. I recorded audio and video of participants, transcribed the data from the audio, and analyzed both the media to construct CDMs.

6.2 Treatment Two

In the second treatment, participants were sampled over a six-day period, attracting114 museum visitors in 38 groups (87% White, 4% Black, 5% Asian, 3% Latinx). This contrasts with the museum-wide attendance demographics of 70% White (difference of +16.61% points), 5% Black, (difference of -0.54% points) and 14% Latinx (difference of -11.32% points). Of the players, 60 were male (51.57%) and 54 (48.43%) were female. Players ranged in age from 2 to 55, with the age distribution skewed to lower ages. The average length of time people played was 387 seconds, as opposed to a museum-wide average interaction time with digital interactives of 105 seconds (as reported by internal museum evaluations). In treatment two, interviewers conducted a pre-post survey, and video and audio recorded the play. From the audio I transcribed the data and analyzed both to construct CDMs.

6.3 Which Changes Are Conceptual Change

Leinhardt and Crowley (1998)'s approach particularly looks at four processes that groups will undertake. If there is learning, after an interaction with an exhibit, a group will:

1. Refer to more items.
2. Include greater detail about those items.
3. Synthesize elements to elements from their prior knowledge.
4. Increase the level of analysis of the phenomena that they discuss.

This approach moves away from focusing on the amount of talk or types of talk while building a strong foundation between amount, type, and the process of learning. Asa result of this review, Leinhardt and Crowley suggest investigating how conversation— as socially mediating activity—acts as a process and an outcome of museum learning experiences.

In the end, the methods of collection and analysis offered by Leinhardt and Crowley, (1998) are interesting, but they are not sufficient. With text capture, and direction capture I can measure the learning more fully than Leinhardt and Crowley.

6.4 Constructing Constructivist Dialogue Mapping

As shown in Fig. 6, CDM tracks the nouns that players mention during play with a system (such as ants); the adjectives that modify those nouns, such as "six-legged" (6b); andthe verbs players use to describe those nouns' actions, such as "follows trails" (6d). To gather the transcripts that I analyzed for CDM, I applied a mixed methods approachto the observations (Clampet-Lundquist et al., 2011).

Researchers coded the transcripts by reading them. When we read a noun, we added a row to the matrix. We then recorded the interview session, noting whether this interaction was during the pre-interview, gameplay, or post-interview, noting who spoke, whattime they spoke, the question being answered, the exact quote of the response, and the node label we coded the quote as. Then we recorded what parent box it modified and counted the node's depth in the hierarchy. Finally, I coded if the observation was an elaboration to an ontological entity or addition of a new one. For example, in Table 1: I ask Briana if she has ever noticed anything about ants. She responds, "They carry fifty times their weight?" As shown in Fig. 6b, I would code this as a second order elaboration hierarchy: 1st order-Ants and 2nd order-Carry 50X weight.

To answer the second research question – *How can we capture visitors' moment to moment sense making while exploring complex systems notions such as emergence* –I deployed this method to construct maps of ontological entities in the game (i.e., agents), their actions, and properties as described by players. The maps were constructed by building a hierarchical map of players' utterances using the coding method shownin Fig. 6 and Table 1, before, during, and after play. The coders proceeded through the transcript utterance by utterance. When an entity was named, such as "ant", the coder added a box. When other entities were named – such as the property “six-legged” or action “carries a lot of weight” – coders determined what that concept modified. If, as in Table 1, they modified “ant” unambiguously, then the coders added a subordinated box below ant. In coding during treatment two, we broke the transcript into a pre-interview, a game play portion, and a post-interview. We analyzed the change in how players thought about the actions ants take, and the functions of those actions. In Fig. 6, look at how the person describes ants as six-legged (Fig. 6a), creatures that carry 50x their weight, and follow trails to find food (Fig. 6d). When these entities, actions, and properties change over time is my measure of learning during the intervention.

In treatment two, I color coded changes in the ontological maps over time so we can visually compare learning. When I added a box between the first sample and the second, I colored it red. When I removed a box, either because the visitor has stopped mentioning the idea, or has explicitly contradicted themselves, I colored the box grey and strike through the text. When it is not clear if the item was there previously, or if it has gone away, I left them blue. I present two times scales of analysis in the results, showing both the micro and macro changes. In treatment one, I analyzed the concepts participants put forward as they emerged during the protocol and highlight the conflicting notions participants demonstrate.

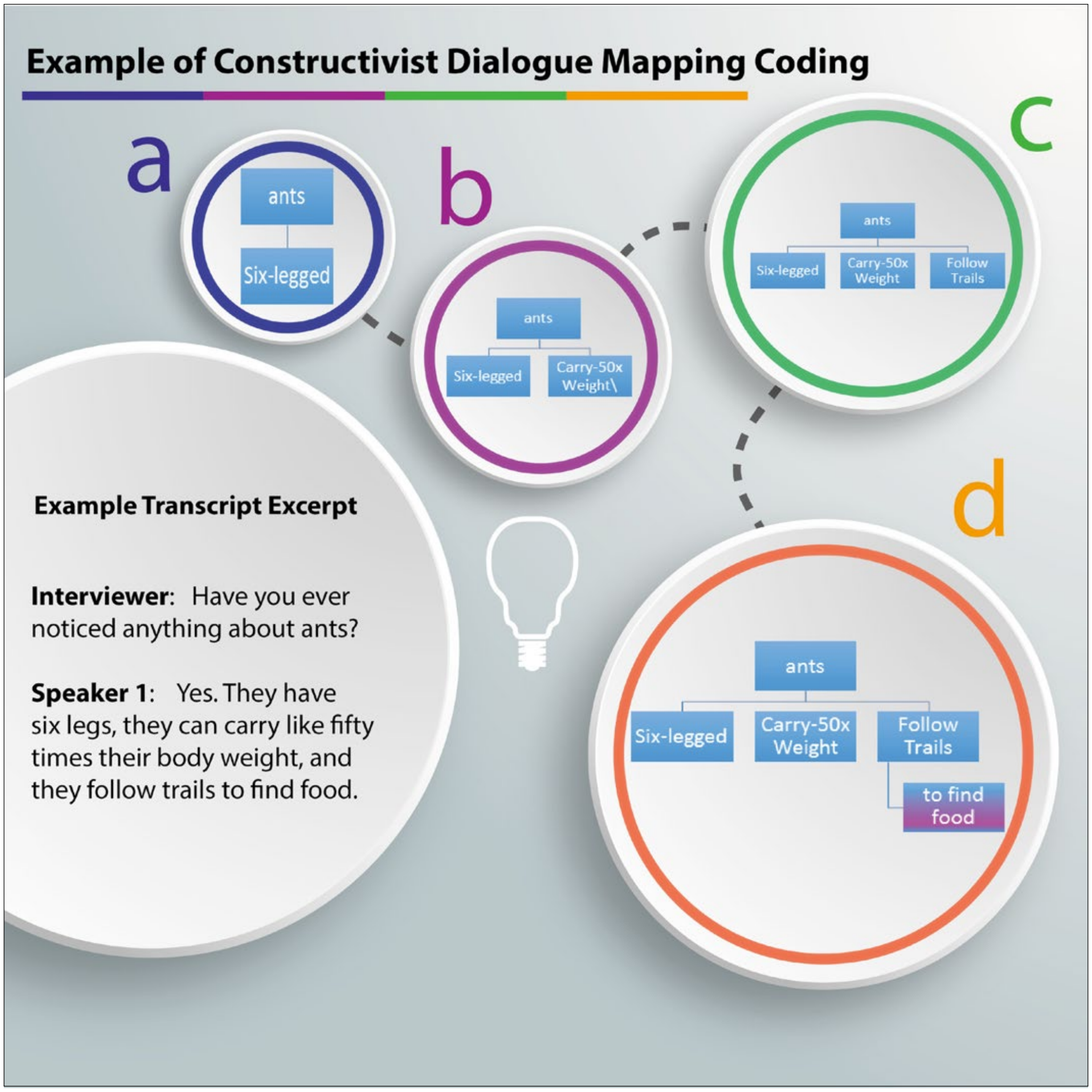

Fig. 6. Constructivist dialogue mapping provides a simple interactive way of mapping ontological entities, to their functions as demonstrated by players in short interactions.

Table 1

Coding Scheme for Constructivist Dialogue Mapping. Captures where in the interview knowledge was mentioned, and by who, allowing for analysis of the development of knowledge as it is mentioned

Session ID	Time	Interaction Type	Sequence	Speaker	Question	Quote	Node	Parent Node	Node Depth
3	10:23	Pre-Interview	0	Kit	"Have You ever noticed anything about ants?"		Ant		1
3	10:33	Pre-Interview	1	Briana	"Have You ever noticed anything about ants?"	About ants? They carry fifty times their weight?	Carry-Six times their weight	Ant	2

The resulting representations have three affordances:

1. Evaluators can count the number of entities players notice during gameplay. This affordance is used in the next chapter. We can note the number of actions and properties they ascribe to those entities. The counting proceeds by adding the total number of 1st order elaborations, then the totalnumber of actions, and then properties users ascribe to that entity.
2. If surveyors sample visitors at more than one point in time, we can track the changes over time of both the number of entities they notice, and the number of actions and properties players' use to describe the agents' actions and properties.
3. Researchers can track the conflicting and changing nature of knowledge as demonstrated through the maps.

In an open-ended environment people learn what is allowed in a system, rather than memorizing. Therefore, I used mapping instead of measuring change through responses in a more rigid classroom-style questioning.

"In this picture, the participants are active theorizers. They gather new evidence and devise methods to test their theories. Instead of accepting classifications as given, they see these classifications as provisional theories that are constantly reassessed and reconstructed in light of the dialogue between theory and evidence" (Wilensky & Reisman, 2006, p. 172). I sought to capture how talk changes based on interactions with the learning environment of *Ant Adaptation*. I propose that learning is demonstrated by what students added to their discussions while playing. This elaboration in their discussion is demonstrated by coding their interactions with CDM.

7 Results

During the activity, users learned about complex systems. I tracked this elaboration of concepts through CDM. I tested CDM in two treatments: (1) in a clinical interview with Rebecca, where she develops ideas about ants during an hour-long interview, and (2) with a group of students who played *Ant Adaptation* in a museum. These samples were chosen because they demonstrated key aspects of constructivist dialogue mapping as a tool.

7.1 Dialogue Elaboration through Play

First, I present two demonstrative cases of the 44 interviews conducted with a complex system model of ants (Martin & Wilensky, 2019) to illustrate how researchers can use CDM to study learning as elaboration.

I coded two people engaged with *Ant Adaptation*. First, conducted under treatment one, I examined how a woman, Rebecca, who is Asian, (all names changed for anonymity), age 28, engaged. Second, I examined a group from treatment two, noting how Thomas, age 10, engaged in a group of three other White youth: Ed, age 12; Mary, age 9; and Sam, age 6. The five players presented were more engaged than the average player of *Ant Adaption*, along several measures: All of them touched the screen and smiled during play (showed their teeth while separating their lips either with or without audible laughter), and all worked to maximize their ants' colony population (stating that they wanted to maximize their population during play). This contrasts with the wider sample where 81% (92 people) touched the screen, 43% (49 people) smiled during play,and 41% (46 people) tried to maximize their colonies' population during play. As there was no guidance on the goal of the interaction, it was surprising that so many of the groups chose maximizing their population of ants as their primary goal. In the open-ended environment, they could just have easily drawn smiley faces with their fingers or planted a flower garden. Perhaps this goal was so popular in the museum because of the competitive arrangement of the exhibit. Yet this also occurred in the clinical interviews. For the clinical interviews (which lacked this arrangement), I am not sure why this happened; however, this pattern arises again in chapter 5.

7.2 Treatment One, Assimilation of Pheromones:

Constructing a Trail Theory from Watching Digital Ants

Sitting on a couch in a quiet room with just Rebecca, I began the interview. I started the interview by telling Rebecca I was interested in ants.

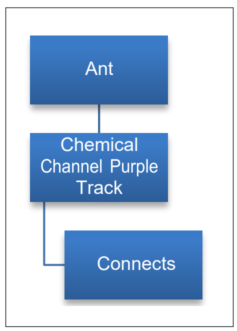

Fig. 7. Rebecca identifies a purple track process that ants do.

I asked her if she would mind if I recorded her while I asked her some questions. After she consented to be interviewed, I asked about her prior ant knowledge from which it was clear she had minimal understanding of ants, or ant colonies. I then coded the interview with CDM to observe moments of cognitive change, or synthesis. This codingis the beginning of making a little mental machine we can use to dissect with transcript analysis. The current process produces maps from transcripts, which are useful as summaries, but the exact connection between cognitive change and the maps is still amorphous. In other words, I have only just stepped into this nebula, and am still finding names for what I am seeing.

In this proto state, to demonstrate CDM, I present the portions of the interview where Rebecca constructs an understanding of the pink lines (pheromone trails), what they do, and how they affect ant colonies in three parts. This construction occurs through addition of ontological identities, such as trails, and ant queens, and synthesis of competing ideas. Over time, it becomes clear she holds two separate notions of the same concept, and then merges them. This is an example of concept formation, instability of concepts, inference from theory, and the conservation of a new ontological entity. Thisis also an example of deep learning as she adds functions and properties onto her idea ofants, and pheromone trails. In the first minute, she presents her confusion over "purple lines." Here she identifies an action she observes, but has not yet conserved into a concept she can operate with. At first, she identifies a mysterious, unknown object. This is pre-conservation:

Rebecca: *I'm looking at... I'm looking at the red one in the middle. The red ant. What are they doing here? What's that line for?*

Interviewer: *What do you think?...*

Rebecca: *This purple line. Look, there is another one. It's dropping here. Thered ants are dropping. Hold on, let me put this. What's that? Is that channel? Chemical Channel?*

Interviewer: *What do you mean?*

Rebecca: *Like when it releases chemical it left – it's a new colony here – it lefta purple track. You see just the purple track you see just purple track disappeared a couple of minutes ago when I ask you and, you asked me 'what do I think'. I thought it was just something like a track. A chemical. Yeah.*

Rebecca first sees a purple line coming out of a red colony in the middle of the screenand wonders what it is and what is it for. She sees that ants leave it. At this time, Rebecca's confusion with the purple lines first comes up. As seen in Fig. 7, this identification helps her identify a novel process: "Chemical channel/purple track" that ants do. She immediately ties chemicals to the lines but does not seem to put forward why. In other words, she creates an ontological entity to see the game, but has not added many functions or properties to it yet. She attempts to test her ideas. When I probed for what she thinks the track does – that is what functions it has – she is seeing the connective property of chemical tracks and then one of the potential impacts of that track. In other words, she starts to hypothesize about the function of the line. This is the beginning ofa conceptual change where she begins to elaborate her understanding through dynamic interaction with the complex system game. I then ask what she thinks it does. She suggests it connects ants:

Interviewer*:* *What do you mean by a track or a channel? What's the channel or thetrack do?*

Rebecca: *It connects these two groups together. This chemical.*

Interviewer: *Okay. It connects them how?*

Rebecca: *Look they are connecting. And I guess it's not a good thing because it looks like it's a war.*

Interviewer: *It's a what?*

Rebecca: *It's a war.*

Interviewer: *Where's the war*?

Rebecca: *Here. [Points at the top right of screen where ants from both coloniesare massing]*

Interviewer: *How do you guess that?*

Rebecca: *Oh, there's no turning back as there's no turning around that red ants back there. They are entering into the black ant colony. And the numberof escape, hold on. The reason it is not dropping because there's flowers.So far. Look they keep going to this black ant colony. And I don't knowwhat they're doing are they trying to steal that food? They are stealingfood. They are stealing food. Come on! [talking to the ants] You have food here [points at the screen where flowers are].*

This is an example of elaboration of her concept. After initially realizing she doesnot know what the chemical does, through testing she constructs the idea that channels "connect" ants and sees that this can cause "war" of some sort. The war seems to be an example of prior knowledge about conflict. But she also begins conceptualizing "war" because of self-organization of ants connected by channels. After answering other questions, she brought the topic up again, confirming that she had kept thinking about it. Sheat first was confused, but then the interviewer provided her with some declarative knowledge, that pheromones are a type of chemical which connects ants together. She then takes that information and identifies a new function, but does not merge it right away with her previously constructed idea of chemical channels. Instead, she holds two categories of activities ants do. In other words, she has constructed two models, one to do withchemicals, and one based on the declarative knowledge provided. About pheromonesshe has no reason to think they are the same concept from her previous experiences, asin the real-world pheromones are invisible, and these purple lines are quite visible. As a result, she holds two concepts at once:

Rebecca: *I'm kind of confused by the purple zone. So, this purple zone issomething like this chemical released?*

Interviewer: *Why do you think that?*

Rebecca: *So, let's see. I just created this chemical channel.*

Rebecca: *If this flower needs to be on this chemical to help. I'm gonna see. It doesn't eat them. Go eat them. Why don't they eat them? It's weird. Itdoesn't work that way. I should just make circles around the corners, so they know there are chemicals.*

Interviewer: *Could you press pause? I have a long question to ask you.*

Interviewer: *So, E.O Wilson, a big ant researcher, discovered pheromones. He realized that they are like perfume placed on the ground to mark the food or when they encounter an enemy to encourage other ants in the colony to come over and fight with the enemy. After hearing about Wilson's findings how might you adjust your explanation of how theseants are getting food?*

With the probe, I wanted to see what the introduction of declarative knowledgedid to the understanding of the model. At first, she took the declarative information in, and made valid predictions based on it. But interestingly, at first, she did not map this onto the purple lines she's been confused by. Instead of conceptual change she held two concepts at the same time, as shown in Fig. 8**.**

Rebecca: *Well when one ant senses the food around them they will leave this pheromone to other ants. And so, they can be attracted to this, to the flowers.*

Interviewer: *Ok. So, what in here do you think is the pheromones?*

Rebecca: *I guess somewhere near this colony. Uhh. No no actually here maybehere.*

Interviewer: *Here what?*

Rebecca: *Here when there is a group of [ants]. There is a huge group of ants.*

Interviewer: *Ok.*

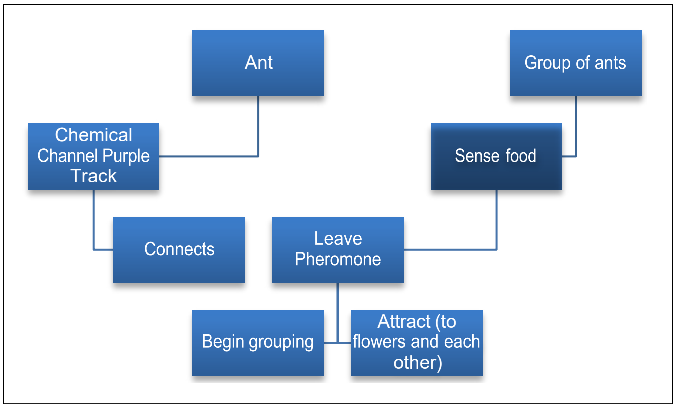

Fig. 8. Rebecca sees two separate processes, Pheromones and Chemical.

Rebecca: *Does that make sense?*

Interviewer: *So, you think this this what wouldn't call it this perfume placed on theground. Do you think that's just where ants are?*

Rebecca: *Yeah.*

Rebecca: *Because where there are ants there are pheromones are a way for themto communicate. If there is somewhere like this [points at blank area]*
... this, there's no ants. How do we know there are pheromones? Thereare not ants so there is no pheromone. This is a chemical releasedby them. So, I guess that the denser a group of ants are, the more communication they will have to with each other, the more pheromone there will be. That's my that's my guess.

Interviewer: *That's your guess. Does that change your way you're explaining how they get food?*

Rebecca: *Can I say I don't know.*

Interviewer: *Yeah that's fine.*

Rebecca: *I don't really know.*

At this point, Rebecca has two ideas. As shown in Fig. 8, first there is a purple chemical trail left by an ant, which connects ants. Second, she also imagines a mechanism called "pheromones" learned from the researcher, that ants leave, and that attracts more ants as communication gets denser. Interestingly, this sort of situation may come up in many learning environments, where a learner may have her idea constructed from her own experience in the world, and a set of declarative facts. For Rebecca, these two are models she is simultaneously using to explain the simulation. Shortly, she will conserve the two ideas into one through concept formation, inferring missing information from that concept. At this point in the interview, she uses these dual models to answer some questions about how congestions and disease work in ant colonies and applies both ideas differently, demonstrating the robustness of the two models. After seeing it employed, I finally ask her if she can also see pheromones in the model. At that point, she seems to assimilate by merging the two understandings into one idea:

Interviewer: *Can you see the pheromones in this model?*
Rebecca: *OH! Is this purple line pheromone?!*
Interviewer: *What do you think?*
Rebecca: *Hold on, let me just... cannot see clearly. My view is blocked by these flowers.* [She examines the model for 16 seconds.]

Rebecca: *Oh, I see. Did you notice that? You obviously noticed that.*

Rebecca conserves her two understandings and then double checks her Eureka moment:

Interviewer: *Noticed what?*
Rebecca: *The denser the ants are the more ants, the more it is, you see that antsare grouping here and the surrounding background color is like, brighwhite. Because they are releasing a lot of pheromones. A single ant justleaves this pheromones trails, like these purple trails, but when they are grouping together there's like a powerball.*
Interviewer: *Like a what?*

I was not sure what a *powerball* was, but assumed she meant a massing andwanted to be clearer on that. She clarifies that it means a lot of ants grouping together.

Rebecca: *It's a powerball.*
Interviewer: *Ok.*
Rebecca: *There's a very strong powerball which means a lot of ants are releasingthe pheromone chemicals together.*
Interviewer: *So, what does that do?*
Rebecca: *Now? I don't know what they are doing. There are just I guess, I knowwhat you mean, look you see these bright white places. that means theyare releasing chemicals aggressively. They are releasing chemicals tolet maybe friends their team players know there are food. 'Come here.'*
Interviewer: *So why is the pink disappearing?*
Rebecca: *Disappearing? Maybe there are no food for them. Like here.*

Rebecca: *Because it's a chemical. It disappears.*
Interviewer: *What do you mean?*
Rebecca: *You release chemicals and chemicals kind of spread in the air.*

As seen in Fig. 9, Rebecca assimilates the two branches of her understanding into one, forming a concept. This is an example of conceptual change. She no longer holds two separate ideas, one from declarative knowledge and one from experience, but instead holds one model of how pheromones lead to agglomerations of ants that help themself-organize to solve daily tasks. This is also an example of deep learning, where she forms a process understanding. Then she adds a beginning state and an ending state to the pheromones process. She also maps the action onto the coloration of the ground process, seeing that whiter increases grouping resulting in a "powerball" where white means higher concentrations of ants, and purple means lower concentrations of ants. This is an example of forming a theory to better see the world. It also is an exampleof discovering the meaning of representation through both declarative knowledge and working with a computer-based model. Finally, she also invokes prior knowledge inher theory to predict the chemical dissipate in air subprocess when released. It's a breakthrough moment. After joining these processes together, she makes several predictions with her newly conserved theory helping her see the world in a new way:

Rebecca: *Now I know.*
Interviewer: *What do you know?*
Rebecca: *This chemical maybe [sends] the wrong messages.*

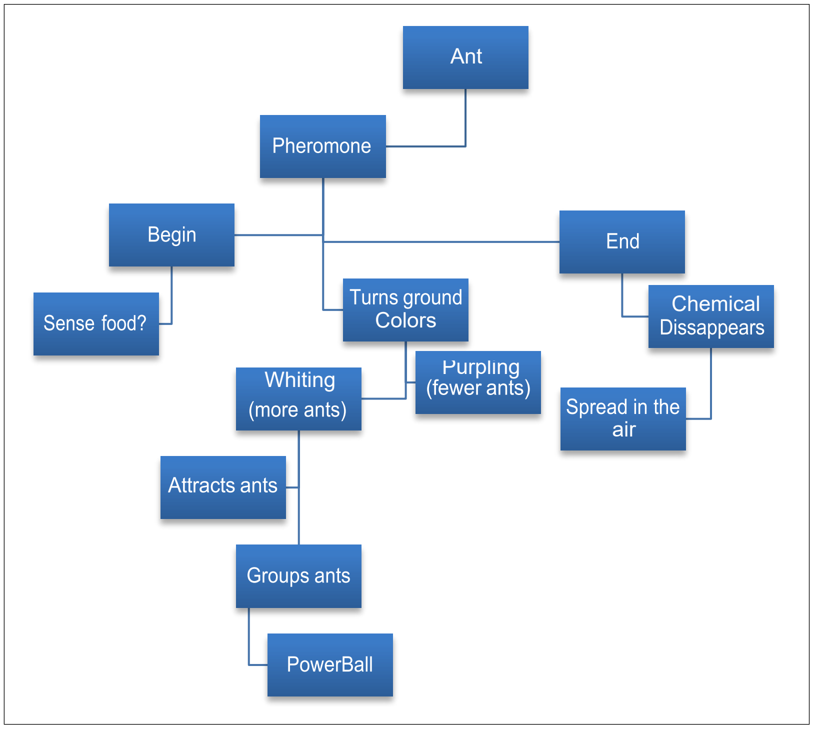

Fig. 9. Rebecca accommodates her observation of chemical trails with her declarative knowledge about pheromone trails. She then adds beginnings and endings to her processes.

Rebecca: *Oh, I see. The stronger the messages the stronger this is, the more antswill be attracted to this place. .... Such a good game, I love it.*

Rebecca: *We're kind of helping them to release the pheromones. Is that right,about this chemical?*

Interviewer: *How would you know that?*

Rebecca: *Because there is a chemical option.*

Interviewer: *And how do you know that's how you do. Why did you guess that?*

Rebecca: *Why do I get that?*

Rebecca: *Yeah.*

Rebecca: *Because once I press this button they are kind of attracted by this track,this pheromone track.*

At the end of the interview, Rebecca has constructed a model that the purple tracks are called pheromone tracks, they are laid down by ants, and help the ants organize for food gathering and war. At this point Rebecca has assimilated the two understandings, observed and declarative, to form a new theory of how ants communicate. This new theory led to conceptual change through play with a simulation that she can use to predict the ants' movements, and trail placement:

Interviewer: *Do you know anything now, after playing this game that you didn't know before?*

Rebecca: *Yeah. Like this pheromone. Chemical I was confused. Like creating channels for them. You know actually not, its leaving chemicals for them to follow.*

Rebecca went from visual confusion to a theory. She now mapped a system of communication onto what was once visually confusing through assimilating while using a simulation through a process of conceptual change.

As a caveat, from this interview there are two key factors I do not know: (1) the length of time this new conception will last. Without reinforcement, or better yet, a delaystudy to query how long the idea persists, I can say nothing about the durability of theseideas. Worryingly, because they arose quickly, they may fade quickly. And (2) I shouldconsider the situationally contingent nature of the learning, it happened in interaction witha model, and an interviewer. If other researchers use the method, some thought should begiven if such situational constraints should be inserted while coding with CDM.

7.2.1 Rebecca's Agent-Based Knowledge Construction

To answer the second research question – *can we see evidence that new knowledge structures emerge through gameplay and how these structures shift through time* – I examined how Rebecca builds her knowledge about ants. I answer this by following through her elaborations during the play. As shown above, Rebecca first identifies a phenomenon she does not understand, namely, purple tracks. Then through observation, she makes a conclusion that that phenomenon attracts other ants. From this prediction she expresses its function, but maintains some confusion. This confusion is the affective state that accompanies her process of accommodation, that pushes her to accountfor greater and greater parts of the complex phenomenon. As she forms the concept, she operationalizes it to make conclusions about excessive aggregation of ants in large reinforcing groups caused by pheromone's attractive qualities categorizing them as "powerballs". She then explains, with here newly constructed theory, how the phenomena end through the dissipation of the chemical through air, like perfume.

This interaction also suggests an answer to my first question, *how can we capture visitor's moment-to-moment sense-making*? Moment-to-moment expressions stick together as they account for what Rebecca sees. She even connects the declarative knowledge I probe with. When they no longer helped her understand, she dropped the ideas. Like her idea of pheromone just being on or off, rather than a continuous variable where more concentration is lighter colored and more attractive. This process of pruning suggested that in future uses of CDM, researchers should set a heuristic where we say something like "if an idea isn't used for X minutes, we can drop it from the tree." Asa result, I used a version of this heuristic in the second treatment. Thus, her knowledge is assimilating over time to accommodate to what she has seen. In other words,she is constructing an idea of the parts of the representation she at first is confused by, and forms trees of processes that account for ever greater complexity of action in the agent-based model. At times, the process looks like unstable concepts. However, as she gathers more evidence, the concepts become helpful, predictive theories, which assist her learn how to see the model.

CDM is a useful way to document this change and stability. The interaction had two main effects that CDM demonstrates. First, through the process of playing in the open-ended environment Rebecca created a theory of how feedback, represented by pheromone trails, in the complex system leads to self-organization. Second, the analysis of Rebecca's use of *Ant*

Adaptation demonstrates constructivist dialogue mapping's utility in tracking learning as concept elaboration. *Ant Adaptation* scaffolds her development oftheory to see how trails work in the microworld. The process of discovery engaged her as she constructed this understanding of pheromone trails.

7.3 Treatment Two:

Constructing a strategy to play Ant Adaptation through Feedback

While CDM can be used to track concepts as they develop, as in Rebecca's case, you canalso use it to monitor what people know before and after an intervention. In this second use, researchers can code the changes, before and after. New additions are coded in red in this treatment.[4] To demonstrate the process, in this section I re-analyze the interaction of Thomas and his family had that I discussed in chapter 3 using CDM. While playing in a large natural history museum with his brothers and sister, Thomas and his siblings were recruited to the study. Often in open-ended learning environments, interactions are not individual. Instead, a whole family can interact with the exhibit simultaneously. Thus, Thomas' case demonstrates what one user developed through interacting with the game and his siblings. This section addressesmy third question: *How do these knowledge structures shift over time, or remain stable emerging as theories explaining the context*?

7.3.1 Thomas Pre-Interview

During the pre-interview, the interviewer established Thomas' prior understanding of ants througha semi-structured interview. As shown in Fig. 10, during the pre-questions he said ants carry 50x their weight. In response to probes about how ants control traffic, he offered the explanation that ants make physical paths, to control traffic, which he later rescindedafter playing the game. Protocol questions on aggressive roles prompted Thomas to guess aggression simply increases an ant's likelihood of getting hurt or even killed – a claim he later amended:

> **Interviewer**: *Okay. And then imagine you're an ant. How do you think being aggressive affects your life?*
>
> **Thomas**: *Uh, if you're aggressive, like, you-you're ... Uh, how do we explain this? If you're aggressive, like, you-since you go for more things you have a greater risk of getting hurt or killed, I guess.*

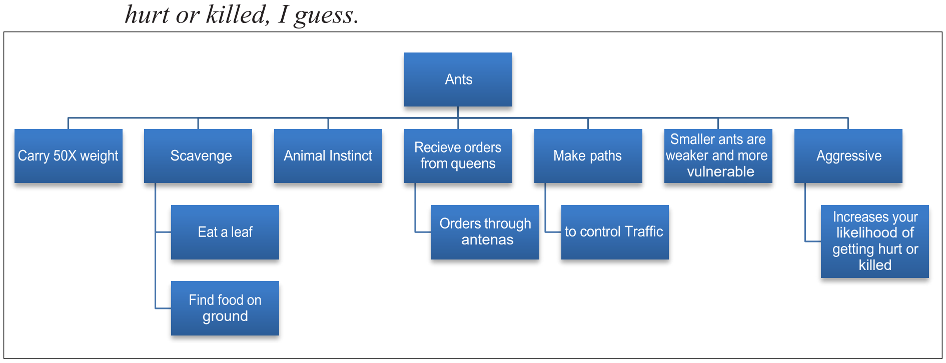

Fig. 10. Pre-play constructivist dialogue map of four players' understanding of ants. For instance, they think ants (entity) can carry 50 times their own weight (mechanism).

---

[4] While in this case we only present before and after, researchers could pictorially represent the changes at finer grain sizes to show more of the construction of concepts, and set heuristics for when ideas no longer inhabit a CDM.

In response to questions about how ants know what to do, Thomas offered animal instinct, or, a command and control understanding when ants get orders through antenna and/or from their queen.

Interviewer: *Okay. That's fair. And then, how do ants know what to do? So, you said they-they pick up leaves, or they scavenge, but how do they know to do that?*

Thomas: *Um, animal instincts.*

Interviewer: *Animal instincts.*

Thomas: *Or the queen tells them to, if there's a queen ant. We don't know.*

Interviewer: *So, what's the diff-How does the queen tell them to?*

Thomas: *It's something with their antennas.*

Interviewer: *Okay.*

Thomas: *I think. Additionally, he said that ants scavenge food from the ground or maybe eat leaves.*

Thomas: *Like, they can like go hunt for food. They can like, um, try like, get to some, like, maybe some food on the ground like in the city or like in a park, or they can just eat a leaf.*

After I explained how the game worked, players broke into two teams, Thomas and Ed versus Mary and Sam. At the beginning, the younger Sam and Mary chose to have maximum-sized, not very aggressive ants (10% aggressive). Thomas and Ed chose to have medium sized though not very aggressive ants (2% aggressive).

Near the beginning, players agreed through conversation on what strategies matter. Soon, Thomas informed the older Ed how to play: "add flowers close to the nest." From this I add a new ontological entity, flowers, into the group's constructivist dialogue map and put a mechanism 'close to the nest' under it because the players started planting flowers in proximity. Notably, they did not mention flowers in the pre-interview, but also did not seem surprised by them. I think this is a case of constructing knowledge through action in the game world, as it is unlikely, they knew that some types of *Attini* ants collect flowers for fungus farming. The players then established the connection between two complex systems ideas – the proximity of food and increasing ant population – through dynamic play with the model.

7.3.2 Post-Interview

Thomas seemed to develop an understanding of the feedback cycles inherent in *Ant Adaptation*. Thomas offered this understanding as the way to play the game.

> **Thomas:** *Yeah, you had to figure it out and the-you have to have some flowers, see, and then you put the chemicals and lead it to there, then they'll bring it back, and like, if you want to get rid of the chemicals you use the vinegar. So, um, you put some sunflowers down, then you get the chemicals and lead it to the sunflowers and if-if there's too much thenthe ants aren't getting the sunflowers and you-then they'll just like, then you use the vinegar and erase it. But if-if you just do one path thatleads to the sunflowers it'll just get the energy and just keep going backand forth and back and forth. And that's how we got 21 [ants].*

As shown in Fig. 11, in the post-test the players' concept maps became more elaborate. Thomas takes on a more cyclical understanding of the role of ants' paths to attract

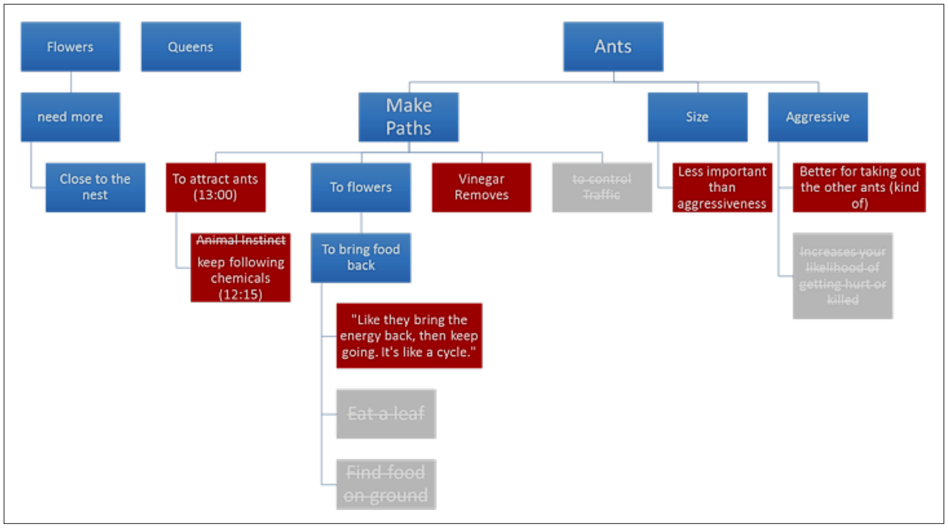

Fig. 11. Post-play constructivist dialogue map of four players. Came to the cyclical understanding of pheromones in food foraging and a more contextual understanding of the role of adaptions utility. Red Boxes indicate elaboration. Grey ideas that were not mentioned again.

each other to flowers. Thomas argued that ants follow chemicals to bring food to "getto 21" ants. He also saw that sometimes ants can get trapped in their own chemicals or "white spots." He used vinegar to clear excess chemicals. One of the primary motivations of the learning environment was to teach that the simple rules of agents couldlead to complex, social patterns that sustain a population of ants. Thomas's descriptions indicate he came to understand the impact of ants' simple rules on complex social behaviors. Thomas leveraged this new understanding to use the macro-level effects of population size to lead his ants to victory by making predictions based in his theoryof the agent-based model's simple rules. Thomas employed complex systems thinking learned in the short interaction to reach his goal of maximizing population. He set this goal in communication with his teammate in the open-ended constructionist learning environment afforded by the design of

Ant Adaptation.

Through play, Thomas learned that entities like ants have a mechanism such as laying trails to attract other ants to flowers in a cycle by recursively following the chemicals. He also realized that sometimes this process can lead ants astray as shown by his use of vinegar to redirect them out of deleterious local optima. He also constructed the concept that the food source's proximity to the colony increased the ant population by increasing food intake. After he stated as much, the other side started placing flowers close to their nest, indicating they also understood agents' actions led to the macro-leveleffect of population growth.

8 Discussion

I developed CDM (Martin, Horn and Wilensky, 2018; Martin and Wilensky 2018)to study learning in Ant Adaptation based on constructionist theory. Piaget presented a notion of knowledge that is constructed. Since he believed knowledge is constructed, Piaget invented the clinical interview to document how people construct these understandings during open-ended conversation and question-asking designed to illuminate the way the child thinks of or explains a phenomenon. The next step in this research for informatics is to automate it. This work poses some challenges, but will be feasible in the near-term and so is worth discussing here.

It is difficult to automatically transcribe what children say. But the barriers are dropping. According to a report by RideOut Media (2017) Most American Children (98%) are now in a home with a tablet or smartphone. These provide ubiquitous access to microphone, voice search enabled devices (Lovato & Piper, 2019). Otter.ai, an AI powered transcription service, along with REV.com, are dropping the cost and increasing the accuracy of automated transcription. As this revolution takes place, the need to understand and identify key learning moments in children's speech will rise.

Consequently, we should focus on natural language processing to identify ontological entities, the functions they employ and their properties. In a naïve search, if all concepts had a specific name, this would be a search for nouns, the verbs those nouns take, and the adjectives and adverbs that that modify them. Unfortunately, language is less clear than that. Not all entities have names. Antecedents are not always clear. Now, bridging these connections is the work of humans.

Previous research, however, suggests that an automated ontological concept mapping can help track theory development. For example, the OntoClean (Guarino & Welty, 2002) methodology supports the construction and evaluation of taxonomic relationships based on the use of a number of philosophical meta-properties– namely, unity, rigidity, identity and (notional) dependence – as well as on constraints limiting relationships that can be established between concepts (types, properties) tagged with these different meta-properties. It was the first attempt to formalize ontology for information systems. Moreover, Newsreader (Vossen et al., 2016) provides aggregation of information over massive amounts of online text to build stories that decision makers can use. In Newsreader there are four steps:

(1) Identification, that extracts what happened to whom, and where.
(2) DeDuplication, which makes sure that similar information across a corpus isavailable only once, while referencing each article referring to it.
(3) Aggregation shows complementary articles across the corpus.
(4) Perspectivation, makes sure different viewpoints and perspectives in the corpusare traceable in the final narrative.

While the previous work with Newsreader and OntoClean examines large corpus understanding, I want to track how children develop their ideas in informal learning environments. They develop their ideas rapidly, and sometimes hold two conflicting notions simultaneously. CDM suggests potential for the context. The advantage of this approach is researchers can present the smallest parts of what we observe with large samples of transcript data. For instance, we can note when users first identify an unknown activity and then how they build out an understanding of the subprocesses of that task. Currently, we read the transcript to see how users construct such operational knowledge on the representations they see.

9 Future Work

As a purely qualitative method, CDM demonstrates learning as concept elaboration over time through the proxy of changes in speech (Leinhardt and Crowley, 1998). Through automating this process, in future work, I will scale the process, and thereby test the methods on larger n-samples of data. Mixed Human-Artificial Intelligence solutions, like Quantitative Ethnography (Shafer, 2017), Ncoder (Cai et al., 2019) or DeepLabCut (Mathis et al., 2018) are methods that I will use in the near term. These approaches mix researchers' hand-coding a small number of bits of text,that can scale to larger corpuses of text or behavior data using regular expressions or machine learning.

10 Conclusion

The system of tracking players' conceptual development illustrated how their thinking changed throughout the interaction. Thomas learned that the chemical trails lead ants through a cycle that feeds the whole colony. He also learned that sometimes this processcan lead ants astray, and how to intervene with vinegar in the self-organized system to optimized it. Meanwhile, Rebecca's knowledge emerged through play. Using CDM, I tracked how Rebecca created a theory of how feedback between ants in a colony leads toself-organization of food foraging. These two examples showed CDM's utility in tracking learning as concept elaboration arising from interaction with technology and a facilitator.The method highlights two parts of the theory:

(1) Elaboration of discussion improves understanding. And by tracking the paths of elaborations and its sometimes-dual conceptions a researcher can observe conservation of concepts in transcripts.
(2) Knowledge fluidly grows through action taken in a complex system. The dynamic interaction with the model enabled users to learn through mediation with the computer system, each other, and a facilitator. This sort of dialectic interaction could be scaled to other learning environments, and tracked through CDM.

The game facilitated its learning objectives by scaffolding theory development. Whenpeople engage part of a complex system, they attempt their best theories in real time, andreceived dynamic feedback from the computer and each other. The overly abundant datagets fit into a theory that people test in mediation with the machine and themselves and finds regularities and breaks in their system of thought.

The constructivist dialog mapping approach discussed in this chapter was used to capture changes in a player's understanding of agents, functions, and properties. I demonstrated the process of learning complex ant systems through playful interaction with the agent-based modelling game, *Ant Adaptation*. In future work, CDM could also be useful in tracking what people learn in other learning environments where they interact with peers, technology and teachers.

Elaboration of discussion improves understanding. Through the presentation of Thomas' case, I show that a player added to his understanding of ant colonies throughthe elaboration of discussion about a complex system. He started using vinegar to eliminate local optima where the ants got trapped. Simultaneously, he used the feedback of pheromone trails to organize his colony's foraging and adjust adaptations to changing circumstances.

Knowledge fluidly grew through action taken in the complex system. From Rebecca,I notice that her knowledge is fluid – namely, she came up with theories on the fly. She took a while to decide on one, and it seems her notions were a bit more fixed and situational. The research question was how users are building their knowledge whenusing the model of ant competition. She understands the rules of the behavior of individuals to make sense and meaning from agent-based models. Thus, I find this agent-based model a felicitous way for students to freely construct. Working with each other and mediated by the tabletop game, they build understanding of the complex systems processes by conserving change and forming operational concepts – like children growing up in a gas cloud.

Effortful problem-solving activity is the process of science, and that is the process constructivist dialogue mapping tracks. The CDM approach discussed in this chapterwas able to capture changes in a player's understanding of agents, functions, and properties (entities' mechanisms), while they learned complex ant systems through playful interaction with the agent-based modelling game, *Ant Adaptation*. CDM captured the changes as utterances occurring during a short interaction. By analyzing changes intalk pre-and post-play, I found that a player learned about feedback, and employed that learning at multiple levels to maximize an ant population. The elaboration took place by forming and testing theories with *Ant Adaption*. This method could be usedin other learning environments to inquire about the association between elaboration of talk about concept formation.

11 Acknowledgements

I would like to thank Paul Won, Gabby Anton, Bill Hoover, Rui Han, and Melissa Perez for their feedback. I would like to thank Francisco López-Bermúndez for his illustrations. I also would like to thank the Center for Connected Learning and Computer Based Models, TIDAL, and TIILT labs for their feedback throughout the project, and the supportive Northwestern Community. Additionally, Marcelo Worsley and Emily Wang provided copious insight. Finally, I would like to thank the National Science Found (NSF Stem+C: Award # 1842375) and the Multidisciplinary Program in Education Sciences (IES: Award # R305B090009) for funding the project.

Chapter 5: Ant Adaptation: Measuring Complex Systems and Model-Based Learning with Multiple Data Streams

By Kit Martin

1 Introduction

Recalling Thomas from chapter 3, we should remember that he, and his generation will have to cope with climate change and how that is upsetting the biosphere. He will have to contribute to the solution to plastic waste. Today, his family must cope with the issue that the semiconductor shortage arising in Taiwan make it harder to buy a car in the United States. The issues he will face arise because the world is interconnected like never before (Bar-yam, 1999; Bryson & Crosby, 2005; Hmelo-Silver & Pfeffer, 2004; Morin, 2008; Wilensky & Rand, 2015; Wilensky & Resnick, 1999). As a result, in recent years there has been a greater focus on teaching and understanding the complexity of the emergent connections between humans, the environment, and the externalities. The interconnectedness means that the problems Thomas and his generation face, from climate change (League et al., 2019) to income inequality (Atkinson, Muro, and Whiton, 2019; Escobar, 2011) to traffic (Barceló, 2010), or pandemic disease (Jenkins et al., 2020), no one is really in charge of these emergent processes (Bryson & Crosby, 2005). To effectively address these issues, we need an informed public along with educational environments that teach how to think about complex systems. Students understanding the complex world we live in will improve the chances of overcoming these pressing problems. With this remarkable interconnectedness, it is clear we are all now planetary citizens; but it is clear our educational systems have not yet prepared us for this situation (Morin, 2008).

In this chapter, I use two methods that have been successful in supporting learning about complexity: constructionism and agent-based modeling. Constructionism, a mnemonic term (Papert, 1986) coined to describe a species of constructivist thought (Piaget, 1983), focuses on the benefits of learning from the external construction of an artifact alongside the internal construction of a mental model. This pedagogical theory argues that learning is the emergence of new knowledge structures, or mental models (Papert, 1980; Papert & Harel, 1991a) aided through building public entities (Bamberger, 2001; DiSessa et al., 1991; Holbert & Wilensky, 2019; Nemirovsky & Tierney, 2001; Papert, 1980; Papert & Harel, 1991a). One area of research that has been emphasized in this field involves the use of agent-based models to create constructionist learning experiences for complexity education (Goldstone & Wilensky, 2008; Klopfer et al., 2009; Sengupta & Wilensky, 2009; Wilensky, 1999b; Wilensky & Rand, 2015; Yoon, 2018). In particular, this work has focused on understanding complex systems of science phenomena, for example: bees (Guo & Wilensky, 2014), material science (Blikstein & Wilensky, 2009), ecology modeling (Wilensky & Reisman, 2006), electricity (Sengupta & Wilensky, 2009), evolution

(Wagh & Wilensky, 2018), and ants (K. Martin, Horn, et al., 2019). With growing work in understanding social systems, like income inequality (Guo & Wilensky, 2018), that may have outsized leverage (Wilensky & Papert, 2010).

In this paper, I explore two aspects: 1) using constructivist dialogue mapping (CDM), I show the knowledge structures participants evince during the interactions, and 2) using automated face analysis (AFA), I demonstrate how they engage while constructing these mental models. I use the latter to recut the data, applying a data reduction technique, to show how engagement impacts construction of mental models.

This paper has two components. First, I will present the Ant Adaptation learning environment introduced in (K. Martin, Horn, et al., 2019; K. Martin & Wilensky, 2019), which I designed primarily for deployment in the informal environment of the museum based on prior complexity science education. In this paper, I use data both from in-person and remote sessions with Ant Adaptation. The goal of Ant Adaptation is to motivate learners to investigate complex systems, which helps users create mental models of these systems. Second, I will implement a mixed-methods research approach, using both constructivist dialogue mapping (CDM), which has been discussed in prior work (K. Martin et al., 2020) and computationally augmented ethnography (K. Martin, Wang, et al., 2019), to measure and assess the growth and engagement of learners during complex systems thinking. Applying a mixed methods approach is necessary because constructionist learning environments encourage open-ended exploration, which has made it harder for education researchers to develop strong assessments for them. While there have been decades of constructionist work, we have still not solved the problem of measuring learning in an open-ended environment (Berland et al., 2014; Ochoa & Worsley, 2016; Papert, 1980; Papert & Harel, 1991b). But also, when we allow students to become the architects of their own knowledge, we have the chance to increase engagement, a key metric we can measure with these approaches. As learning is no longer one-size-fits-all, it has become harder to develop uniform, standardized assessments. The lack of evaluation slows down their implementation in educational environments. This is unfortunate: this type of thinking is exactly what society needs to deal with core problems in a complex world. The hope is that these two methods will grant greater insight into student learning and engagement with Ant Adaptation, and will be generalizable to other learning environments.

1.1 Overview

In this paper, I present users' interaction with Ant Adaptation via in-person and remote sessions. I collect data via pre-post-delay tests, interviews, video, audio, field notes, and screen capture with the data, I use CDM to measure learning and AFA to show moments of engagement with the exhibit. In this chapter, I first explore theoretical underpinnings of this work.

In the literature review, I start with the theory of complex systems and its use in teaching complexity, then introduce museum studies, focusing on concept elaboration around interactions with exhibits. I conclude with a review of the affective computing literature to focus on a means of measuring engagement through affect: facial coding using automated face analysis (AFA). After the literature review, I recapitulate the design of Ant Adaptation laid out previously in detail (Martin, et al., 2019), which builds on the constructionist learning environment design (Kafai, 2006; Papert, 1980; Papert & Harel, 1991b), and museum studies literatures (Dierking & Falk, 1992; Humphrey & Gutwill, 2005). Following that I discuss methods of measuring engagement and learning including constructivist dialogue mapping, affective state tracking with sensors, and the triangulation of these measures. I then introduce the methods of this chapter.

The chapter proceeds as follows: in section 2, I cover complex systems, including the promise of complexity science and challenges people face with learning complexity. Then I introduce restructuration theory and how we can design ideal environments for students to learn this difficult topic. Finally, I discuss constructionist learning environments that teach complexity—focusing on computer aided microworlds of STEM phenomena.

In section 3, I introduce museum studies, and specifically, active prolonged engagement, and the concept of elaboration through speech as a means of measuring learning. This concept is particularly useful for designing this informal learning environment. In this section. This work also motivates my use of tracking elaboration in speech with constructivist dialogue mapping.

In section 4, I look at affective computing. Affective computing shows many physiological measures that can be tied to human emotion. I will review the history of research into emotion and some of these measures. I will focus on one method of measurement in affective computing: facial tracking using facial action units. Then I will discuss what research says about the role of affect and intelligence and learning, measuring mood, use of these tools in research, manual coding of faces and its limitations, and how that builds into AFA. I will conclude this section by asking about whether emotion is the same as affect, and the role of affect in human choice. This will conclude the background.

In section 5, I recapitulate the design of Ant Adaptation and its learning objectives, and how it is used to situate the results in the environment. I also cover how it can be used for learning about complexity with models and games, and how this overcomes the twin challenges of complexity education: level slippage and overlooking the rules of individual agents in processes like the particles of ink in water.

In section 6, I introduce the measures I use in this study, including constructivist dialogue mapping, measuring engagement, and the ways in which I triangulate these measures to study learning more deeply.

In section 6-7, I go over the methods of the study, the setting of the study, data sources, and my analysis framework. In section 8, I present my findings. In section 9, I discuss the findings, and in section 10, I describe my findings, the limitations and next steps. Then I conclude.

1.2 Research Questions

This chapter poses three research questions:

- Does the technique of CDM add new abilities to gain new insights into how learners, in group conversation, can advance their learning?
- What can we learn from physiological measures of people's affective states while they engage learning in informal learning environments, such as a museum exhibit or a learning game, where people choose to come and learn?
- Is there a relationship between high stimulation and learning as measured by CDM and affective computing signals? Within moments participants elaborated, is there a relationship between positive affect and learning?

Answering question three involves looking at how much students update their understanding during moments of high affect, as compared to other times while using Ant Adaptation. Theory shows that moments of high stimulation (D'Mello & Graesser, 2012a; McGaugh, 2003, 2004) are associated with learning. Affective states pathways, specifically the engagement-confusion-delight-engagement pathway, has been hypothesized to facilitate advanced problem solving if the confusion state does not overwhelm the learning (D'Mello & Graesser, 2012). Andres and colleagues (Andres et al., 2019) found while researching informal learning with *Betty's Brain,* a computer-based learning environment that utilizes the learning-by-teaching paradigm to engage students in learning about science topics, that the only emotional pathway associated with learning gains was sustained delight. Horn et al. (2016) found that post-test gains signified with all of the affect words they considered, such as "*wow*" and "*cool*," after a museum exhibit. Each of these uses of affective state tracking point to greater integration of affect detection to improve design, interactivity, and analytics in learning technology. Question three will take these findings one step deeper and will investigate both the process of learning in moments of stimulation and the role of positive affect in that process of complex system's thinking. The proposed mechanism is as follows: moments of high stimulation are related to things people remember more. I want to know what leads up to those moments, which means investigating whether good design impacts positive affect or those learning moments. What is prompting positive affect? Is it related to the design or the way the social interactions play out? It is a system of mediation? How is it connected? This work aims to underscore the connection between learning and affect.

A large body of research has informed the present study including constructionism, constructivism, teaching complexity science, museum studies, designing Ant Adaptation for open-ended learning, and the challenges of measuring open-ended learning environments in a rigorous way. In place of a lengthy review on the full body of work done studying informal learning and evaluation, I choose to emphasize the literature of complex system (Bar-yam, 1999; Davies, 1989; Morin, 2008; Wilensky & Reisman, 2006), the use of affective measures (Picard, 2000) to study people's responses, the design of digital interventions to restructure education (Wilensky & Papert, 2010), and the use of carefully tracking talk in order to study learning around museum exhibits settings (Leinhardt & Crowley, 1998). I begin the review by first reviewing the literature of complex system. Then I review on representations and learning in a constructionist paradigm through restructurations (Wilensky and Papert, 2010) as it guides the discussion of the design of digital learning environments to teach complexity ideas in educational settings. Then I will discuss museum studies to situate the work in the context of the study. Then, I review affective computing to provide a full over-view of this mixed-methods research project.

2 Complex Systems

Since Weiner's definition of cybernetics (Wiener, 1948) and von Neumann's early work during the Manhattan project on cellular automata (Von Neumann, 1951), researchers have studied systems in which simple parts organize themselves into complex wholes. Self-organization has been defined as making meaning out of chaos (Atlan, 1986) the integration of elements perceived as disorder into more encompassing, often larger, organizations. These are important structures for understanding science and society. We see these systems in many structures, including minds, social insects, crystals, flocking, climate change, income inequality, evolution, and many other domains. These complex systems also touch some of the deepest issues in science and philosophy, including randomness vs determinacy and order vs chaos (Wilensky & Resnick, 1999).

2.1 Decentralization in Scientific Models

Another direction that is fascinating is complexity in scientific models. For 300 years Newtonian physics dominated the world of science and even more so people's perception of science (Bar-yam, 1999; Capra, 2016; Davies, 1989; Laughlin, 2005; Morin, 2008; Pagels, 1989; Wolfram, 2002). Newton offered an image of the universe as a machine, a clockwork mechanism. Newton's universe is ruled by linear relationships of cause and effect. A gear turns another gear creating linked causal chains. In the Newtonian world mutual interaction is not emphasized. When we think of interaction and the Newtonian universe, we think of one object acting on another object. One object is in control. For example, Jupiter controls its moons. The other is acted upon. In that view, the moons do not control Jupiter. Most of the focus goes to two laws of motion with a focus on how force influences the motion of object (Resnick, 1994). Less attention goes to the intersubjective action, underlined in Newton's third law, that focuses on the reaction that accompanies every action. In the 20th and 21st century, the Newtonian worldview has been challenged on many fronts.

One of the more serious challenges comes from the growing interest in this kind of thinking called complex systems. In many fields, scientists have shifted metaphors about the nature of science, viewing the clockwork mechanisms as antiquated and instead employing a metaphor of complex ecosystems (Bar-yam, 1999; Capra, 2016; Davies, 1989; Laughlin, 2005; Morin, 2008; Pagels, 1989; Wolfram, 2002). Davies saying "The new paradigm emphasizes the collective, cooperative, and organizational aspects of nature; its perspective is synthetic and holistic rather than analytic and reductionist" (1989, p. 2). Davies is describing a move away from the classical scientific worldview toward complexity. Rather than looking at the world through one agent or an individual acting on another in a neat causal chain, researchers view the world in terms of decentralized interactions and feedback loops and looking at how complex emergent patterns emerge from interactions among these simple components. This focus on emergence can sometimes feel almost pseudo religious or mystical. In machine learning, when we talk about node structures in a neural network, we often discuss in in terms of a black box, inputs and outputs, because it is difficult, if not impossible to inspect what happens inside. When we need to talk about structures that emerge from social interaction, we oftentimes throw up our hands saying it is too difficult to study. Throughout his career Morin proposed means to study these social interactions trying to abolish the separation between studying humans and nature (Montuori, 2004; Morin, 2008). The Newtonian worldview of causal change is not wrong. Instead, these models are simply inappropriate for trying to make sense of certain types of phenomena (Bonabeau et al., 1999; Gause, 1934; Resnick, 1994; Wilensky & Rand, 2015). And therefore, I argue we need new models to operate at the different level from the Newtonian physics focusing on the behaviors of groups of individuals, not the actions of individual agents. This echoes Wolfram's *New Kind of Science*

(2002). The book contains a systematic, empirical study of computational systems such as cellular automata. Wolfram defines these systems as *simple programs* and argues that the methods and scientific philosophy for the study of simple programs are relevant to other fields of science. In the book, Wolfram takes a programmatic approach. He takes a sequence of simple programs and runs them to see how they behaved. To his surprise, what he found, despite the simplicity of the programs, was that the behavior was far from simple. He argues this is a profound moment in the history of science:

> "It took me more than a decade to come to terms with this result, and to realize just how fundamental and far-reaching its consequences are. In retrospect there is no reason the result could not have been found centuries ago, but increasingly I have come to view it as one of the more important single discoveries in the whole history of theoretical science. For in addition to opening up vast domains of exploration, it implies a radical rethinking of how processes in nature and elsewhere work." (Wolfram, 2002, p. 2)

Beyond the bombast, the work to explore these processes and understand how that reshapes our epistemology are exciting directions in the future of complexity research. Below I present a few of those explorations to illuminate the work of complexity.

2.2 The Mind: A Complex System

These ideas that we need to look at, groups and interactions, have also influenced our thinking about how the mind works. This shift in thinking affects the methods of this paper in a way: I am looking at decentralized units of thought inside individuals and how they share between groups using the method of constructivist dialogue mapping and affective state tracking. This shift underpins a fundamental change of worldview that has occurred in the last half century. Descartes argued, *Je pense, donc je suis (Descartes, 1637)*. Though nothing seems more obvious to most people than the singular nature of mind and a unitary understanding of the self, we experience self-consciousness. In the book *Descartes' Error* (1994), Antonio Damasio aims to restore our appreciation of the role of the rest of body, and the shared balance of powers from which we emerge as conscious persons. We feel like we have a single unified presence in this world. However, as Donald McIntosh — a major intellectual foundation of Nobel laureates in behavioral economics, Kahneman and Tversky — pointed out: "The idea of self-control is paradoxical unless it is assumed that the psyche contains more than one energy system, and that these energy systems have some degree of independence from each other" (1969). If there was a single unitary self, there would be no problem of self-control. Additionally, as the philosopher of mind Daniel Dennett (1993, 2017) argued, each of us imagines our own mind as *I*, not a collection of *we*. The idea of a unified centralized mind has fallen by the wayside in the past century, with accelerated motion over the last three decades. Resnick describes in his book (1994) that people began to study the mind from a multi-agent, complex system perspective. Freud's construct of the subconscious, which obviously could be used to explain issues of self-control, challenges the idea of a single executive in charge of the mind. Freud argued the unconscious is an equal participant with the conscious in the workings of the mind. Freud is not saying the subconscious is a repository of forgotten ideas that comes up merely in dreams, but is a lively agent actively organizing thoughts. He further subdivided the mind into the ego, the super ego, and the id, which pull the ego in different directions. Objects relations theory proposes a collection of inner agents within the mind that primarily seek relationships. With that view, the self emerges from interactions among internalized objects.

Marvin Minsky and the field of artificial intelligence in the 1980s studied the mind from a perspective very different from that in psychoanalysis, but it too moved towards a studying the mind as a complex system 1980s (Minsky, 1988). Marvin Minsky, the pioneering computer scientist, developed an elaborated theory of how the mind works, describing it as composed of many computational agents that comprise a society of mind. In the middle of the 1980s, and again in the 2000s, there was a renewed enthusiasm for neural networks. In Minsky's book, *Society of Mind* (1988) an image of society of mental agents working together to do things that no mental agent could do on its own is put forward. Minsky's later work argued that intuitions, emotions, and feelings are not distinct, but instead different ways of thinking. Through examining these different forms of mental activity, Minsky says, we can explain why thoughts sometimes takes the form of carefully reasoned analysis and at other times pivot on emotion (Minsky, 2007).

Philosopher and cognitive scientist Daniel Dennett has argued for a model with multiple drafts of consciousness. The idea builds off the Uexküll and Sebeok work on *Umwelten*, usually translated as the self-centered world view, of the 1960s. Dennett's work argues that there is no single stream of consciousness in the mind. Instead, consciousness is a manifest image constructed. I interpret Dennett to mean the mind is like multiple news feeds on Twitter or Facebook. Multiple narratives are created and edited in different parts, the dominant ones are trending and lead to consciousness at a particular time. The idea of a single stream of consciousness, Dennett argues (Dennett, 1993), implies a single functional unit that controls it all, where it all comes together, like on one news feed. He says, such a centralized narrative does not exist. He goes on to say the ideas of a special center in the brain is the most tenacious, bad idea, bedeviling our attempts to think about consciousness.

Quantum mechanics is a fundamental theory of physics which provides a description of physical properties of nature at the scale of atoms and subatomic particles. Quantum mechanics arose to describe nature smaller than the macroscopic (ordinary human) scale. Quantum mechanics has three key differences from classical mechanics 1) that quantities of a bound system, like energy and momentum are restricted to a discrete value, 2) objects can have characteristics of both a particle and a wave (wave-particle duality), and 3) the uncertainty principle that there are limits to how accurately a value can be predicted prior to measurement. The third difference is the principle of *indeterminacy.* This idea has become a key metaphor in the study of the mind as a complex system. The advent of the field of quantum mechanics brought with it new metaphors for thinking about the mind. For example, Lambert-Mogiliansky and Busemeyer (2012), colleagues of Nobel laureates in economic sciences, Kahneman and Tversky, formalized this situation in their paper *Quantum type indeterminacy in dynamic decision-making: Self-control through identity management*. Indeterminacy is like an ambiguous picture. For instance, the duck rabbit, in one way you look at it, it is clearly a duck, but by another, it is rabbit. The tricky part is it cannot be both at the same time. This situation is analogous to Neil Bohr's complementarity in quantum physics: objects sometimes have properties that appear contradictory. It is possible to switch between two sets of properties and see the duck and the rabbit, but, theoretically, not at the same time. Heisenberg (1942) and Bohr (1991) recognized the similarities between human societies and quantum mechanics. The fundamental similarity is that in both human societies and quantum mechanics: "the object of investigation cannot – always – be separated from the process of investigation (Lambert-Mogiliansky and Busemeyer, 2012 p. 98). For Lambert-Mogilansky and Busemeyer, this makes it fully justifiable to study human behavioral phenomena through the mathematical formalism of quantum mechanics. Such investigations hold out the potential to discuss the effects of multiple fields' effect on multiple selves within the individual that each are vying for their own utilities. In other words, we can discuss the limits imposed in multi-voiced situations, and theoretically, the interplay of multi-self-individuals' interactions in one field, and potentially multiple fields.

2.3 The Ants: A Complex System

Early work on social insects motivated work on agent-based modeling. Early work on agent-based modeling was inspired by the behavior of social insects (Resnick and Wilensky, 1991; Wilensky and Resnick, 1993, Langton, 1997). Ant behavior has inspired games. SimAnt (McCormick and Wright, 1991) is based on Hölldobler and Wilson's (1990) *The Ants*. The collective behavior of ants has been simulated using agent-based models many times. StarLogo was used to model the collective behavior of social insects (Resnick and Wilensky, 1992, 1993). Wilensky (1997a; 1990) modeled food source preferences resulting from pheromones as well as the formation of ant trails (Wilensky, 1997b). Bonabeau investigated the role of agent-based model in pattern formation (1997), and more broadly, looked at swarm intelligence (Bonabeau, Dorigo and Theraulaz 1999). Prat (2005) modeled collective nest selection of *Temnothorax albipennis* also using an agent-based model. Sumpter and Pratt's joint work explored the importance of group decision making with quorums (2009). Their work showed that when choosing a destination together, cooperation reduces the probability that an individual will suffer predation. Robinson, Ratnieks, and Holcombe (2008) used an agent-based model to explore attractive and repellent pheromones in pharaoh ants. Likewise, frameworks, such as Anthill, have been used to support the design, implementation, and evaluation of technical systems, such as peer-to-peer networks (Babaoglu, Meling, and Montresor, 2002). Their work drew on examples of complex adaptive systems to justify engineering and user applications because complex adaptive systems exhibit resilience, adaptation, and self-organization that are seen as valuable in social applications.

2.4 Complex Systems for Education: promise and challenges

While understanding science and society in the twenty-first century requires students to understand conceptual perspectives about complex systems that arose in the field of complexity (Bar-yam, 1999; Goldstone & Wilensky, 2008; Jacobson & Wilensky, 2006; Morin, 2008; Wilensky & Jacobson, 2014), a review by Perkins and Grotzer (2000b) found students tend to explain complex phenomena with simple causal explanations. There are well documented examples that people resist conceiving of the world in self-organizing terms (Resnick, 1994; Wilensky & Resnick, 1999).

Learning about complex systems can be difficult (Chi et al., 2012; Jacobson et al., 2011; Wilensky & Resnick, 1999). Researchers have documented the difficulties learners have in understanding complex systems. For examples, in interviews conducted in the 1980s and 1990s, Wilensky and Resnick describe people's resistance to ideas of explanations of self-organizing systems (Resnick, 1994; Resnick & Wilensky, 1992; Wilensky & Resnick, 1999).

Before proposing designs to ameliorate the difficulties, here I outline two threads: first, the promise scholars' see in teaching complexity, and second, I describe the scholar's views on the difficulties humans have in learning complexity.

2.4.1.1 Promise of Complex Systems Learning

Our world is full of complex interactions that self-organize. For instance, in the introduction to *An Introduction to Agent-Based Modeling* Wilensky and Rand (Wilensky & Rand, 2015) ask us to consider an ant. The ant wakes in the colony, goes outside, wanders randomly until it finds food. When she finds food, she leaves a trail as she returns the food to the colony, bringing energy back to the colony. When the next ant comes out, she follows the trail to food. As ant after ant repeats the loop of following simple rules the colony gains food. Additionally, a trail network emerges that the ants move along. Each ant, following simple rules leads to a complex, emergent pattern that feeds the colony without any central control. Resnick introduces his 1994 book, *Termites,* *Turtles and Traffic Jams,* by discussing the choreography of flocking birds. While many people assume bird flocks are led, he describes the coordination emerging. He asks, how did the movements become so orderly and synchronized? There is no leader-bird. Rather the flock is an example of what people call a self-organizing system. Each bird follows simple rules, reacting to movements of the birds nearby it—namely, distance and proximity functions. Orderly bird flocks emerge from these simple, local interactions. Bird flocks and ant colonies are not the only thing that follow these rules, minds, along with market economies, traffic jams and immune systems follow patterns not set by some centralized authority of a queen or president, but local interactions among several centralized components. In ant colonies, as ants forage for their food, the trail patterns are determined not by the organizational machinations of some Machiavellian queen, but the local interactions among thousands of workers. Patterns of traffic on the 5 in Los Angeles emerge from the interactions among drivers in individual cars making local decisions.

2.4.2 Difficulties in Learning Complex Systems

Starting with interviews that Wilensky and Resnick conducted in the 1980s and 1990s, they describe people's resistance to ideas of explanations of self-organizing systems (Resnick & Wilensky, 1993, 1998; Wilensky & Rand, 2015; Wilensky & Resnick, 1995, 1999). During the interviews, Wilensky and Resnick met with people from diverse backgrounds and asked them to explain emergent patterns. The results suggested a cognitive pattern called the deterministic centralized mindset (DC mindset). The pattern emerges from two empirical findings. First, people do not see the role of randomness in creating structures they observe in the world. Instead, the participants describe randomness as a destructive force to patterns, rather than a force that creates patterns. Second, subjects describe patterns that arise from the actions of a single centralized controller or leader, like a conductor in an orchestra. When asked how birds got into formation, interviewees typically responded the leader bird in front guides his lieutenants, or a mother bird in front tells her children what to do, or perhaps the biggest bird pushes back the air and the next strongest follows. Further, when asked to explain traffic jams, interviewees hypothesized that there was perhaps a broken bridge, radar trap, or even an accident that was causing the problem ahead. All these ways of understanding complex systems reflect a DC mindset where something is in control. Ideas revolve around leadership: a leader that orchestrates, agents moving into formation because of orders, or some specific centralized cause driving the effect. Counter to these intuitive understandings most traffic jams arise from random entry and exits of cars onto the freeway. These merges result in a statistical distribution of cars and speeds. Birds do not stay at the apex of a V formation; rather, the formation changes over time, when the current leader drops back and another bird flies in the front. The formation emerges.

Second, in further analysis Wilensky and Resnick identified key components of this DC mindset that pose obstacles to thinking — namely the need to think in levels. Emergent phenomena can be described at two levels: at the level of the individual and at the level of the system (such as the colony of ants). Many people fail to distinguish between these levels, and instead slip between levels to give properties from one level, such as the individual, to the entire colony. This ability makes people think the stability is in the flock, whereas they move in the formations, as individuals. This becomes described as a bias, these kinds of descriptions persist, even when the system is self-organizing.

This description persists, even when the system is self-organizing. Interviewees want to describe it in terms of leaders and seeds (Resnick, 1994). For example, Resnick writes of Marvin Minsky walking into his office at MIT and watching slime molds on a monitor and saying immediately, this is not self-organizing. There must be localized food that is drawing these slime molds together. Even when Resnick explained, Minsky persisted for some time. He also describes two high schoolers creation of ant cemeteries, and their resistance to the use of self-organization in modelling ant behavior.

These biases are not just for inexperienced learners. Wilensky and his colleagues (Jacobson & Wilensky, 2006; Sengupta & Wilensky, 2009) have shown that even expert researchers find emergent levels difficult to understand. In other words, it is not just a misconception of the scientifically naive. Instead, it seems to affect the thinking of nearly everyone.

Perkins and Graezer (2000b) argue that in contrast to everyday lived experiences of people, many scientific models involve more than causal explanations of ordinary events. These models introduce invisible entities like rule systems such as Ohm's law that govern the global behavior of systems, electrons, and large-scale patterns of action that are "emergent" from small-scale interactions, as with the gas laws. Wilensky and Resnick (1999) describe the difficulties people have in “thinking in levels,” exhibiting “levels confusion” and difficulties with distributed control and stochastic processes. Other researchers have noted that not only do learners have a hard time thinking across multiple levels such as disease of the whole body resulting from microscopic pathogens, but they tend not to think about phenomena such as the flow of ink dropped into water as the processes of collectives of agents interacting (Chi et al., 2012). If the glass of water changing color is explained by the individual parts of ink interacting with H2O, then the process becomes more intuitive. Coming from a non-constructionist background, Chi et al., (2012) argued that most people, are not familiar with ink particles. These two points of view have created two strands in the literature (Sengupta & Wilensky, 2009). Next, I will describe each of these perspectives.

2.4.2.1 Strand 1: Emergent vs Process Schemas

For Chi et al. (2012) processes, can be categorized in two ways – sequential or emergent.

Sequential processes can be subdivided into a sequence of events, like an assembly line where the metal is rolled, pressed, stamped, and smoothed before being turned into cans and filled with tomato sauce. This is a process with multiple agents: the machines, their operators, and a manager. It makes sense to speak of sequential processes resulting from a single agent's actions. For example, we could say the manager increased the efficiency of the assembly process by increasing the speed of the conveyor belt. Even though all the agents participated in the process, we can focus on the goal-setting manager's actions to account for the change. As a result, Chi et al. (2012) say one can give special controlling status to the agent that caused the change in a sequential process. This means if someone thinks of a process, like ant colonies, as sequential, people oftentimes assume interactions at the agent level are done to reach the goal of the higher level, whether that be producing cans of tomato soup at the managers, or the queen ant giving orders. The causal mechanism that leads to the result comes from the summing of the assembly line's outcomes directed by the manager. Many people confuse this control when thinking about emergent processes.

Emergent processes, like ants searching for food, marching in orderly paths, or getting stuck in a doorway are slightly different. These processes result from each ant taking actions and some of these actions are random. The results emerge from the repetition of the action, but no agent is in control. These processes are encountered in school standards such as osmosis and diffusion, electrical current, and buoyancy.

Chi and her colleagues argued that misconceptions about emergent phenomena in general result from an incommensurability (Chi, 2005; Slotta & Chi, 2006) or incompatibility (Chi et al., 1994) of amateur and expert schemas. Chi et al. (2012) argued that all processes can be categorized into two types: sequential and emergent schemas. This stands in contrast to the constructionist perspective led by Wilensky and colleagues, who argue that the difference between amateurs and "experts" may not be so distant for their explanations of emergent phenomena. Moreover, researchers—even experts—sometimes find emergent levels counterintuitive and difficult to understand, which I will cover next.

2.4.2.2 Strand 2: Difficulties Learning Complexity, Regardless of Experience

The second strand argues amateurs and experts can both find it hard to reason about these complex emergent systems. As Resnick (Resnick, 1994) would argue, this "centralized mindset" is one of the major impediments to teaching complexity. Wilensky and colleagues argue that the difference between amateurs and "experts" may not be so distant for their explanations of emergent phenomena. Moreover, researchers—even experts—sometimes find emergent levels counterintuitive and difficult to understand. When a person sees a pattern in the world, they often assume centralized control, even when it does not exist (Wilensky and Resnick, 1999). The assumption leads people to tend to look only for sequential chains of causes and effects even when systemic, emergent patterns are at play (Perkins & Grotzer, 2000b). Jacobson (2001) examined expert–novice comparisons of complex systems thinking by examining interview transcripts of undergraduate students and complex systems experts. He found that students favored simple causality, central control, and predictability. In contrast, complex systems experts' explanations demonstrated decentralized thinking, multiple causes, and the use of stochastic and equilibration processes. This data indicates a bias in reasoning about complex situations towards simple causality, but that with experience, some people can learn to view situations from a decentralized mindset.

Thus, in the two strands, we see a disagreement over whether people can learn a to think in complex system terms. Regardless, it is crucial for people to understand complexity to understand important issues. According to current research, species are emergent entities of the processes of evolution over time (Perkins & Grotzer, 2000b). Emergent processes also help explain serious issues for humanity like climate change, nuclear arms proliferation, and income inequality.

Previous research suggests ants are an exemplary way for people to come to understand complex systems (K. Martin, Horn, et al., 2019; Wilensky, 1997b). In ant colonies, we can see the emergence of ants searching for food. They wander somewhat randomly until they come across a morsel. They pick it up and bring it back to the colony, leaving a pheromone trail as they do. Then other colony mates move toward the food, filing in a straight line along the scent trail made by each ant leaving pheromone as it returns. When the flow of ants returning from the food fades, the ants discover and construct new trails through their random walks and feedback. The pheromone trail is a self-organizing process that organizes ants without a central agent controlling their actions. This process requires randomness. The organization is hard for some people to accept because of a "deterministic mindset" in line with Einstein's erroneous statement that "God doesn't play dice" (as in Wilensky, 1997c). In fact, randomness is central to many phenomena. As Taylor (2003) writes, "disorder does not simply destroy order, structure, and organization, bus is also a condition of their formation and reformation." As people come to understand the randomness of how ants find food and lay trails, they form an intuitive understanding of complex environments.

2.4.3 Constructionist Learning Environments to Teach Complexity

Learning about complex systems can be difficult (Wilensky and Resnick, 1999; Chi, Roscoe, Roy and Chase, 2012; Jacobson, 2011). Computer mediated learning may help with complexity learning. We designed a museum game to teach complexity in short interactions in a museum. Inserting the complex system content into video games may be a useful design because 97% of children in the United States play video games at some point in their lives (Lenhart et al., 2008).

One type of game, constructionist video games, seems particularly promising. By encouraging open-ended engagement and exploration, games can support learning across a wide variety of topics and contexts, providing a powerful way for learners to construct new knowledge and understanding (Kafai, 2006; Kafai & Resnick, 1996; Papert & Harel, 1991a). Constructionist learning games try to strike a balance between open-ended play and targeted treatment of learning content through providing learners objects to think with (Holbert & Wilensky, 2019; Weintrop et al., 2012). Constructionist video games employ traditional game structures infused with constructionist ideals to create a game experience that both encourages exploration and delivers targeted content. These games mediate open-ended learning.

These difficulties motivate the design of Ant Adaptation to facilitate learning about complex systems with social insects. In this proposal, we seek to study the use of an agent-based model, Ant Adaptation, to improve learning about complexity.

2.4.3.1 Learning from Insect Systems

In the earlier work, I showed that some participants can shift schemas through the use of complex systems models (K. Martin, Horn, et al., 2019), lending weight to Wilensky and colleagues strand. In the first two chapters, (K. Martin et al., 2020; K. Martin, Horn, et al., 2019), I presented visitors with the opportunity to create an interactive digital ant ecosystem by exploring a microworld (Edwards, 1995; Papert, 1980). Doing so allowed learners to better understand the world of ants, a complex system, by immersion. The work done promises to overcome the twin-challenges of complex systems education, including: level slippage (Wilensky & Resnick, 1999) and not accounting for the individual particles in a system (Chi et al., 2012). This work builds on past work with agent-based models of social insects microworlds that have been effectively used for classroom teaching of complex system ideas. Examples of past work includes *Ant Food Grab,* a yearlong block-based programming curriculum with ants (K. Martin et al., 2018; Sengupta et al., 2015), and *Beesmart*, a curriculum about the hive-finding behavior of honeybees (Guo & Wilensky, 2014). These examples highlight that observing and/or interacting with insect microworlds allow students to construct their own understandings of complex systems through exploring and adapting models in a classroom setting. Like other complex systems interventions, the students build their understanding of complex systems in natural systems. The literature has depicted learning through diverse simulations of natural systems, including: particle simulation of water transitioning from ice to steam in mixed reality (DeLiema et al., 2019), bees (Danish et al., 2011), material science (Blikstein & Wilensky, 2009), multi-level modeling (Wilensky & Reisman, 2006), electricity (Sengupta & Wilensky, 2009), and evolution (Wagh et al., 2017).

2.5 Constructionist Microworlds

Constructionist microworlds (Edwards, 1995; Papert, 1980) have often been long term explorations of motivating problems. In the museum setting, however, we needed to deploy a system to explore, but that provides an experience in as little as two minutes, which is the average playtime in museums. This design constraint motivated some modifications in my use of constructionism to teach complexity. With *Ant Adaptation*, we examined the type of learning that occurs between a user, and a complex system model. In previous work, we provided an example of a learning environment that scaffolds individual meaning-making through touch, discussion and play for museum patrons (K. Martin, Horn, et al., 2019). In the current work, we will investigate the roles of engagement and concept elaboration in learning among groups of museum patrons.

Building on Wilensky and Resnick's interviews in the 1980s and 1990s, in the book *Turtles, Termites and Traffic Jams*, Resnick (1994) outlines a three good points that motivate this paper. First, he develops new tools—at the time using StarLogo to investigate ants, termites, traffic jams and many other self-organizing phenomena. Second, he puts forward an argument that the best way to promote thinking in a decentralized way is having learners construct and play with such systems. Third, to make that possible he argues as did Feurzeig and Papert (Feurzeig et al., 1970; Turkle & Papert, 1992) for programming languages with low floors and high ceilings, such as Logo and NetLogo (Wilensky, 1999b), where people can simply program the actions of thousands of computational objects to follow simple rules to create elegant, emergent patterns. The three parts together are a powerful restructuration (Wilensky and Papert, 2010) of learning in this topic which I employ in this project.

2.6 Restructurations

One recent outgrowth of constructionism is Wilensky and Papert's theory of restructuration (Wilensky, 2020; Wilensky & Papert, 2010). A restructuration (Wilensky & Papert, 2010) is a new way of doing something that affords additional abilities. For example, Wilensky and Papert discuss the advent of Hindu-Arabic numerals. In that transition the Hindu-Arabic numerals supplanted the romans numerals, first in accounting, but later through computation because they were easier to perform operations with. Wilensky and Papert position the process as follows: "Hindu-Arabic numerals were not invented with an educational intent. But they could have been, and that allows us to show the need for a new branch of the learning sciences with the mission of understanding, facilitating and even designing shifts similar to the shift from Roman to Hindu-Arabic numerals" (2010, p. 2). We could call the new branch of learning sciences 'the science of designing new better ways of thinking embedded in our tools.' But Wilensky and Papert present a better name—namely structuration. They define the term as "the encoding of the knowledge in a domain as a function of the representational infrastructure used to express the knowledge" (2010, p. 2).

In a restructuration, Wilensky and Papert argue there are 5 properties researchers should attend to: power properties, cognitive properties, affective properties, social properties, and diversity properties.

Powerful Ideas: A restructuration must afford what could be done before, but for it to spread, it also must be able to do more. In the case of Hindu-Arabic numerals this is the power of multiplication and division massively expanded. When we open up education for students to explore computational microworlds on their own recognizance, this is a powerful example of what could be done before — understanding physics and biology for instance — but with a huge new set of ideas in complexity, emergence and self-organization. This notion undergirds the complex systems as it lowers the floor for exploration of these multiagent, self-organizing systems.

Cognitive properties: a successful restructuration is more easily learned, while still providing the augmenting power of the old tools. In the current case, this is the idea of understanding ecology, while adding the ability to explore the role of emergence and complexity in how ant colonies ecology sustains.

Affective properties: a new structure can be more or less engaging. Computation media offer one example of more engaging properties. One question is how to measure this engagement, a field the second component of this paper takes a deep dive into.

Social properties: As a structuration is more shareable, it spreads through a group more readily. In the case of Hindu-Arabic numerals, their use in business made them highly sharable. In the case of Ant Adaptation, I have tried to design it to encourage folks to want to draw friends and family into its use. The properties of this sharing are key to the restructuration's spread.

Diversity Properties: with the power of design and iteration, a restructuration spreads, and as it does so, it matches people's styles of use ever more closely.

In short, Wilensky and Papert argue novel restructurations have powerful impact on learners. They describe the move from Roman numerals to Arabic numbers, operations such as multiplication and division became significantly easier because of the new representation's affordances. Importantly, this power comes as the community wide literacy with the tool, reduces the difficulty of performing certain sums. This results from the fact that the various new representations are the focus of both internal reflection and external action that foster shared meaning, positively mediating groups' sense making about actions taken in the world. These restructurations can move science education past rote memorization of the traditionally identified core elements of disciplines to study (Clark et al., 2009). They argue well-designed digital games and simulations help learners build accurate intuitive understandings of the concepts and processes embedded in the games "due to the situated and enacted nature of good game play" (p. 3). These types of games provide excellent objects to use to think about scientific concepts and processes (Holbert & Wilensky, 2019). It is a version of these deep restructurations that I study in this paper, as they engage learners and shift their thinking in interaction with each other, and Ant Adaptation.

This restructured education will better align with students of the future. Papert (1993) argued that the advent of the restructuring of digital worlds that children can explore in computers will create less patient, accepting students. "Children who grow up with the opportunity to explore the jungles and the cities and the deep oceans and ancient myths and outer space will be even less likely than the players of video games to sit quietly through anything even vaguely resembling the elementary-school curriculum as we have known it up to now" (p. 9). As a result, Ant Adaptation restructurates the learning of complex systems to make it more intuitive, play-based, and occurring in a group.

To describe the system Wilensky and Papert introduce three examples of restructurations: the turtle, agent-based models, and the tick model.

2.6.1 Agent Based Models

Agent-based modeling (ABM) is a powerful structuration that has emerged from complex systems theory (Epstein & Axtell, 1996; Grimm & Railsback, 2005; Wilensky & Resnick, 1999). Agent-based modeling use simple computational rules as the fundamental modeling elements, in contrast to more traditional mathematical modeling. When modeling with equations we observe a phenomenon and try to fashion an equation that fits the observed data. In the Lotka-Voltera equations, used to model the change in predator and prey population levels over time (Hastings, 1997), the core elements of the model are variables that refer to population level parameters. In the Lotka-Volterra equations, the core elements are H, the population level of the hare prey; L, the population level of the lynx predators; and K, the interaction constant that describes the average predation. To understand the state of the system at a future time *t*, you solve the equations at time *t*. In contrast, as shown in Figure 1, in the agent-based modeling game, the core elements are computational objects or "agents" that represent individual agents, such as ants. Each of these agents has state variables that describe its particular state including, variables such as age, energy level, hunger, or location. The behavior of the agents is determined by the computational rules that tell each agent what to do at each "tick" of a clock. As the modeler, we frame the rules from the agent's perspective. For example, if the agent is an ant, the rules might say: each tick, move a step in the direction you are headed; if you find food, pick it up, if you have food, return it to the nest, if you are returning to the nest, leave a trail other ants will follow, if not turn randomly, etc. To determine the state of the system at future time *t*, you run the system for *t* clock ticks. This is a set of simple rules, but it also is a complex system.

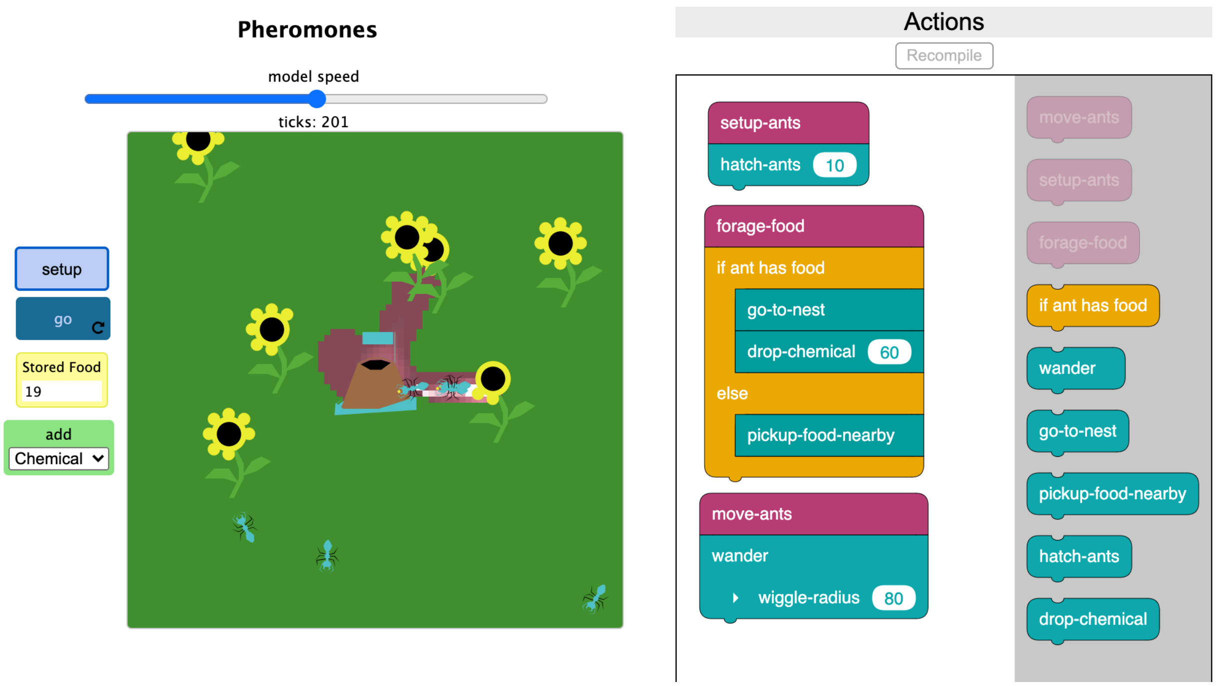

Figure 3: Ants collecting food in Ant Adaptation, following the simple rules of forage-food and move-ants. try it yourself: https://antomology.netlify.app/

3 Museum Studies

Next, I review literature of museums studies to provide context to the project, as well as support the use of constructivist dialogue mapping to analyze constructionist learning environments. CDM builds on the museum studies' perspective of elaboration as learning (Leinhardt & Crowley, 1998) during active prolonged engagement (Humphrey & Gutwill, 2005) as groups interact around exhibits (Vom Lehn et al., 2001). This literature review motivates the design and analysis of Ant Adaptation.

Since Leinhardt and Crowley (Leinhardt et al., 2003; Leinhardt & Crowley, 1998) presented their vision for studying elaboration as learning, museum studies has been looking for a way to operationalize studying elaboration. The framework makes it clear that it is essential to attend to the richness of visitor conversation around an exhibit. In this paper we offer a tenable means to track these elaborations using constructivist dialogue mapping (CDM) (K. Martin et al., 2020) and use it to compare between two museum experiences. To ground the work, I review literature on museum studies that looks at participant engagement with exhibits. First, I review Active Prolonged Engagement (APE) (Humphrey & Gutwill, 2005), which allows museum studies to modify what it means to be engaged in museums. Second, I look into the origins of measuring learning in museums through tracking elaborations (Leinhardt & Crowley, 1998). Using CDM, I then operationalize these two frameworks to present findings from exhibits that show divergent interaction types: deep behavior observation and wide recognition. Depending on what we aim to study, we should design our exhibits accordingly.

3.1.1 Active Prolonged Engagement

Thomas Humphrey and colleagues argue that hands-on interaction with science museum exhibits is not the same as active, prolonged engagement (Humphrey & Gutwill, 2005). They argue that designers should look beyond attracting and initially engaging visitors to focus more on creating situations that invite prolonged exploration. APE was built on the hands-on revolution stemming from the impact of constructivism on informal education. Thomas Humphrey, a researcher at the Exploratorium in San Francisco, outlines how from the Exploratorium's early days, as part of the hands-on revolution of the 1970s, exhibits were labeled with signs of how to engage with them, with signs such as “What to notice?”, “What to do next?”, and “What is in fact going on?” When the experiences follow this formula, the Exploratorium calls it a planned discovery (PD). Building on the PD hands-on exhibit, Humphrey outlined the APE framework to aggressively promote self-discovery and to turn over authority to museum patrons to grant epistemic agency. APE uses the following five steps to achieve these twin goals: (1) minimize instruction and explanation, (2) encourage visitors to initiate play, (3) encourage experimentation, (4) encourage observation, and (5) encourage speculation.

Subsequently, Humphrey's team proposed their first principle: museums should shift authority from the museum to their visitors. This principle would, in effect, allow the user to try all the gears in the activity and sequence their own action. This principle was a part of the pedagogical revolution of constructivism.

Humphrey wondered whether APE required focused hypothesis testing, visitors engaging in inquiry like behavior including informal hypothesis testing, or if it could involve more free-form exploration at exhibits, which is a question I also am investigating. Among the patterns of visitor interaction identified by the APE team, two were directly relevant to our current research: *investigative* interactions—analytically well-formed lines of thought when investigating the system—and *exploratory* interactions—visceral chains of action and attempts to arrive at an aesthetic or Gestalt result.

The team went on to find that visitors asked more questions at APE exhibits than they do at normal exhibits, though the average number of questions per minutes was the same. This means that the visitors engage longer at APE exhibits as compared to normal exhibits.

The kinds of questions they asked also changed. They asked more action and exploration questions, such as "Can you turn it faster?", "What if we connect this to this?", "How does that work?" or, "Why did that happen?" Additionally, they answered the questions using the exhibit or by asking each other as opposed to referencing the signs or talking to museum staff. They also engaged longer—three and a half minutes as opposed to one minute and six seconds—which led to more questions asked. Finally, they left most often due to external factors such as their family pulling them away, instead of completing everything in the exhibit, like in PD.

3.1.2 Learning as Elaboration

Learning Conversations in Museums (Leinhardt et al., 2003) proposes means to study how learning actually occurs in museums. Their work in museum learning motivates their attempt to solve a core issue: what constitutes learning in museum research, and specifically what terms or actions do we track to measure it. From these problems, the authors propose three outcomes, the last of which is important for CDM. They argue that their approach will provide a novel, stable, and disseminatable methodology to conceptualize, collect and analyze conversations as a process and as an outcome of learning in the museum context. The qualitative method I use in this study, constructive dialogue mapping, has the same aim and could solve core operations issues with the original framework, which I will demonstrate in this paper.

The core issue surfaced by Leinhardt and Crowley, how to broaden the definition of learning beyond idiosyncratic and implicit definitions of learning, created a healthy dialogue on the subject (Leinhardt & Crowley, 1998, p. 3). This dialogue pivoted the definition of learning to be dynamic for each study rather than a static definition, as they see static definitions as taking insufficient account of the social contexts of learning. They make clear they are not arguing for a major debate over what constitutes learning, but instead they argue for accountability and clarity in whatever definition of learning is used in a particular study. In this paper we present an operationalizable definition of learning in the museum context based on the CDM method of analysis. After introducing the problems and outlining them, Leinhardt and Crowley (1998) describe a pragmatic approach to study learning in museums: define learning as Conversational Elaboration. For their pragmatic approach and operational definition of learning they chose how visitors elaborate conversations.

They focus on conversational elaboration because it is both a naturally occurring part of the museum experience while also being a product of the experience. By elaborations they mean a particular kind of talk that occurs within a group both during, and surrounding a museum visit. Conversation is important because it is a reflection of the “inter-twining of social with cultural processes” (White, 1995, p. 1). Sociocultural theories of voicing (Rogoff, 2003; J. V. Wertsch, 1997) emphasize that inter-twining of voices in the Bakhtinian sense (1981) as the primary activity through which knowledge is constructed and appropriated across people and generations. This process is important as conversation is a natural process and consequence of an enjoyable, shared experience of visiting a museum as a group.

The approach particularly looks at four process that groups will undertake. If there is learning, after an interaction with an exhibit, a Coherent Conversation Group (CCG) will:

1. Refer to more items,
2. Include more greater detail about those items,
3. Synthesize elements to elements from their prior knowledge, and
4. Increase the level of analysis of the phenomena that they share about.

This approach moves away from focusing on the amount of talk or types of talk while building a strong foundation between amount, type, and the process of learning. The big takeaway here is that the part of the field has argued that it is essential to attend to visitor talk as both a process of learning and a learning outcome. This motivates CDM as a methodological tool that I use in this comparative study. I expand on the work of Crowley and Leinhardt while offering new tools for researchers and practitioners working in museums.

3.1.3 Design with Alternative Learning Technologies for Informal Learning

Different exhibit designs elicit different interactions. Alternative design technologies include: augmented reality (Sutherland, 1965; Yoon et al., 2012), immersive technologies (Snibbe & Raffle, 2009), tangible user interfaces and touch interfaces Horn et al. (In press), and gigapixel imaging to bring science to the public (Louw & Crowley, 2013).

For this paper I used Ant Adaptation. Due to Covid-19, for some of the interviews I had to change the activity for remote learning. In choosing this technology I needed to restructure the activity from the touch interaction in the first iteration, for remote learning, using a mouse. These adaptations were technically straightforward, but could impact the use of Ant Adaptation as result of allowing all users to reach all of the screen, as well as giving one user, that holds the mouse, outsized authority.

Because positive affect and engagement is associated with better learning in museums Horn et al. (2016), I am proposing to study positive affect more closely, I will review this literature next to motivate the use of affective measures of mood with this refactored interaction with Ant Adaptation.

4 Affective Engagement

Affective computing allows for characterization of users' affective states, which can be used for analysis, as in this chapter, or to allow the computer to interact based on a user's affect. Technologies are deployed both to sense affect and impact users' affect. For example, Mattel released a short-lived project, *Hello Barbie.* Produced in collaboration with *ToyTalk*, a San Francisco-based company, the project worked toward the fulfillment of an ancient dream, toys that can talk to us (Vlahos, 2015). In the past, the idea of talking toys had been a bit of a party trick; toys used recordings to fool the user. But that is all changing. In the project, Mattel released an affect detecting, speaking Barbie. The project was built on the past decade of breakthroughs in artificial intelligence and speech recognition that gave devices around us—such as smartphones, computers, and toys—the ability to engage in conversation to generate intelligent responses to users. As the technology improves, it may become the primary way people engage computers (Vlahos, 2015). The success of the *Echo* and *Alexa* augur this future. Though it did not last long, *Hello Barbie* was discontinued only months after release with 51% of reviews on Amazon at 1 Star. The main complaint was that the technology was not ready: the battery would go flat, or the Barbie would not understand the child. Nonetheless, the promise, and perils of the technology persist. The technology was specifically designed to appeal to children by creating a friend that remembers and gets to know the user. And while the implementation is not their just yet, affective agents and detection are getting better.

These systems are not just theoretical, there have been practical means to use affect detection. For example, designers have used it to alleviate user frustration increasing the probability of users liking software agents (Bickmore & Schulman, 2007) as *Hello Barbie*. Moreover, affect detection also assists in measuring voters' candidate preference based on affective responses to election debates (McDuff et al., 2013). Furthermore, as outlined in the section below on affect and learning, affective states pathways, especially the engagement-confusion-delight-engagement pathway, have been hypothesized to facilitate advanced problem solving if the confusion state does not overwhelm the learning (D'Mello & Graesser, 2012a). Andres and colleagues (2019) found while researching informal learning with *Betty's Brain* that the only emotional pathway associated with learning gains was sustained delight. Horn et al. (2016) found post-test gains associated with positive affect words like "*wow*" and "*cool*" after a museum exhibit. Each of these uses points to greater integration of affect detection to improve design, interactivity, and analytics in learning technology.

4.1.1 History of Affective Research

Since ancient times, emotion has been researched (Healey, 2014). In ancient China, the Yellow Emperor, the source for traditional Chinese medicine for more than two millennia, argued emotion resided in the body, and excesses of emotion could damage a person's life-energy. In ancient

Greece the physician Hippocrates theorized that the body was composed of four humors. These humors drove a person's physiological and behavioral patterns, and imbalances could lead to deleterious outcomes. Imbalances would lead to choleric, melancholic, phlegmatic, or sanguine outcomes in the blood, respectively. Aristotle also had a physiological view of emotion. He viewed them as "passions", comparable to physical states like hunger, thirst, or desire. These ideas still influence our thinking about emotion (Healey, 2014).

Further, moving into modern times, Hans Eysenck (1947) cited humors as the inspiration for his dimensions of personality such as neuroticism. Additionally, William James (1893), the first theorist of modern times to put forward a theory of emotion, viewed the physical response as primary in the feeling of emotion. James believed the stimulus triggered the activity or the response in the autonomic nervous system (ANS), which caused an emotional response in the brain. For instance, he would argue people feel sad because they frown or cry. Similarly, Carl Lange proposed a theory that was very similar to James, and so this theory has become known as the James-Lang theory of emotion. Continuing his earlier work, James described the physiological activities that accompany emotion. Simultaneously, Charles Darwin also started documenting the observable physiological responses in both animals and people (1872). He focused on fear reactions and used it to classify and differentiate species based on affective states. These descriptions of physiological states and patterns are the theoretical underpinning for using physiological signs and patterns in order to recognize emotion and affective computing, which I use to examine how people respond while the playing a simulation.

Walter Cannon, a Harvard professor, significantly criticized the James-Lange theory along five lines, including that the autonomic nervous system or “visceral change” (1927, p. 112) was too slow and nonspecific to be unique to each emotion. Therefore, Cannon argued emotion must primarily be a cognitive event (Cannon, 1927).

Building on mounting evidence for a purely physiological or cognitive source for emotion, Schachter (1964) spelled out the implications of a cognitive-physiological formulation of emotion. Schachter (1964) proposed a compromise between James-Lange and the Cannon theories, i.e., his *two-factor* theory of emotion. In his studies, he injected epinephrine into subjects, but was not able to detect emotion after the introduction of this chemical. However, when he added a situational context, such as scaring the person, he could get an enhanced response after injecting epinephrine. Therefore, he concludes feelings such as fear, anger, or happiness have a physiological part. In other words, physiology was part of emotional experience, but emotion is made up of the combination between physiological changes and cognitive interpretations. Though this two-factor theory was important, Schachter informed but did not entirely explain complex interactions between physiological and cognitive influences of emotion. Physiological responses to emotion are an interaction of cognitive responses and physiological responses that create complex interactions we normally call emotions. Notably one major criticism was that epinephrine is too coarse of a measure to study this complex (Zajonc, 1994).

4.1.2 Affect and Intelligence

In her seminal work, Rosalind Picard, an MIT researcher, studies the promise, challenges, and potential reward of augmenting artificial intelligence with a core component of human intelligence—namely, emotion (Picard, 2000). The ability to recognize emotion is among the key aspects of emotional intelligence, which is a facet of human intelligence. Furthermore, emotion plays a key aspect in perception (Picard, 2000). There are many examples that point to an intervening role for emotions in perception. Studies show how mood influences participants' perception of ambiguous stimuli. As shown in Figure 1, this results from the fact that brains have two perception pathways, quick and dirty stimulus processing through the limbic system, and a slower, more accurate system through the cortex (Picard, 2000). The interaction of the two systems drives intelligent systems.

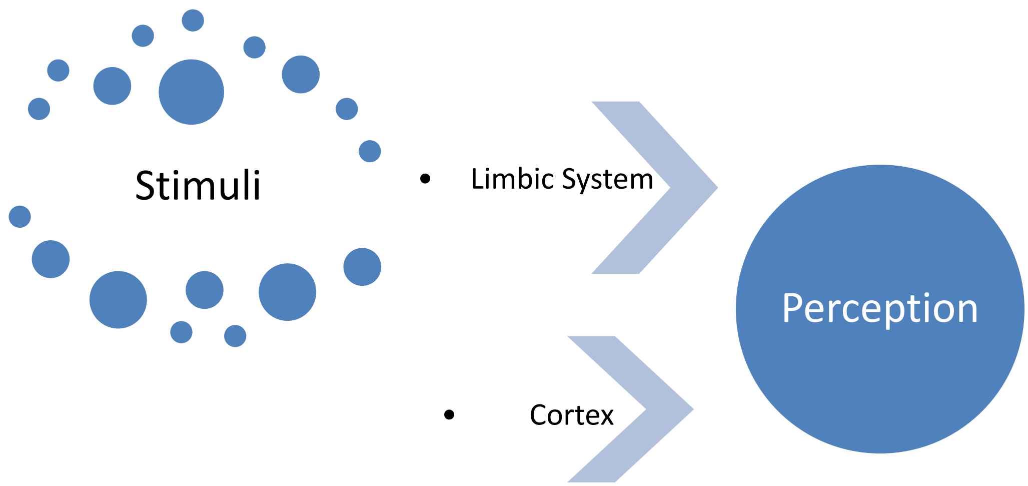

Figure 1: Brain has two pathways for perception, Limbic and Cortex.

This dual role inserts passions into rationality. Picard is not disturbed by this paradox, instead, she fully embraces it: “Even reasonable behavior is neurobiologically directed by these so-called passions” (Picard, 2000, p. 8). As a result, we need to understand affect’s role in that equation in order to understand noise in human decision making (Kahneman et al., 2021), or what is more, build artificial intelligence (Picard, 2000), or study learning (Calvo & D’Mello, 2011; D’Mello & Graesser, 2012a; Singh et al., 2002).[15] This suggests my research questions that explore the role of affect in learning and engagement in this chapter.

How do researchers track that? To understand the connection, they need to track physical

---

[15] See *Descartes Error,* Damasio, 1994, Thaler and Ganser’s *Misbehaving* (2015) for additional powerful arguments of the role of emotion in rational decision making.

responses. The connection between limbic and cortex thought modulation, such as facial expressions, posture, and voice inflection, are physical means by which an emotional state is expressed. Humans primarily communicate emotion through these embodied forms (Picard, 2000). Ekman (Ekman, 1992; Ekman et al., 1993) argues “basic” emotions have particular musculature, movement pathways and social display rules associated with them. His theory led him to innovate a facial action coding system. While recent research (Barrett, 2019) has put the one-to-one mapping between emotions and affect display in doubt, how humans recognize emotion in others is still primarily through physiological communication. As a result, researchers can measure mood by measuring physical responses. Mood, or affective state, are associated with processes of learning regulation, which I will turn to next.

4.1.3 Affect and Learning

Increasingly, the affective processes associated with learning are of most interest for researchers who are trying to understand how students regulate their learning processes (Calvo & D’Mello, 2011; D’Mello & Graesser, 2012a; Singh et al., 2002). There are affect-sensitive or affect-aware educational systems that dynamically tailor their instructional activities based on student affect (D’Mello & Graesser, 2014b). Reviewing 24 studies having 1,740 participants, D’Mello and Graesser’s take-home message was that happiness, engagement/flow, confusion, boredom, curiosity, and frustration were the major affective states shown by learners (2014). These states are good candidates for modeling affect in learning. Furthermore, happiness was the only basic emotion to feature in education environments frequently. They found this not entirely surprising because there is inadequate theoretical justification to expect notable levels of some of the basic emotions, such as disgust and fear, during short one-on-one learning sessions with computers. Several researchers have studied how students cycle between emotions during learning (Baker et al., 2007; McQuiggan et al., 2010; Ocumpaugh et al., 2017). In their influential 2012 work, D'Mello and Graesser argued that flow states, punctuated by confusion, may contribute and assist in complex problem solving (D'Mello & Graesser, 2012a). Specifically, they hypothesized that there are two main pathways that students follow when transitioning between affective states: one that promotes learning (the engagement-confusion-delight-engagement pathway) and one that reduces learning (the engagement-confusion-frustration-boredom pathway). In should be noted that while educational research has explored the role of student's affective states in learning, they may not have captured the frequency or meaning of transitions between states (Andres et al., 2019). As a result, Andres' colleagues expand the number of pathways studied. Their results suggest that increases in the frequency of two affective state patterns, engagement-confusion-delight-engagement, and sustained delight, relate to an increase in post-test scores. Their results should be interpreted carefully because they used non-sensor-based detection methods of affect detection. Nonetheless, the results are promising as they indicate that delight, and the affective states occurring around it are related to learning measured on post-tests. Curiously, the only pathway associated with learning gains was sustained delight (Andres et al., 2019). This finding is interesting as it connects to the museum research that positive affect words are associated with post-test gains (Horn et al., 2016). Even though they are relatively infrequent, the two hypothesized pathways from D'Mello and Graesser (2012) appear to influence knowledge measures. Therefore, the evidence supports the hypothesis that moments of high stimulation, especially of delight, would correlate with gains on post-tests, because high stimulation is implicated in memory formation (McCaugh, 2004). In short, affect plays a role in learning.

4.1.4 Measuring Mood

To study these physiological displays research needs to choose bio-markers. Creating a system to measure physiological signals is the first step to measure and automatically recognize physiological patterns associated with emotion. This step will be crucial to measure affective signals during use of *Ant Adaptation*. Consequently, I will review what bodily signals to measure. In the widest view, all bodily signals could be viewed as physiological signals, including brain signals, voice patterns, facial expressions, or body chemistry. While the descriptions of physiological responds that Darwin, James, and Lange described make sense to humans, computers need quantifiable metrics to detect features for their analyses. As a result, affective computing researchers use digital recording devices and electronic sensors to calculate such features as heart rate acceleration and skin conductivity metrics to classify emotions (Ekman et al., 1983; Levenson, 1992). However, all the nuance of emotion may not be captured with these sensors.

Two of the most promising technologies for examining human response using quantifiable metrics are automated facial analysis (AFA (Cohn & De la Torre, 2014) and skin conductance (Healey, 2014). I will review AFA below, as that's the measure I use in this paper.

Emotions are degenerate, meaning they have more than one bio-marker to determine them (Barrett, 2019). As a result, researchers should look to more than one, time variant signals. There are several bio-markers which vary with time (Picard, 2000), including the following:

- Voice intonation
- Facial expression
- Posture
- Heart rate
- Diastolic and systolic blood pressure
- Pulse
- Pupillary dilation
- Respiration
- Skin Conductance and color
- Temperature

Physiological measures have long been associated with emotion (Healey, 2014). In this study, I will collect one of these bio-markers, facial expression (Cohn & De la Torre, 2014). The signal is important because they show high utility and signal fidelity when participants are active, such as when playing in a museum. Below, I will review the technique I use in this study, facial expression coding using AFA.

4.1.5 Facial Expression

Automated face analysis for affective computing is maturing (Cohn & De la Torre, 2014). The face conveys information about a person's age, background, sex, and identity (Bruce & Young, 1998; Darwin, 1872; Ekman & Rosenberg, 2005). Facial expression regulates face-to-face interaction, indicates interpersonal repulsion and attraction, reciprocity, and communicates subjective feelings between members of various cultures (Bråten, 2006; Fridlund, 1994; Tronick, 1989) among other details. As a result, behavior scientists have focused on the face (Cohn & De la Terre, 2014). Kanade (1974, 1977) began computer scientists' interest in the face as a potential biomarker. This interest led to work with computer vision to automatically analyze facial expressions (Ekman et al., 1993; Parke & Waters, 1996). Infrastructurally, the facial action coding system (FAC) has enabled this work (Ekman et al., 2002; Ekman & Friesen, 1978).

Recent research with facial expressions is struggling with several problems (Cohn and De la Torre, 2014), including the following: (1) Challenges with detection in naturalistic settings, (2) low base rates, (3) partial face occlusion, (4) pose variation, (5) rigid head motion, and (6) lip movement associated with speech. As researchers have overcome these challenges, automated face analysis (AFA) is beginning to realize the goals of advancing human understanding (Ekmam et al. 1992; Cohn De la Torre, 2014).

4.1.5.1 Manually Coding Faces

When researchers code faces, there are three common approaches: (1) message, (2) sign, and (3) dimensional measurements. Message-based measurement is where human coders make inferences about human emotion or affective state. This is based on Darwin's work (1872) where he coded 30 emotions. Ekman and colleagues refined these (Ekman & Friesen, 1978) to what they called basic emotions. An appealing assumption of this method is facial analysis provides a direct readout of emotion (Buck, 1984). But this appeal is problematic. Emotions are contextually dependent. There is a reason to be dubious of one-to-one mappings between expression and emotion (Cacioppo & Tassinary, 1990). Second, sign-based measurement starts with appearance. It is a descriptive sign-based system that you follow up with experimentally or observationally to discover, or assign, the relationship between signs and the emotion. The facial coding system (FAC) is the most popular (Cohn et al., 2007) (Cohn, Amador, Ekman 2007). Third, dimensional coding is a way to use human coders. While the first and second options for coding emphasize the differences between emotions such as joy and confusion, an alternative dimensional measurement emphasizes the similarities (Schlosberg, 1952, 1954). The approach posits that facial expressions can conform to a circular surface, where for instance pleasant-unpleasant (valence) and attention-rejection, are principal axes.

An attendant issue with each of these measurement methods is the code reliability (Cohn & De la Torre, 2014). Reliability is how repeatable, consistent and free from various errors a measurement is (P. Martin et al., 2007). This affects all measurement approaches. What researchers need is first *agreement* and *consistency*. Agreement means each coder assigns the same score to a measurement. Consistency means the degree to which ratings from different sources are proportional when expressed as deviations from their mean scores.

Combining these three methods, and upholding reliability, researchers can begin to use automated face analysis (AFA). AFA seeks to detect the movement of one or more of these measurement types, the process requires several steps, including: (1) face detection, (2) tracking, (3) feature extraction, (4) recognition, and (5) learning. In studying learning it is common to use supervised methods where label data of a limited set of observations trains a classifier for work on a larger corpus of data. There are several difficulties with AFA, including registration, in which non-frontal angles on cameras and head motion can cause serious problems. I will address this issue by mounting a camera in the center of the tabletop display to capture faces as they look down to play the game. Additionally, action units (AU), the movements of muscles on the face that create facial expressions, modify the appearance of the face. These changes interfere with computer vision's detection. Also, since facial actions are subtle, I need to be aware of these micro-actions. Furthermore, the non-standardized face shape of people can undermine generalizations of facial detection across people. Moreover, the temporal dynamics of facial motions need to be attended to. Additionally, variable classification can be overfit, where the training data, the people in our study, do not generalize to a larger population.

4.1.5.2 Automated Face Detection: Steps

The difficulties and challenges of automating facial detection have created a plethora of algorithms and applications. These steps and the work to avoid these difficulties has led to the exponential growth of solutions (De le Torre & Cohn, 2011). Regardless of the application, AFA begins with face detection. For frontal detection of faces, Viola and Jones' (2004) algorithm is the most used. There are two basic ways of going about this detection. The first one is sparse, where observers just look for minimal features such as eye contour. The second way is dense, which involves eye contours and other permanent facial features. This allows us to get good yaw, pitch, and roll for 3D face inference and detection (Cohn an De la Torre, 2014).

AFA has several parts (Cohn an De la Torre, 2014). First, we have face detection, where we use a camera or a set of cameras to detect the face through contour analysis. Then we have registration, where we register the face of the primary parts we are going to use to follow it through the video. This is followed by feature extraction. For instance, we could use geometric, the appearance, or the motion for the feature extraction. We then go through a process of data reduction and selection to reduce the dimensionality of our data. This is important because facial recognition data is often high-dimensional and through reduction, we can get more generalizable data. Researchers put this data through supervised learning usually. Supervised learning has input label data on a small set that we then use to detect similar features in a larger set of data. Though sometimes we do use unsupervised learning in these cases (*See*, De la Torre and Cohn, 2011). In state modeling, it used to be that neural networks were the primary means by which we analyze the data, but research has been moving towards support vector machines (SVMs) which the best open-source facial detection library, OpenFace, uses. A crucial part of these projects are reliable databases, there are two that stand out: DISFA (Mavadati et al., 2013) and the Binghamton-Pittsburgh 4D database (Zhang et al., 2013).

From this pipeline there have been many exemplary applications (Cohn and De la Torre, 2014). Applications have included action unit detection, intensity monitoring, checking for physical pain, depression and physiological distress, deception detection, interpersonal coordination, and expression transfer between subjects in a test. Other applications have included marketing, distinguishing subtle expressions, drowsy driver detection, autism aids and, important for our project, instructional technology (Craig et al., 2007; Whitehill et al., 2011). As a result, facial detection provides an interesting affective computing signal to look at subtle differences of affective state, and a powerful collection of research that maps affective state to emotional state in participants. As such, I use AFA to track expressions while participants engage.

4.1.6 Limits

The approach has powerful affordances for tracking learning and engagement, but it also unearths a challenge with incorporating AI in education research. AI systems have been shown to be less accurate at identifying the faces of dark-skinned women, to grant women lower credit-card limits than their husbands, and to erroneously predict Black defendants will commit future crimes more than whites. When applying these systems to learn about education, many of the solutions require having a human in the loop. As a result, a key domain of my research is investigating the diversity challenges posed by AI, while leveraging the technologies' benefit for education. The connection between affect and learning and the diversity challenges of applying AI to education will be my double-barreled research program going into my career.

4.1.6.1 Is Emotion the Same as the Signal

I think I should pause, however, and ask, can computers detect emotions? Lisa Feldman Barrett (2019), an esteemed researcher of human emotion at Northeastern University, works with the Harvard Medical School and MIT on the use of emotion detection in computing. Her work covers the excitement of the advent of emotion reading gadgets. She asks, "Is it possible for machines to read emotion in a body?" (2019). Companies are claiming to do it, founded on the idea that emotions can be read in bodies (Picard, 2000). She offers a warning and claims that those who say they are reading emotions successfully are exaggerating what they can accomplish. It is not that machines cannot interpret emotions, but that researchers misunderstand what emotions are and how they are generated. As a result, there is the chance to solve this fundamental problem by drawing on the unique advantages and breakthroughs of affective computing in bio-marker detection.

Affective computing, especially facial recognition, is based on modulations of sentic signals, such as the face or electrodermal conductance. These lead companies and researchers to say they detect happiness, when really, they are detecting smiles. There is a difference between our perception of facial movements versus the actual facial movements. The common stereotype is that a person will frown when sad, smile when happy, scowl when angry, etc... However, individuals are different. Some people may laugh when they are afraid, or cry when happy. In fact, only 30% scowl when they are angry (Barrett, 2019). This means scowling is a weak predictor of anger, but anger does not cause scowls. In other words, software can detect scowling, under ideal conditions, well. But that detection is different from detecting an emotion. The detection of emotions from looking at a face varies by culture, situations, and even among people in the same situation (Gendron et al., 2018). Grendron and colleagues argue while the scientific evidence suggests people sometimes smile when they are happy, frown when they are sad, and scowl when they are angry, the ways people communicate anger, disgust, fear, happiness, sadness, and surprise varies substantially across cultures, situations, and even across people within a single situation. Therefore, while it has been claimed configurations of facial movements can be universally recognized as emotional expressions because they provided information in situations critical to gene propagation for our hunting and gathering hominin ancestors, experiments have called this particular evolutionary hypothesis into doubt (Gendron et al., 2020).

One of the long-established ways to see how people predict emotions is having them read a story and then either pick a face that most aligns with the protagonist's feelings or to describe the feeling. This method may be flawed: the experimental method of reading a story and then assigning the emotion the characters have manufactures the evidence of universal emotions (Barrett, 2019). Barret suggests that if we want to analyze emotions in humans using computers, the evidence is not strong enough for real-world outcomes such as legal determinations. She describes how people move their faces in different ways with the same emotional state and that variation is the norm. So, you cannot simply measure a face to monitor human emotion. The faces we use for emotion detection are basically stereotypes, or “a science of emojis” (Barrett, 2019). We do not want to build AI analysis on stereotypes but on real, non-impoverished emotional episodes. A feeling like anger, or another emotion, is not static. Emotions are dynamic and use hundreds of bio-markers. Emotions are degenerate, meaning they have more than one bio-marker to determine them.

For instance, even if we can make determinations of facial expression, there are several complicating factors of emotion detection and their connection to displayed affect. Here are seven examples (Picard, 2000):

8. The intensity of emotion;
9. Type of emotion, i.e., there are many types of love (See *Symposium*);
10. How the state was induced, i.e., from a film or being a genuine experience, such as a conflict;
11. Social display rulers: whether a person was encouraged to express or suppress the emotion. In short, the context of the detection matters;
12. Hormone, diet, and medications;
13. Mood-state cues memories associated with that mood: positive moods tend to cue positive memories. Ledoux’s work shows emotion can hijack the cognitive centers of the brain; and
14. Person-independent emotion detection is difficult because people display emotions differently, “given that a particular emotion is felt, a variety of factors influence how the emotion is displayed” (Picard, 2000, p. 32).

Consequently, the use of affect detection should be used based on Negro Pontes’s suggestion which is in a person-dependent way, and I would add, a context dependent way. I can measure a person within a context to measure moments of high facial expression display. At the moment, these may not be generalizable beyond the individual. As such, I measure relative changes of stimulation.

4.1.7 Role of Affect in Human Choice

Emotion is critical for forming memory, attention, and rational decision making (Picard, 2014). Not only does mood influence judgment of seemingly objective perceptions, it also affects memory retrieval (Picard, 2000; McCaugh, 2004). Picard argues convincingly the following:

> Whether or not affective computing is an area in which you conduct research, you are using emotion when you choose to read this. You are involving your emotion system when you make a decision where to spend your time—when you act on what matters most to you. Affective computing researchers have a chance to elucidate how emotion works: how to build it, how to measure it, how to help people better communicate and understand it, how to use this knowledge to engineer smarter technology, and how to use it create experiences that improve lives.
>
> (Picard, 2014, p. 19)

Once a research team determines how to measure affect, they can use it to study its role in human behavior. Affect impacts decision making. Researchers have found several interesting impacts of affect on behavior. Positive affect facilitates creative problem solving (Isen et al., 1987). After attending a movie, moviegoers' decisions about people changes depending on if they watched a scary, thrilling, or comedic film (Forgas & Moylan, 1987). Good typography induces a good mood while reading leads to higher performance on relative subjective duration and certain cognitive tasks (Larson & Picard, 2005). Picard argues that emotions' influence on cognition may happen primarily "through emotion's influence on memory" (2000. p. 40). This would be exciting, as it would tie to McCaugh's work (2004) that strong emotional states lead to strong memory encoding. Work shows that negative emotions outweigh positive ones in affecting lives (Baumeister et al., 2001; Rozin & Royzman, 2001). They present evidence that bad parenting, or an insult, impacts behavior much more than good events. Trauma is one of the most impactful events in most people's lives and there is not even a word in the English language for the positive equivalent (Tierney & Baumeister, 2019). Further, positive modes have been shown to lead to more bad decision making (Kahneman et al., 2021).

5 Ant Adaptation

The goal of designing Ant Adaptation (K. Martin, Horn, et al., 2019; K. Martin & Wilensky, 2019) is to create an agent-based modeling environment for learning complex systems. Social insects provide a compelling context to explore and learn about complex systems. Social insects, like minds or the climate, are self-organizing systems. This means an overall order arises from local interactions between parts of an initially disordered system. This self-organization offers affordances to learning about complex systems, including reasoning about local actions leading to global outcomes. A branch of complex systems theory, Swarm intelligence (SI) in artificial life, shows simple creatures repeating simple rules can display surprising amounts of efficiency and complexity (Beekman et al., 2008; Bonabeau et al., 1999). Social insects, like ants, are excellent examples of these emergent principles (Resnick & Wilensky, 1992; Wilensky, 1997b; Wilensky & Rand, 2015). Though individual ants are quite limited in their functions, ant colonies construct overall order—including bridges, farms, highways of food, and information through local interaction—from an initially disordered system. Ants are good exemplars of the principles of self-organization. In this work, I have designed a digitally embedded curriculum based on ants' self-organizing behavior, and in this section, present the research that informs the design.

5.1 Learning about Complexity with Models and Games

Learning about complex systems can be difficult (Wilensky and Resnick, 1999; Chi, Roscoe, Roy and Chase, 2012; Jacobson, 2011). Computer mediated learning may help with complexity learning. I designed a museum game to teach emergent schemas in short interactions in a museum. Inserting the complex system content into video games may be a useful avenue of intervention because 97% of children play video games at some point in their lives (Lenhart et al., 2008). One type of game, constructionist video games, seems particularly promising. By encouraging open-ended engagement and exploration, games can support learning across a wide variety of topics and contexts, by providing a powerful way for learners to construct new knowledge and understanding (Kafai, 2006; Kafai & Resnick, 1996; Papert & Harel, 1991a). Constructionist learning games try to strike a balance between open-ended play and targeted treatment of learning content through providing learners objects to think with (Holbert & Wilensky, 2019; Weintrop et al., 2012). Constructionist video games employ traditional game structures infused with constructionist ideals to create a game experience that both encourages exploration and engages desired content (Egenfeldt-Nielsen, 2006). These games mediate open-ended learning.

I built Ant Adaptation to be engaging while building off the literature of designing digital interactives for museums to improve learning (K. Martin & Wilensky, 2019). Next, I explain how I measured engagement and learning around the use of Ant Adaptation.

6 Measures Collected

In this chapter, I use a mixed-methods approach, employing two complementary sources of data: close examination of transcript data using constructivist dialogue mapping (CDM) and automated face analysis (AFA) to determine affective state as a means of tracking engagement. CDM offers a rich connected view, as conceptual understandings are elaborated in real time. This view enables studying how concepts are elaborated during model use. If researchers want to study learning as it unfolds, they will need to track the ideas as they develop and are elaborated on. Humphrey and Gutwill (2005) point towards this when they say, "visitors seem to be constructing a conceptual understanding." (p. 20). This would support the use of CDMs inside of museums. It also fits into the work by Gaea Leinhardt and Kevin Crowley (1998, 2003) on learning as elaboration.

Because using agent-based models has been found to be an effective way to learn about complex systems, and complex systems underpin the solutions to some of the world’s most pressing problems, educators need more effective ways to teach the subject and measure the learning while they do. But researchers and policy makers have the challenge of understanding what students are taking away from open-ended learning about this complicated way of thinking. We need to evaluate the effectiveness of open-ended learning environments. However, as Jonassen argued: “perhaps the thorniest issue yet to be resolved regarding the implications of constructivism for learning is how to evaluate the learning that emerges from those environments” (Jonassen, 1991, p. 1). In general, the literature has documented the difficulties in evaluating learning in informal settings delivered in a constructivist way. Previous empirical studies (Segers, 1996) have shown constructivist environments do not always result in the expected outcomes. Possible explanations have been proposed to explain why these environments do not live up to their promise (Delva et al., 2000). One problem with these studies is that they treat constructivism as a narrow pedagogical approach rather than a broad-based theory of how people learn (Renkl, 2009). For a review of several constructivist evaluation methods and their shortcomings, see Rikers and colleagues (Rikers et al., 2008). This boardbased theory should include how people regulate learning, which is an affective question; ideas are not just thought—they are felt. Even reasoned behavior, such as learning, is neurobiologically directed by the affective states of the learners (Picard, 2000). As a result, I will also collect affective state data in order to study how people are regulating as they learn.

To address the paucity of evaluation in constructivism, and take account of affect’s role in this regulation, I employ two principal methodologies, which can be applied broadly to analyze learning and engagement in open-learning environments: constructivist dialogue mapping (CDM) and user engagement tracking with affective computing measures. Then, I will triangulate these measures to identify moments where users (a) elaborate their ideas, and (b) engaged.

6.1 Constructivist Dialogue Mapping

Martin, Horn and Wilensky (2020) offers a methodological innovation that is useful for the interactions typical of constructionist learning environments. Using CDM, researchers can track the construction of conserved concepts through the available proxy of change and elaboration of speech. The method is differentiated from age specific representations as semantic maps have been. I do not attempt to use the maps to improve student learning of the connections between material as much of the concept mapping has. Instead, they to track learners' moment to moment manifestation of conceptions as they construct them into less variable forms

In this paper, I code CDMs of players elaborating their ideas about agents or the items through observation and interaction at different points including, the pre-interview, during the interaction, and the post-interview that accounts for the changing nature of ideas throughout the activity. Maps visually depict the ideas players share through what they say and how they interact with the exhibits. I can research what people conserve (that is, learn through accommodation demonstrated by what they say) by filling in a map with what people say and do during play. I then count the number of 1st level items, such as ants or bear brains, and then use constructivist dialogue mapping to track how people's words and actions indicate learning through how they elaborate on those items, such as ants lay trails or bear brains are small. These later elaborations I call 2nd, and 3rd order elaborations on an item. In our findings I present graphs of total elaborations and elaborations by their depth (1st, 2^{nd,} 3rd, 4th, and 5th order elaborations respectively).

However, these maps, are only a proxy, built from the externally observable action of speech. The advantage of this approach is I can present the smallest parts of what I observe. For instance, I can note when users first identify an unknown entity, such as "purple line" and then how they build out an understanding of the subprocesses of that task such as "attracts ants" or "fades away."

As a result, I read the transcript word by word to see how users construct and elaborate knowledge on the representations they see. CDM demonstrates learning as concept elaboration over time through the convenient proxy of changes in speech.

6.2 Measuring Engagement

Second, measures of engagement in informal learning have been rather pedestrian, like numbers of minutes engaging, but I have new technologies that more directly measure engagement, through computer systems, to collect biometric and facial expression data, that can give us more direct measures of physiological measures to study degree and kind of engagement (K. Martin, Wang, et al., 2019). Facial / biological measures of affect can help us understand the relationship between positive affect and learning. In this study I will track facial expression using videos of faces (Cohn and De la Terre, 2014).

6.3 Triangulation of Measures

Finally, I triangulate these measures of engagement with CDM to identify moments participants elaborated their ideas with sentic modulation of the affective computing signals. This step may provide insight into the role of affect in learning in informal learning and will provide rich data to analyze the processes visitors go through during constructivist learning with an agent-based model of ants. Finally, while this step is the most experimentally untested, it also could provide useful insight to evaluate the impact of the design on learning outcomes of the museum exhibit. I use these methods to study the effectiveness of Ant Adaptation. The methods themselves provide an example of how we can use social insects to teach complex systems ideas. And the evaluation methods I am deploying apply to museums more broadly, but definitely to informal learning in remote environments.

7 Methods

I used two means of measuring learning in the designed, "open" environment: (1) a transcript analysis with constructivist dialogue mapping (Martin, Horn and Wilensky, 2020), to monitor what people say and elaborate on, and (2) affect tracking as a proxy for monitoring the internal cognitive machines driving engagement in learning (K. Martin, Wang, et al., 2019). In this paper I present an experiment. The experiment concerns studying how people understand an agent-based model I developed as a game for learning. I study they understand (1) how ants gather food (algorithmically), and (2) the complex system that helps ants self-organize using pheromones to build a trail network that leads to a self-organized superorganism (Hölldobler & Wilson, 2009; Wilson, 1984). The main part of the study is two case studies of participants playing the game. I explore whether the participants construct an understanding of agent-based modeling, complex systems, and ant organization. Simultaneously, I study their affective states while the play the game. This paper provides an example of how we can study learning and affect while participants play a game. To study this, I collect pre-post clinical interviews, follow users as they play, write field notes, and audio-video record to examine what people learn and how their affective states and engagement change while they participate in the program. The examination consists of three parts. First, I code the transcript data, following Leinhardt and colleagues, to count how many elaborations users had. Second, I use constructivist dialogue mapping to rank these elaborations as 1st, 2nd, 3rd etc., to order these elaborations and count how many of each type. Third, I count the presence of affective states and, using time-series analysis, and present the order chronologically. This allows the research to answer how much elaboration is occurring, what kind of elaboration it is, and what affective-states participants are evincing during the interview.

This analysis of the Ant Adaptation environment is the culmination of a multi-year design-based research (DBR) (Brown, 1992; A. Collins, 1992; W. Sandoval, 2014), and thus combines the methods of the previous iterations (K. Martin et al., 2020; K. Martin, Horn, et al., 2019). In this iteration, I am specifically looking at whether moments of high stimulation as measured through AFA identification are associated with moments of learning, which I identified in earlier iterations through constructivist dialogue mapping as moments participants elaborated on a concept. The experiment will shed light on the use of multiple data streams to evaluate learning in informal environments.

7.1 Multimodal Synchronized Data Collection

To measure how children in environments learn, I use multimodal learning analytics to collect the source data for facial tracking, video, and audio (Ochoa & Worsley, 2016; Oviatt et al., 2013; Schneider & Blikstein, 2015). As seen in Diagram 1, I collect camcorder data, that I then split into audio recording and video recording. I process the audio into transcript data that I qualitatively code using constructivist dialogue mapping. I process the video using OpenFace, a computer vision processing application, to perform Automatic Facial Coding, to create affective state determination. OpenFace provides a determination 30 times per second (30 Hz). This process has the affordance of synchronization. In data science, a feature is a set of data used to study a construct of interest. For instance, in an Excel file, these are usually columns, with each row being a unique entry. For example, in my data each feature is an action unit (AU) of an affective state, such as the intensity of a smile (AU12 + AU06) measured 30 times per second. Consequently, a feature is the running affective state determination along the rows. In multimodal learning analytics (MMLA), synchronized data is crucial to assure that identified constructs and features of interest can be studied with multiple data streams. When collecting multiple data streams, synchronized data collection assures that data collection starts at the same time. For example, if a child learns a construct of interest, such as how pheromone trails self-organize an ant colony, at time point *i*, I use synchronized data collection to know what affective states came before and after the point she first mentions that construct. Additionally, if I know when the idea emerges in her speech, I identify which tools were used, and from there, which types of mediation that are occurring. In this study I evaluate these interactions by integrating multimodal data about the affective state and student elaborations.

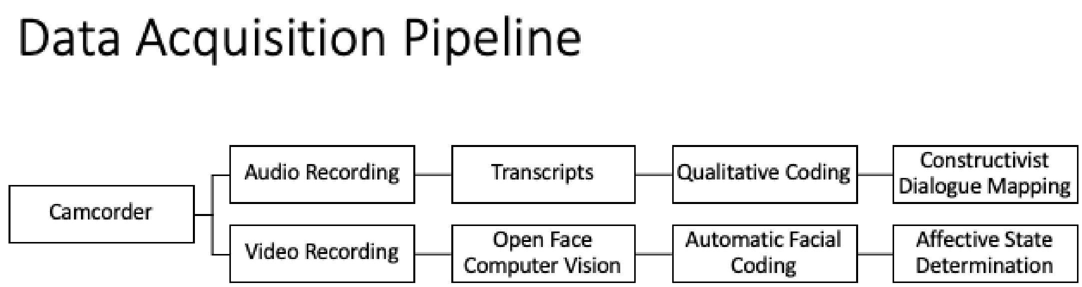

Diagram 1: Data recording is broken into two streams, Audio and Video, and then analyze the data streams independently.

7.2 Setting

I conducted the experiment via Zoom and in person. There were six groups of participants, and there were a total of eight people participating in the experiment. I present case studies of two of these interviews, one with an individual and the other with a group of two people. I could not use the remainder of the interviews because of technical difficulties with their affect determination, due to collection via Zoom during the pandemic.

The first case is with Mar, a science teacher. Mar is a 36-year-old White woman. She has a strong background in science, education, and test-prep, and loves to write. I interviewed her remotely at the beginning of the COVID-19 pandemic while using a web-based version of Ant

Adaptation[16]. This was done at the beginning of the 3rd phase, to gather data for the third iteration, she was selected to study how a science teacher would understand the environment. I recruited her via email to understand how educators would understand the environment.

The second case is with two college students, Mel and Bob. Bob and Mel are undergraduate engineering students. Bob is a 22-year-old African American woman. Mel is a 22-year-old Latinx woman. They both have a strong engineering background and love to code. I interviewed them in person, while using a touch enabled version of Ant Adaptation. This was done between iteration 2 and 3, and was implemented to gather pilot data on the third iteration, while still using interview protocol in Appendix 1. As such, they stood side by side over a large touch display with Ant Adaptation on it. I recruited them via list serves at a midwestern university.

7.3 Data Sources

I collected pre-post tests and semi-structured interviews with the questions outlined in Appendix 1. The interview consisted of a pre-test of participants' knowledge of complex systems and ants. I then had them interact with the ant-based learning curriculum (Ant Adaptation), recording what was discussed. All this information was collected synchronously so that constructs of interest could be triangulated.

7.4 Analysis

I used OpenFace, a computer vision software for automatic facial detection to identify facial action units (AUs). Facial action units are movements of muscles on the face that constitute expressions. For a good reference to see the AUs I would recommend iMotions GIF based visual library:

[16] Try the version out yourself: https://antomology.netlify.app/ It is a set of six scaffolded models, culminating in the final model, the same model tested in earlier museum work on a touch screen.

https://imotions.com/blog/facial-action-coding-system/. As seen in Diagram 2, I then synchronized the AU data to moments of identification to create the primary data for analysis. From the synchronized data I conducted three analyses: (1) I calculated the correlations between AUs and constructs of interest, chiefly moments participants elaborated. (2) I conducted time series analysis, and mapped moments of elaboration to AUs of interest. (3) I implemented three predictive models, which I explain below: simple rule, random forest, and XGBoost to predict moments of elaboration from the AUs. Next, I introduce how I identified elaborations, and hypothesizing, and how I synchronize the that data to the affect data. Then I discuss the methods of analysis including principal component analysis. I then describe the three predictive models, one rule, random forest and XGBoost.

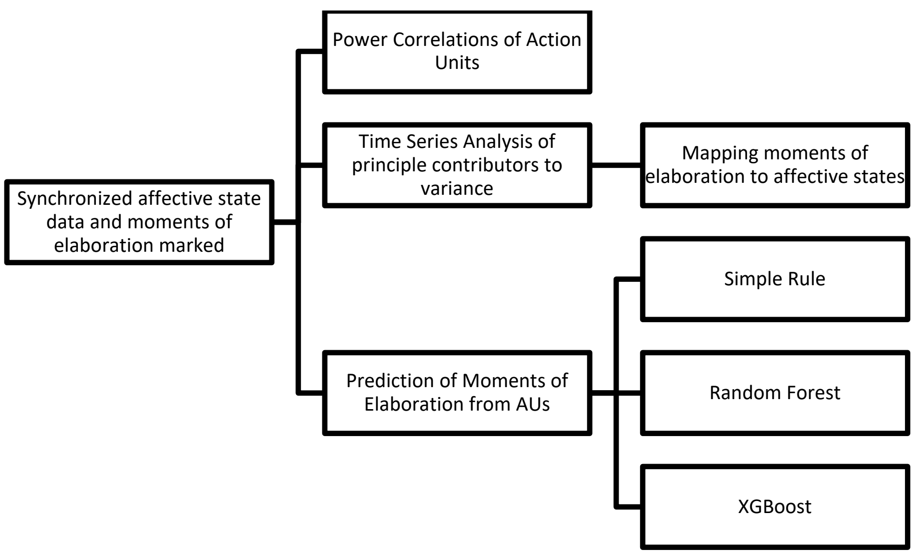

Diagram 2: Analysis of moments participants elaborated identified with CDM and AUs.

7.4.1 Elaborations

I used CDM to show how ideas elaborate during the interaction. In table 1, to demonstrate what is an elaboration in this method, I present an example of a single row of CDM data. This is an example of one row of the CDM: a row includes the time the turn of talk starts, who is speaking, the transcript of what was said, the node in the CDM that corresponds to, and the depth of the node in CDM. I code a row as an elaborated if the participant added a new node, or added onto a previous one. Pragmatically, in this approach I can simply code the moments with an item in the "Node" column: because if there is any text in the Node column, that means the person is elaborating either a new idea or expounding on an old one. In excel this is coded as (=IF(Node11<>"",1,0), which simply checks if there is text there. This is the most parsimonious approach, a moment where a person elaborated is each time, they verbally added to what they had said before. What this approach presents is the qualitative coder of the CDM has marked a moment a participant elaborated on an idea, or concept. This is noted by adding any text in the node column.

7.4.2 Hypothesizing

I also coded when questions arose. Hypothesized is simply coded as =IF(ISNUMBER(FIND("?",Quote11)),1,), which means, if there is a ? in the quote column, mark this column as a 1, otherwise, mark it as a 0. In other words, if there are any question marks in the Quote place a true in the hypothesized column. What the questions are about is not taken account in this version, as it was not a key part of my investigation.

Table 2: An example of one Row of the CDM: a Row includes the time it starts, who is speaking, the transcript of what was said, the node in the CDM that corresponds to, the depth of the node in CDM. Then I code for elaborations and hypothesizing

Time	Speaker	Quote	Interaction Type	Node	Parent Node	Node Depth	Elaborated
4:43	**Mar**	**it changes how much it wiggles while it walks.**	**Model 1: Single ant motion, probe 1**	**Wiggle**	**Changes ant motion**	**3**	**1**

I broadcast the elaborated codes to the affective states by aligning their time stamps and merging the data using python. Specifically, I code over all the action units iterating over ~85,000 rows and 715 columns per interview. To facilitate this process, I used the GPU accelerator option in the Google Collab notebooks. The code to merge the data was as follows:

```
for r in range(1,len(DF)):
    # set previous row index
    p = r - 1
    # if there is no entry in df2 for a certain timepoint, use
      the previous entry
    # one entry stops when the next one begins
    if DF.iloc[r, 715:].isnull().all():

        DF.iloc[r, 715:] = DF.iloc[p, 715:]
    else:
             DF.iloc[r, 715:] = DF.iloc[r, 715:]
```

After merging the data by aligning milliseconds—715 columns from AFA, and 17 columns from CDM—I created a dataset to analyze for each of the two case studies presented in this paper. For the first case study of Mar the procedure created a dataset of 83,719 rows and 732 columns that results in 61,282,308 cells. For the second case study of Bob and Mel the procedure created a dataset of 84,624 rows and 732 columns that results in 61,944,768 cells.

I use this data to investigate the relationship between AUs and elaboration, and then to predict moments of elaboration from AUs. To triangulate I study what affective states appear and use those for predictive model, time series analysis and correlations. Using time series analysis, I analyze when in the interaction they occur. During the time series analysis, I conduct principal component analysis (PCA). I also used pathway analysis to calculate what emotional states sequences lead up to moments of elaboration. I present how they connect to moments of elaboration identified via CDM using a correlation matrix. I conclude the triangulations section by using three models: one rule, random forest, and XGBoost to predict moments where participants elaborated from affective states. Next I explain each of those analyses.

7.4.3 Predictive Models

As you see in diagram 2: In a machine learning model, we provide input data, in my case that is AUs. Then we split the data, in my case we train the model on 70% of the data, and then try to predict the outcome from the remaining 30%. This is called a train test split. I use this to train a model to predict moments marked as Elaborated (1) vs not Elaborated (0). The models then output a confusion matrix, of how many 0s were correctly labelled as 0s. How many 0s were incorrectly labelled as 1s. How many 1s were correctly labelled as 1s. And, how many 1s were incorrectly labelled as 0s. From this data we know how accurate the model is at predicting 1s and 0s as well as the precision and recall. In my data, this is from the training AUs how accurately the model can predict rows marked as elaborated from the testing AUs.

7.4.3.1 One Rule

Another way of doing this, is frugal models with simple rules. One Rule, is a simple, yet accurate, classification algorithm that generates a rule for each predictor in the data, then selects the rule with the smallest total error as its "one rule". To create a rule for a predictor, we construct a frequency table for each predictor against the target. The approach is often used as a first pass model: data scientists start the one rule, as they have surprising predictive power, and are readily understandable. I use the model to predict moments participants elaborated in the CDM data from affective state data (AUs).

7.4.3.2 Random Forest

The random forest is a classification algorithm consisting of many decisions trees defined by the modeler. The model uses bagging and feature randomness when building each individual tree (Burkov, 2019; Koehrsen, 2018). The random forest takes the idea of a single decision tree, and creates an ensembl*e* model from hundreds or thousands of trees to reduce variance. Each tree is trained on a random set of the observations, and for each split of a node, only a subset of features is used to make a split. The model attempts to create an uncorrelated forest of trees whose prediction by committee is more accurate than that of any individual tree. I use the model to predict moments participants elaborated in the CDM data from affective state data.

7.4.3.3 XGBoost
The XGBoost algorithm has been dominating applied machine learning applications for structured or tabular data. XGBoost is an implementation of gradient boosted decision trees designed for speed and performance. In other words, it's a random forest on steroids. In prediction problems involving unstructured data (images, text, etc.) artificial neural networks tend to outperform all other algorithms or frameworks. I use the model to predict moments participants elaborated in the CDM data from affective state data.

7.4.4 Principal Component Analysis (PCA)
PCA is the simplest of the true eigenvector-based multivariate analyses. Often, its operation can be thought of as revealing the internal structure of the data in a way that best explains the variance in the data. If a multivariate dataset is visualized as a set of coordinates in a high dimensional data space (1 axis per variable), PCA can supply the user with a lower-dimensional picture, a projection of this object when viewed from its most informative viewpoint. This is done by using only the first few principal components so that the dimensionality of the transformed data is reduced. PCA has two uses, it is used in exploratory data analysis and for making predictive models. It is commonly used for dimensionality reduction by projecting each data point onto only the first few principal components to obtain lower-dimensional data while preserving as much of the data's variation as possible. PCA can be used to see the primary ways that data moves when dealing with multi-dimensional data. I use it to see how AU data moves together after

dimensionality reduction in time series analysis.

To summarize my methods: if minds are complex systems in congress with other minds, we need methods of analysis that study these systems in vitro to research their learning.

8 Results

I use this data collection method for participants Bob, Mel, and Mar while they explored a constructionist learning environment, Ant Adaptation (Martin and Wilensky, 2019), to get a more complete picture of the context of learning.

Research Questions

- Does the technique of CDM add new insight into how learners, in group conversation, can advance their learning?
- What can we learn from physiological measures of people's affective states while they engage learning in informal learning environments, such as a museum exhibit or a learning game, where people choose to come and learn?
- Is there a relationship between high stimulation and learning, as measured by CDM, and affective computing signals? Within moments participants elaborated, is there a relationship between positive affect and learning?

8.1 Overview

In the findings I cover three key parts. First, as shown in section 8.1.1, I present frequency counts of elaboration of items found through constructivist dialogue mapping. Constructionist dialogue mapping (CDM) demonstrates the ontological entities that participants mention during play, and connects them together to show levels of elaboration. In this section I first show how many entities are mentioned in the pre-interview, during the play, and after the interview, and then compare the pre-post change. I find that some students reduce the number of items they find. In the second portion of this part, I compare the two case studies, based on the order of their elaboration. In

CDM, the researcher categorizes items as 1st, 2nd, 3rd, 4th, or 5th degree elaborations. Thus, I can count the hierarchical level of items, and get an estimate of the depth of the elaboration the participants make while playing. I find that students elaborate deeply, while playing Ant Adaptation, with 33% of the total elaborations being the deepest level—4th or 5th order elaborations. To dig into these two findings, I then present the CDMs of two cases, Mar, a science teacher, and Bob and Mel, undergraduate students of engineering. I demonstrate their hierarchical maps using CDM and show how they came to understand the complex systems of ant colonies.

Within each case study, after showing how participants understand the complex system, I then review their affective states during that interaction to start to shed light on what the physiological measures of affect can tell us about learning and how people are regulating their emotions during these moments participants elaborated. I conclude the finding sections with my results about triangulation, applying 3 models to predict moments participants elaborated identified by CDM based on action units, identified through automatic facial tracking. I find that while the simple rule model does not provide accurate predictions because of no single AU providing much predictive power, more complex models, random forest and XGBoost, do provide high accuracy, precision and recall when predicting moments of elaboration from AUs. Next, I turn to frequency counts of different kinds of elaborations and their depth, resulting from CDM.

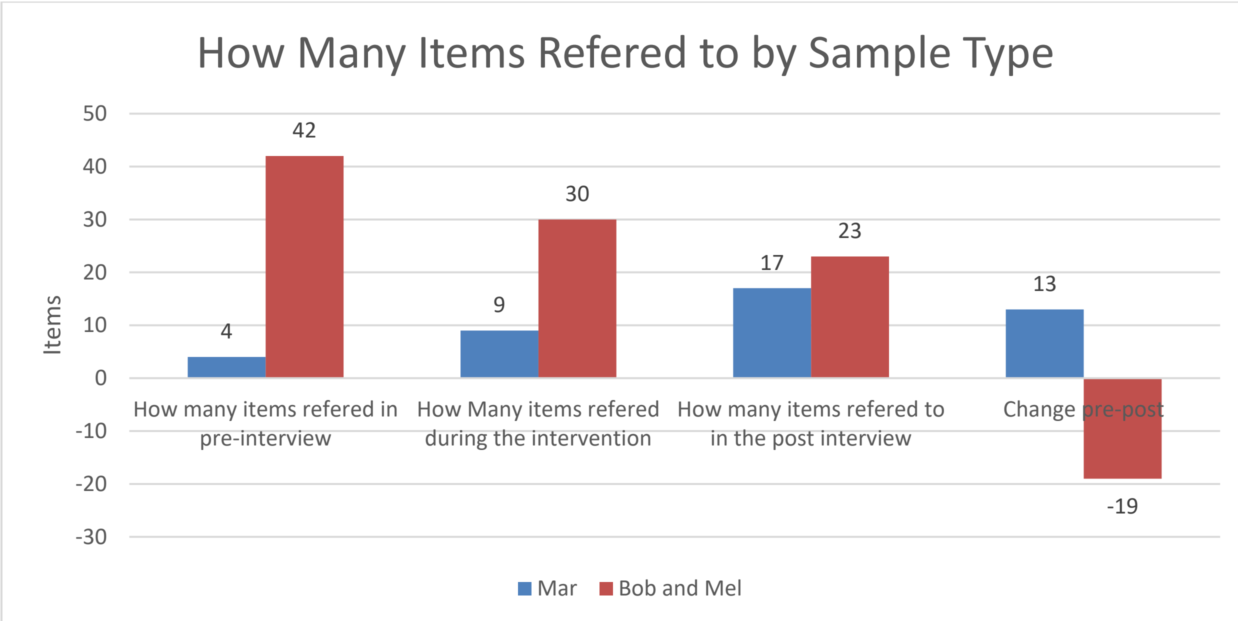

Figure 4: Mar, Bob, and Mel used the agent-based model (ABM)

8.1.1 How much detail did participants add to the referenced items?

I found participants elaborate their ideas richly when using the dynamic agent-based model — demonstrating deep behavior observation. In Figure 2, I count how many items participants referred to by the portion of the interview. As shown in Figure 2, participants name 40 (Mar = 17; Bob and Mel = 23) items when playing Ant Adaptation by the end of the post interview. They named 46 in the pre-interview, 39 during the intervention. This results in a net loss of 6 items (Mar = 13, Bob and Mel = -19).

A seen in Figure 2, these Ant Adaptations users make many elaborations: a total of 40 elaborations in the post interview. However, some of the change was negative. In Bob and Mel's interview, they mentioned 42 items in the pre-interview, but synthesized that down to 23 entities by the post interview, resulting in a net -19 items pre to post. I interpret this to mean, the process systemized knowledge, for Bob and Mel, they mention less items but understand the ant colony, for Mar, she expands what she thinks about in the ant colony. Though that Mar gained ontological entities Bob and Mel lost items might seem strange; however, as I explicate below, the change is the complexity of their understanding. As I show below in Bob and Mel's case study, they take a key moment to review their knowledge, and in this process systematize what they have learned. They simplified their explanation of what they saw, as they reduced the number of items mentioned. As Morin argued, "Rationality is a play, it is the incessant dialogue between our mind that creates logical structures, applies them to the world, in dialogue with the real world. When the world is not in agreement with our logical system, we must admit our logical system in insufficient, that it encounters only part of reality" (2008, pp. 47–47). Sometimes, encountering a situation leads to more ontological entities, and sometimes less. For example, as seen in Figure 2, while playing the game Mar adds 13 elaborations, but Bob and Mel lost 19. Since we see much progress in both cases, the number of just number of elaborations does not suffice to establish progress. We can use it as one measure, but we need others to address the quality of the elaborations.

Next, I introduce a measure, complex elaborations. A third (13 / 40) of participants elaborations were complex elaborations. To recall, in CDM, a 4th order elaboration is one where they describe the situation in four levels, such as "Ants (1^{st}) collected food (2^{nd}) by following trails (3^{rd}) to feed the colony (4^{th}). Higher order elaborations build on more prior elaborations, and thus represent deeper, more complex, behavior explanations that account for more of what the user is seeing. As shown in Figure 3, users of Ant Adaptation named 23 3^{rd}, 4^{th}, and 5^{th} order elaborations. Mar elaborated 17 times (11 > 2^{nd} order), and Bob and Mel 23, (12 > 2^{nd} order). Each elaboration introducing or expanded on an ontological entity the participants had previously discussed. Especially notable was the ample presence of complex elaborations, elaborations of 4^{th} and 5^{th} order (13, 33% of the total elaborations). I interpret the findings of Figure 3 to mean players added details about the rules of agents during the intervention, building on what they knew through playing with the game by constructing complex elaborations.

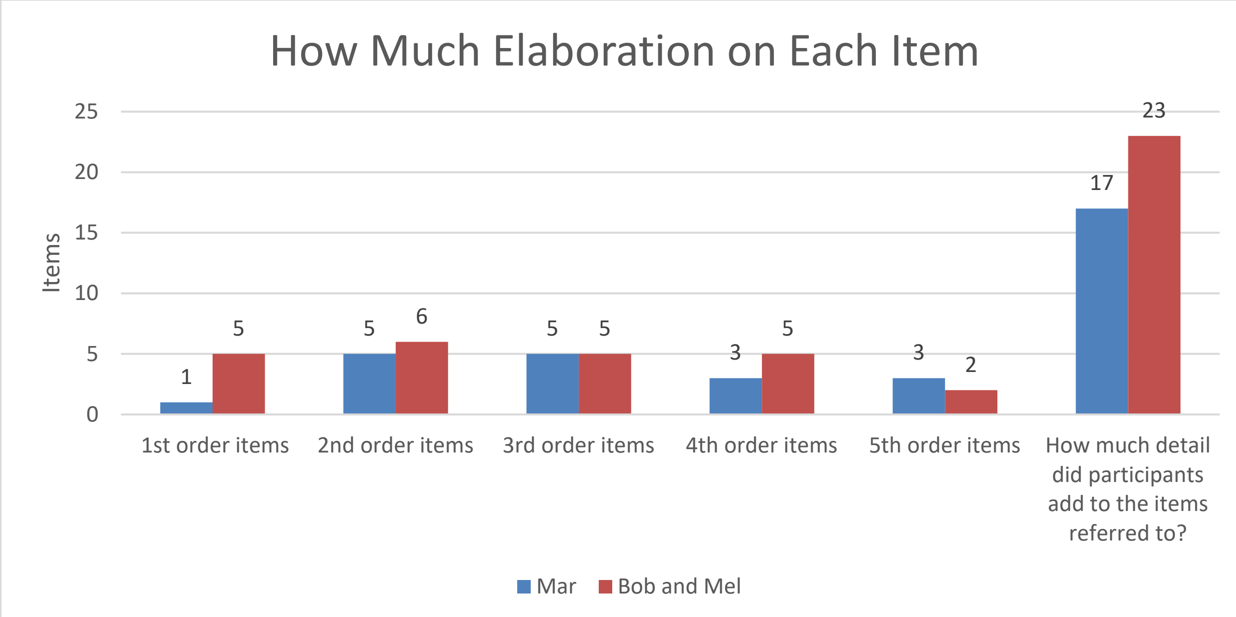

Figure 5: Users of Ant Adaptation named 40 elaborations. Many of these were complex elaborations (4th or 5th order), building on what they talked about prior. There were 23 3rd, 4th, and 5th order elaborations. There were 13 4th and 5th order elaborations (33% of the total elaborations).

To demonstrate the complexity of the elaborations, I present two cases below using CDM. In each case I present the participant's understanding of ants before playing the game, then remap it for while playing the game, and account during the post interview. This shows the change throughout the interaction. At the end of each case, I analyze the affective states the participants undergo during the interaction. Next, I present Mar's case.

8.2 Case 1: Mar Engages in Deep Behavior Observation

Mar is a 36-year-old science teacher who is White. She has a strong background in science, education, and test prep, and loves to write. I interviewed her remotely at the beginning of the COVID-19 pandemic while using a web-based version of Ant Adaptation.

During a one-hour interaction, Mar learned that size does not negatively affect an ant's chance to reproduce. But what led to her discovery was more complex: she first had to learn that an ant colony is a complex system where only the queen reproduces, and all the individual insects work towards that collective's betterment.

As we see in Transcript 1, at the beginning of the interview, Mar had a purely observational perspective on ants.

Time	Speaker	Question	Quote	Interaction Type	Node	Parent Node	Node Depth	Elaborated	Hypothesized
0:37	Mar	Have you ever noticed anything about ants?	And they walk around and sometimes they appear in big groups walking around and sometimes they carry things from place to place. Like food stuff and, and house, house building materials.	Pre-Question 1	Carry things	Ants	2	1	0
0:56	Kit		How do ants collect food and then	Pre-Question 2				0	0
1:02	Mar	How do ants collect food and then	Lift did with their tiny feet above their heads where they then carry it.	Pre-Question 2	Lift feet	Carry things	2	1	0
1:13	Kit		Sir answer collecting food and doing house repairs and such how to ants know what to do.	Pre-Question 3				0	0
1:20	Mar	how to ants know what to do?	Um, because they followed an older and and oh, maybe and also they're pre programmed genetically.	Pre-Question 3	follow orders	Ants	2	1	0
1:20	Mar	how to ants know what to do?	Um, because they followed an older and and oh, maybe and also they're pre programmed genetically.	Pre-Question 3	Programmed genetically	Ants	2	1	0
1:37	Mar		Dude, dude					0	0
1:41	Mar	how to ants know what to do?	um yeah. So I'm I'm guessing without knowing this, that they're somewhat evolved to collect food in a certain way and possibly also learn socially from one another For example, how to find the food source in their territory. I do not know this whatsoever.	Pre-Question 3	Evolved to collect food	Programmed genetically	3	1	0
1:41	Mar	how to ants know what to do?	um yeah. So I'm I'm guessing without knowing this, that they're somewhat evolved to collect food in a certain way and possibly also learn socially from one another For example, how to find the food source in their territory. I do not know this whatsoever.	Pre-Question 3	Learn socially from one another	Ants	2	1	0

Transcript 1:Mar's Pre-interview answer to how ants know what to do.

When the interviewer asked if she'd ever noticed anything about ants she said, "Yeah, they walk around and sometimes they appear in big groups walking around and sometimes they carry things from place to place. Like food stuff, and house building materials." Figure 4 shows how I visualized Mar's concept of ants in a constructionist dialogue map, each box is an elaboration, or in Piaget's theory ontological entity that the participant has mentioned in the interview. As seen in Figure 4, Mar had one 1st order elaboration (ontological entity named) (Ants) and three 2nd order elaboration (Walk, Carry thing, Group up). She also mentions a hypothesis at 1:41 that ants evolved to collect food the way they do, and at 1:02 that that ants carry things by lifting tiny feet above their heads. I did not include these in the CDM, as they were conjecture, and she did not elaborate on them further. That being said, a researcher could.

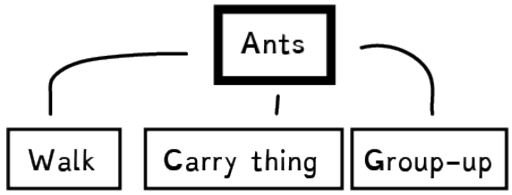

Figure 6: Mar's Initial Observational based concept of ants. Here there is one 1st order ontological entity named (Ants) and three 2nd order ontological entities named (Walk, Carry thing, Group up).

In Transcript 2, we see Mar elaborating on her ideas further. Notably I did not add in the walk-in-lines here, because she seemed to be echoing me, but in a minute, she has picked it up by exploring the model.

Time	Speaker	Question	Quote	Interaction Type	Node	Parent Node	Node Depth	Elaborated	Hypothesized
5:37	Kit		We're trying to get a get the best guesses you know nothing about it has nothing to do with you know any answer. These are not school questions there. I forgot to mention there are no right answers. You're not supposed to know any of this. Hmm. So how do you think an ants size would affect an ants life?	Pre-Question 6	Size	Ants	2	1	1
6:12	Mar	So how do you think an ants size would affect an ants life?	I think and and that is relative to its immediate peers probably lives a long and healthy life compared to its peers.	Pre-Question 6	Bigger ants live longer	Size	3	1	0
6:27	Mar		Large relative to its immediate peers.	Pre-Question 6				0	0
6:31	Kit		Big ants live better.	Pre-Question 6				0	0
6:33	Mar		Yeah, because probably they're eating well. How tall humans don't live as long as short humans. But a) I'm not sure this is true and b) I just think the lifespan of an ant is not long enough for like blood supply to become an issue. So probably eating well and plenty would be good.	Pre-Question 6	because probably they're eating well	Bigger ants live longer	4	1	0
6:56	Kit		Have you ever heard of pheromone trails	Pre-Question 7	Pheromone Trail		1	1	0
7:04	Mar		No.	Pre-Question 7				0	0
7:10	Kit		Have you ever seen ants walking lines?	Pre-Question 7	Walk in line	Ants	2	1	1
7:14	Mar		Yep. And as soon as you said the term and made sense to me, but I hadn't previously heard that term.	Pre-Question 7	Walk in line = pheromone trail	Walk in line	3	1	0
7:19	Kit		Okay. Do you know why ants	Pre-Question 7				0	0
7:21	Mar		walk in lines? Yes, because they follow the pheromone trail laid by the ants in front of them.	Pre-Question 7	because follow the pheromone trail laid by the ants in front of them	Walk in line	4	1	1
7:30	Kit		That makes sense. Mm hmm.	Affirming				0	0
7:34	Kit		Okay, I need to stop this thing from recording all of my slack messages. Oh, wow. So now we're going to thank you for that pre interview. Very good guesses about how things work. Now we're going to go look at a series of models.	Introduction to Modelling Activity				0	0

Transcript 2: Mar continues to elaborate on what she knows about ants, specifically, how size impacts and ant's life-span.

In Figure 5, I visualize the remainder of the pre-interview based on the node depth and parent nodes following the protocol from earlier work (K. Martin et al., 2020). Through the interview, I drew out the rest of her concept of ants. In Figure 5, Mar has one 1st order elaboration (Ants); five 2nd order elaboration (Size, Walk, Carry thing, Group up, Know what to do; and three 3rd order elaborations (Live longer, Genetic Programming, Social Learning).

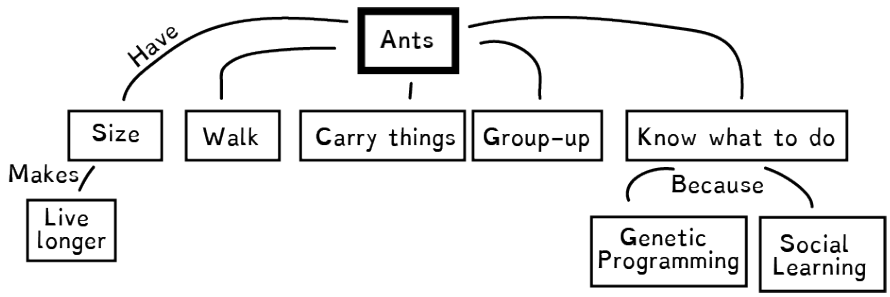

Figure 7: Mar describes the how an ant's size impacts its life, and how ants know what to do as a function of nurture and nature. Here there is one 1st order ontological entity named (Ants); five 2nd order ontological entities named (Size, Walk, Carry thing, Group up, Know what to do); and three 3rd order ontological entities named (Live longer, Genetic Programming, Social Learning).

While engaging with Ant Adaptation she moves to a much more decentralized understanding of ants' action. The shift occurs because through observing the agent-based model, she expands on how ants walk, carry things, and group up. As shown in Figure 6, she forms the idea of three functions: (1) 'Changes of Ant Motion', (2) 'Forage', and (3) 'Lay Pheromone Trails'. For example, in Transcript 3[17], she changes how she thinks about ant motion by manipulating the ants wiggle radius.

[17] Mar's Zoom interview was interrupted when she had to restart her computer. Consequently, the transcripts are broken into two parts. The first video was 16 minutes and 24 seconds long. We immediately resumed after the computer booted back up. therefore, to get the total interview time add 16 minutes and 24 seconds to the time column of transcripts 3 and 4.

Time	Speaker	Question	Quote	Interaction Type	Node	Parent Node	Node Depth	Elaborated	Hypothesized
2:25	Mar	What happens when you change wiggle radius?	A little straighter line?	Model 1: Single ant motion	Straightness of Lines	Walk in line	3	1	1
2:28	Kit		Why do you think maybe it was the straighter line?	Model 1: Single ant motion				0	1
2:30	Mar		because I made the wiggle radius really small and now that changed things. I mean, is it the wiggle of the ant itself? Yeah. Okay, can I make it bigger again? How Can I go?	Model 1: Single ant motion	Smallness of wiggle radius	Straightness of Lines	4	1	1
3:01	Kit		I think it's a circle. So up to 360. Obviously that doesn't, that's just a full circle. Okay.	Model 1: Single ant motion	Wiggle Radius = Circle of 360	Smallness of wiggle radius	5	1	0
3:16	Mar		Gotcha. Are we trying to optimize this? And what so it is the end is wiggling too much. And because it is wiggling too much, oops, that's not random. It is covering more ground, though it also is not going to the nest as often.	Model 1: Single ant motion	Optimize?	Walk in line	4	1	1
3:16	Mar		Gotcha. Are we trying to optimize this? And what so it is the end is wiggling too much. And because it is wiggling too much, oops, that's not random. It is covering more ground, though it also is not going to the nest as often.	Model 1: Single ant motion	Wiggling too much covers more ground	Walk in line	4	1	1
3:56	Kit		So what's that looks like here.	Model 1: Single ant motion				0	0
4:01	Mar		Mmm	Model 1: Single ant motion				0	0
4:04	Mar		this looks like an effectively and efficiently is searching and to me that is covering more ground and also not completely getting lost in the picture.	Model 1: Single ant motion	Searching efficiently	Wiggling too much covers more ground	5	1	0
4:15	Kit		Make sense? So in programming, we call that blue box, a function that basically takes in an input and it does some output. So if you had to describe what that blue box function does, what do you think, wander does?	Model 1: Single ant motion, probe 1	Programatic Function		1	1	1
4:35	Mar	what do you think, wander does?	It changes the motion of the ant?	Model 1: Single ant motion, probe 1	Changes ant motion	Ant	2	1	1
4:41	Kit	in what ways	in what ways	Model 1: Single ant motion, probe 1				0	0
4:43	Mar		it changes how much it wiggles while it walks.	Model 1: Single ant motion, probe 1	Wiggle	Changes ant motion	3	1	0

Transcript 3: Mar changes how she thinks about ant motion by manipulating wiggle radius.

In Figure 6, the three function are labelled in red, magenta, and blue, respectively. In Figure 6, Mar has one 1st order elaboration (Ants); five 2nd order elaborations (Size, Walk, Carry thing, Group up, Know what to do); five 3rd order ontological entities named (Live longer, In lines, If at pant, nibble, Genetic Programming, Social Learning); three 4th order elaborations (Straightness of lines, If has food, return to nest, Pink chemical tells ants other ants have been there); and three 5th order elaborations (Smallness of Wiggle Radius, Optimizes searching efficiency Pink chemical tells ants other ants have been there).[18] These additions show she has formed the idea, a complex systems notion, of how small, simple agents following simple rules lead to longer-lived ants, because they have plenty of food.

First, she expands her idea of how ants walk to derive the function Changes of Ant Motion (Figure 6, Transcript 3). The shift begins when she reprograms the ants to wander a little less. After she does this, I ask if she notices a difference. Mar responds, “A little straighter line?” And when probed to account for it, she responds by asking and answering her own question with the model: “Because I made the wiggle radius really small and now that changed things. I mean, is it the wiggle of the ant itself? Yeah.” She then wonders, “Are we trying to optimize the wiggle?” Tinkering with the parameter, she notices that increasing the radius increases how much ground an ant explores, but also decreases how efficiently the ant returns food it finds. From these

[18] The last one, pink chemical tells ants other ants have been there, is counted as both 4th and 5th order, because it falls under two different trees in the CDM shown in Figure 6. Multi-branching in a CDM leads to problems with counting. If researchers seek simplicity in counting elaborations, employ single branching when constructing a CDM so that every box has at most one parent node.

observations she postulates that it's a balance that leads to efficient area search.

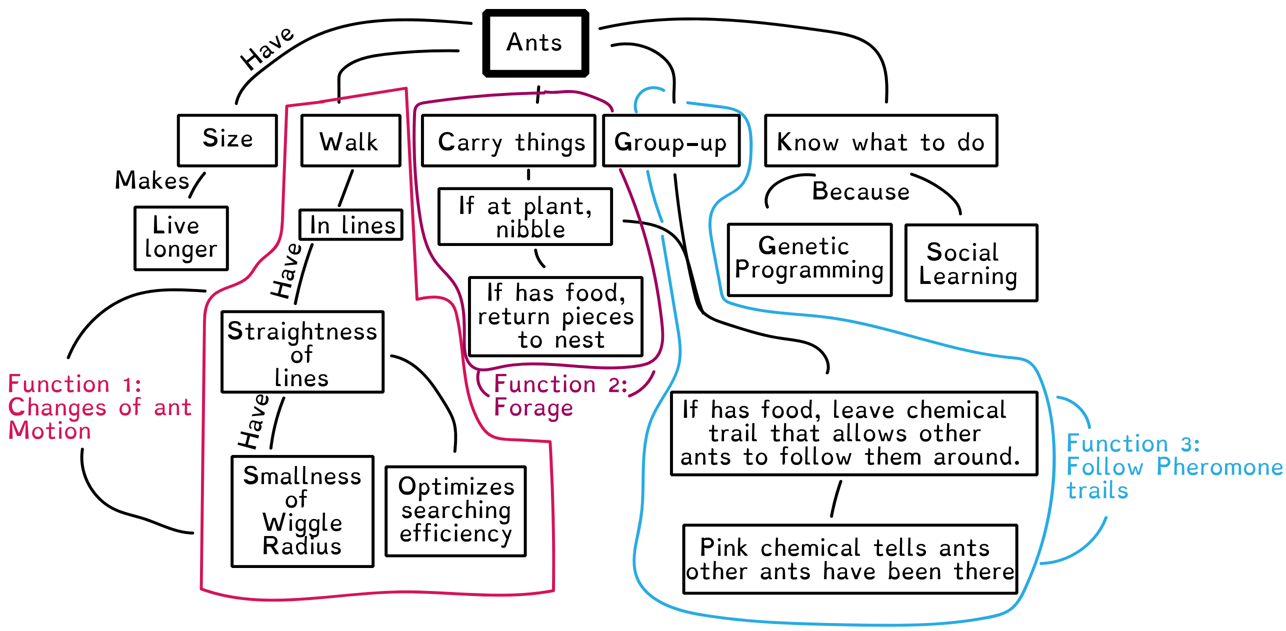

Figure 8: Change of ant motion, foraging, and pheromone trails explains how ants explore space. Here there is 1 1st order ontological entity named (Ants); five 2nd order ontological entities named (Size, Walk, Carry things, Group up, Know what to do); five 3rd order ontological entities named (Live longer, In lines, If at pant, nibble, Genetic Programming, Social Learning); 3 4th order ontological entities named (Straightness of lines, If has food, return to nest, Pink chemical tells ants other ants have been there); and three 5th order elaborations (Smallness of Wiggle Radius and Optimizes searching efficiency, Pink chemical tells ants other ants have been there) The last one, Pink chemical tells ants other ants have been there, is counted as both 4th and 5th order, because it falls under two different trees in the CDM.

Second, as shown in Transcript 4, she elaborates her understanding of how ants forage for food and build and use pheromone trails.

Time	Speaker	Question	Quote	Interaction Type	Node	Parent Node	Node Depth	Elaborated	Hypothesized
25:37:00	Mar	How do ants collect food?	how do they collect food? they walk	Report on Learning	Walk	Ants	2	1	1
25:41:00	Mar		to it, and then nibble on plants and then return the plant pieces to their nest leaving chemical trails that allow other agents to follow them around.	Report on Learning	If at plants, nibble	ants	2	1	0
25:41:00	Mar		to it, and then nibble on plants and then return the plant pieces to their nest leaving chemical trails that allow other agents to follow them around.	Report on Learning	If has good, Return pieces to nest	Walk	3	1	0
25:41:00	Mar		to it, and then nibble on plants and then return the plant pieces to their nest leaving chemical trails that allow other agents to follow them around.	Report on Learning	If has food, leave chemical trail that allow other agents to follow them around.	Walk	3	1	0
25:56:00	Kit		Nice. So what's the pink chemical, do	Report on Learning	Pink chemical			1	0
26:01:00	Mar		It just tells ants that ants have been there	Report on Learning	Pink chemical tells ants ants have been there	Pheramone Trails	3	1	0
26:07:00	Kit		what features of this game did you like best?					0	1
26:17:00	Mar		The vinegar	Design Question				0	0
26:19:00	Kit		ones was that	Design Question				0	0
26:21:00	Mar		lets me mess with the happenings	Design Question				0	0

Transcript 4: Mar describes how ants forage for food and how they group up.
In this elaboration she started to take on the language of conditional logic, describe, if at plant, nibble, and, if has food [from nibble], return the pieces to the nest. Notably here, she has noticed the food as little agents and started to discuss them as individual pieces.

Third, starting at 25:41 of Transcript 4, she elaborates her understanding of how and why ants group up. Here she combines functions 1 and 2 to say that if the ant has food, it lays the trail to allow other ants to follow it to food, arguing the trail simply tells other ants where ants have been. Combining all this together, she comes to her big learning moment that more, small ants collect more food faster, which implies a system understanding of how food flows into the colony. As she said prior at 25:21, "You know, from all I know so far smaller and less aggressive ants are going to reproduce at a faster rate." Here she has understood all the organizational and gathering processes at the ant level, and she has accounted for the individual actions of the collectives of agents (ants, flowers, and food).

8.2.1 Mar's Affective State

Next, I look at what Mar's affective state was while she came to this understanding. I analyze Mar's affective state during the interaction. For the distribution of AU, I plot a box plot as shown in Figure 7. What this means for Mar is that for most AUs, most of the time she is not holding that facial expression. In other words, AUs are only occasionally detected.

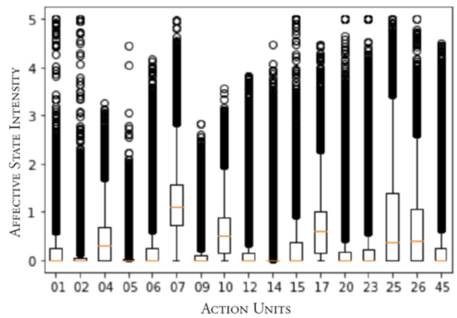

Figure 9: Intensity of emotional action units.

I find that the action unit scores are far from their mean. This indicates the data is sparse. With sparse data, most rows are not readings. Examples would be counts of cars by time on an abandoned highway. For most time periods, no cars would be detected. In my sparse data, many 0s drive down the mean. In addition, many features' mean is close to 0 and only 6 features' mean is not equal to 0. The closer the mean is to 0 and, the more outliers one observes, the more prevalent the affective state is.

Next, I plot the correlation coefficients of AUs in Figure 8, which demonstrates the AU connections. There are many ways to conceptualize a correlation coefficient. One that is intuitive is as follows: the correlation between two variables is their percentage of shared determinants. To unpack this, think of genetics. If some trait was entirely genetically determined, we would expect to find siblings with 50% of their genes in common and .50 correlation. Likewise, we would find .25 correlation between first cousins that share 25% of their genes. We could read a correlation like this as .25 of the causal factors that determine one also determines the other. In Figure 8, I present the positive and negative correlations between AUs and moments Mar elaborated. The heat map provides an overview of how the data is connected by presenting the correlation coefficients between the data (customarily abbreviated as *r).*

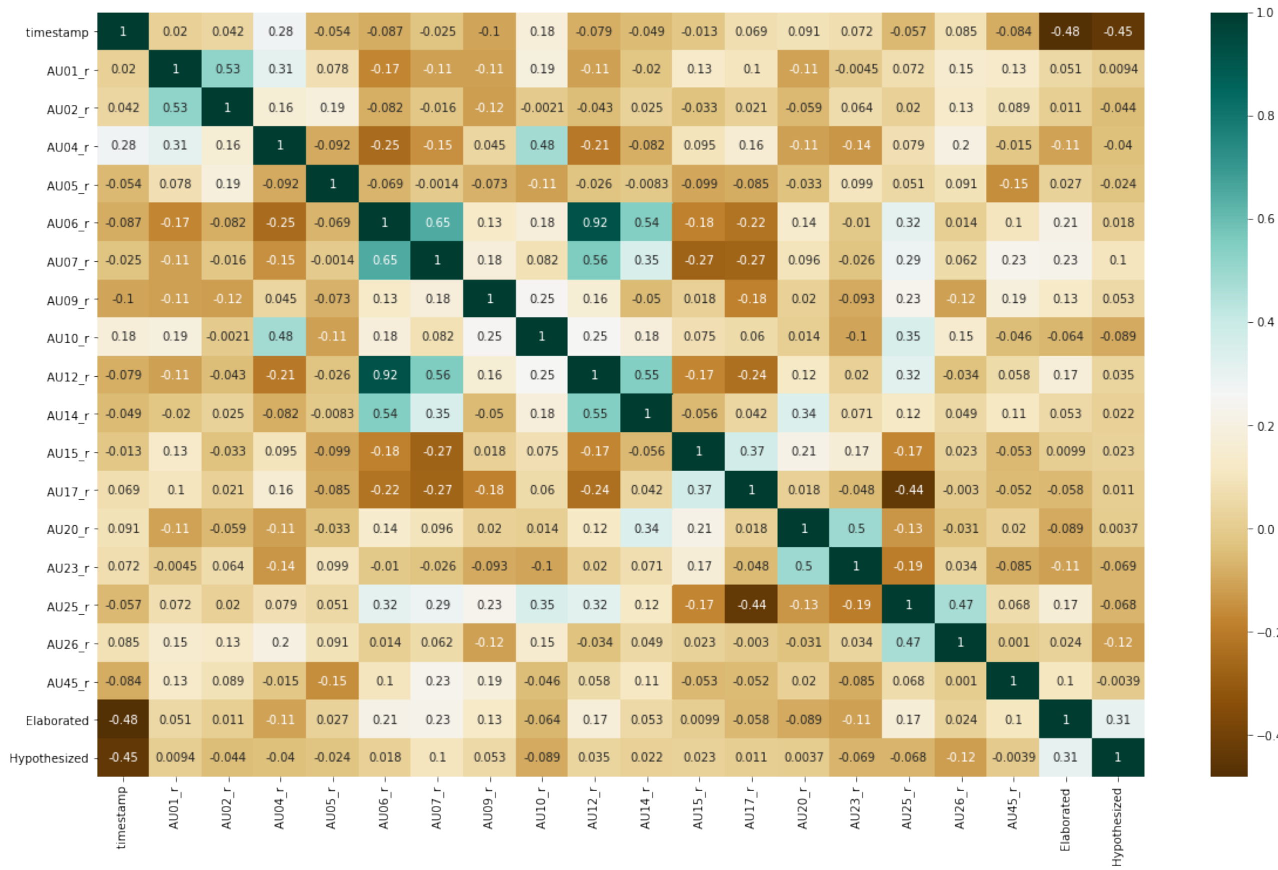

Figure 10: Correlations between facial action units and CDM data, elaborated and hypothesizing.

In Figure 8, we observe strong correlations between AU06 and AU012 (r .92 $p < 0.0$) and between AU26 and AU25 (r .47 p < 0.0). AU06 and AU12 are the principal parts of a smile, showing Mar does both when she smiles. AU25 and AU26 are dropping the jaw and separating the lips, which are the principal parts of speech. Notably, Elaboration is not strongly correlated with speech (AU25 r 0.17, p < 0.0 and AU26 r 0.024, p = 2.34). In other words, elaboration as coded by CDM is not merely a proxy of participants talking, but instead measures a particular kind of talk. I also note that no AU is highly correlated with Elaboration. The highest being between AU07 and Elaborated (r .23 p < 0.0). However, positive states are significantly correlated with Elaboration: for Mar AU06 and Elaborated have r 0.21 p < 0.0 and AU12 and Elaborated have r

0.17, $p < 0.0$. See the data analysis Appendix for a complete overview. This suggests they may be useful in a predictive model.

8.2.1.1 Time Series Analysis

Next, I wanted to know the timing of Mar's engagement. In other words, I do not only want to know that she curls her lip up most of the time when she arches her eyebrow, I want to know when she is smiling. As a result, next I examined which of these were primary AUs in the data. For feature selection, I used the function SelectKBest from SciKit-Learn, with the score function chi2. SelectKBest is one of the single variable feature selection methods. The principle of these methods is to calculate a statistical index of each variable separately, and determine which variables are important according to the index. After finding the important features, it excludes those that are not important. In this experiment, I take the score from chi2 as the single variable.

As shown in Figure 9, I found the resulting five most important features were Lips Apart (AU25), Lip Corner Pull (AU12), Cheek Raise (AU06), Blink (AU45), and Lip Stretch (AU20). The combination of Cheek Raise, AU06, and Lip Corner Pull, AU12, is classified as a smile. In general, as shown in Figure 9, analyzing the features, I find Mar is mainly happy during the interaction, with a noticeable plateau near the end of interaction.

Next, to examine how much the variation in the data was determined by the smiling, I used principal component analysis (PCA). PCA compresses many features into two-dimensional data. This is useful when you want to examine the variation of the data. Shown in Figure 9, after constructing the PCA from all the AUs, I determined through the time series analysis of the PCA that the time series data of smiling is like the series for the PCA. This similarity of the graphs indicates AU06 and AU12 drive variance shown in the PCA, I.E., smiling is important to the overall movements in the data.

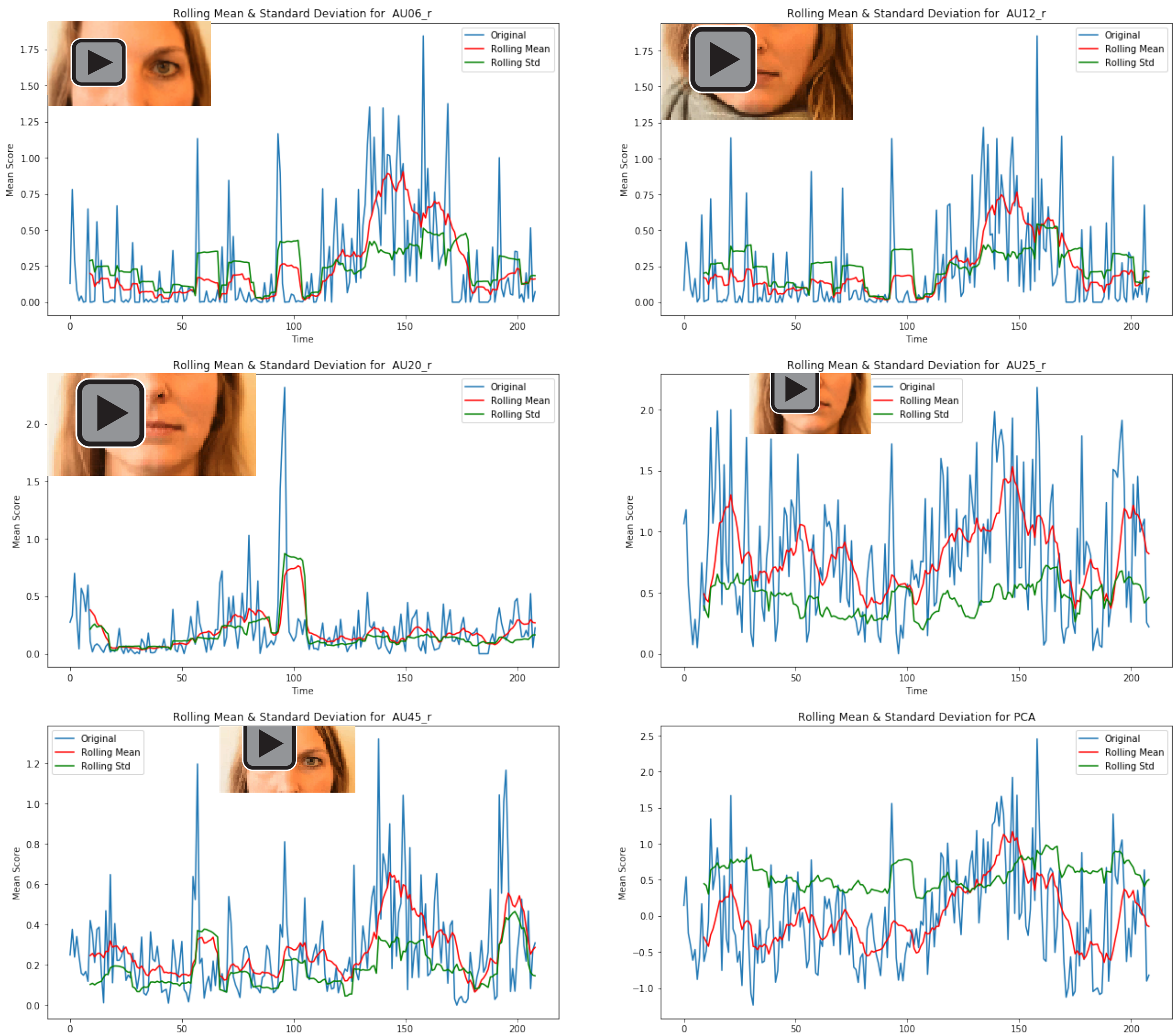

Figure 11: five primary action units in the interaction and the principal components of Mar's data. GIFs from iMotions visual library: https://imotions.com/blog/facial-action-coding-system/

To test the prominence of these AUs, I ran a PCA. I find that the two main drivers of the engagement are AU06 and AU012. By following the top two graphs, we can see initial engagement with the learning activity and then, just after time 100, a major ramp up in engagement. The PCA shows that these results were the major contribution to the analysis of the interaction. Notably, we could use AU25, controlling for smiles, as a proxy for talking.

Finally, I identified the affective pathways that occurred prior to moments participants elaborated in the CDM. To recall, in Leinhardt et al., elaboration is learning. Using this pragmatic definition, I find that in the 30 seconds directly before the moments participants elaborated during the one-hour interview, the most common affective state pathways were delight-contempt-contempt, disengaged-contempt-delight, and disengaged-delight-delight. Afterward, the sequences were delight-delight-contempt-delight, and disengaged-disengaged-delight-delight.

These were not common sequences in the data. When I examine the data, I find the most common pathways were delight-delight, contempt-contempt, delight-delight-delight, delight-delight-delight-delight and contempt-contempt-contempt-contempt (☹). As summarized in Table 2, of the total affective identifications in the one-hour interview, only 22 show the delight-contempt-delight combination.

Table 3: Counts of Pathways

Pathway	Obs	Pathway2	Obs
Delight-Delight	9,843	Contempt-Contempt	903
Delight-Delight-Delight	9,830	Contempt-Contempt-Contempt	897
Delight-Delight-Delight-Delight	9,817	Contempt-Contempt-Contempt-Contempt	891
Delight-Delight-Delight-Contempt-Contempt-Contempt	2	Contempt-Delight-Delight	10
		Contempt-Contempt-Delight	10

8.2.2 Lessons Learned from Mar’s Case

What can we learn from physiological measures of people’s affective states while they engage with an informal learning environment? In studying Mar’s case I see two new insights.

Delight was the predominant emotional pathway, accounting for 29,490 of the pathways around learning moments. The next two important pathways were contempt and contempt-delight pathways, with 2,713. Delight-to-contempt or contempt-to-delight pathways also constituted an interesting subset of 22 moments. Second, as seen in Figure 9, the most important component of the PCA in the interaction when looking holistically are the components of a smile, action unit 12 (AU012) and action unit 6 (AU06).

In summary, smiling constituted a large part of the interaction and delight surrounded most moments of learning. As seen in Figure 8, the strongest correlation between any two affective AUs (r 0.92) was between AU 06 and AU 12. AU 25 and AU 26 were next most highly correlated (r .47). The latter are Lips Apart and Jaw Draw, respectively, which are a proxy for talking. That I can track talking and see when participants are engaged is a benefit of affective gesture analysis.

In terms of Mar’s knowledge building, I see the following aspects: first, Mar elaborated functions while she worked with Ant Adaptation. The three functions together describe an understanding of how an ant colony can collectively feed itself and grow the population, without the need for a central controller. She explicated the rules of how ants group up, how ants pick up and carry food, and how the rules of how ants walk drives their ability to find new food, and exploit food patches they have already found using pheromone trails. This function expanded on her prior knowledge about ants by giving her a way to explain what she knew before playing the game: that ants walk, carry things, and group up. In short, I have evidence that Mar happily learned how ants self-organize their own colony to grow the population and integrate into a complex ecology.

8.3 Case 2: Bob and Mel develop a wide understanding of ants' behavior

Bob and Mel are undergraduate engineering students. Bob is a 22-year-old African American woman. Mel is a 22-year-old Latinx woman. They both have a strong engineering background and love to code. I interviewed them in person, while using a touch enabled version of Ant Adaptation. This was done between iteration 2 and 3, and was implemented to gather pilot data on the third iteration. As such, they stood side by side over a large touch display with Ant Adaptation on it.

During their 53 minute and 52 second interview, Bob and Mel expanded on their extensive prior knowledge about ants, but reduced the number of entities they use to explain it. During the pre-interview, they discussed 42 elaborations, but in the post interview they mentioned only 23, a reduction of 19. This reduction occurs sixteen minutes into the interview when they re-organize through a short discussion to 'summarize what they've learned." This shows key moments of knowledge architecture are an important process in knowledge management. I find, in moments participants elaborated, discussion serves a vital role in synthesizing ideas.

During the pre-interview, as shown in Transcript 5, Bob and Mel discussed their prior knowledge of ants, covering topics such as how they carry food, their strength in respect to body size, if and how they reproduce, and whether larger ants naturally expend more energy.

Time	Speaker	Question	Quote	Interaction Type	Node	Parent Node	Node Depth	Elaborated	Hypothesized
0:32	Mel	Have you guys ever noticed anything about ants	Have I ever noticed anything?	Pre-Question 1	Ants		1	1	1
0:35	Mel	Have you guys ever noticed anything about ants	About ants. They hurt when they bite you? Or like behavior wise? I don't know.	Pre-Question 1	Bite	Ants	2	1	1
0:40	Emi		anything.	Pre-Question 1				0	0
0:41	Bob		They're obnoxious.	Pre-Question 1	Obnoxious	Ants	2	1	0
0:42	Mel		They come in two different colors	Pre-Question 1	Two-colors	Ants	2	1	0
0:44	Bob		because they like	Pre-Question 1				0	0
0:48	Bob		cuz there's usually a lot of them at one time and like once they enter a space they just keep coming unless you spray Raid. Then they stop.	Pre-Question 1	In groups	Ants	2	1	0
1:00	Kit		How do you think ants collect food?	Pre-Question 1	Collect Food	Ants	2	1	1
1:06	Mel		I've seen them like carry breadcrumbs before. Like, I don't know if he like I've never seen like the actual ant I just seen like the breadcrumbs or something like that. So I assume they're like carrying it.	Pre-Question 1	carry breadcrumbs	Collect Food	3	1	0
1:18	Bob		I don't know	Pre-Question 1				0	0
1:19	Bob		I think I have seen an ant carry a leaf, is that a thing?	Pre-Question 1	Carry leaf	Collect Food	3	1	1
1:22	Mel		Oh, I've never seen that.	Pre-Question 1				0	0
1:24	Bob		They're kinda strong. But you know, I have to get there people things that eat. I mean, I know they eat but I don't really think about it.	Pre-Question 1	Strong	Ants	2	1	0
1:31	Kit		fair. So how do you say that? How do you think ants know what to do?	Pre-Question 2	Know What to Do	Ants	2	1	1
1:39	Mel	How do Ants know what to do	I assume someone like tells them maybe like they watch other people doing iy. other people. other ants [laigh]. Like, I don't know where ants come from. Are they like born? like when they're born, I'd imagine they probably get	Pre-Question 2	Through Watching other people	Know What to do	3	1	1
1:54	Bob		How do they multiply?	Pre-Question 2	Multiply	Ants	2	1	1
1:55	Mel		Yeah, I don't know. But like...	Pre-Question 2				0	0
1:57	Mel		Do they have sex [whisper]?,	Pre-Question 2	Have Sex?	Ants	2	1	1
1:59	Kit		That's a good Question.	Pre-Question 2				0	0
2:01	Mel		I guess you do like when they're too small to like do other stuff, they probably like get fed somehow. And then once they're like big enough to do stuff	Pre-Question 2	Ants Grow?	Ants	2	1	1
2:09	Bob		Do they grow?	Pre-Question 2	Grow?	Ants	2	1	1
2:10	Mel		I don't know, that's the point!	Pre-Question 2				0	0
2:12	Bob		just always like the same size.	Pre-Question 2				0	0
2:13	Mel		Do they they just spawn? [all laugh]	Pre-Question 2	Spawn	Ants	2	1	1
2:19	Kit		Alright, so the answer on ants how ants know what to do? These were all thoughtful...	Pre-Question 2				0	1
2:24	Bob		I feel like they were born,	Pre-Question 2	were born (inate)	Ants	2	1	0
2:25	Mel		they just watch other people	Pre-Question 2	Just Watch other people	Ants	2	1	0
2:27	Bob		ants	Pre-Question 2				0	0
2:28	Mel		other ants. You know what I mean...	Pre-Question 2				0	0
2:31	Bob		Also I feel like we're colored really well for this kind of thing. You're like wearing red	Pre-Question 2				0	0
2:34	Mel		Yeah, that's true.	Pre-Question 2				0	0

Transcript 5: Bob and Mel discuss what they know about ants.
Figure 10 is a CDM map that includes all the information they talked about.

During the pre-interview, I gathered their understanding of ants. In that interview, they were asked if they had ever noticed anything about ants. As seen in Figure 10, they have 42 elaborations; for instance, that ants bite, and they're obnoxious, that ants come in two colors, and they have groups. Bob noted, "because there's usually a lot of them at one time and like once they enter a space that just keep coming unless you spray raid, then they stop." I then asked how you think ants collect food. And Mel at 1:06 noted her prior understanding: "I've seen them like carry breadcrumbs before, like I don't know if you like ever seen like the actual ant I just

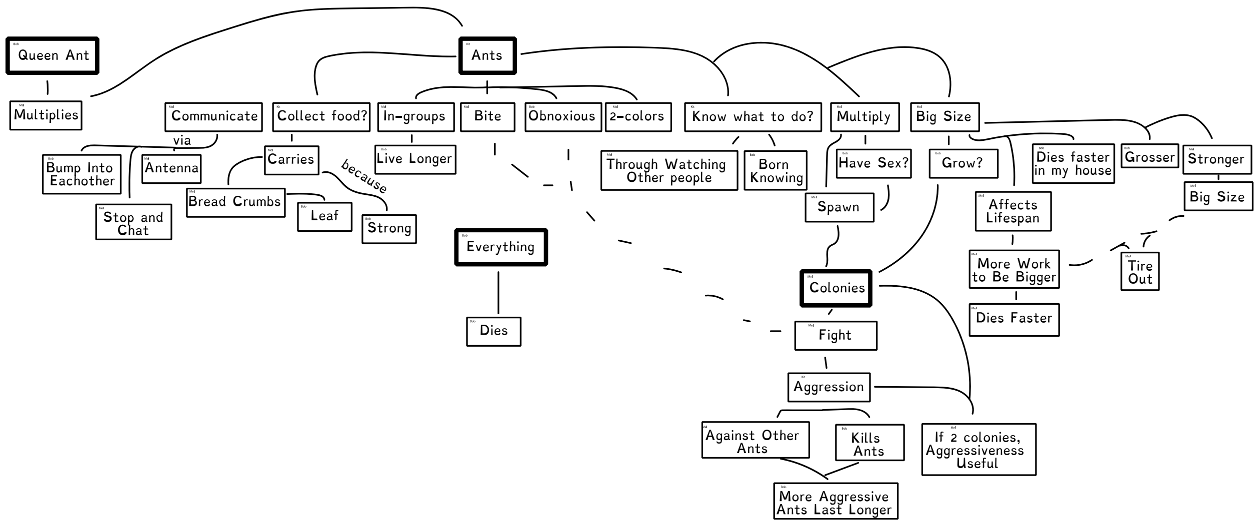

Figure 12: Pre interview, Mel and Bob Prior Knowledge about ants

have seen like the breadcrumbs or something like that so I'm assuming they're carrying it." And Bob said she didn't know, she thinks she's seen them carry a leaf, and wondered, "Is that a thing?" They both noted that ants are strong, and Bob said she didn't know exactly what they eat. because she's never really thought about it. Using CDM over the entire pre-interview, it becomes more obvious that they have a very wide understanding with four 1st order elaborations, and thirteen 2nd order of the 42 total moments they elaborated. Through the pre-interview they mention that ants communicate via their antennas and perhaps by bumping into each other, showing they have some prior ant biology knowledge. And Mel noted, "They may stop and chat." Mel noted that the ants carry the food. And Bob mentions a leaf, and synthesizing that the carrying is happening because they're strong.

When I asked, 'How do ants know what to do?" After some back and forth, Mel responded at 2:25 of Transcript 5, that they know by watching each other (though she calls the ants Humans). Bob then corrects her that the ants watch other ants, not people. Bob later adds that ants have some sort of innate knowledge: "They're born knowing." Mel noted that ants also multiply. And Bob wondered, not distinguishing between real and digital ants, "Do they have sex?" And Mel concluded that they spawn. They also had a conversation about size and ants living in groups. Mel noted, ants live in groups, and Bob hypothesizes that that might make them live longer, at 4:10 of Transcript 6, saying "I think it's I feel like they would be will live longer because like there's more of them. I don't know like." When I asked them how size affects their life. Mel said, "they have big and small size," and Bob asked, "do they grow, or do they start fully formed?" Mel wondered how it affects the lifespan and thought about how more work leads to being bigger, or perhaps being bigger makes them die faster. And a big size also contributes to them being grosser. I see here a lot of prior knowledge, and some questions about how an ant's life works.

Along that trend of thinking about size, Bob and Mel also thought about how an increase in size could lead to an increase in the ants' strength. They then hypothesized a connection between more work and being bigger, and between bigger size and strength, and wondered if bigger ants tire out faster, inserting a connection between size and energy use. The participants also already knew that ants from different ant colonies fight each other. When I asked about how aggression might affect their life, they hypothesized that "more aggression, could make the ants last longer." And Bob wondered how that would lead to killing ants. They conclude from this idea of colonies, and more aggression that if there's two colonies, aggressiveness is useful. This kind of wide understanding drew on the previous understanding of humans and their observations of ants to conclude a series of concepts about ants. As seen in Transcript 6, this statement was made at 4:05, which then kept going in the conversation about size, grouping and mortality.

Time	Speaker	Question	Quote	Interaction Type	Node	Parent Node	Node Depth	Elaborated	Hypothesized
3:22	Mel	How would an ant's size affect his life?	Well, it depends on where it is probably because a bigger ant is probably more likely to be killed in my house. So like it will have a shorter life. But like a smaller and my house might be able to get away with it. Depending on how	Pre-Question 4	Killed in my house	Size	3	1	0
3:38	Bob		how many of them they are there are I feel like size makes them more intimidating like eww gross like higher gross factor. But again, if it's one big ant I mean, I'm probably still gonna kill it. I feel like I'm going to kill it regardless of its size, but the sizing bigger makes it easier for me to kill,	Pre-Question 4	Higher Gross Factor	Size	3	1	0
3:58	Mel		but also like in terms of lifetime. Like, it's like being big....	Pre-Question 4	Affects Life Span?	Size	3	1	0
4:01	Bob		I mean, if a human doesn't kill it, then	Pre-Question 4				0	0
4:05	Mel		Probably	Pre-Question 4				0	0
4:05	Bob		Oh yeah, I mean everything dies.	Pre-Question 4	Dies	Everything	2	1	0
4:07	Mel		I don't think size affects that though. Like, I think they have the same lifespan	Pre-Question 4	Does not affect lifespan	Size	3	1	0
4:10	Bob		Really? I think it's I feel like they would be will live longer because like there's more of them. I don't know like,	Pre-Question 4	live long	Grouping Together	3	1	1
4:18	Mel		but wouldn't that make it look.... like its more work to like, be bigger isn't	Pre-Question 4	more work	Size	3	1	0
4:22	Bob		I suppose, its a good way to look at it	Pre-Question 4				0	0
4:27	Bob		Or they are just stronger	Pre-Question 4	Stronger	Size	3	1	0
4:29	Mel		so if anything, the bigger ones die faster, if anything is	Pre-Question 4	Die Faster	More work	4	1	0
4:33	Bob		If anything but they're also stronger so they can do more?	Pre-Question 4	Do more	Stronger	4	1	1
4:35	Mel		I guess the tire themselves out more.	Pre-Question 4	Tire out	Do More	5	1	0
4:39	Mel		So like it's just like there's more factors going into the big ones	Pre-Question 4	More Complicted	Size	3	1	0
4:42	Bob		you saying small people last longer?	Pre-Question 4	Smaller lasts longer	People	2	1	1
4:44	Mel		Yes.	Pre-Question 4	Smaller lasts longer	People	2	1	0

Transcript 6: Bob and Mel Discuss size and how it effects life-span.

In these interviews, ideas happened rapidly, as the two participants bantered, and got to understand the Microworld. I affirmed this quality during the interview when I transitioned and introduced them to play Ant Adaptation. I said, "You guys are hypothesis generating machines. I love to keep that going. Now I'm going to do what I said I was gonna do. As I was saying, you have some choices in this game, how you set these determines the kind of the size and the aggressiveness of the ant when it's born, not the ant you use currently, but the new ones. When you change this [slider], that's how much energy they have made it like a gas tank when they're born, how much they have in their gas tank before they need to feed themselves, so they can go further. You have a couple options in this game you can restart the game whenever you want, but probably talk to a co-partner." After this short introduction they started playing.

As shown in Transcript 7, during the beginning portion of the gameplay, Mel and Bob took this wide understanding, and re-organized it through interaction with Ant Adaptation, under the rubric of the game itself. They came to understand a new entity important to the functioning the ant colony, their population: at11:24 noting how many of each kind of ant there are. They then wonder about how the ants snatching food leads their ants dying.

Time	Speaker	Question	Quote	Interaction Type	Node	Parent Node	Node Depth	Elaborated	Hypothesized
10:24	Bob		That is so interesting	Gameplay				0	0
10:24	Mel		So we can just do this	Gameplay				0	0
10:25	Mel		what's interesting Bob?	Gameplay	Block ants	Flowers	2	1	1
10:27	Bob		that like the flowers block them?	Gameplay				0	1
10:28	Mel		They also can't get back in right? Or Are they getting back in	Gameplay				0	1
10:31	Mel		No they're going in and out. So wait so then so we owe the flowers like got killed over here too.	Gameplay				0	0
10:38	Bob	Did they eat flowers?	Oh, did they eat it?	Gameplay				0	1
10:39	Kit		Did They eat what?	Gameplay				0	1
10:41	Mel		flowers?	Gameplay	Get eaten by ants	Flowers	2	1	1
10:43	Mel		Oh look the petals are missing here. So they like probably like deactivate them or something like they're like, I don't know the computers deactivating the flower's or petals or like whatever?	Gameplay	Missing petals	Flowers		1	1
10:53	Bob		Yeah, no over here. Oh, I just saw them pick it up.	Gameplay	Pick Up	Ants, Flowers	2	1	0
10:56	Mel		Oh you did?	Gameplay				0	1
10:56	Bob		Yeah. Oh, wait.	Gameplay				0	0
10:57	Mel		Oh yeah, it disappeared.	Gameplay	Disappeared	Pick up	3	1	0
10:59	Bob		It's like oh, little yellow dot See? Mr. Queen, Ms. Queen.	Gameplay	Turns into yellow dot	Disappeared	4	1	1
11:04	Mel		So what were we doing them a service or disservice by like adding all these flowers here they look like they're like	Gameplay	Service	Ants		1	0
11:10	Bob		potentially a service but if we got rid of flowers...	Gameplay				0	0
11:14	Kit		Kit What are you doing?	Gameplay				0	1
11:16	Kit		Good. Let's go ahead.	Gameplay				0	0
11:19	Bob		If we got rid of flowers then they would die.	Gameplay	Get rid of flowers ants die	Ants, Flowers	2	1	0
11:22	Mel		You'd think so? Kit can we liek take them out?	Gameplay				0	1
11:24	Bob		wait though Okay, there are only like 12 Red Ants and there is 27 black ones so maybe something about the flowers being in the way is a problem?	Gameplay	In the way, relative size	Flowers, Population	2	1	1
11:33	Mel		what if wee?	Gameplay				0	1
11:34	Bob		Yeah, like, why is there less of you? Oh, we're taking from your flower stash. Oh, look at that.	Gameplay	Taking from Flower Stash	Ants	2	1	1
11:40	Mel		Yeah, I guess they can like come out and like snatch it and then like the red ones are just like,	Gameplay	Snatch	Taking from Flower Stash	3	1	0
11:45	Bob		It looks like it looks like a mission.	Gameplay				0	0
11:48	Mel		These ones are dying or no	Gameplay	Dying	Ants	2	1	0
11:49	Bob		they are. wait no.	Question the unseen				0	0
11:50	Mel		We can't Change that though.	Question the unseen				0	0
11:52	Bob		Yeah, I mean, we could	Question the unseen				0	0
11:54	Mel		Where is where is the other one here? Where's the the black queen?	Question the unseen				0	0
11:57	Bob		I don't know. I haven't seen it.	Question the unseen				0	0
12:00	Mel		So many questions that need answers.	Gameplay				0	0

Transcript 7:Bob and Mel organize their understanding of how ants collect food by engaging with the microworld.

They go on to connect the number of ants to the food source. As seen in Figure 11, while playing the game, they came to understand that ants fight between the two colonies, and that this leads to colony death. They state that colony death can happen for a couple of reasons: for instance, in Transcript 8, they discover ants could lose their way and die.

12:41	Mel		Who killed the black ants?	Interogative	Vinegar	Population		2	1	1
12:43	Bob	Does Vinegar Kill ants?	Was it the vinegar? Here? Wait, let me add vinegr	Interogative					0	1
12:46	Mel		because they couldn't find their way back	Interogative	Lost their way	Dying, Population	3,3		1	0
12:47	Bob		Can you add vinegar?	Interogative					0	1
12:48	Mel		Oh, yeah, sorry.	Interogative	Vinegar does not kill	Dying		4	1	0
12:49	Bob		Yeah, you got let me see. Do I kill ants with this? Oh, no,	Interogative					0	0
12:52	Mel		you don't kill them [by adding vinegar] but like, I think it's like they have to find their way back or something. See cuz these guys get like oh, See, but they didn't do that going back like it'd be more useful for them to leave the trail. That way they can follow it back.	Gameplay	Follow trail back	Chemical		2	1	0

Transcript 8: Bob and Mel discover ants losing their way can kill them.

They go on to find a scarcity of flowers could kill them. By either wandering around too much or not having enough food to eat, the ants could reduce their population, Mel argued, the ants also move, and this is a new line of thought that the direction or the way the ants moved was important somehow. They weren't talking about moving before the game, when they were only thinking of a static representation of ants. Mel introduces this idea of moving, and ants going in and out of the colony. And as they're going in and out of the colony, they could create signals everywhere they go, which would create local information the ants' colony could self-organize around. But then they elaborate the idea, adding chemicals that could disappear over time through evaporation, and in the representation of the game, the chemical has white parts where it's strongest and pink parts where it's weaker, and this gradation of chemical strength allows ants to follow trails back to the nest.

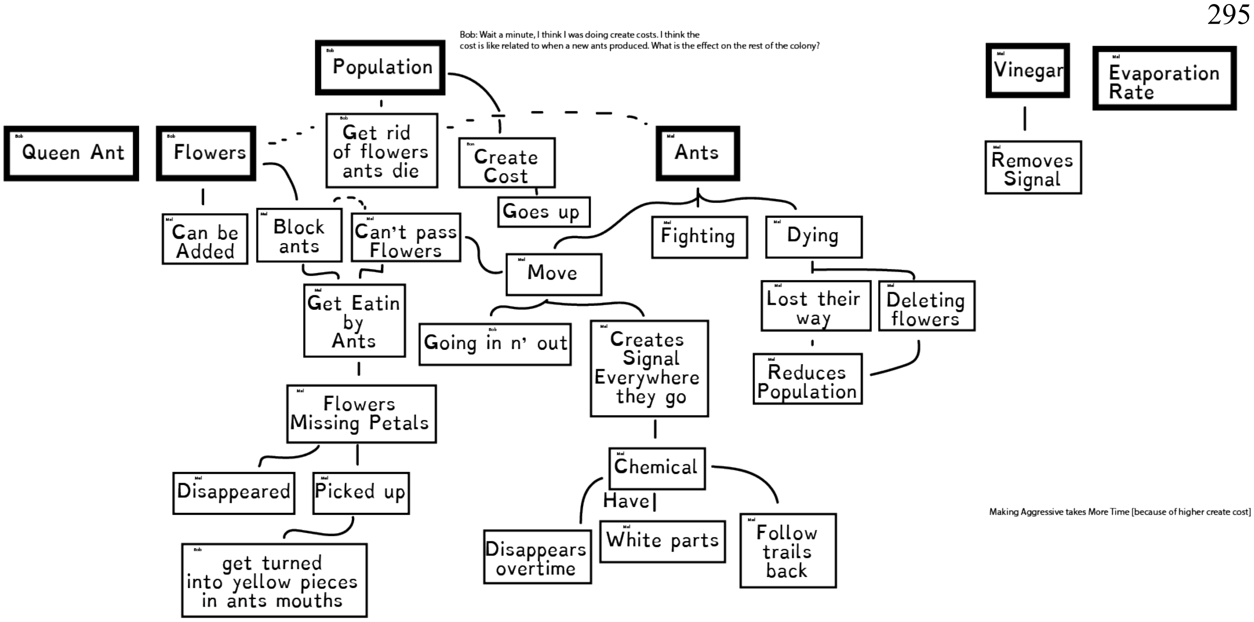

Figure 13: Early Game Play, Mel and Bob come to refine thinking with Agent Based Model

They elaborate after noticing a new ontological entity: flowers represent the food source in this game. Because flowers can be removed by the ants as they forage for food, they have a blocking property: the ants get blocked by the flowers; when they run into the food they return to the nest. So, the flowers act as a wall. Mel, discovering this, said, “They can't pass the flowers." But these flowers are important because they get eaten by ants. And we can see this in the representation, Mel argued, because the flowers begin to have parts go missing, and eventually they disappear. But as the flowers disappear, they get picked up by the ants. And Bob noticed they get "turned into yellow pieces in the ants’ mouth," which they bring back to the colony. This cycle created a connection from the flowers to the ants' population, they noticed. Bob found that the population is affected by getting rid of flowers: more flowers, more population. Less flowers, less. In this way, the participants connected the micro interactions of the flower and ant agents to this macro concept of how many ants there were (the colonies population). Bob noticed that the population number was very much dependent on the “create cost,” which is displayed in Ant Adaptation as a number that designates the food “cost” to create a new ant and is one of the game details that participants can modify—as the create cost goes up, the population goes down. The participants also noticed a couple of other entities, but hadn't yet elaborated them: they saw a queen flying around, which they had previously thought of only as a bio factory to make more ants. They also discover two other things: first, they notice vinegar removes the signal, and second, they start thinking about evaporation rate. If the ants are following the chemical back to the nest, the evaporation of those trails became important to them.

During this time, Bob asked "I think the [create] cost is…related to when a new ant is produced. What is the effect on the rest of the colony?" By tying together these different levels of the agent-based model, Bob started coming up with the complexity model. However, the group of two had not yet turned this into functions, as Mar had; instead, they noticed several things about the agent-based model. This led to the big breakthrough, a time of reflection, which seemed to come from earlier educational practices they were both familiar with.

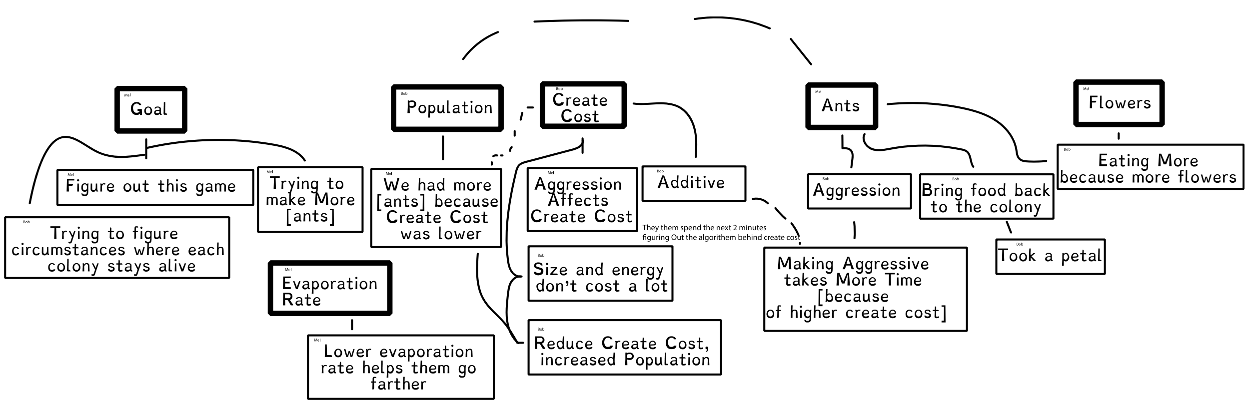

Figure 14: A goal setting moment, Mel and Bob recap what they've learned so far.

As shown in Transcript 9, sixteen minutes into the gameplay, they took a moment: they started recapping what they had learned to understand what the goal of the interaction was. Bob and Mel synthesized their knowledge by recapping what they knew.

Time	Speaker	Quote	Interaction Type	Node	Parent Node	Node Depth	Elaborated	Hypothesized
16:03	Bob	I'm gonna pause okay [breaks to deliberate].	Gameplay				0	0
16:08	Bob	I don't know how much we figured out.	Gameplay				0	0
16:10	Mel	nothing.	Gameplay				0	0
16:11	Bob	Maybe the goal is to figure out this game.	Recap What they've Learned	Figure out this game	Goal		1	0
16:12	Mel	We figured it out the flower thing	Recap What they've Learned				0	0
16:14	Bob	Yes,	Recap What they've Learned				0	0
16:15	Mel	the evaporation rate is the lowest how they the trail evaporates.	Recap What they've Learned				0	0
16:19	Bob	Also, we figured out that like a chemical to another colony actually does something a bit like chemical to anything else doesn't really do anything. They don't care. They're going to naturally go towards flowers. But if we want new ones that come out of a colony to go towards the right one so basically I think chemical only helps ants that are coming out. Not ones that are already out.	Recap What they've Learned				0	0
16:41	Bob	Uhhhmmm	Recap What they've Learned				0	0
16:44	Bob	also Yeah, flowers are important. Flowers block. So like and dying happens, but they take the pedals. Yeah, they do. super interesting. So it's like but they blocked but they also like a sourcing or something. also create cost is interesting so I feel like for every ant that's big like takes more I don't know what like cost we don't have maybe it takes longer is that seconds?	Recap What they've Learned				0	1
17:12	Mel	Oh, but we definitely waited longer than 14 seconds for everything to die and nothing spawns and nothing dies.	Recap What they've Learned				0	0
17:19	Bob	That's true.	Recap What they've Learned				0	0
17:21	Bob	Also, I don't know where the big guy went.	Recap What they've Learned				0	0
17:24	Bob	I don't know if he died. He could have died. Maybe he's weaker because he's bigger.	Recap What they've Learned				0	0
17:30	Mel	But nothing... but See, but they're still alive.	Recap What they've Learned				0	0
17:33	Bob	But he is a tiny one. yeah,	Recap What they've Learned				0	0
17:35	Mel	he was like a regular guy	Recap What they've Learned				0	0
17:36	Bob	oooh I think the aggressive one's not as good as we think.	Recap What they've Learned				0	0
17:40	Mel	Well, maybe the big aggressive one.	Recap What they've Learned				0	0
17:42	Mel	Maybe Small aggressive ones.	Recap What they've Learned				0	0
17:47	Mel	It's like pent up.	Recap What they've Learned				0	0
17:49	Mel	I'm like projecting all of my human things onto ants	Recap What they've Learned				0	0
17:56	Bob	I mean I feel like there are some issues we need to talk about after this	Gameplay				0	0
18:00	Bob	Okay, so what do we want to try	Gameplay				0	0

Transcript 9: Bob and Mel synthesized their knowledge by first, recapping what they knew.

In Transcript 10, they conduct rapid testing of their ideas in interaction with the microworld with breathless speed, asking and answering their own questions.

Time	Speaker	Quote	Interaction Type	Node	Parent Node	Node Depth	Elaborated	Hypothesized
18:01	Mel	We want to take control	Gameplay				0	0
18:03	Bob	this true. I think we could click that	Gameplay				0	0
18:06	Bob	a refund. Oh ticks, what's a tick?	Gameplay	Ticks		1	1	1
18:10	Bob	normal speed?	Gameplay				0	1
18:13	Bob	Faster. Oh, is it faster? slower? Do it fast	Gameplay				0	1
18:17	Bob	too fast. A hide controls and then restart. I think everything's gonna. Okay start faster zero ticks. We started zero.	Gameplay				0	0
18:26	Mel	Yeah, but that's how long it was like that's how many ticks it is?	Gameplay				0	1
18:28	Bob	Oh, you right. All right, do we wanna? What do we want do. Wait, do we want to change my ant?	Gameplay				0	1
18:34	Mel	Yeah, make it small and aggressive.	Gameplay	Size, Aggression	Ants	2	1	0
18:37	Bob	Okay, small aggressive. Do we want the start energy to be tons or are we building not care?	Gameplay				0	1
18:44	Mel	Well the create cost is 10. So what if we just make really aggressive ants? They're small and they don't have that much energy. So like,	Gameplay				0	1
18:52	Bob	They they are still mad about everything. That's interesting, aggressive, really ups this create cost. Because like I put energy all the way like low	Gameplay	Aggression Ups	Create Cost	2	1	0
19:00	Bob	Mines 9	Gameplay				0	0
19:02	Mel	Oh, put that back one more	Gameplay				0	0
19:06	Bob	But you can only go to three centimeters	Gameplay				0	0
19:16	Unk	red ones	Gameplay				0	0
19:19	Bob	are stronger. are like substantially more aggressive. Also Why is my guy so big	Gameplay				0	0
19:26	Bob	naturally more?	Gameplay				0	1
19:27	Kit	If you want to change the guy to be small press restart again.	Gameplay				0	0
19:32	Bob	Oh, oh eww	Gameplay				0	0
19:35	Mel	but they're very aggressive all the little things for their high aggression. Let's see what see what happens. Woah	Gameplay				0	0
19:43	Bob	Woah	Gameplay				0	0
19:43	Mel	they immediately went on the attack and also this is a lot faster let me say faster. Yeah that's true. Okay so four ants	Gameplay	Go on the attack	aggression	3	1	0
19:52	Bob	I'm making more	Gameplay				0	0
19:53	Mel	like they make paths to...	Gameplay	Make Paths	Chemical	2	1	0
19:55	Bob	You know your three, I am six so maybe	Gameplay				0	0
20:00	Kit	What do you think it's impacting it?	Gameplay				0	1
20:07	Mel	everything else is the same except besides	Gameplay				0	0
20:11	Bob	well not technically because your size is smaller a centimeter and your body	Gameplay				0	0
20:14	Mel	Yeah,	Gameplay				0	0
20:15	Bob	yeah. and your energy's different.	Gameplay				0	0
20:18	Bob	Your energ...	Gameplay				0	0
20:19	Mel	Oh yeah, my start energy is high. Whoa, and aggression is higher. And the size is higher on the red ones.	Gameplay				0	0
20:26	Bob	I'm going to slow this down.	Gameplay				0	0
20:27	Mel	Yeah, good idea.	Gameplay				0	0
20:28	Bob	I can't figure I don't	Gameplay				0	0
20:30	Mel	Oh, that's good. This is a good one. You're kind of watching	Gameplay				0	0
20:34	Bob	Oh my god, there's a big	Gameplay				0	0
20:35	Mel	ohh Wait,	Gameplay				0	0
20:37	Bob	that didn't happen last yeah Well, why didn't we have a queen?	Gameplay				0	1
20:44	Bob	So I think we're gonna die.	Gameplay				0	0
20:46	Mel	Wait my colony me Oh, it did die.	Gameplay				0	0
20:54	Mel	Haha.	Gameplay				0	0
20:57	Mel	Red is Victoria.	Gameplay				0	0
21:00	Mel	I don't think that's the point.	Gameplay				0	0
21:02	Bob	No, I really think it is	Gameplay				0	0
21:02	Mel	I'm starting to think that the point is that we're supposed to be able to answer those pretest questions	Gameplay	Answer Pretest questions	Goal	2	1	1

Transcript 10: Bob and Mel rapidly test their ideas with rapid fire questions.

This interaction laid out in transcript 9 and 10 led to synthesis of their knowledge: their CDM shrank, from 42 items in the pre-interview to 23 in the post, showing they organized their knowledge through active knowledge architecture. This is important, because Ant Adaptation is designed to be a microworld, so this moment incapsulates a goal setting moment, important in any team, that improved their understanding through simplification. Here, I see the moment their learning is crystalizing. As seen in Figure 12, this sequence started with them talking about their goal. Mel asked aloud what their goal was, hypothesizing that perhaps "the goal of this interaction is to figure out this game." But Bob tied it back to the life histories of ants, arguing their "goal was trying to figure out the circumstances where each colony stays alive," and Mel simplified this to: trying to make more ants. They also then simplified and connected the goals to the representation. They saw that create cost impacts the population. Mel argued that the pair had more ants because create cost was lower. They theorized, as a create cost to make one more ant drops, the population goes up. And they also noticed the connection between aggression affecting the create cost. Once they had this connection, they spent the next two minutes figuring out the algorithm behind create cost, experimenting with different slider parameters, moving them up and down. Bob noted, "This is what happens when you have engineers play a game." An exciting performance of her identity as an engineer. They summarize this by saying the create cost is additive, by changing the sliders up or down to raise and lower the create cost. They noticed that size and energy don't cause the create cost to increase by much, and concluded, "reduce create cost, increase the population." Thus, they're making a new synthetic connection between population and create cost. These new connections led to a reduction in the size of their CDM. The ability to observe this moment and count their named ontological entities is a key affordance of CDM. CDM affords the ability to see their knowledge expand and simplify as they engage in the learning environment.

They noticed that ants being aggressive takes time because the higher create cost slows down population growth. In other words, they also noticed how the adaptations serves as a dampening effect on the increase of the agent-based model.

Next, Bob and Mel noticed the connection between ants, the population, and flowers. Bob noted, “[Ants are] eating more because there are more flowers," and then pointed out the functional mechanism evaporation rate. Meanwhile Mel noted that "lowering the evaporation rate helps ants to go farther because the trails are more useful.” The summative moment, starting 16 minutes in, really galvanized the creation of their own knowledge, as they recapped what they knew and shared it between each other as a group reduced the size of the CDM from 42 to 23 (a 54.8% reduction). This is notable because I did not see this kind of seminal moment when Mar was playing on her own with just the interviewer. This is significant for two reasons. First, as educators, we should try to build these sorts of recapping moments into the use of microworlds. Second, this moment is interesting from a methodological perspective on the use of CDM, as their way of talking about the interaction radically shifts, CDM demonstrates this change better than simply reading the transcript.

As shown in Figure 13, in the post interview, this recapping became very important as they summarize what they knew. They came to understand what queen ants do: Queens produce more ants. And when I asked them again how ants knew what to do, they still talked about inherent behavior as Bob had mentioned before, but they also talked about the way paths work to lead them to where they need to go. Mel said, "They only make trails when useful, because of the evaporation rate, like how to go back to the flower to get more food." and Mel continue, "the trail created when they are coming back from a flower that creates a path to follow up for other ants." Bob summarized the thought by postulating that the "back and forth along trails [creates] a feedback loop." And further analogized, “it's like a compass, it's like breadcrumbs, showing us where to go in the larger world." They summarize the use of flowers in the model when Bob said, "flowers allow ants to eat." And then they brought in the bigger picture, they've done their recapping of the goal. The goal is, Bob says, "to keep parameters optimal for my ants," and Mel says, "[To] get colonies to an equal level experiment on green to see how they get equal with red." And Bob goes on, saying a "successful heuristic is moderate create cost to increase the population." These all summarize the complex system they have been experimenting with.

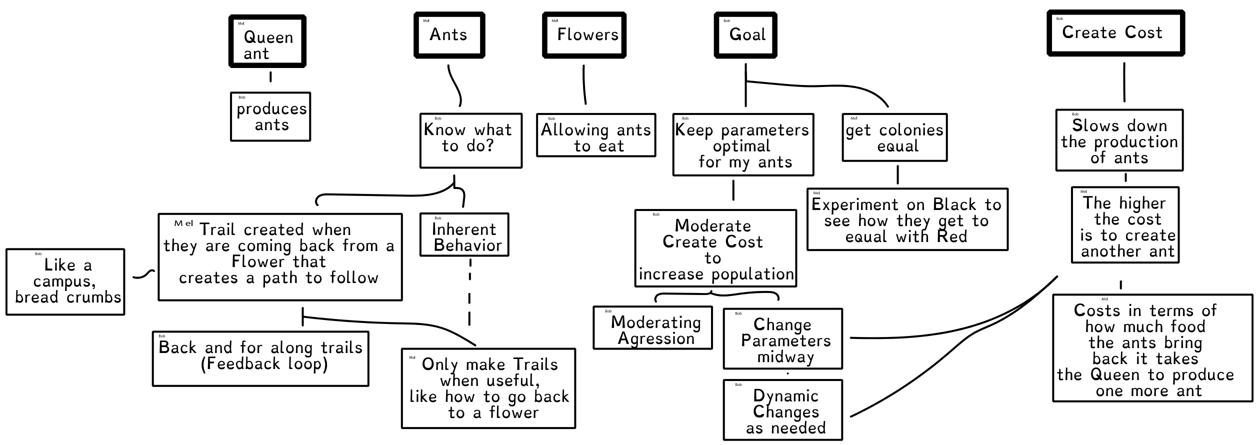

Figure 15: Post Interview, a systematic understanding of the dynamics of the complex system.

The pair of participants also stated that aggression should be monitored, and perhaps in a different ecological situation, ants may need different parameters, suggesting a player should change parameters midway through the game, creating dynamic changes in the system as needed for the health of the colony. This would be like role playing evolution, a key design idea of the game. In this way, they explain the role of random mutation in evolutionary history to keep up with selective forces. They also now have a better understanding of the algorithm behind create cost. They offer that the create cost slows down the production of ants, and that the higher the cost is to create another ant —the cost in terms of how much food the ants bring back it takes the Queen produce another ant. In this way, they have come to an understanding of how grouping of ants, ants' connection to flowers and queens, and the way they make trails, allows for a healthy colony. In other words, they have come to an understanding that adaptations in an evolutionary system assist organisms to persist in a changing ecology. They have also performed an idea of equilibrium—that is, the need for a moderate create cost to have enough ants, a moderate aggression to keep the population going, and the need to change the ants' adaptations' midway to allow for the dynamic changes needed for the health of the colony as the ecology changes. This they see through the paradigm of create cost, which both slows down the production of ants, but also may lead to a more successful colony depending on the environment.

8.3.1 Lessons Learned from Bob and Mel's Case

In terms of Bob's and Mel's knowledge building, I see the following aspects: first, the pair took a key moment during the interaction to reorder their knowledge, demonstrating a shift from what they brought into the interaction, and what interacting with the game engendered. The restructuring led to them explicating how ants know what to do, the role of flowers, and how ants interacting led to population growth. They explicated the rules of how ants know what to do through trail formation, and the connection between create cost and food intake. This allowed them to describe their different goals in the game. Bob wanting to optimize parameters in a changing ecosystem, role-played evolution in this system. Whereas Mel set herself the goal of discovering how to make the two colonies equal, expressing an equity lens. She set about running experiments to try to figure out how to achieve her goal. This different goal setting played a tacit, but crucial step, in their interactions. In short, I have evidence that Bob and Mel learned how ants can control their own colony, grow the population, and integrate into a complex ecology and that from this new information, the participants then set their own objectives in the system. When they became active architects of their knowledge, they synthesized their understanding. The CDMs of their interaction demonstrate they reduced the number of entities as they better organized their knowledge.

8.3.2 Mel and Bob's Affective State(s)

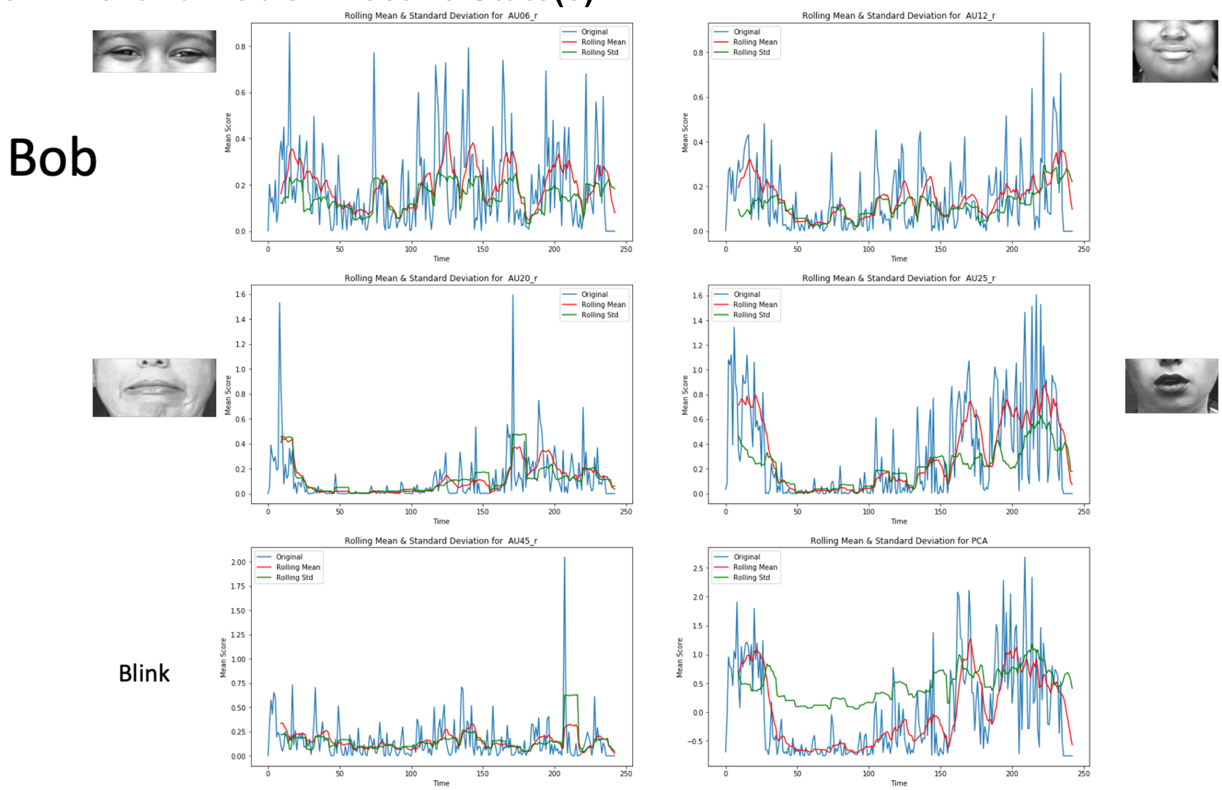

Figure 16: Bob Affective Engagement: AU06 and AU12 are constituent parts of a smile, AU20 is flat lip, AU25 is part of talking, and AU45 is blinking. The PCA shows more correlation with Talking than Mar's case, but still largely driven by smiling AUS.

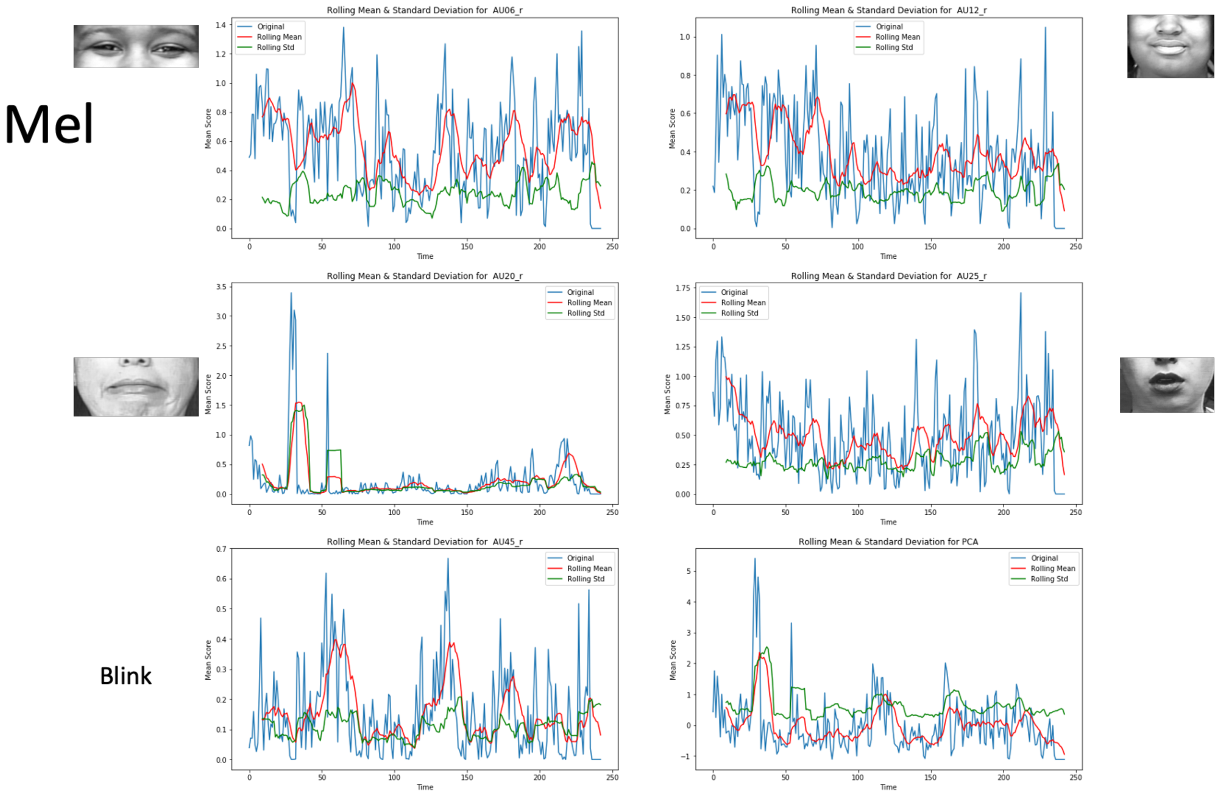

Figure 17: Mel Affective Engagement: AU06 and AU12 are constituent parts of a smile, AU20 is flat lip, AU25 is part of talking, and AU45 is blinking. The PCA shows more AU20 than Mar or Bob's case, but still largely driven by smiling AUS.

Bob and Mel played the game together over a touch screen, like the interaction in chapter one, but done in a university lab. While they played the game together, their affective engagement does not show much similarity. There are three features when comparing Figures 14 and 15 that outline the main features presented in Bob and Mel's interaction.

First, they both smiled within the .02-.14 range, but with Bob's average intensity lower than Mel's. I would describe this trend as a lower rate of smiles for Bob, which marginally picks up in the latter half of the interview.

Second, Bob spoke less in the first half of the interaction, with her AU25 showing very low amplitude between time 50 and 100. Then she began talking as much as Mel. Mel on the other hand spoke continuously

Third, Mel's PCA is impacted by one large contemptuous expression at time 25. The expression had a max intensity of 3.5. Besides that extreme value, the PCA is constituted by the smiling. Bob's PCA is driven by her smiling and talking. In other words, Mel engaged right away and spoke and smiled throughout the interaction. With the most noticeable feature in Bob's being the paucity of smiles or speech between time 50-100. In other words, we can see from Figure 14, that at first Bob did not engage as much, but approximately halfway through the interview, she quantitively smiled more and spoke more.

So, what can we learn from this data? With affective analysis, I can determine which affective gestures are most common and most rare in our data. In Bob's and Mel's cases, I observe Mel did most of the talking and smiling at the beginning of gameplay, but Bob began interjecting more about halfway through. I provide a range of a values for these: they smile between intensity scores of 02-.14. Additionally, I quantitatively determine when a participant begins to engage and discuss their absence from a conversation. From this I can ask: what was Bob doing while she remained less affectively engaged? While using transcript analysis, the tool itself affords noticing when a participant is speaking, analyzing affective engagement with time series analysis makes those silences apparent in the participation graphs.

Next, I attempted to do pathway analysis on Bob and Mel, but it provided inconclusive results. This was because, the method posed a methodological issue when jointly studying two participants. The data was of the two participants at the same time. It did not make sense to display the joint pathways. This issue was compounded by the synchronization of CDM along milliseconds, so when viewed this way, it became unclear whose affective pathways were being mapped to what moment of elaboration. Therefore, for group analysis this means my approach to pathway analysis is not sufficient. In future work, I update the methodology to better integrate pathway analysis and CDM. This limitation also led to the analysis shown next in the triangulation section.

8.4 Triangulating the Data

I draw insights from how much AUs can tell us about moments participants elaborated.

As seen in Figure 8 no AU is highly correlated with moments Mar elaborated. The same is true for Bob and Mel.

8.4.1 Triangulating

Beginning to answer question two, *What can we learn from physiological measures of people's affective states while they engage learning in informal learning environments, such as a museum exhibit or a learning game, where people choose to come and learn?* Recutting the video data with AUs, the time series analysis of physiological signal affords the ability to analyze the data from a new light: What was going on during spikes in the five principal components? For example, with the method I ask, looking at the peak in AU20, what's happening? To determine where the peak is, I find the max value of AU20.

Then I find the time stamp of the interaction and I find the important moment = 343.1 seconds (that is, 5.71 or 5:42 minutes). As shown in Figure 16, at that time, Mar prepares to take a drink of water, and thus flattens her lips. At the time, she's nonverbal, so reanalyzing the data at this moment does not grant any greater insight. Nonetheless, the approach allows us to investigate the situation from a new angle.

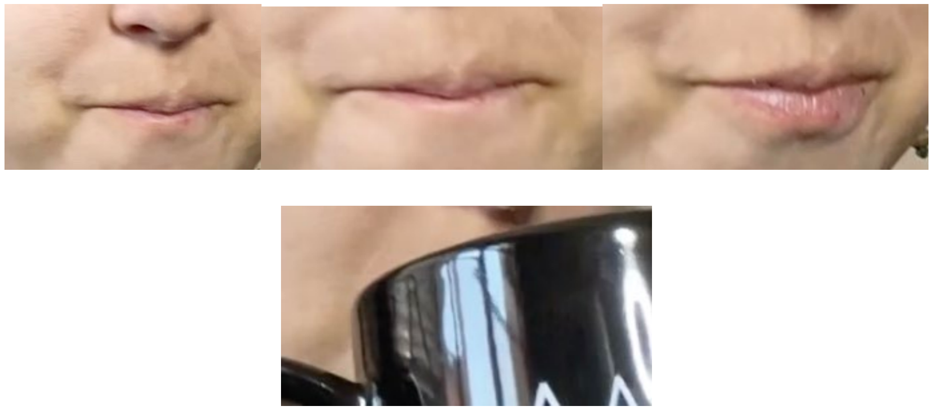

Figure 18: Mar takes a sip of water at the peak value of AU20.[19]

Mar's sip of coffee highlighted researchers need to pay attention: not all the moments expression analysis brings to light are salient. We do not want to get in the business of finding every sip of coffee. Therefore, instead of examining every single moment of expression, I built three predictive models.

8.4.2 Predictive Modeling

I used the moments participants elaborated determined in CDM as label data. I then use a train test split to cross validate, and find we can predict moments participants elaborated from AUs using random forest and its cousin XGBoost.

8.4.2.1 One Rule

Not all models perform as well. One Rule is not a good model for this data. As summarized in the confusion matrix of Bob and Mel, it does not predict 1s at all. Instead, it predicts 66% of the 0s correctly, and incorrectly predicts 34% of the 0s as 1s. The one rule model literally could not predict any 1s at all, which leads me to argue that that the relationship is just too complex for the one rule model. This means no single rule provides predictive power to sort moments participants elaborated from the rest of the data.

[19] *The IRB for this study allows showing participants face in scientific literature if it furthers the research goals.*

These findings indicate the relationship between AUs and elaboration is complex, and likely expressions are personal. If it was a simple relationship, I would expect the one rule model to have detected something. There would have to be something about the dimensionality of the data, that makes the data not amenable to a simple rule. In other words, the prediction is complex.

Thankfully, more complex models predict moments of elaboration more successfully.

Table 4: Confusion Matrix using One Rule Model on Bob and Mel.

```
Confusion matrix (relative):
          Actual
Prediction    0    1  Sum
       0   0.66 0.34 1.00
       1   0.00 0.00 0.00
       Sum 0.66 0.34 1.00
```

As a result, next we move to more complex models, and get better accuracy.

8.4.2.2 Random Forest

When running the random forest, I get much better accuracy. As seen in Table 4, for Bob and Mel (left) the random forest model correctly predicts 0s 88% of the time and 1s 96% of the time. These results lead to an overall accuracy of 90%. For Mar (right) the random forest model correctly predicts 0s 96% of the time and 1s 96% of the time. These results lead to an overall accuracy of 90%.

Table 5: Confusion Matrix of Bob and Mel (left) and Mar (right). The random forest model does an outstanding job at classifying moments participants elaborated from action units.

```
pred1    0    1
    0 0.88 0.12
    1 0.04 0.96
> accuracy <- (t[1,1]+t[2,2]) / sum(t)
> precision <- t[2,2] / (t[2,1]+t[2,2])
> recall <- t[2,2] / (t[1,2]+t[2,2])
> round(accuracy,2)
[1] 0.9
> round(precision,2)
[1] 0.96
> round(recall,2)
[1] 0.73
>
```

```
pred3    0    1
    0 0.96 0.04
    1 0.03 0.97
> accuracy <- (t[1,1]+t
> precision <- t[2,2] /
> recall <- t[2,2] / (t
> round(accuracy,2)
[1] 0.96
> round(precision,2)
[1] 0.97
> round(recall,2)
[1] 0.81
```

These results are fantastic. And with fantastic results comes due caution. Being cautious, there are two main threats to credibility. First, it looks too good to be true: with a 90% (Bob and Mel) and 96% (Mar) accuracy of predicting moments Mar elaborated as identified by CDM with just facial expression. When we look under the hood, the model has a precision of 96% and 96%, respectively, and a recall of 73% and 81%, respectively. In absolutely terms, the model performed very well. To ground the results, this means for Mar the model correctly predicts 20,503 of true 0s in the test data set (i.e., moments she did not elaborated). The model correctly identified 3,655 of 1s in the data set (i.e., the moments Mar elaborated). On the downside, the model falsely identified 840 0s as moments Mar elaborated, and 115 moments she elaborated as 0s. For Bob and Mel, the story is similar. If this held up across more participants, I could likely skip the CDM step and identify moments participants elaborated directly from facial expression. Unfortunately, when I used the model trained on Bob and Mel to predict Mar, the model performed dismally, with precision falling to 21% and recall plummeting to 16%. This finding has two implications. First, Facial expressions in a group of two are not good predictors of groups of one. Second, comparing facial expressions between people is likely a poor idea. As mentioned in the introduction, Barrett (2019) argues we should not get in the game of baselining expressions between people, but instead focus on intra-participant study. As a result, these findings could be used to track moments an individual elaborated over a longer period, but without much larger datasets, we should not use these findings to generalize a learning moment elaborations detector.

In making this prediction the random forest model made use of several AUs. As seen in Figure 17. In predicting moments participants elaborated the AUs doing the work, in ranked order, are as follows: for Bob and Mel (left) AU14, AU17, AU01, AU06, AU02, AU 45, AU04, AU15, AU12, and AU07. For Mar AU10, AU15, AU26, AU17, AU20, AU07, AU45, AU01, AU04, AU02, and AU25. The following AUs appear important in both cases: AU17, AU01, AU45.

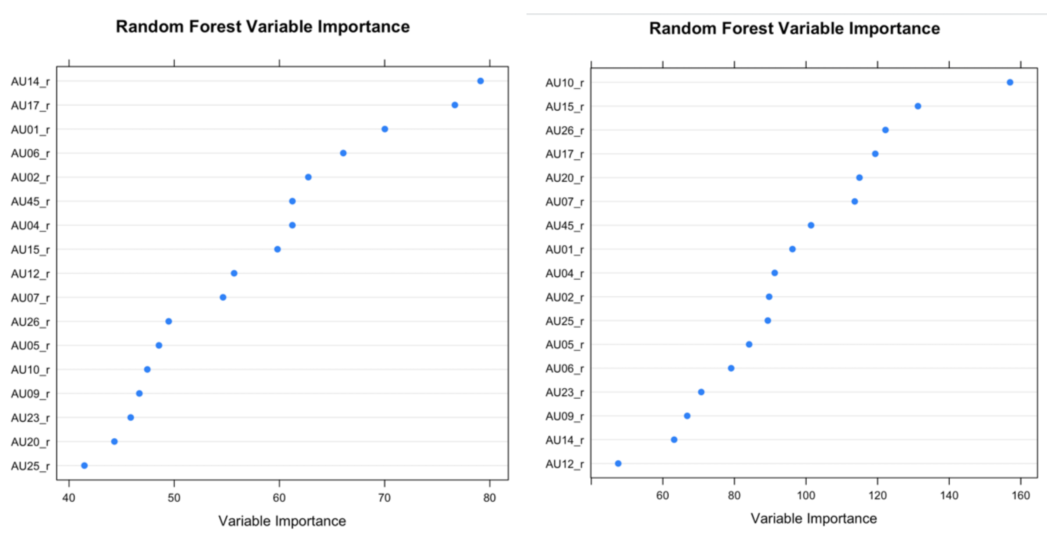

Figure 19: AUs in Ranked order in random forests prediction of moments participants elaborated: Bob and Mel (left) Mar (right).

8.4.2.3 XGBoost

Lastly, I used XGBoost to predict elaborations from AUs. While not being as accurate as random forest, it still provides excellent outcomes verifying the trend of the findings from the random forest. As seen in Table 5, in Bob and Mel's case, in the test data, XGBoost correctly identifies 84% of the 0s and 70% of the 1s, with a total accuracy of 79%, 70% precision, and 70% recall. For Mar's case, it performs even better, correctly identifying 97% of the 0s and 76% of the 1s, with a total accuracy of 93%, precision of 76%, and recall of 89%. I would summarize the result as follows: XGBoost is not bad. Generally, with these models, this sort of accuracy is what I expect to see. The modeler in me says, yes, we're getting some 1s wrong, and we are missing some 1s, but we're doing decent. In other words, counting the results, the model did very well identifying moments Mar, Bob and Mel elaborated. For perspective, in Mar's test data, on the one hand, the model correctly identified 3,989 of the moments Mar elaborated (1s). It incorrectly identified 1,285 moments she elaborated as non-elaborative moments (0s). On the other hand, the model correctly identified 19,333 moments she did not elaborate (0s), but miscategorized 506 moments she did not elaborate (0s) as moments she did (1s).

Table 6: Confusion Matrix for XGBoost Bob and Mel (left) Mar (right)

	RESPONSETEST	
xg.pclass	0	1
0	0.84	0.16
1	0.30	0.70

	RESPONSETEST	
xg.pclass	0	1
0	0.97	0.03
1	0.24	0.76

As shown in Figure 18 compared to figure 17, XGBoost used different variables to makes its prediction than random forest. This indicates there are many ways to identify moments participants elaborated from AUs, and a custom model combining the two may perform better than both. As shown in Figure 18, they are ranked as follows in predictive power for Bob and Mel: AU17, AU06, AU12, AU04, AU01, AU07, AU45, AU26, AU25, and AU14. Importantly, AU14 has moved from the most important AU in random forest, to the middle of the pack using XGBoost. Interestingly, AU06 and AU12 (the components of a smile) have moved much higher up. For Mar the predictive importance is as follows: AU06, AU04, AU07, AU10, AU15, AU09, AU20, and AU25.

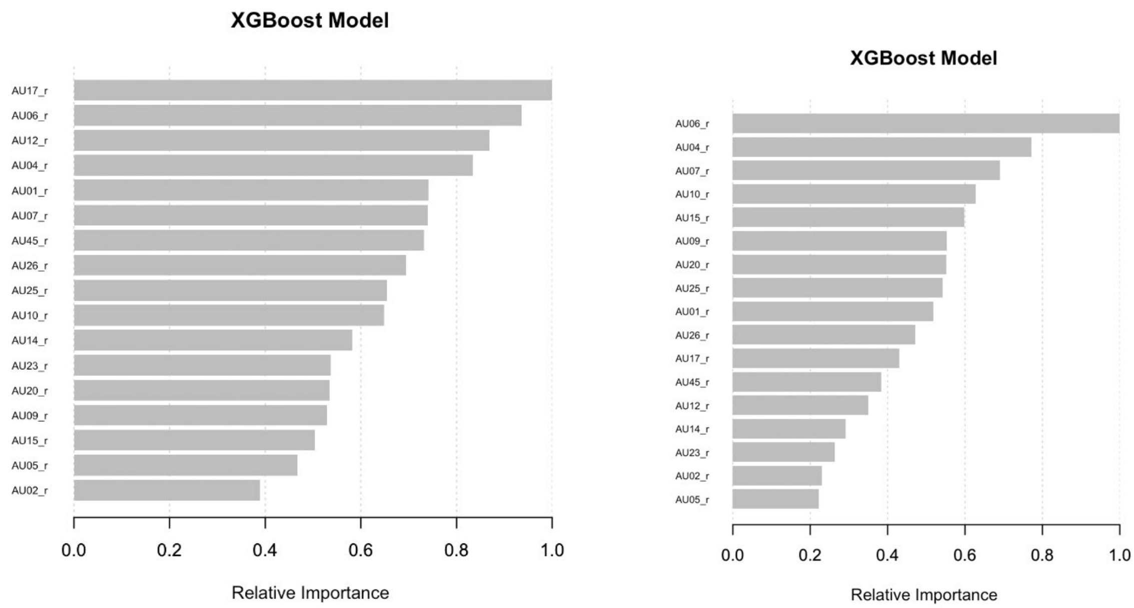

Figure 20: AUs in Ranked order in XGBoost Prediction of Moments of Elaboration Bob and Mel (left) Mar (Right).

In Figures 19 and 20, I plot the top 3 AU's from each model importance ranking using time series analysis. Then, I overlay the moments participants elaborated. Using this the results the AUs important to prediction and time series analysis triangulation we can triangulate each of the top predictive AUs interactions with moments where participants elaborated. As seen in Figure 19, predictive affective states peak near moments Mar elaborated. Likewise, as seen in Figure 20, Bob and Mel's do as well.

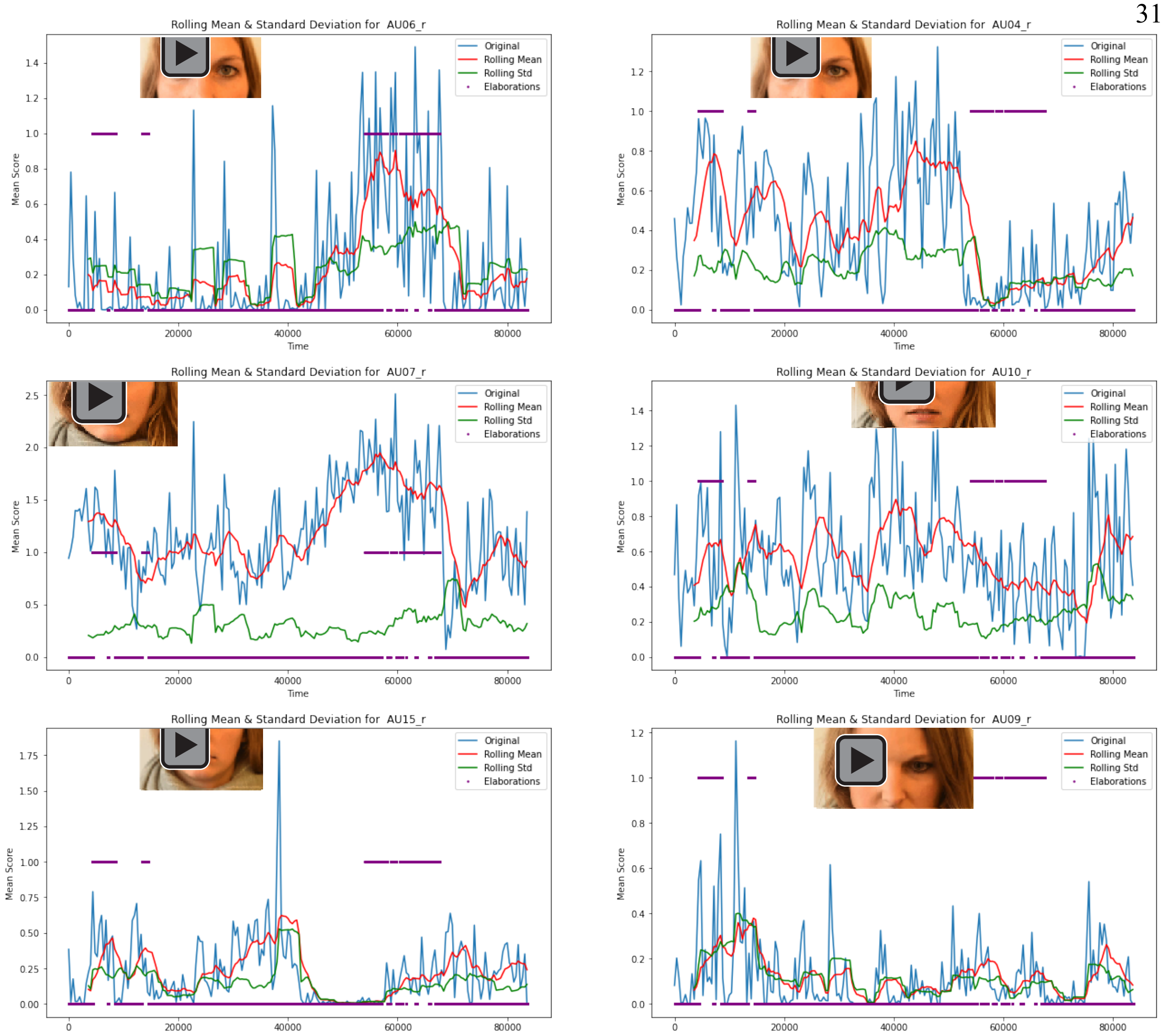

Figure 21: Mar's top predictive affective states and moments of elaboration on Mar's Data. GIFs from iMotions visual library: https://imotions.com/blog/facial-action-coding-system/

Figure 22: Bob and Mel's top predictive affective states and moments of elaboration in their own data. GIFs from iMotions visual library: https://imotions.com/blog/facial-action-coding-system/

9 Discussion

9.1 Is this Practical for Assessing Learning Environments?

From these findings, a key question arises: the interactions in this chapter are short, and I am arguing the approach could assess what people are learning. Admittedly, I am only tracking learning through the pragmatic research method of studying what people elaborate as demonstrated through a close reading of the transcript using CDM. Thus learning, as defined in this chapter, is only a proxy measure. Nonetheless, its important if we mean to implement these methods to assess learning more broadly to scope how long would theory suggest these things would persist. And this is more of a question of practicality. For example, if I am a schoolteacher, and I am trying to understand what my kids are knowing, can these types of in the moment explanations help bridge the gap and understand their learning in an open-ended way? Yet further, could we replace standardized tests with these methods to assess what kids are learning? If the cognitive machines indicated by these methods and especially the ones happening around moments of high affective stimulation— where memory is more probable to form—than the answer is yes. We could replace other assessments with this one. But do they persist? And at first blush, I might say I don't know, sometimes they are persistent, and sometimes not. As discussed in chapter 2, in Sherin's descriptions of in the moment explanations, nodes, and modes (Sherin et al., 2012), if the situation was persistent, we are likely to see this persistence. But is there a better answer to the question? This is important, because we will need to demonstrate if this is practical for understanding what people are thinking, and remembering in open-ended learning environments.

On second blush, while I don't know if there's a better answer, one could ask the question for any kind of teaching at all. That is 'If someone studies something, how long would you expect it to persist?' And this being just yet another little intervention, it may seem futile. We could refine that thought though: from a knowledge in pieces framework, if any knowledge is touched at all by the interaction, then it at least adds to the plausibility that there is some learning with respect to that little bit of knowledge. It is simply we don't really know how long it is going to last. But if it touches it, then there might be some impact on knowledge. As a result, this frames the learning potential more hopefully as 'every little bit helps.'

Moreover, an interesting question arises from the work of knowledge in pieces (diSessa, 1993; Sherin et al., 2012a) that we couldn't answer before employing these methods: I at least have the intuition that there could be these kinds of critical constructions, where a student builds an explanation, and gets this little *click*, where they say, "oh!", and it kind of fits together nicely, and gets a little mental bookmark, because it pushed the ‘satisfying explanation’ button. I have an intuition that those moments exist.

But from cognitive methods, there is no way you can get at that. This has been a perennial topic for future research, whether there are special kinds of moments in these interviews that are particularly important for persistent effect.

In this chapter, I have begun a program of inquiry trying to use affect to look at these *click* moments where participants faces' display a lot of emotional affect of one kind or another—either confused or smiling a lot, for instance. When we look at those moments of deep expression—because theory says high emotional stimulation is tied to post-test gains and means of assessing learning (McCaugh, 2004)—we are studying potential critical *click* moments. In other words, people's brains check in when the stimuli are high. So, in this work I have iterated some on that approach. This approach is a way to improve research of those special moments where a click happens. In this chapter I showed several of those. For example, in Figures 19 and 20, I depict the connections: the purple line is the moments participants elaborated stuff as marked by CDM, and the AUs are the type of affective state they are in when they elaborate. Triangulating them, we have a new way to investigate the connections between affect and learning. I used XGBoost and random forests to try to predict those *click* moments and took the top predictors of the moments of elaboration and map them from the two models from affect data. We see in the top right of Figures 19 and 20, smiling is important around moments of elaboration. And we see an inquisitive eye, demonstrated by AU 7, is important here too.

To recap, in this chapter am using Ekman who worked with several groups to do what's called automatic facial analysis (AFA). AFA used a large sample of faces and then used a vector machine to determine how the motion on the face, the lips, or the eyes moving in a certain way attaches to label data about facial expressions. In other words, it is shape analysis over time. AFA gives you both presence, absence, and intensity scores of affect. We see on the on the left of the time series analysis of this chapter, intensity scores. Though, we must recall they are not meaningful between individuals because people have different expressions but within each individual you can say Mar is not smiling so much at the beginning and the intensity of her smile from her eyes is increasing near the end.

So, I took the video input through OpenFace because it's free, and it outputs 17 action units off the face 30 times per second. Then I merge the AUs to elaborations, the hand coded behavior of when participants are elaborating using CDM. I merged the two through a python script. Next, I use the action units and behavior data to run random forests and XGBoost to predict these behavioral constructs from facial expressions. Importantly from a computational ethnography approach, I could code anything I want: for example, moments of inquiry-based learning, epistemic agency, cognitive load, model base thinking, or other constructs meaningful to the learning sciences. With the approach of this chapter, we can then merge it to sensor data, such as skin response or sentiment analysis and get synchronized, multi-modal analysis. Consequently theoretically, if I argue moments of high stimulation are going to be moments that are remembered, I can say, I am predicting these moments and I can spend time to find the ones that are around moments of elaboration.

In this way, I have begun a process of inquiry to study *click* moments where learners' cognitive and affective machines put a bookmark that gets stored in their memory. From this inquiry, the learning sciences could begin the process of design that affords these clicks in order to improve how learning environments feel while improving how students learn.

9.2 Thoughts on Integrating Machine Learning in Educational Environments

Researchers are developing machine learning solutions to improve student engagement and saving time (Campa, 2021). For example, ProJo is a program aimed at assisting students with spotting their math and science mistakes. ProJo is encased in a small humanoid robot, but instead of trying to replace the instructor, ProJo acts as peer saying things like “Let’s take turns, I am not very good at this sort of problem.” ProJo is one of a variety of teaching aids in development that deploy machine learning, which researchers say could boost classrooms of tomorrow. There are fears that such tools may replace teachers (Campa, 2021). Instead, they could help teachers pick up patterns in behavior that busy humans could overlook. Lilitha Vasudevan, managing director of the Digital Futures Institute at Teachers College, Colombia University recently argued, “So much classroom learning is still dependent upon, the teacher says something, the student responds, and then the teacher is able to form an observation and an assessment of that child.” Machine Learning can assist in this process when capable of analyzing multiple communications cues, along with academic performance, and social dynamics. This approach could facilitate many more students being assessed. The approach has several ethical and privacy concerns, and these tools will certainly need at least the following two features: the ability for users to edit the algorithms, and adaptation to many different hyper local contexts and demographics (Campa, 2021).

Diana, a second product being developed to bring AI into the classroom, is a teaching assistant designed to respond to students non-verbal and visual cues. When a student could use some help or gets distracted it responds by engaging them in a conversation or prompting a teacher. The researchers are now building out the ability to reliably recognize faces, focusing on diverse skin colors, a perennial problem in the technology. Jill Watson, a third emerging technology, is an AI assistant built on top of the IBM's Watson. It currently works in dozens of Georgia Tech classrooms to answer student questions, and relieve time pressures on instructors of record. The chatbot worked so well, it was nominated as the best TA its first semester. These are each examples of machine learning approaches to solving education issues.

Each of these relies on rapid identification of affective states, followed by socio-culturally aware solutions. If we can identify real time learning moments, these findings would allow us to offer recommendations and respond to learners. With further affective analysis, with these methods researchers could identify moments where students get stuck, and provide real time feedback in an instructive learning environment. This opens the possibility of doing the analysis at scale, as we develop more accurate models.

There is a tension here however, in what ways do near immediate feedback and the slower, more open-ended explorations of constructionism coexist? There should be more work on what kinds of feedback assist learners in their explorations, and which overtly coerce the learning journey. If a system points out that a student is elaborating, or confused, who should get that report, and at what time-scale? Will that summative report interrupt the knowledge architecting constructionist environments aim to engender?

With these open questions, to describe the contribution I focus on the process: the process of synchronizing qualitative methods and learning analytics in this chapter provides an example of plugging in sensor measures as needed to provide real-time learning analytics in an open-ended constructionist learning environment. The process also refocuses on outcomes of learning such as learning engagement and groups. This is a powerful means of understanding the impact of our learning designs.

My process of dedicated design of a restructurated unit to teach complexity along with the innovation of learning methods is already promising. In future work, I will continue to validate the findings.

10 Answering the Research Questions

In this paper I answered three research questions:

1. Does the technique of CDM add new abilities to gain new insights into how learners, in group conversation, can advance their learning?
2. What can we learn from physiological measures of people's affective states while they engage in informal learning environments, such as a museum exhibit or a learning game, where people choose to come and learn?
3. Is there a relationship between high stimulation and learning as measured by CDM and affective computing signals? Within moments participants elaborated, is there a relationship between positive affect and learning?

To answer these questions, I used CDM to compare how much participants update their understanding during moments of high affect, as compared to other times while using Ant Adaptation. We knew three key points from previous work. First, moments of high stimulation (D'Mello & Graesser, 2012a; McGaugh, 2003, 2004) are associated with learning. Second, affective states pathways, specifically the engagement-confusion-delight-engagement pathway, has been hypothesized to facilitate advanced problem solving if the confusion state does not overwhelm the learning (D'Mello & Graesser, 2012). Each of these uses of affective state tracking point to greater integration of affect detection to improve design, interactivity, and analytics in learning technology. My third research question took these findings one step deeper and investigated both the process of learning in moments of stimulation and the role of positive affect in that process of complex system thinking.

What follows are my findings, based on my three research questions. For question one, *Does the technique of CDM add new abilities to gain new insights into how learners, advance their learning?* I found overall that CDM demonstrates learning as concept elaboration over time through the proxy of changes in speech. Supporting findings of earlier work (K. Martin et al., 2020), tracking elaboration in discussion improves researchers' understanding of the learning. In that work, we saw with participants, we can observe conservation of concepts in transcripts more clearly. Using CDM, researchers can capture participants' moment-to-moment sense-making. In this chapter we extend and deepen the finding that CDM affords tracking how Mar, Mel, and Bob come to understand the functions — such as ants grouping up, gathering food, or laying pheromone trails — that drive agent-based models, because we can see the participants' discussion. This ability to track the discussion so closely allows me to use the proxy of speech to track the ontological entities they employ to understand what they are seeing. Further, as we see 16 minutes into Bob and Mel's interaction, CDM affords identifying moments where players stop and then restate and reformat their knowledge. Giving participants a chance to be architects of their own knowledge, deciding what to explore and learn, was key to the refining we saw in Bob and Mel's understanding of the complex system. In short, CDM adds the ability to see moments of learning in the transcript data and identify times when participants actively restructure their knowledge through discussion.

For question two, *What can we learn from physiological measures of people's affective states while they engage with a museum exhibit?* I obtained three key findings for question two. First, in Mel and Bob's case, I found that joint users display widely disparate readings of affective state intensity. Participants did not engage in the activity the same way. Second, I found that we can describe who does most of the talking. In multi-user cases, the ability to notice when users engage and disengage is important. By noticing the silence, we can ask what Bob was doing during that time. Time series analysis makes that possible. Second, in all cases, we can describe the frequency of affective expressions. For example, we can say quantitatively that while Bob in the first half of the interview did not engage as much, she became more ebullient and loquacious towards the latter half. This approach to studying affective engagement affords this insight. A looming question is, is it worth it? From merging the qualitative and the sensor data we can analyze minute changes, and use one to understand the other. While the methods are more laborious, they are also more meticulous. The approach opens the possibility of synchronizing more instrument data to qualitative data, and thus bringing together two methods of analysis to better understand learning environments.

For question three, *Is there a relationship between high stimulation and learning as measured by CDM and affective computing signals? Within moments participants elaborated, is there a relationship between positive affect and learning?* I show that in Mar's case, the main emotional pathway around moments of learning was the delight pathway. In all cases, the primary affective expression was smiling. Moments participants elaborated are accompanied by a lot of smiling. There is a positive correlation between learning identified by CDM and positive affective states. During Mar's interaction around learning moments, delight is the predominant emotional pathway, accounting for 29,490 of the pathways around learning moments. The next two biggest emotional pathways are contempt and contempt delight pathways, account for 2,713 occurrences. As Mar's sip of coffee highlighted, researchers need to pay attention: not all the moments expression analysis bring to light are salient. We do not want to get in the business of finding every sip of coffee. Across the data sets there is a significant correlation between smiling and elaboration.

This analysis method affords the possibility of doing the analysis at scale, as we develop more accurate models. Using a random forest model, I got an accuracy of 90% and 96% when predicting moments participants elaborated from action units on the face. Using the XGBoost I got 79% and 93%. accuracy. The methods each surfaced critical AUs used in the detection of moments of elaboration, but, smiling, AU06 or AU12 not necessarily the only variable of important when identifying moments of elaboration from facial expressions. These findings would point towards identifying learning by studying affect. In future work, I will corroborate across a larger set of interviews. I will keep in mind that intra-participant comparison is more important than generalization at this resolution of data collection. The process I used to synchronize qualitative methods, and learning analytics provides an example of plugging in sensors as needed to provide real-time learning analytics, and refocus on outcomes of learning such as learning engagement. This is a powerful means of understanding the impact of our learning designs. My process of dedicated design of a restructurated unit to teach complexity along with the innovation of learning methods is only just getting started, but the results are already promising.

This project has revolved around learning to understand how to study a core set of questions I ask of people: *Do you know how ants collect food? How do evolutionary adaptations impact the ants and the colony's success? How do ants know what to do? How are you feeling while talking about this? How do pheromones work, what would you like to be different about the game? Why did you like the game? In what ways can a game be scientific?* These are questions to explore complex systems, self-organization, the design of a game to enhance that exploration, and the feelings of the user. At the heart is a user-centered approach to designing a learning environment. The core of the effort is designing an environment where users are architects of their own knowledge and exploring how the come to understand it, and the facial expressions they make while they do it.

I designed this exploration based on prior theory of teaching complexity, understanding users' engagement, and creating users that are architects of their own knowledge. This work has been aimed at overcoming the twin challenges of complexity systems learning: level slippage and not accounting for the parts of the system. To develop the next generation of environments in this domain, I reviewed the literatures of learning complexity through constructionist methods, museum explorations, and the role of affect in learning. From this starting point, I have found the role that positive affect plays in game-based learning, and have discovered that participants can be architects of their own knowledge through interaction with each other and a situated microworld.

I used Ant Adaptation to implement two new ways to research how users come to understand complex systems: 1) constructivist dialogue mapping (CDM) (Martin, Horn & Wilensky, 2020), and 2) affective engagement. The two approaches measure and assess the learning—using CDM—and engagement—using computationally augmented ethnography — of users during complexity science education.

New methods to study participant learning and engagement are needed because constructionist learning environments encourage open-ended exploration. This shift is a restructuration (Wilensky & Papert, 2010) of education resulting from the impact of the advent of powerful computation. In this restructuration, the learning environments are more cognitively, and affectively engaging. But the adoption does not come without growing pains. We can see these properties' operation as learning environments grapple with the incorporation of more affectively and cognitively engaging curricula that students want to share. These constructionist curricula attend to the unique accommodations and diversity of students while lowering the barriers to understanding powerful ideas.

This chapter shows how we can study these engagements more meticulously, to design affective and cognitively engaging, and enriching, learning environments. As mentioned in the opening quote, in 1960, Ivan Sutherland, the creator of the first mixed reality display, argued we need to provide children with computational learning environments that allow them to explore mathematical wonderlands. Though the evolution of the restructuration is happening, we need to remain humble; microworld exploration in a mathematical wonderland of computational environments has not become mainstream yet. While we could say it's only a matter of time, there is one elephant in the room holding back the spread of open-ended individually designed learning: assessment. What are students learning? Though there have been decades of constructionist work, we have still not solved the problem of measuring learning in an open-ended environment. The singularity of learning trajectories afforded by the exploration of these mathematical wonderlands has made it harder for education researchers to develop strong assessments for constructionist learning environments. When learning is no longer one-size-fits-all, it becomes more difficult to develop uniform, standardized assessments. The lack of evaluation slows down their implementation and adoption in educational environments. This is unfortunate because this type of thinking—complex systems and the outcomes of self-organizing systems—is exactly what society needs in order to deal with core problems in a complex world. This paper shows how two methods can be interwoven to study the affective engagement and the cognitive construction that takes place during gameplay, through the restructuration of an agent-based modeling game. The approach shown here quantifies participants' affective and cognitive engagement with Ant Adaptation. The approach is applicable to studying learning environments more generally, and especially useful when studying open-ended exploration of mathematical wonderlands and microworlds.

I would like to reflect on what this means at scale. For me, the important part of constructionism is supporting learners with different goals/interests/learning pathways and their expression of that throughout the experience may be very different than others. I am not discussing affective expression here but instead how they engage more broadly. This study has brought up a feeling for me: I wonder the difference between asking or interviewing someone about that affective engagement directly compared to doing it through tracking/big data and if you lose anything about what is personally relevant or meaningful in that moment as related by each participant. It's not that the methods used here aren't useful, focusing on what to count has often been helpful in the process of science, but it feels like the methods may remove me from the participants being able to express what and why it is important, which is kind of the motivation of constructionism. As we move the methods to scale, I seek to keep this core part of constructionism, while allowing for these broader evaluations. And on reflection, I think it's important not to be reductionist, understand each person on their own terms. The divergence of the predictive models between participants pointed in this direction. We can make models of individual's expression, but joining leads to lowering precision and lower recall. These methods are trying to examine minutely what a person is doing, saying, and feeling, as they go through a learning environment. But even with that goal, on concluding the project, I am reminded of constructionism being so personal to each learner. The main thing I am trying to do is show how we could evaluate constructionist environments at scale on their own terms. I hope maybe with all the techno-centrism of this piece, that goal of evaluating each person on their own terms did not get lost a bit.

10.1 Limitations

This chapter has several limitations, some include data collection issues, inter-rate reliability, issues with affective pathway analysis, identification of learning moments, predictive models with small datasets, issues of studying groups of people, the lack of human computer interaction analysis, the surface level affective analysis, and learning retention. Next, I turn to each of these in turn.

This chapter has suffered from data collection issues. Paramount among them has been difficulty with collecting videos of faces while playing the game. Several younger, less advantaged participants had the issue that running Ant Adaptation while screen recording was too much for their computer system. To overcome this, I tried methods with off computer cameras, but these too posed difficulties with file management. Additionally, only using Zoom to record resulted in unreliable videos, that unexpectedly shifted perspective, which can cause issues of analysis. The issues arise from the fact that analyzing single participant videos is more accurate and faster than multi-person videos. Consequently, knowing if more than one face is in the video is a crucial step before using OpenFace. Each of these issues, that predominately effected less advantaged participants, impacts the quality and quantity of data in this chapter. This is concerning, because the methods of study have a systematic bias towards more advantaged, older participants with access to me as a privileged researcher at an elite university. It could go without mentioning at this point, but the COVID-19 pandemic acerbated this bias. Had it not been for the pandemic I could have simply collected more data in person, like Bob and Mel’s interview, or taken the study to other study locations. In future work I will expand on this data, and mark this chapter as demonstration for the next iteration.

Inter-rater reliability (IRR) should be more closely studied. When conducting IRR for CDM, trees arise that are the same depth, but different coders will reorder the connections between boxes, and visually inverse the hierarchical maps. As a result, establishing IRR may not be the ideal way to validate CDM, because in a way it is a visualization technique to summarize the branching explorations participants undertake. That being said, developing validity checks for CDM is still a necessary work still in progress.

I conducted affective pathway analysis to understand the preceding and post-ceding affective states participants have around learning moments. Unfortunately, the way I coded CDM and how I joined it to affective states meant pathway analysis did not work when multiple participants were in one video. This needs to be further studied, as while CDM is useful for coding an individual clinical interview, one of the method's biggest affordances is studying groups learning as a congress of minds that jointly learn. As such, until pathway analysis is updated, the method does not work with this aspect of CDM, thus weakening the approach's utility in studying affect and learning synchronously.

There are limits to identifying moments participants elaborate arising from the methodology. The methods of this study are predicated on the idea that the antecedents of the elaborations are moments of complexity learning. In other words, the study rests on the warrant that what I am measuring with CDM is meaningful to study learning in a larger sense. This is an empirical question, that will require more study. When I identify moments participants elaborate, I use what they say, and note when they elaborate on an idea, or introduce a new ontological entity. However, the question of if the methods of study, a clinical interview, and the use of the game, somehow effect the underlying cognitive system they employ. As noted above, quantum physics has noted that we cannot be certain of our measures, because the act of measuring impacts the situations of study. Though I have built the argument of this chapter on these moments of elaboration, researchers should recall these are constructs arising from theory, and demonstrated through the measures.

Usually, the predictive models used in this study, XGBoost and random forest, employ large datasets. While the data by some measures is large, with more than 60 million observations per hour, with just one video to train each model on—one for Mar and one for Bob and Mel—we should remain skeptical. While the findings show that if we used these models on more videos of Bob and Mel, or Mar, respectively, we could identify their own moments of elaboration adequately, that is the limit. The models trained do not have enough examples to be generalized to a scaled 'elaboration detector'. As such, for next steps this approach can be used to train up a model for each learner in a webcam enabled classroom. Then once we agree on behaviors of interest, we can synchronize the data and study each learner over time across a longer course or project. But what we cannot do is train a model from one person and identify others' learning moments. In other words, these models used in this way are context and person specific. As such, they fit into an ethnographic approach of study better than they do a big data approach where we place all the data in a hopper and determine the outputs. In summary, while the models, with enough examples, likely can adapt to find patterns in larger, multi-participant datasets, this hypothesis has yet to be demonstrated.

In the chapter, I studied Bob and Mel as a co-joint system. In other words, I did video analysis of multiple people. This idea of minds in congress jointly learning is counter intuitive for some. As the complexity research has shown, we tend to think Bob is a unitary individual in charge of her own direction, not emergent co-define in interaction with Mel and her surroundings. Until this view is further broken down, it will continue to seem odd to study people from a socio-technical system perspective. In other words, it will continue to be difficult to discuss learners as components in interaction in a socio-technical system. This is unfortunate, as the research of the complexity of the mind is showing we are complex systems, in interaction with other people, who themselves are complex systems. Like an ant colony, this interaction at a social level reproduces itself through time recursively. However, our methods of study are still at an early stage, and that limits the scope of the discussion and findings of this project. To go beyond this limitation, learning sciences needs to embrace the work of using indeterminacy to study people and groups, as shown in the work by Lambert-Mogiliansky and Busemeyer (2012). For infrastructural reasons, this approach has not yet caught on.

In this project, a big part of what occurs results from the mediating influence of the microworld. However, in the analysis there is a dearth of analysis of user interaction with the game. I studied what they said, how their cogitation changed, and their affective states, but I did not study the key moments in the game that impacted that change. From a human computer interaction standpoint, this approach is inadequate. Hopefully, the interaction analysis in chapter 3 can ameliorate this shortcoming; nonetheless, the paucity is a significant limitation of the current chapter.

The approach I took to affective states is only skin deep. A predominate amount of cogitation is subdermal. Though the affective state determinations of this study are made on dermal folds resulting from muscle contortions, these expressions are a window into those immediate affective cognitive responses, that may point to those indeterminate cognitive machines. There are, however, other methods to study and identify affective states that are a) not as much of a privacy concern as cameras, and b) still provide rapid insight to the cognitive machines that account for our consciousness, learning, and memory.

While I can show people elaborated in this study, I take no attempt to study retention, recall or transfer. A large part of the evaluation of our children, such as tests, is designed to test this longer-term retention. This results from the higher goals of an education, to make better citizens. This limit makes the methods of this study less comparable as they are studying proxies of in the moment thinking and feeling. One could ask, as I do in section 9.1, given how short the engagement times were and how short the interviews were, how stable do you think thinking shifts were? In a knowledge and pieces perspective, these could be momentary expressions cued by previous experience, that would not necessarily be retained. Especially considering what we know about knowledge-in-pieces? While it's an open question how this knowledge is retained, re-employed and in what ways it restructures participants' future knowledge use, in section 9.1 I provided some ideas of how these approaches begin an inquiry into *click* moments where participants cognitive and affective machines bookmark moments to be stored in memory.

11 Conclusion

Since Leinhardt and Crowley (Leinhardt et al., 2003; Leinhardt & Crowley, 1998) presented their vision for studying elaboration as learning, museum studies has been looking for a way to operationalize the study of elaboration. Additionally, research has shown that learning has a lot to do with the emotional landscape at the time it occurs. Increasingly, for researchers who are trying to understand how students regulate their learning processes, the affective processes associated with learning are of interest (Calvo & D'Mello, 2011; D'Mello & Graesser, 2012a; Singh et al., 2002). Therefore, the evidence supported the hypothesis that moments of high stimulation, especially of delight, would correlate with moments of elaboration (my pragmatic definition of learning) because high stimulation is implicated in memory formation (McCaugh, 2004). In short, affect plays a role in learning. In this study, I first show that using CDM to mark moments participants elaborated allows us to synchronize qualitative data and quantitative sensor data, which can open new avenues for exploring learning and affective engagement. Second, I show in Mar's case, the main emotional pathway around moments of learning was the delight pathway. In all cases, the primary affective expression was smiling: moments of learning as identified by constructivist dialogue mapping are accompanied by a lot of smiling. There is a positive correlation between learning identified by CDM and positive affective states. For Mar AU06 and Elaboration is r 0.21 $p < 0.0$ and AU12 and Elaboration r 0.17, $p < 0.0$. During Mar's interaction around learning moments, delight is the predominate emotional pathway, accounting for 29,490 of the pathways around learning moments. The next two biggest emotional pathways are contempt and contempt delight pathways, accounting for just 2,713 occurrences.

Third, I show that participants playing with the complex systems game about ants learn to form an understanding of the complex systems model. This builds on two properties of restructurations (Wilensky & Papert, 2010): affective and cognitive engagement. This finding allows us to explore when new mental models form, and how engaged people are during formation.

Lastly, I show predictive models of classification, random forests and XGBoost, work exceedingly well at classifying moments of elaboration from facial expressions.

This work together provides an operationalization of the work of Leinhardt and Crowley, demonstrating how we can study learning in constructionist environments with CDM, while also demonstrating a way to study user engagement quantitatively with expression analysis that led to the ability to predict moments participants elaborate with a high degree of efficacy. That the learning and engagement in the study are so rich is because I attended to the design of the learning environment, employing lessons learned from constructionism, restructuration theory, and complexity. I would suggest to the community to also follow this design pathway to design games for learning.

12 Acknowledgements

I would like to thank the Center for Connected Learning and Computer Based Modeling (CCL), TIDAL, and TIILT labs for their feedback throughout the project, and the supportive Northwestern community. I would like to thank Uri Wilensky, my adviser, for keeping me focused on the main objectives of the project and not getting too far in the weeds. I would also like to thank Sarah Griebel for all her thoughts on the writing, and Paul Won and RJ Schechner and the Northwestern Writing Place for their edits. I would like to thank Kelsey Edwards and Wencheng Tang for invaluable feedback on and assistance with the machine learning and time series analysis. Finally, I would like to thank the National Science Foundation (NSF STEM+C: Award # 1842375) for supporting this work.

13 Appendix 1: Interview Questions

Appendix I: Pre-Post Survey

- Name
- Location
- Birthday
- Grade
- Gender
- Race/Ethnicity (checkbox)
- What language (or languages) do you speak at home?
- Who did you come to the museum with today? (check box: class, family, other__________

Pre-Interview

- Q0: How are you feeling today? (Emoji slider going from :(to :D)
    - What exhibit did you just come from? How did you like it?
- Q1: Have you ever noticed anything about ants? (Rational: Learning about bugs)
    - **Response:** some description of an ant and what they do.
-
- Q2: How do ants collect food? (Rational: complex system/learning about bugs)
    - **Response:** some description of what they've seen before when watching ants
- Q3: How do ants know what to do? (Rational: complex system)
    - **Response:** some description of ants' government
- Q4: With So many ants in a colony, how do ants protect public health of the colony? Why is public health important in an ant colony?
- Q5: With so many ants in a colony, and confined to tunnels, how do you think ants deal with traffic congestion? Why could congestion be a problem in an ant colony?
    - **Response:** some description of ants' coordination
- Q6: How would an ant's size effect its life?
- Q7: How would an ant's aggressiveness effect its life?
- Q8: How do pheromone trails work?

Post Interview

- Q1: Which would give the colony the best chance of success, bigger ants or more aggressive?
- Q2: How do ants collect food? (Rational: complex system)
- Q3: What's the pink chemical do?
- Q4: How are the ants in the computer game different from ants in the real world?
- Q4: Can a game be scientific?
- Q5: What features of the game did you like? (R: Design)
- Q6: What would you like to be different about the game? (R: Design)
- Q7: How do pheromone trails work?
- Q8: Retrospective: What was the most exciting part of the game? What was your biggest disappointment? How did you feel about working with the other (players/your family etc.)? What did you learn?

Chapter 6: Conclusion

1 Introduction

This dissertation contributes theory, and has specific lessons from each study. In this final chapter, I will proceed as follows: first, I will discuss each chapter considering its contribution. Here I provide an overview of the project to refocus and show the contribution, units of analysis, methods, and goals of the design. Then, I provide answers to the six research questions that I asked during this dissertation. The evidence of each answer is broken down into answers to questions I asked in the introduction and explored in each of the intervening chapters (chapters 3-5). Finally, I zoom out, and try to see the forests for the trees. That is, I present my contributions, their significance, the limitations, the next steps, and new directions.

1.1 Overview

My research into complexity learning uses design to drive powerful moments both in museums and online to change how people learn. I use the lenses of constructionist learning and the creation of restructurations (Wilensky, 2020; Wilensky & Papert, 2010) that lower the barriers of entry for students through agent-based models to reform and improve education. I envision a future where every student understands a Lotka-Volterra (predator-prey) equation and how the simple rules ants follow can self-organize into large societies that include complex behaviors including pheromone networks, agriculture, bridge building, and cattle rearing. Understanding the simple rules that organize these complex systems can unlock a set of intuitions about complex systems.[20] These mental models are key to coming to grips with global issues, such as climate change or income inequality. If students construct their understanding of these issues, they will be the architects of their knowledge and design their understanding of the world.

My research into learning analytics, such as the machine learning powered learning analytics, I have demonstrated in chapter 5, is a powerful means of reaching this future by tracking learning in self-directed environments. Through developing cutting-edge learning environments, iterative user-centered design research, and novel learning analytics, I developed methods to assess cognitive and affective engagement with the learning environments and improve outcomes.

My dissertation builds on my past research to demonstrate the connection between affect and learning through qualitative methods and machine learning approaches. The research is threefold: First, in chapter 3, I developed and tested a thinking and learning intervention, the agent-based modeling simulation Ant Adaptation (Martin and Wilensky, 2019). I showed that the intervention can shift people's schemas from a process schema to an emergent schema during 10-minute museum interactions (Martin, Horn, and Wilensky, 2019). I then used Ant Adaptation throughout the rest of the project. Second in chapter 4, to track the conceptual change common in Ant Adaptation, I developed a novel form of concept mapping — constructivist dialogue mapping (CDM) — which is particularly useful as a learning analytic of informal learning environments. Through CDM, I analyze participants' spoken elaborations in small subsets to study how people develop their understanding of a system or museum exhibit over time. Third, in chapter 5, building

[20] This phrase asks the reading to consider the connection between complex systems, and how learning about one, can aid in solving issues in other complex systems. This has been the key point of this project. While the siloed nature of knowledge makes moving between the fields big leaps, the parallels of the complex systems are striking.

on the work of computational ethnography (Martin, Wang, Bain and Worsley, 2019), through video analysis, I developed a method of affect detection to identify how participants are engaged across 45 facial action units. I map those emotions to moments of learning using AI/ML methods of detection: CDM and computational ethnography. Because the data source is video, the method has outsized potential for scale to predict unseen data. In short, in my work, I have applied advanced methods to the design and evaluation of educational interventions to teach complexity.

Throughout this dissertation, I have explored one set of questions before (and while) participants played a game about ants. I asked: How do ants collect food? How do pheromones work? How do evolutionary adaptations impact an ant and her colony's success? How do ants know what to do? How are you feeling while talking about this? What would you like to be different about the game? Why did you like the game? In what ways can a game be scientific? These are questions to explore complex systems, self-organization, the design of a game to enhance that exploration, and the feelings of the user while they learn. At the center is a user-centered approach to designing an education. In each iteration of this project over the past 4 years, I have sought to improve the design and measure users' cognitive and affective engagement with the lessons embedded in the digital tool, Ant Adaptation, and the discursive practices, competitive discussion occurring around the game. Each of these together drives the learning observed in this dissertation. The project is also a demonstration of how to design a game while focusing on the user, to improve learning about a difficult to learn subject: complex systems. Learning about complex systems has a lot of promise, from solving climate change to traffic and income inequality. Unfortunately, there are two roadblocks to people thinking about complex systems, levels slippage and the inability to think about how the parts of a system emerge to form the macro-outcome. These challenges mean that most kids now do not learn about complex systems, often not coming to them until after advanced mathematics courses in graduate school. This project has provided evidence that participants can overcome both barriers in a short amount of time by interacting with Ant Adaptation. Building on this approach could mean more kids form the mental models and intuitions necessary to take on big challenges like climate change and global poverty.

2 Answering the Questions

In this project, I had six formal research questions. Next, I discuss what I have learned about each in turn. Now we are going look at specific research questions.

2.1 Answering Question 1

For question one, *What do students learn about complex systems by engaging with a social insect-based curriculum?* While playing Ant Adaptation, an insect-based microworld that affords open-ended constructionist learning, students form mental models that help them create intuitions about complex systems and their constituent agents. In chapter 3, through play, Thomas learned that entities like ants have a mechanism such as laying trails to attract other ants to flowers in a cycle by recursively following the chemicals or wandering. By forming the mental model of self-organized foraging, he was able to intuit the connection between ants, pheromones, and food collection. Then, as shown in chapter 4, Rebecca first identifies a phenomenon she does not understand: namely, purple tracks that she later understands are pheromone trails. Through observation, she concludes that pheromones attract ants, forming a self-organizing foraging system. From this prediction, she expresses its function but maintains some confusion. This confusion is the affective state that accompanies her process of accommodation.

Jean Piaget developed the term to describe what occurs when new information or experiences cause someone to modify her existing mental representations. According to Gary Drescher (Drescher, 1991, 1987), the schema mechanism controls the actions someone takes in a microworld; based on these interactions with the world, a mechanism accumulates and organizes knowledge. It uses the knowledge to guide activity in pursuit of goals, the foremost of which being the acquisition of more knowledge (Drescher, 1987). The mechanism's data has three parts: 1) items, binary state variables; 2) actions, that are the elements that transition states; and 3) schemas, which are the effects actions on states cause. As a result, schemas are defined by the items and actions constructed by the mechanism. The mechanism constructs other schemas, including, 1) spin-off schemas, as it finds correlations between items and actions; 2) composite actions, which are defined by their goal state that other schemas organized to reach the goals state implement; and 3) synthetic schemas that designate an item that is novel that are not expressible as any function of earlier extant states. The construction of new schemas is important because it affords discovering regularities when many things, haphazardly, happen all at once. Synthetic items afford the invention of new, sometimes radically new, concepts. They organize Piagetian conservation, where an individual recognizes something stays constant in the microworld even though its manifestation changes, or even stops. Syntenic items designate "*whatever unknown condition in the world* governs the schema's success or failure" (Drescher, 1987, p. 293) .

As I described in chapter 4, for example, consider if the child grew up in a gas cloud in some remote universe. At first, the twirling nebular shapes would appear to have no order, and she would have no ontology for them. Therefore, when a particular eddy returned to the developing child's perception, she would not attach to it any permanence because the learner would not have *conserved* its continuous change into an identified concept. As the learner lived there longer, however, Piaget's theory predicts she would combine the gas cloud eddies into predictable movements based on experience to create concepts of gas cloud objects through the process of assimilation and accommodation This process would be conservation: seeing an object despite its changing features, or ever-changing atomic arrangement as a continuation of its prior existence. In this way, all matter is more or less variant, but the child can construct an invariant version out of this directionless, senseless mass, to form a theoretical frame. With this frame, the child can operate on their constructed ontology to make predictions and act. This is like a biker, who sees a group of people walking, and can bike past them very closely because she can assume they will continue to move the way she has observed them up to now. This prediction of the pedestrians' operation is prone to error, but as a working hypothesis for riding a bike, is close enough. In childhood, we are thus able to see enough rhythm and repetition around us to create meaningful concepts.

In my data, I found this process entails some confusion as the mental model is challenged, extended, or replaced. Rebecca's process of accommodation pushes her to account for greater and greater parts of the complex phenomenon. The confusion that she experienced when observing the ant-pheromone relationship generated a new mental representation, which created intuitions about the complex system as she noted the items that stayed the same, and the items that stopped. This process happened slowly as she interacted with the complex system of the ants competing.

Similarly, the participants in chapter 3 demonstrated a moment of confusion and subsequent accommodation. In this case, one participant, Ed, experienced confusion throughout the game. Ed's younger brother Thomas, meanwhile, set a goal for himself of maximizing population, which was set in communication with Ed. By working towards this goal aloud and sharing observations along the way, Thomas gave Ed the architecture of mental models so that Ed could begin to construct his own. While engaging with Ant Adaptation, Thomas used goal setting and exploration to engender learning by sharing what he noticed with the rest of his family, ameliorating their confusion. The process facilitated Ed's and the family's accommodation, the process of turning the observable world into new mental models they hold, and may use in future situations.

As well as learning about agents and their effect on a complex system, As shown in Figure 1, players learned to predict the structure of complex system. As shown, our world is full of complex interactions that self-organize. Ants might be a good way to teach complexity. Consider an ant's cycle. The ant wakes in the colony, goes outside, wanders randomly until it finds food. When she finds food, she leaves a trail as she returns the food to the colony, bringing energy back to the colony. When the next ant comes out, she follows the trail to food. As ant after ant repeats the loop of following simple rules, the colony gains food. Additionally, a trail network emerges that the ants move along. Each ant, following simple rules leads to a complex, emergent pattern that feeds the colony without any central control through a recursive process.

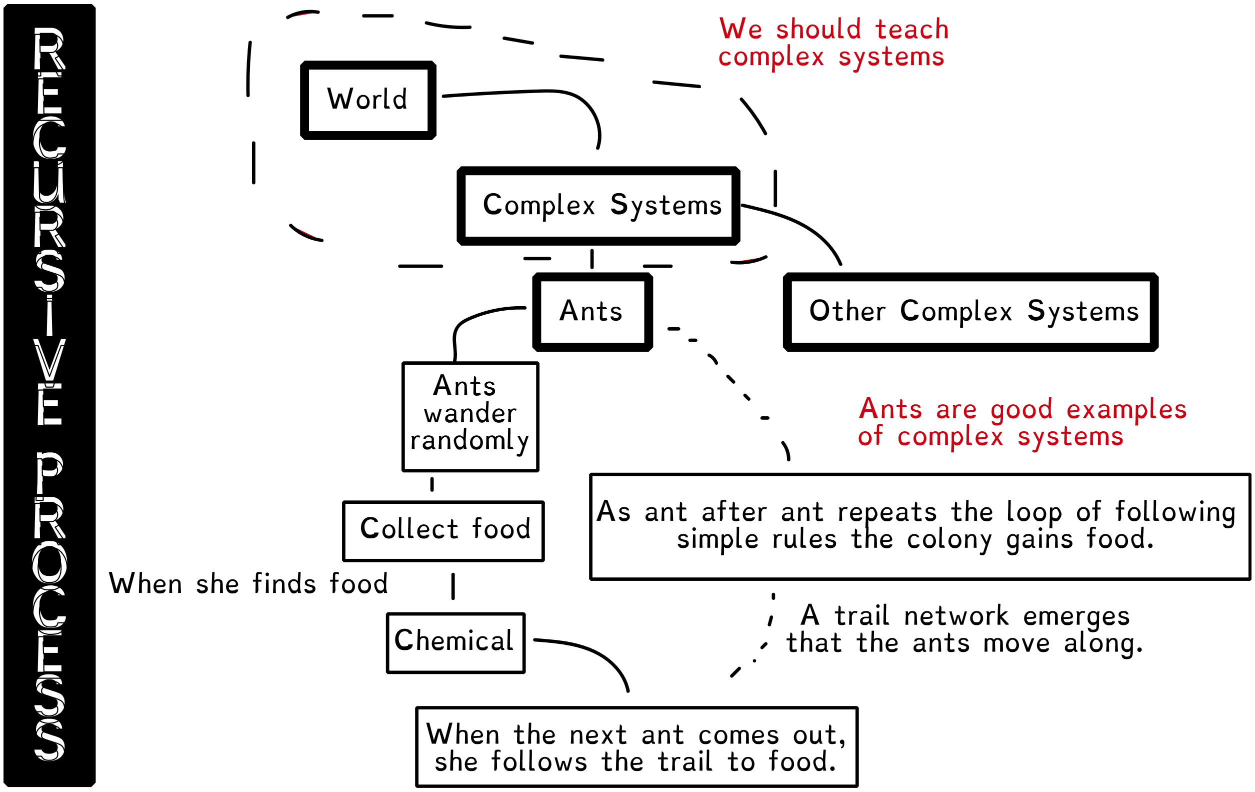

Figure 23: There are lots of complex systems in the world, using CDM I depict the generalized plan of how Ant Adaptation helps students learning the complex system of ants that helps them think about others.

Thomas realized this process: that sometimes because no one oversees complex systems, ants did not follow his goal of population maximization. Instead, they sometimes got caught in a portion of the pheromone network and fatally circumambulated local optima.[21] While he did not learn the term 'local optima,' Thomas tried to reverse the resulting deleterious effects by using vinegar to free the ants, evincing his understanding of the structure of the system that local optima

[21] In a complex system, local optima are the extreme value of the objective function for a given region of the input space. In the case of pheromone networks, the local optima are the strongest point of pheromone accumulation within an ant's perceivable space; for example, when ants get in a fight, each ant lays down more pheromone when they join, leading to a situation like a black hole: all ants nearby are eventually sucked in. In nature, if ants lay pheromone trails in a circle, they can lead to a march of death where each ant follows and lays a trail the next behind follows *ad nauseum*.

constitute (i.e., the pheromone network). Furthermore, as shown in chapter 5, Mar learned three functions that structure the ant colonies' food collection: how ants 1) group up, 2) collect food, and 3) use pheromones. Learning these functions allowed her to better understand what ants do because she better understood self-organization. Later in chapter 5, Bob and Mel used the system to understand two parts of an insect social system: 1) Mel established the need for equilibrium for colonies to keep from tipping into instability, but 2) Bob found the need for an insect society to deploy various adaptations depending on the mercurial shifts of an ecosystem. In this way, Bob argued for the use of adaptation in an evolving ecosystem.

In chapter 3, some users shifted their schemas from process to emergent schemas by using Ant Adaptation. For example, three participants in two groups changed their schema from direct to emergent. For example, Ed changed from a direct schema to an emergent schema after playing. Some people did not shift their schemas. Both Stacy, a seven-year-old girl in Group 20, and Pri, a sixteen-year-old girl in Group 27 held their original direct schema. Finally, when playing Ant Adaption, the learning is one-directional: people did not shift from emergent schema to process schema by interacting with the complex systems about ants.

In short, when playing a game about ants in a colony collecting food and competing, participants contested ideas about how ants self-organize. Through the contest, players understand micro-level workings of ants lead to a macro-level outcome of a higher population. As a result, when playing in the somewhat familiar domain of ants — animals most people have seen on the sidewalk — participants formed intuitions of self-organization through feedback and moved away from a more determinist centralized mindset. In this way, through playing with an ant game, some participants overcame the twin challenges of complexity education including level slippage

(Wilensky & Resnick, 1999) and not accounting for the individual particles in a system (Chi et al., 2012). Thus, I can now say, though some people struggle with a deterministic mindset, through playing social insect-based games and holding discussions about what players are seeing, educators impact the prevalence of such ways of understanding systems. This is a hopeful finding, as we need more complex systems thinkers to face the challenges of today where no one is really in charge, such as climate change, income inequality, or traffic.

2.2 Answering Question 2

For question two, *What process of iterative design can result in a system like Ant Adaptation?* I find that providing consequential decisions in the game creates the tensions that players explore in order to understand complex systems. For instance, in chapter 3, when Ed says, “I don’t want to be too aggressive” as the two teams stare at the touch screen, it shows a burgeoning realization that adaptations affect his ant colony’s life cycle. This leads to the interaction of the younger brother, Thomas, teaching his older brother, Ed, that the direction of pheromones leading successfully toward food increases the ants’ population through the process of feedback. This is an example of the discursive structures leading to better learning. This is like Mel’s decision to create a balance between her two colonies. She sought to set the adaptations of ants to balance the outcomes of the gameplay.

As designers, we should attend to the kinds of talk that result from the learning environments we design. As shown in the answer to question 1, as a result of engendering this kind of discovery talk, players of the design discover the principles of complex systems. In Chapter 3, when I was describing Ant Adaptation, I put forward three aims of the game: enjoyment, comparability, and productive conflict. Each of these goals is situated within play. This play is a discursive practice of making changes in the complex system and discussing its outcome. From my review of literature, I hypothesized that designers should encourage discussion and comparison in a competitive game mode where the biology and complexity science weave into the experience. From user testing, I found that the environment accomplishes these goals.

To design this experience, we created interactive, agent-based, complex system tabletop games for museum settings that expedite learning in these short interaction times. For the design we had three hypotheses:

1. Designers should encourage discussion and comparison in a competitive game mode where the biology and complexity science weave into the experience.
2. A body in motion, talking out ideas would create a rich discussion and a problem-solving mindset around the game interface.
3. Varied motivation of participants affects their learning.

I used engagement measures, such as affect tracking tied to cognitive engagement in a group to measure this variety. Now that I have shown how much smiling and the kinds of deep explorations of complex systems the game affords, I can say that the hypotheses bear out. This is shown by comparing between groups and individuals at play. While I did not perform a formal comparison, we can see in the group play sessions of Thomas, Ed, Mel, and Bob that the discussions are richer, the bodies engage more fully, and varied motivations are obvious. In other words, to create consequential learning, designers should focus on designing enjoyable games that promote players to compare between options through competition are consequential decisions. This play should happen in a group and be designed to elicit discussion.

There are two design features of the game that seem to support the learning the game engenders: forming their own mental models are consequential decisions and the rapidity of the response to their actions:

First, constructionism argues kids learn best when they make stuff that matters to them because they form deep mental models of the phenomena. I find that players also enjoy forming their mental models. I found that for the design to work, play-based learning needs to be consequential. Users appreciate in the design of Ant Adaptation that you have to figure it out. While scaffolding at the start enables beginning the exploration, players want to be architects of knowledge (their own and peers). As evidence of this, in chapter 3, Thomas says why he liked the game, “yeah, you had to figure it out and the-- you have to have some flowers, see, and then you put the chemicals and lead it to there, then they’ll bring it back, and like, if you want to get rid of the chemicals you use the vinegar.” He says this as if he is teaching his siblings, telling them what they should think of the situation. Here, he is taking on the role of a peer-teacher. In other words, the game is enabling the social properties of the restructuration (Wilensky & Papert, 2010) as kids want to tell other kids about what they are finding and build knowledge.

Second, the timing of the response is critical. The close timing of players’ actions with changes — such as adding pheromones to the model, changing how ants move — accelerated learning about pheromone trails. This makes sense. In digital game design, Gregory (2018) argued that when actions and results happen within one frame, or approximately 10 milliseconds, players connect between cause and effect more readily. We see this occurring when players can touch the screen to interact with the complex system. Designers should include immediate response when designing future complexity learning games to hasten learning because when we design close connections, it causes proximity between user actions and effects.

Purposeful engagement with peers and the mediating object drives the learning in Ant Adaptation. The design of Ant Adaptation allows us to identify and describe learning like in earlier microworld (Edwards, 1995) work by examining how players reconstruct provisional theories considering dialogue between theory and evidence (Wilensky & Reisman, 2006). The decision built into the main action of Ant Adaptation — whether to peacefully collect food to increase population by employing feedback cycles or go to war to eliminate their opponent – sets up a crucial engagement where the uncertainties make the testing immediate and productive (D'Mello & Graesser, 2012a). The decisions players make allow their colonies to thrive or to die. As I present in chapter 3, the design also hastens the learning, compacting it into play sessions under 10 minutes.

To summarize, we should design ways to facilitate complexity learning with three design principles:

1. People engage part of a complex system, attempt their best theories in real-time.
2. People receive dynamic feedback from the computer and each other as immediately as practical. Ideally feedback should come in under 10 milliseconds. This time scale best ties cause to effect.
3. Learning moments happen most when players notice limits in their thinking (ruptures) and accommodate the new situations; these are situations where they have to break early intuitions in order to get out of their confusion (D'Mello & Graesser, 2014a). When they do, learners engage in an intense, purposeful, psychological effortful, problem-solving activity (D'Mello & Graesser, 2012b).

In the museum, Ant Adaptation impacted users in four ways: First, in competition with an opponent, construction of their colonies afforded comparisons that allowed for dynamic theory validation and imitation. For example, Thomas placed food close to the nest and drew ants to the food that showed he understood the proximity and connection of flowers to the nest population growth. We saw this again with Bob and Mel after they refactored their knowledge. This moment galvanized the creation of their own knowledge, as they recapped what they knew and share it with each other as a group. This interaction was notable because we did not see this kind of seminal moment when Mar was playing on her own with just the interviewer. These are examples of learning to make micro- to macro-level connections through an agent-based model in a team.

Second, the other team copied Thomas's strategy. Sharing strategies allowed players to update their operating theory. This finding suggests we should provide comparison models when using agent-based models to teach. Third, taken together, these scaffolds facilitate players' exploration and learning about the complex system.

Within less than a quarter of an hour of play, the game facilitated one player to switch from a direct schema to an emergent schema during a conversation with his brother. Thomas teaches Ed that the direction of pheromones leading successfully toward food increases the ants' population while they play the game. As he traces his fingers across the screen to draw the pheromone trail to the flowers closest to the nest, he says, "See, they are adding more energy," indicating that the ants gathering food increases the colony's energy in the game monitor. Ed says he will just watch, arguing, "You know what you are doing, so I will just watch," and leans back a bit, setting his hands on the edge of the table the screen is sitting on. Thomas then crosses his arms and says, "All right, let's just see it go on." Meanwhile, Sam has also learned through watching these two to use pheromone trails to direct his ants, drawing a long pheromone trail from Thomas and Ed's colony to his own, potentially confusing Thomas and Ed's ants. Thomas realizes this and says, "Hey, hey, hey!" demonstrating annoyance with the first aggressive move of the game. Through scaffolding, Thomas taught both his older and younger brother how to draw pheromone trails so that the ants interact with them and how to use them to have the ants initiate a feedback loop of grabbing food and returning it to increase the population. When Sam tries to add another trail, Thomas hits his hand, "And that's when you ruined our plans," Sam responds. From interactions like this, Ed learned ants are self-organizing, and that annoyed him. He had an expectation of the video game that as the player, he was in charge, and through the intervention, came to understand ants as autonomous, non-goal-oriented agents in the system.

I find in chapter 5 that designing for affective engagement creates a happy hypothesizing interaction where users smile as they contest these immediate and productive uncertainties. As we see in chapters 4 and 5, we see this process through affect tracking and follow this elaboration through CDM. In short, as designers, if we turn over knowledge authoring to the user of the game, provide enough scaffolding at the start, and allow for players to address uncertainties through discussion and competition, players can learn how the system works. If that system is about powerful ideas such as self-organization and reasoning between individual agents and macro-scale outcomes like population growth, they will overcome the twin challenges of complexity learning and understand those through immersion in the microworld.

2.3 Answering Question 3

For question three, *Are social insects a helpful means of teaching about complex social systems*? *Do these thinking tools help them think about core global problems?* In terms of Bob and Mel's knowledge building resulting from interaction with the social insects, we see the following six aspects: first, they took a key moment during the interaction to re-order their knowledge, demonstrating a shift from what they brought into the interaction, and what interacting with the game engendered. This key moment was driven by the fact that prior knowledge of ants was insufficient to account for what they were seeing in the game. As a result, they tried to organize their knowledge to account for the observations they were making about the game.

Second, the restructuring led to them explicating how ants know what to do and how the interaction of flowers and ants leads to population growth. In other words, they explicated how self-organization works and how agent actions, such as food collection, affects population growth rates. That they were interacting with social insects—a somewhat familiar, but not entirely known group of organisms—allowed researchers the opportunity to observe players while they update their understanding. While this rapid growth would not be available in an ideal future, where every schoolchild learns to understand complex systems through playing Ant Adaptation, at the time of writing, the context of ant colonies is a fecund ground to primarily introduce people to reasoning about the connection between the simple rules of agents, and the macro-outcomes of the collective.

Third, they explicated the rules of how ants know what to do. We saw this when participants described trail formation, or in their description of the connection between the cost in food for the colony to make one more ant and its connection to food intake. In the game this cost is called *create cost*. Ants provide the perfect environment to reason about feedback's role in social formation. Additionally, the use of create-cost in Ant Adaptation provides a way for players to compare the benefits of different adaptations in a social organism. While the mechanism is not perfect, leading to some confusion around what a 'create-cost,' is, such as when Bob and Mel in chapter 5 spent seven minutes solving for the equation of create-cost by systematically varying the sliders, this mechanic does allow for questioning. That the question asking is bound into the ant social reproduction supports the idea that ants are a model system for thinking between individual agents and the macro-outcome. Still, there is room for improvement on using create-cost.

Fourth, learning about the complex system with a digital ant colony allowed users to describe their different goals in the game. Bob wanted to optimize parameters in a changing ecosystem, thus role-playing evolution in this system, whereas Mel set herself the goal of discovering how to make the two colonies equal, expressing an equity lens. She set about running experiments to try to figure out how to achieve her goal. This different goal-setting played a tacit, but crucial step, in their interactions. The game provides very little description of goals. As a result, users at first try several different things, playing down trails, adding food, adjusting sliders. As they engage with ant colonies, they develop intuitions about how the interaction with individual ants leads to population growth.

Fifth, Bob and Mel learned how ants can control their own colony, grow the population, and integrate into a complex ecology. All the users started the interaction thinking merely about the ants, but through interacting with the digital ant colonies, users started to discuss a whole host of additional ontological entities present in the game, including flowers, pheromone trails, and flying ants that found new colonies. As a result, users begin to discuss the ecology of items in the system and their co-joint impacts on the emergent outcomes of the model. The insect social system provides an awe-inspiring environment to explore and learn about complex ecologies made of only just a few entities.

In terms of Mar's knowledge building, we see the following three features: first, Mar elaborated functions while she worked with Ant Adaptation. The three functions together described an understanding of how an ant colony can collectively feed itself and grow the population, without the need for a central controller. Second, she explicated the rules of how ants group up, how ants pick up and carry food, how the rules of how ants walk drives their ability to find new food and exploit food patches they have already found using pheromone trails. This function expanded on her prior knowledge about ants by giving her a way to explain what she knew coming in: that ants walk, carry things, and group up. In short, I have evidence that Mar happily learned how ants can control their own colony, grow their population, and integrate into a complex ecology. Third, notably, Mar did not share much about her goal setting. This might be because she played by herself, so had less reason to explicate her goals. In short, by learning from the complex system of an ant social system, users learned deeply about the roles that adaptation, equity, and ant behavior serve in the system. They were able to explore these interests in the sandbox mode that Ant Adaptation affords.

Sixth, this new information allows players to set their own objectives in their learning activity. This process was aided by their prior familiarity with ants and grew through interacting with the ant game. They didn't have to ask about ants, because they were familiar to them, and as a result, went about the task of understanding the complex system without questioning if they knew enough.

2.4 Answering Question 3.1

For the second part of question three, *Do these thinking tools help them think about core global problems?* Design should incorporate Vygotsky's methods of describing play. For Vygotsky (1966), play has four characteristics: (1) it is the fulfillment of an activity that the player(s) cannot engage in in real life, such as a young girl playing as a powerful queen, because she is bullied at school; (2) it contains rules that players agree to and maintain through interaction; (3) it is a partial recollection of something that actually happened; and (4) it is some sort of record. He argues play has six functions for a child. She can (1) create imaginary circumstances and (2) change affect and circumstances. (3) Play also provides a transitional state to operate on meanings, (4) that leads to the ability to think as movement in a field of meaning. (5) Play also subordinates action to meaning (6), which allows a child to operate a zone of proximal development (ZPD) to live up to her best imagined version of herself, for instance, two twin sisters playing "sister" where they enact their fullest version of their social role. Vygotsky built on two theorists, Piaget and Spinoza, to argue these functions to abide by rules developed during the first few months of life and that the child builds on these functions to develop moral rules. Piaget's work on the moral development of the child through the play of marbles (1933), showed the rules and social regulation of rule generation and maintenance. In play, young boys contest, create, and continue rules to govern their games. Spinoza pointed out that incentives influence behavior. He posited that people affect what they affect unless the incentive to do otherwise is stronger. As a result, Vygotsky (1966) argues that moral development, through the development of the ability to act in a mental, non-visible field (cognitively) through play undergirds the development of morality: "Why does the child not do what he wants spontaneously at once? Because the rules of the play structure promises much greater pleasure from the game than the gratification of an immediate impulse" (p.14). Through play the child develops a new set of desires, teaching her desires through interaction with a fictitious, best-imagined self. This ZPD provides two interesting features:

1. The imagined individual guides the child's future self.
2. This ZPD often forms around an imaginative pivot: a broomstick pivots in the child's mind to be a horse that gives the girl the ability to ride over the desert sands. Behind the couch becomes a "robber's den."

Through these features, games like Ant Adaptation (mediating objects) allow the child to pivot into imagined space, where a child forms rules that organize life and plays with ideas about situations. These situations are often ones she does not have the ability to change in her everyday life. In this way, computer-based models and games may provide an outlet for this wish fulfillment and moral development. For example, when Greta Thunberg tells the world to do a better job of dealing with climate change after sailing across the Atlantic, she cannot directly affect the outcome of slowing climate change. Instead, she plays in the game of international relations that could lead to new policies to control greenhouse gas emissions. Through interaction around games and peers, a person can fashion a best self and future directions for their development through working in a field of meaning, such as a game. If Vygotsky and Piaget were right, that play develops self-control in children, to reason about situations they cannot do anything about in their actual lives — issues like poverty, traffic, or climate-change —then in the future play-based, imagination inspiring games can provide a space to moralize and reason about issues outside youths' control. These considerations are key to the structure and design of play-based learning environments. As designers of these environments,

we need to provide learning spaces that allow students the ability to think through these consequential, often complex, situations in their everyday life.[22]

Unfortunately, there is no evidence from the dissertation that Ant Adaptation helps people understand other core problems. This is a result of the study design. I designed the interviews to ask about ants, and hypothesized users would bring up analogies in other, human settings. I chose this design, because I saw, in an earlier field trial —a summer program — that the game elicited analogical reasoning about core human problems.

During a summer program in the summer of 2018, I helped with a short course about complex systems. Members of the Center for Connected learning, including myself, held an eight-day summer enrichment program on programming complex, socially relevant, NetLogo (Wilensky, 1999) models in a large midwestern city. A non-profit focusing on work experiences for at-risk youth recruited the participants. The students of the implementation were 10 high school students living in low-income areas. On the first day, we demonstrated how NetLogo works by using Ant Adaptation. During that demonstration, I found a group of youth used a game about two ant colonies fighting over food as a pivot to reflect on, and then provide solutions to violence. Sean

[22] This kind of design would be critical in supporting a design that affords discussions about core global problems. Core global problems are often dialectic.

noticed violence—"Oh no! They're killing him!"—Juan confirming, "Yeah, he got beat up. He ran home." This anthropomorphism was the first projection of the human world onto the ants. It reminds us of a Vygotskian (1966) pivot mentioned above. The teams continued exploring the model, constructing a mental model of how the game worked. First, they worked on getting some food. They discussed how aggression takes up a lot of energy. Victoria summarized this idea, "Aggression takes up a lot of energy, so we'd need twice as much food." Juan also pointed out that the teams do not have to fight. He said, "We are not enemies. When one team increases the other team does too. More food helps all of us.", noticing the populations are tied to total food in the system. From these realizations they decided together that they should keep aggression low. They also found that pheromone trails attract ants. As a result of these three realizations, they decided to focus on increasing their own ant colony's populations. They did this by tying micro level actions to macro level population level changes. Hector said by "Putting the flowers close to the cities. Agriculture. Shortens how far our ants have to walk. It increases their population." This striking realization about how agriculture close to the population increases population caught me by surprise; it shows an unprompted leveraging of agent's rules to make sense of and reason about population level behavior. Using this realization both teams raised their populations of ants to 300 each through placing food close to the colonies. Once they did, they began to fight with each other. The fight started when Juan erased Victoria's trail to food. Victoria accused Juan of being aggressive, "You tried to destroy our bridge [pheromone trail] when you were attacking." Juan responded "Well, we did it only because you whispered you would attack us. Once you used vinegar you just started whispering all your stuff anyways." Bria said, "You didn't hear Victoria whisper anyways..." Admitting they had been planning to attack.

They then drew pheromone trails towards each other's colonies to draw their ants into a fight. Their ants engaged in a protracted fight. Their populations declined but they continued to use pheromones to draw their ants towards each other to fight. Sean chanted "fight, fight, fight!" In the end, both colonies' populations collapsed as fewer ants collected food, and more fought what seemed a tit-for-tat war.

The students then began to pivot the field. They began to explore issues that matter to them by anthropomorphizing ants. For example, Juan then analyzed the activity: "This is like gun violence. They killed us. So, we killed them. It just happens back and forth. Starts with a stupid thing. Then the system takes over. It's just like it happens in my neighborhood. Kids just killing each other." Victoria went on, "But how could have we stopped the killing?" Bria said, "Vinegar, which kind of erases the memory of the way we got to fighting." Juan summarizes, "Yeah we have to forgive and move past the memory of the fight to move on. Or we could have talked about it beforehand. We could have used diplomacy." Juan here suggests a higher order solution to violence, leveraging their analogy to gun violence to explore warring ant colonies. The two sides could talk it out and reach the same result as just forgetting with time or a "vinegar" intervention to wipe away the trails to violence. Here students used context to understand bidirectionally: the players pivoted the game to understand their own situation while using their own situation to understand the game.

In this case, Latinx and African American youth in a constructionist learning environment focusing on ant colonies, made use of a game of ants competing over food to pivot the ecology model into being a representation of neighborhood violence. Once in the new imagined space, they used the game to narrate how violence gets started for trivial reasons. They saw that the reciprocation of violence between the colonies in the model organized a system and related that back to their own neighborhoods. Here, play with a constructionist learning game sparked multilevel thinking about social issues, and Bria saw the role memory and forgiveness play in ending cycles of violence. A play-based, imagination inspiring game provided a space to moralize and reason about issues outside youths' control. Juan saw that through the development of a higher moral order, of talking out problems, the situation could have been prevented. Here the youth used play, in an unrelated complex system model, to reason about issues they face in their own neighborhood. Unbidden, these children made parallels between the different complex systems, and discerned an emergent pattern, as well as solutions. As a result, in this case, Ant Adaptation led them to think about these kinds of problems.

However, users did not bring up these analogies in the work in this dissertation. I do not know if this was a one-off occurrence, or an affordance of Ant Adaptation, that simply is not appearing more often in my data due to sampling bias. As a result, I suggest if researchers want to hear about what people think about core human problems, during a confined research period, the study should be designed to ask about them. While I have no direct evidence of users talking about analogous situations in the dissertation data, the type of deep observations participants showed about understanding structures of a system suggest additional uses. As a result, it is left to future work — with a more targeted research design — to explore the potential for analogical reasoning between complex systems.

2.5 Answering Question 4

For question four, *Does the technique of CDM add new abilities to gain new insights into how learners, advance their learning?* I found overall that CDM demonstrates learning as concept elaboration over time through the proxy of changes in speech. As shown in chapter 4, tracking elaboration in discussion improves researchers' understanding of the learning. With Rebecca, we can observe conservation of concepts in transcripts more clearly. Using CDM researchers can capture visitors' moment-to-moment sense-making. In Rebecca's case, moment-to-moment expressions stuck together as they accounted for what Rebecca saw. While she connected the declarative knowledge I probed with, when they no longer helped her understand the ants, she dropped the ideas. In this way, her knowledge fluidly grows through action taken in a complex system. Rebecca drives growth by choosing what to explore, and what questions to answer. In this way, Rebecca is the architect of her own knowledge. The dynamic interaction with the model enabled users to learn through mediation with the computer system, each other, and a facilitator. CDM allows us to track this dialectic interaction in the learning environment.

Hegelian dialectic introduced the power of contradiction a transformation in the maintenance of an ontological entity. Hagel pointed out that contradiction are transformation are the core parts of identity (Morin, 2008). In other words, what defines these problems is the complex system of the material world and discursive process that maintain them as issues. In Ant Adaptation they two colonies vying over the food sources creates a conflict, that as their networks grow, causes transformation. This isn't an ordered equilibrium, but an emerging, transforming system. CDM track as participants learn during the interaction as they come to identify the dialectic, self-organizing complex systems entities.

CDM is applicable to studying other learning environments, especially open-ended ones like we find in museums or maker spaces. Effortful problem-solving activity is the process of science, and that is the process that constructivist dialogue mapping tracks. The CDM approach introduced to capture learning, common in this project, was able to capture changes in a player's understanding of agents, functions and properties (entities' mechanisms) while they learned about complex ant systems through playful interaction with Ant Adaptation. CDM captured the changes as utterances occurring during a short interaction. By analyzing changes in talk pre- and post- play, we found that players learn about feedback, and employed that learning at multiple levels to maximize an ant population. As we saw in chapter 5, CDM affords following how people come to understand the functions — such as ants grouping up, gathering food, or laying pheromone trails — that drive agent-based models as they discuss the behavior of ants in the agent-based model. Additionally, CDM is suitable for studying how one person or a group learns. As we saw in chapter 5, CDM affords identifying moments where two players stop and then restate and reformat their knowledge, as we saw 16 minutes into Bob and Mel's interaction. CDM allows new perspectives on how learners fluidly advance their knowledge during these informal learning environments.

There is a tension though: how long do we think these knowledge structures persist? On the one hand, there are short explanations that people come up with based on their prior experience, and on the other hand long term, stable cuing explanations, resulting from networks of p-prims. As discussed in chapter 4, knowledge in pieces is an approach that primes this dichotomy. Knowledge in pieces (DiSessa, 2018, DiSessa, 1993) addresses the following questions: what the elements of knowledge are, how do they arise, what level and kind of systematicity exists, how does that system evolve, and what can be said about cognitive process that underly the system and its operation. The study of this framework looks to study moments of learning that are consistent with constructivism. The theory proposes many fine-grained bits of knowledge (p-prims). P-prims are micro generalization that people abstract from experience. They are small knowledge structures that get enacted by being recognized and cued to an active state on the basis of the perceived configuration. P-prims are mid-level cognitive elements, neither the low-level sensory information, nor the high-level named concepts or categories. P-prims are activated based on a cuing priority. DiSessa does not focus on the origins of p-prims, but instead focuses on their life histories, especially how they might “become embedded in more physics sophisticated thinking” (DiSessa, 1993, p. 114). In this way of thinking, one might see the expressions documented by CDM as unstable explanations, perhaps not learned during the interaction, but instead expressions arising from prior-experience. In that perspective, participants may not retain much of what they are talking about during these interactions. This is an empirical question: in what ways do the examples of CDMs developed during play inform, or reorder p-prim networks related to complex systems heuristics? Thankfully, the method of CDM allows us to study the stability overtime, and is designed to study the evolution of the ideas. To conduct the study requires only a larger, longitudinal study.

2.6 Answering Question 5

For question five, *What new can we learn from physiological measures of affective states, while people engage with a museum exhibit?* I found four uses. First, in Mel and Bob’s case in chapter 5, I found joint users display widely disparate readings of affective state intensity. The variance can be measured with time series analysis, which provides a way to go back to the video and transcript data and inspect what is happening. This is a valuable analytic tool for researching with qualitative data.

Second, we can describe who does most of the talking. In multiple user cases, the ability to notice when users engage and disengage is important. For example, during silence, we can ask what Bob was doing during that time. Time series analysis makes that possible. If applied to other museum learning activities, we could track when, and how much users talk after walking up to an exhibit. In a classroom setting, we could see when different people participate and how much they talk. If elaboration is learning (Leinhardt et al., 2003), measuring how much elaboration is happening becomes a key measure of accessing learning environments.

Third, when this is coupled with CDM, we can investigate how much people participate, and the content of that elaboration at the same time. In other words, affective expression tracking can grant us insight into the entire interaction through the copious data. Using that data-stream and synchronizing it with qualitative measures of learning is an exciting new application.

Fourth, in all cases, we can describe the frequency of affective expressions. For example, we can say quantitatively that Bob in the first half of the interview did not engage as much, but became more ebullient and loquacious towards the latter half. This approach to studying affective engagement affords this insight.

One issue the study of human analytics raises is the fear of a surveillance society. This work is trying to keep that genie in the bottle, while allowing us to use the deep insights of close qualitative analysis to bring out the human side of learning. Recent critique of automatic facial tracking has pointed out that Silicon Valley is over-promising on what facial tracking can do. Barrett (2019) explicated the ways that facial action units cannot track emotion, and that emotions are culturally construed. Crawford, in *The Atlantic* (2021), recently laid out the ways firms are misrepresenting what Ekman's (Ekman & Friesen, 1978; Ekman & Rosenberg, 2005) approach affords. To be clear, action units cannot say anything about emotion. Instead, they allow fine detailed analysis of the frequency and patterns of facial expressions. While facial expressions, such as a genuine smile, do accompany internal states, currently we don't have means by which to disambiguate a genuine smile from an affirmative smile confirming attention. As a result, what we can study is limited to areas of when smiling or talking occur, and then offer insights from there. We can test whether qualitatively identified actions co-occur with expressions and which expressions come before or after moments, such as elaboration. The richness of the affect data, and the possibility to synchronize with interaction analysis data, suggests several new avenues for research. As a result, through computational ethnography we can enable ways to assess larger scale learning environments while attending to the intricacies of individual learner pathways.

2.7 Answering Question 6

For question six, *Is there a relationship between high stimulation and learning as measured by CDM and affective computing signals? Within learning moments, is there a relationship between positive affect and learning?* I find there is. In chapter 5, in Mar's case, the main emotional pathway around moments of learning was the delight pathways. When I identified moments of learning as times in the transcript where Mar elaborated a concept during game play drove the finding. In all situations like this, the primary affective expression was smiling. A lot of smiling accompanies moments of learning. There is a positive correlation between learning identified by CDM and positive affective states. During Mar's interaction around learning moments, delight is the predominate emotional pathway, accounting for 29,490 of the pathways around learning moments. The next two biggest emotional pathways are contempt and contempt delight pathways, account for 2,713 occurrences. However, as Mar's sips of coffee highlighted, researchers need to pay careful attention because not all the moments that expression analysis brings to light are salient. In other words, expression analysis simply highlights an otherwise unremarkable sip of coffee. The sensitivity of the method speaks to both its power, but also the caution we need to exercise with its use.

What otherwise forgotten moments come to light when we change our epistemology, or ways of knowing? In chapter 5, I shifted our way of seeing by performing predictive modeling, using facial expressions to predict moments participants elaborated. In chapter 5, through random forest modeling, I achieve a high the degree of predictive power using expressions to identify moments participants elaborated. For the first case, Mar's, obtaining 96% accuracy, 97% precision and 81% recall. For the second case, Bob and Mel's, I obtained 90% accuracy, 96% precision, and 73% recall. To confirm the pattern, using an XGBoost model I find support as well. Using XGBoost to identify moments of elaboration for the first case, Mar's, XGBoost identifies 97% of the 0s and 76% of the 1s, with a total accuracy of 93%, precision of 76%, and recall of 89%. For the second case, Bob and Mel's, XGBoost correctly identifies 84% of the 0s and 70% of the 1s, with a total accuracy of 79%, precision of 70%, and recall of 70%. With XGBoost top predictors of moments participants elaborated are smiling (AU06 and AU12) suggesting smiling may be a meaningful component in identifying moments participants elaborate as they play in Ant Adaptation, and learn about complex systems. The findings show, using a train test split, and the extensive data provided by affect detection, we can develop algorithms to predict moments participants elaborated identified through CDM. If we can identify real time learning moments, these findings would allow us to offer recommendations and respond to the learners within the critical 10 millisecond response range discussed in chapter 3. Furthermore, with further affective analysis, with these methods researchers could identify moments where students get stuck, and provide real time feedback in an instructive learning agenda. This opens the possibility of doing the analysis at scale, as we develop more accurate models. The process of synchronizing qualitative methods and learning analytics in this dissertation provides an example of plugging in as many sensors measures as needed to provide real-time learning analytics, and refocus on outcomes of learning such as learning engagement in group setting. This is a powerful means of understanding the impact of our learning designs. My process of dedicated design of a restructuration of a learning environment to teach complexity along with the innovation of learning methods is already promising. In future work, I will continue to validate the findings.

3 Contributions, Limitations and New Directions

In this section I zoom out, and try to see the forests for the trees to focus on why this research is important and questions it raises for future work. I discuss what I have introduced and the contributions. In each part, I highlight what questions remain, the limits, and exciting new directions.

3.1 Contributions

This project has five major contributions: Ant Adaptation, contributions to assessment in constructionist, open-ended, learning environments, methodological innovations, and the use of predictive models to assess these complex situations. Next, I will turn to each.

3.1.1 Ant Adaptation

I would be remiss not to highlight Ant Adaptation. The environment has various forms. I developed it for museum exploration, where groups of students play around a tabletop interactive display. This approach has affordances in facilitating group learning as they compete, contest approaches, and develop an idea of how ants self-organize. I developed a blocks-based, scaffolded version of Ant Adaptation to allow users the ability to uncover and reprogram the activity of the ants using domain specific blocks. This version facilitates computation thinking about wiggle-radii, trail formation, and deformation, and ants' food foraging through combined group efforts. It also serves as an excellent set of scaffolds for engaging with the full model. Further, as I will discuss more below in the new directions portion, I've developed a 3D version of the environment called The Ant Game, to be released on Android and iOS devices that uses the methods learned from testing in the other two forms to scale the constructionist learning opportunities. These forms have been used in many different settings from museums to remote learning to classroom embedded curriculums to teach biology, complex systems, agent-based modelling, and social dynamics.

Sometimes, learning sciences students do not foreground the design contributions of their environments. As a result, I want to highlight Ant Adaptation itself in its various forms and its usable in different settings. But what kind of contribution is it? It is a created a design artifact that blends, 1) influences from constructionist microworlds, 2) games and learning, and 3) complex systems and agent-based modeling. The environment goes a significant way to intersecting these approaches in a generative way. Each approach provides a way to address some shortcomings of the other as we more fully engage students in immersing in microworlds through play, we can better enable complex systems learning through open-ended player centric approaches to learning. In future work I will continue to explore this contribution. There are some early limitations to the approach, such as, players have expectations of games that Ant Adaptation does not fully address, they want set goals, and rewards, and video game players especially, want these built into their environments. Learners, educators, and parents also have expectations. Parents want their kids to spend less time on screens and to be involved in real nature learning outdoors. Educators have standards they need to teach, and timetables to keep, and open-ended play with a microworld does not always fit into that. While I have shown in chapter 3 that the environment can teach important lessons in under 10 minutes, some of the explorations, especially those in chapter 5, lasted an hour, as users engaged with the game and each other. What classroom has an hour to spare for such self-directed, open-ended play? Learners also expect learning to be hard, not fun. They think they need to memorize, not explore. Additionally, learners have resistances to doing things in new ways. The highest performing students in a standard classroom have a vested interest in continuing the status quo. As a result, we have to ask who are the key change makers in each of these groups that have undercome these understandable reservations with new approaches?

3.1.2 Assessment in a Constructionist Way

I contribute to constructionist theory and methods. There are acknowledged difficulties in the constructionist community of assessing open-ended types of environments. In this project I am building tools to address these difficulties of assessing in a much more open ended, and in the United States, designing for a context where there are very specific targets. For me that is a big contribution to the work, but it goes beyond this thesis; these methods could be picked up by other people who are interested in constructionist designs and advocating changes to the learning environments.

So, while there is an acknowledged issue of assessing, as a subheading to that, there has been an issue of trying to think about groups as a unit of analysis, and attempting to evaluate and understand what happens. The approaches used here might be especially suitable for informal settings like museums. In this perspective, CDM adds to the list of tools that allows us to think about the thought of a group—as opposed to as an individual—and have some way of comparing, documenting, and categorizing it. This is a gap in the constructions literature that I am filling with my approach.

3.1.3 Research Methods

I have contributed methods of research that when combined are useful in conceptualizing learning. It is not just, I made this tool, but instead this tool has power in it and the tool can be taken further. In that way the contribution is the start of better methods for studying groups and moments where people have a click moment. The approach is coming from a point of view where I am shifting a little bit the perspective of learning away from individuals, and towards thinking about groups as a primary unit of analysis. That shift is a direction for the learning sciences, that my work is demonstrating some first steps into. In other words, if minds are complex systems in congress with other minds, we need methods of analysis that study these systems in vitro in order to research their learning. I am trying to focus the field more on groups, and I am influenced by cultural, historical theories, but also AI theories about whether the nature of mind and learning is right. And that leads me to need a tool like CDM.

3.1.4 The Use of Models for Assessing Constructionist Learning

There is a contribution of the use of machine learning models to study the moments participants elaborate while learning complex systems through Ant Adaptation. The use raises some quandaries for me: does it fit with constructionism and personalized learning. An important question my thesis surfaces is to what extend are machine learning methods fully translatable. How do you square this new sensor methodology, with lots of data capture, with constructionist methodology where we are trying to understand how the person themselves makes meaning out of it? These are two, dual ways of knowing. It may be reaching, but humans themselves are complex systems not amenable to averaging. If minds are complex systems in congress with other minds, we need methods of analysis that study these systems in vitro to research individual and group learning.

This will involve more qualitative research studies of constructionist learning environments that take more account and grant greater agency to participants in their own data, and learning.

Furthermore, to take the study further we will need additional research into the mechanism of affect and learning. To verify the methods showcased in this dissertation will require at the very least larger scale implementations that account for these questions of causation and validity. Below I map out a prototype study to further understand the mechanism.

In that research, we will need to study the causal question of interest: We want to understand the effect of elaborations, conceived of as the activity of cognitive and affective machines in people and among groups, on the emergence of alternative affective-states that influence student behaviors. To test this, we need to design experiments. If I were to design the ideal experiment that would be conducted if ethics and money was not an issue, I would promise two groups of students (representative samples randomly assigned) a reward. I would also have longitudinal behavioral data, for instance, school records, future earnings (in hourly wage), any labor market outcomes, student achievement, family and community context, school records, social peer group and interaction. After letting both groups believe they will receive treatment, that is an ant-based learning environment, I will let the control group play Ant Adaptation, and arbitrarily provide an alternative comparative learning environment to the treatment group. Then research would track impact/difference in affective and elaborations metrics over time for treatment vs control. This would account for omitted variable bias and would give us a perfectly representative population for external validity. Next up we would develop an identification strategy.

Even with this design, the research could have threats of internal validity: such as cutoff, when comparing students between groups, are they the same besides random assignment? Additionally, there is a threat to internal validity because of spill-over, students who would be the best comparison may be in the same schools. If that were the case, this would raise the spectre of contamination: students in one group talk to students in the other passing on the effect through social networks. We would need to address this. Finally, student attrition in the program may cause change. Such attrition could lead to lower or higher representation of affective displays among participants.

There are also threats to external validity to such an experiment. Chief among them, the data is only representative of students that are in the program instead of the whole body of students. We should consider when we make claims about affective expression formation—when it is a contextually bound individual experience—that this issue is especially acute. Addressing this remains an open threat to external validity and generalizability of the research.

In short, to study the methods showcased in this dissertation will require at the very least larger scale implementations that account for these questions of causation and validity.

3.1.5 The Use of Social Insects to Teach Complexity

As shown in Figure 1, this project was inspired by the idea that social insects, like ants, were a good vehicle in order study learning about complex systems. This feeling got buried a little bit in this thesis. But, while it’s backgrounded a bit, I think one of the reasons why my data turns out so nicely and people have these wonderful conversations is because they are using complexity and they are using a familiar yet not familiar environment, there's enough prior knowledge that they load in, like in wolf sheep predation model. Because they have seen ants their entire lives, the context is familiar. Above in answer to question 3, I mentioned this is why social insects are a useful phenomenon to learn complexity from. I think the familiarity plays an important role, but also their alienness. The fact that they are constructing new knowledge. I think that my CDM maps look so cool is because participants are sitting there and watching, and they make discoveries about pheromone trails. Before the play, they do not know what that is. They do not really know how ants forage for food or what wiggle radius does, or how a random walking optimizes foraging or any of that sort of stuff. Regardless, players answer questions, and they surface that stuff every single time. I think that contributes greatly to the richness of my data, which then gives me these cool effects. I think if I was studying like engine blocks, or something less dynamic with these methods I think I would get much less rich data. Participants would not talk about this stuff in this kind of exciting way. They would not take on these competitive poses where identify with their ants, and really take the colonies welfare to heart. I think these are affordances of the system. But since chapter 3, I have moved away from trying to prove it works because a) I thought chapter 3 proved it works nicely and then I went on to develop new methods and keep going with the study and enrich it. And while chapter 5 continues that same kind of work, there is a limitation, neither provide evidence of efficacy compared to other options. I do not say the richer data results from the environment being social insects, because I did not conduct a comparative study.

As I said in answering question 3, the role of social insects and complex systems can be useful vehicles for learning complexity. I would suggest in making these environments, make comparisons facile, allow the self-organizing properties of social insects to take the lead. I would ask how we can get these systems into more kids' hands, and the facilitate the immersive self-organization of these environments. What sort of design decisions will lower the barriers to entry? How can we facilitate families playing these games together, and what discussions and practices best facilitate the interaction? The features that promote smiling, like the heads-up display of the attitude of the ants in the game make people laugh, this feature does not do much at all for understanding complexity directly, but they make a lot of people smile. Features like this, which improve the mood during learning are important. Additionally, being able to draw pheromone trails, and add food to the game facilitates rapid testing of how the ants work. Copying these features will be important in next environments like this and could be generalized to other systems and games. Danish and Guo have already shown bees help understanding complexity. My work contributes a significant extension to this work using ants. When we tend to think about biology, our first thoughts are humans and large animals, where complexity is more hidden, but when we look at ants, they are familiar enough, but they are alien enough learners can shift the levels readily. This affordance raises more questions and continuing the development will allow this helpful context to be one of the primary points for thinking about complex systems.

3.2 New Directions

Conducting this project has inspired several new directions including integrating social design research and a new version of Ant Adaptation. Next, I will discuss them.

3.2.1 Implications for future Ant Adaptation: Social Design

In one view, Ant Adaptation seeks to provide an understanding of how society could work without a central controller for the health of all inside. This kind of mental model is the counter to structures and systems that lead to climate change or patriarchy.

The lowly ant is resilient. Ants feature as important characters in parables in many societies, and by using them to talk about the strength of society in the face of challenge, we have a platform that engenders a discursive practice, where we can talk about social change, structure, and systems during play. I could see Ant Adaptation as a primer to understand social functions that are strongly structural and systemic, which can lead to understanding situations perceived as unjust by non-dominant communities. In this view, Ant Adaptation is a restructuration (Wilensky & Papert, 2010) engaging learners cognitively and affectively, to understand the powerful ideas of complex system, to see systemic structure, and to revalue system and structure more holistically, reforming our view of the complex system that leads the injustices as viewed from minority communities' eyes. I aim to understand and change these uneven attributions of our society. In this way, I could use Ant Adaptation as a mechanism to enable the diverse forms of expertise of non-dominant communities, leaning on something we all are familiar with, the lowly ant, to create a possibility for consequential learning opportunities on complex systems and structures. Vygotsky's descripts of play outline why this method of learning will work, and the anecdotal evidence from the after school implementation of Ant Adaptation shows that at least sometimes, this sort of serious play emerges.

While it has not reached its goal, Ant Adaptation aims to teach history and equity in mutually reinforcing ways. Through creating historically aware actors that understand the structure and the system they live within; Ant Adaptation is a system for consequential learning through contradiction and transformations where students' ecologies of learning are designed to lead to social reflection and change.

3.2.2 The Ant Game

As seen in Figure 2, I developed a version of Ant Adaptation in Unity 3D that incorporates the design lessons from this project. This is release via the Apple and Google play store (www.gettheantgame.com). I hope providing the game at scale will facilitate making these kinds of complexity learning moments ubiquitous. In the game, played on a mobile device, two colonies compete for food. The user can touch their mobile device, and learn how the ants' mechanisms of self-organization, pheromone trails, impacts the colony. Users click on their ant colony's menus to change parameters of size and aggressiveness. This game has an added feature of persistence, the game is automatically saved, so the user can come back to their colony day after day, and learn through immersion in the microworld. In this sandbox game, players learn about ideas of complexity through raising their ant colonies. I am working with teacher networks, and youth programs to study these digital ants, and then have kids do citizen science to better understand their real-life cousins, biological ants.

Figure 2: The Ant Game, www.gettheantgame.com, a just released 3D version of Ant Adaptation for mobile devices.

4 Conclusion

I structured the dissertation to take us to the next steps of this research. I used the detailed coding of interviews about complexity and methods of automatic detection to analyze moments conducive to learning while using an agent-based modeling intervention. As a result, the dissertation has two parts: first, through CDM and other qualitative methods, we coded for complexity learning, then I examined both expert and novice understandings of ant colonies' self-organization. This portion showed how learning with an ant colony improves complexity learning.

Second, the dissertation re-cut the data and focused on moments of high stimulation to find out whether affectively high moments can tell the same story with less data. These approaches may allow us to train a classifier to use video data as input to identify moments of learning on the copious data arising from affective state detection. The approach has powerful affordances for tracking learning, but it also unearths a challenge: that of incorporating AI in education research. AI systems have been less accurate at identifying the faces of dark-skinned women, to grant women lower credit-card limits than their husbands, and to erroneously predict Black defendants will commit future crimes more than whites. When applying these systems to learn about education, many of the solutions require having a human in the loop (K. Martin, Wang, et al., 2019). As a result, a key domain of my research is investigating the diversity challenges posed by AI, while leveraging the technologies' benefit for education. The connection between affect and learning, and teaching complexity through computational environments will be a double-barreled research program going forward in my career.

Through these two parts, I find that we can use a complex systems game to teach properties of complex systems through discussions. The game engenders joy and mental model formation. Following the practices of design-based research, going forward, we can develop these system-games to create more affectively and cognitively engaging learning. These games can one day make understanding the self-organizing properties of the lowly ant as prosaic as the restructuration to Hindu-Arabic numerals has made dividing 66 by 11.

5 References

Ackermann, E. (1996). Perspective-taking and object construction: Two keys to learning. Constructionism in practice: designing, thinking, and learning in a digital world. *Lawrence Erlbaum, Mahwah, NJ*, *25*, 35.

Ackermann, E. (2001). Piaget's constructivism, Papert's constructionism: What's the difference. *Future of Learning Group Publication*, *5*(3), 438.

Anderson, J. R. (1976). *Language, memory, and thought*. Psychology Press.

Andres, J. M. A. L., Ocumpaugh, J., Baker, R. S., Slater, S., Paquette, L., Jiang, Y., Karumbaiah, S., Bosch, N., Munshi, A., & Moore, A. (2019). Affect sequences and learning in betty's brain. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 383–390.

Ant colony optimization algorithms. (2021). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Ant_colony_optimization_algorithms&oldid=1032225199

Atkinson, Muro, and Whiton. (2019, December 4). The case for growth centers: How to spread tech innovation across America. *Brookings*. https://www.brookings.edu/research/growth-centers-how-to-spread-tech-innovation-across-america/

Atlan, H. (1986). *A tort et á Raison*.

Ax, A. F. (1953). The physiological differentiation between fear and anger in humans. *Psychosomatic Medicine*, *15*(5), 433–442.

Babaoglu, O., Meling, H., & Montresor, A. (2002). Anthill: A framework for the development of agent-based peer-to-peer systems. *Distributed Computing Systems, 2002. Proceedings. 22nd International Conference On*, 15–22.

Baker, R., Rodrigo, M. M. T., & Xolocotzin, U. E. (2007). The dynamics of affective transitions in simulation problem-solving environments. *International Conference on Affective Computing and Intelligent Interaction*, 666–677.

Bakhtin, M. M. (1981). *The dialogic imagination: Four essays by MM Bakhtin (M. Holquist, Ed.); C*. Emerson and M. Holquist, trans. Austin: University of Texas Press.

Bamberger, J. (2001). *Turning Music Theory on Its Ear: Do We Hear What We See; Do We See What We Say? Multidisciplinary Perspectives on Musicality: The Seashore Symposium*. University of Iowa Press Iowa City, IA.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*, *13*(1), 1–14.

Barceló, J. (2010). *Fundamentals of traffic simulation* (Vol. 145). Springer.

Barrett, L. (2019). *Can Machines Perceive Emotions? | Lisa Feldman Barrett | ACII'19—YouTube*. Affective Computing and Intelligent Interaction. https://www.youtube.com/watch?v=gPPHgeJHRF8&feature=youtu.be

Bar-yam, Y. (1999). *Dynamics Of Complex Systems* (1st edition). CRC Press.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323–370.

Beekman, M., Sword, G. A., & Simpson, S. J. (2008). Biological foundations of swarm intelligence. *Swarm Intelligence*, 3–41.

Berland, M., Baker, R. S., & Blikstein, P. (2014). Educational Data Mining and Learning Analytics: Applications to Constructionist Research. *Technology, Knowledge and Learning*, *19*(1–2), 205–220. https://doi.org/10.1007/s10758-014-9223-7

Bickmore, T., & Schulman, D. (2007). Practical approaches to comforting users with relational agents. *CHI'07 Extended Abstracts on Human Factors in Computing Systems*, 2291–2296.

Blaho, A., Kalas, I., & Tomcsanyi, P. (1993). Comenius Logo: Environment for teachers and Environment for learners. *Proceedings of the EUROLOGO*, *93*.

Blikstein, P., & Wilensky, U. (2009). An atom is known by the company it keeps: A constructionist learning environment for materials science using agent-based modeling. *International Journal of Computers for Mathematical Learning*, *14*(2), 81–119.

Block, F., Hammerman, J., Horn, M., Spiegel, A., Christiansen, J., Phillips, B., Diamond, J., Evans, E. M., & Shen, C. (2015). Fluid grouping: Quantifying group engagement around interactive tabletop exhibits in the wild. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 867–876.

Blumer, H. (1969). *Symbolic Interactionism: Perspective and Method*. University of California press.

Bohr, N. (1991). *Physique. In Atomique et Connaissance Humaine* (Edition Gallimard: Paris, France,).

Bonabeau, E. (1997). From classical models of morphogenesis to agent-based models of pattern formation. *Artificial Life*, *3*(3), 191–211.

Bonabeau, E., Dorigo, M., & Theraulaz, G. (1999). *Swarm intelligence: From natural to artificial systems*. Oxford university press.

Boucsein, W. (2012). *Electrodermal activity*. Springer Science & Business Media.

Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*. Harvard Univ Pr.

Bråten, S. (2006). *Intersubjective communication and emotion in early ontogeny*. Cambridge University Press.

Brennan, K. (2015). Beyond technocentrism. *Constructivist Foundations*, *10*(3), 289–296.

Brown, A. L. (1992). Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings. *Journal of the Learning Sciences*, *2*(2), 141–178. https://doi.org/10.1207/s15327809jls0202_2

Bruce, V., & Young, A. (1998). *In the eye of the beholder: The science of face perception.* Oxford university press.

Bryson, J. M., & Crosby, B. C. (2005). Leadership for the common good.

TRANSFORMATIONAL TRENDS IN GOVERNANCE AND DEMOCRACY, 185.

Buck, R. (1984). *The communication of emotion*. guilford press.

Burkov, A. (2019). *The Hundred-Page Machine Learning Book by Andriy Burkov*. http://themlbook.com/

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. *Handbook of Emotions*, *2*, 173–191.

Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist*, *45*(1), 16.

Cai, Z., Siebert-Evenstone, A., Eagan, B., Shaffer, D. W., Hu, X., & Graesser, A. C. (2019). nCoder+: A semantic tool for improving recall of nCoder coding. *International Conference on Quantitative Ethnography*, 41–54.

California Department of Education. (2021). *Mathematics Framework—Mathematics (CA Dept of Education)*. https://www.cde.ca.gov/ci/ma/cf/index.asp

Calvo, R. A., & D'Mello, S. (2011). *New perspectives on affect and learning technologies* (Vol. 3). Springer Science & Business Media.

Campa, S. T. and A. (2021, August 6). Schools Look for Help From AI Teacher's Assistants. *Wall Street Journal*. https://www.wsj.com/articles/schools-look-for-help-from-ai-teachers-assistants-11628262062

Cannon, W. B. (1927). The james-lange theory of emotion: A critical experiment and an alternative theory. *Am J Psychol*, *39*(10–124), 42.

Capra, F. (2016). *The Systems View of Life: A Unifying Vision* (1st edition). Cambridge

University Press.

Chi, M. T. (2005). Commonsense conceptions of emergent processes: Why some misconceptions are robust. *The Journal of the Learning Sciences*, *14*(2), 161–199.

Chi, M. T. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In *International handbook of research on conceptual change* (pp. 89–110). Routledge.

Chi, M. T., Hutchinson, J. E., & Robin, A. F. (1989). How inferences about novel domain-related concepts can be constrained by structured knowledge. *Merrill-Palmer Quarterly (1982-)*, 27–62.

Chi, M. T., & Koeske, R. D. (1983). Network representation of a child's dinosaur knowledge. *Developmental Psychology*, *19*(1), 29.

Chi, M. T., Roscoe, R. D., Slotta, J. D., Roy, M., & Chase, C. C. (2012). Misconceived causal explanations for emergent processes. *Cognitive Science*, *36*(1), 1–61.

Chi, M. T., Slotta, J. D., & De Leeuw, N. (1994). From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction*, *4*(1), 27–43.

Chu-Carroll, M. (2006). *Roman Numerals and Arithmetic | ScienceBlogs*. https://scienceblogs.com/goodmath/2006/08/16/roman-numerals-and-arithmetic

Clampet-Lundquist, S., Edin, K., Kling, J. R., & Duncan, G. J. (2011). Moving teenagers out of high-risk neighborhoods: How girls fare better than boys. *American Journal of Sociology*, *116*(4), 1154–1189.

Clark, D., Nelson, B., Sengupta, P., & D'Angelo, C. (2009). Rethinking science learning through digital games and simulations: Genres, examples, and evidence. *Learning Science: Computer Games, Simulations, and Education Workshop Sponsored by the National Academy of Sciences, Washington, DC*.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design Experiments in Educational Research. *Educational Researcher*, *32*(1), 9–13. https://doi.org/10.3102/0013189X032001009

Cohn, J. F., Ambadar, Z., & Ekman, P. (2007). Observer-based measurement of facial expression with the Facial Action Coding System. *The Handbook of Emotion Elicitation and Assessment*, 203–221.

Cohn, J. F., & De la Torre, F. (2014). Automated face analysis for affective computing. In *The Oxford handbook of affective computing* (pp. 131–150). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199942237.013.020

Collins, A. (1992). Toward a design science of education. In *New directions in educational technology* (pp. 15–22). Springer.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *8*(2), 240–247.

Conant, J. B. (1947). *On Understanding Science an Historical Approach*. Yale University Press G. Cumberlege, Oxford University Press.

Craig, S. D., D'Mello, S., Witherspoon, A., & Graesser, A. (2007). Emote aloud during learning with AutoTutor: Applying the Facial Action Coding System to cognitive–affective states during learning. *Cognition and Emotion*, *22*(5), 777–788.

Crawford, S. by K. (2021). Artificial Intelligence Is Misreading Human Emotion. *The Atlantic*. https://www.theatlantic.com/technology/archive/2021/04/artificial-intelligence-misreading-human-emotion/618696/

Damasio, A. R. (1994). *Descartes' error*. Random House.

Danish, J. A., Peppler, K., Phelps, D., & Washington, D. (2011). Life in the hive: Supporting inquiry into complexity within the zone of proximal development. *Journal of Science Education and Technology*, *20*(5), 454–467.

Darner, R. (2019). How can educators confront science denial? *Educational Researcher*, *48*(4), 229–238.

Darwin, C. (1872). *The expression of emotions in men and animals*.

Davies, P. (1989). *Cosmic blueprint: New discoveries in natures ability to order universe*. Simon and Schuster.

Davis, P., Horn, M., Block, F., Phillips, B., Evans, E. M., Diamond, J., & Shen, C. (2015). "Whoa! We're going deep in the trees!": Patterns of collaboration around an interactive information visualization exhibit. *International Journal of Computer-Supported Collaborative Learning*, *10*(1), 53–76. https://doi.org/10.1007/s11412-015-9209-z

DeLiema, D., Enyedy, N., & Danish, J. A. (2019). Roles, Rules, and Keys: How

Different Play Configurations Shape Collaborative Science Inquiry. *Journal of the Learning Sciences*, *28*(4–5), 513–555.

Delva, M. D., Woodhouse, R. A., Hains, S., Birthwhistle, R. V., Knapper, C., & Kirby, J. R. (2000). Does PBL matter? Relations between instructional context, learning strategies, and learning outcomes. *Advances in Health Sciences Education*, *5*(3), 167–177.

Dennett, D. C. (1993). *Consciousness explained*. Penguin uk.

Dennett, D. C. (2017). *From bacteria to Bach and back: The evolution of minds*. WW Norton & Company.

Descartes, R. (1637). *Discours de la Méthode Pour bien conduire sa raison, et chercher la vérité dans les sciences*. https://gallica.bnf.fr/ark:/12148/btv1b86069594/f7.image

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, *32*(1), 5–8.

Dierking, L. D., & Falk, J. H. (1992). *The museum experience*. Whalesback Books.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, *10*(2–3), 105–225.

diSessa, A. A. (2018). A friendly introduction to “knowledge in pieces”: Modeling types of knowledge and their roles in learning. *Invited Lectures from the 13th International Congress on Mathematical Education*, 65–84.

diSessa, A. A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *The Journal of the Learning Sciences*, *13*(1), 77–103.

DiSessa, A. A., Hammer, D., Sherin, B., & Kolpakowski, T. (1991). Inventing graphing: Meta-representational expertise in children. *Journal of Mathematical Behavior*, *10*(2), 117–160.

D'Mello, S., & Graesser, A. (2012a). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145–157.

D'Mello, S., & Graesser, A. (2012b). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145–157. https://doi.org/10.1016/j.learninstruc.2011.10.001

D'Mello, S., & Graesser, A. C. (2014a). *Confusion.*

D'Mello, S., & Graesser, A. C. (2014b). Feeling, thinking, and computing with affect-aware learning. *The Oxford Handbook of Affective Computing*, 419–434.

Drescher, G. L. (1991). *Made-up minds: A constructivist approach to artificial intelligence*. MIT press.

Drescher, G. L. (1987). A Mechanism for Early Piagetian Learning. *AAAI*, 290–294.

Duit, R., Treagust, D., & Widodo, A. (2008). Teaching science for conceptual change: Theory and practice. In *International handbook of research on conceptual change* (pp. 629–646). Routledge.

Edelson, D. C. (2002). Design research: What we learn when we engage in design. *The Journal of the Learning Sciences*, *11*(1), 105–121.

Edwards, L. D. (1995). Microworlds as representations. In *Computers and exploratory* *learning* (pp. 127–154). Springer.

Egenfeldt-Nielsen, S. (2006). Overview of research on the educational use of video games. *Nordic Journal of Digital Literacy*, *1*(03), 184–214.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3–4), 169–200.

Ekman, P., Friesen, W., & Hager, J. (2002). Facial action coding system: Research Nexus. *Network Research Information, Salt Lake City, UT*, *1*.

Ekman, P., & Friesen, W. V. (1978). Facial action coding system. *Consulting Psychologist Press*.

Ekman, P., Huang, T. S., Sejnowski, T. J., & Hager, J. C. (1993). Final report to NSF of the planning workshop on facial expression understanding. *Human Interaction Laboratory, University of California, San Francisco*, *378*.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, *221*(4616), 1208–1210.

Ekman, P., & Rosenberg, E. (2005). *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Encoding System (FACS)*. Oxford University Press.

Escobar, A. (2011). *The problematization of poverty: The tale of three worlds and development*. Princeton University Press.

Eysenck, H. J. (1947). *Dimensions of personality* (Vol. 5). Transaction Publishers.

Feurzeig, W., Papert, S., Bloom, M., Grant, R., & Solomon, C. (1970). Programming-languages as a conceptual framework for teaching mathematics. *ACM SIGCUE Outlook*, *4*(2), 13–17.

Fleck, R., Rogers, Y., Yuill, N., Marshall, P., Carr, A., Rick, J., & Bonnett, V. (2009). Actions speak loudly with words: Unpacking collaboration around the table. *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, 189–196.

Forgas, J. P., & Moylan, S. (1987). After the Movies: Transient Mood and Social Judgments. *Personality and Social Psychology Bulletin*, *13*(4), 467–477. https://doi.org/10.1177/0146167287134004

Franklin, E. L. (2014). The journey of tandem running: The twists, turns and what we have learned. *Insectes Sociaux*, *61*(1), 1–8.

Fridlund, A. J. (1994). *Human facial expression: An evolutionary view*. Academic Press.

Gardner, H. (1987). *The mind's new science: A history of the cognitive revolution*. Basic books.

Gause, G. F. (1934). Experimental analysis of Vito Volterra's mathematical theory of the struggle for existence. *Science*, *79*(2036), 16–17.

Gendron, M., Crivelli, C., & Barrett, L. F. (2018). Universality reconsidered: Diversity in making meaning of facial expressions. *Current Directions in Psychological Science*, *27*(4), 211–219.

Gendron, M., Hoemann, K., Crittenden, A. N., Mangola, S. M., Ruark, G. A., & Barrett, L. F. (2020). Emotion Perception in Hadza Hunter-Gatherers. *Scientific Reports*, *10*(1), 1–17. https://doi.org/10.1038/s41598-020-60257-2

Goldstone, R. L., & Wilensky, U. (2008). Promoting transfer by grounding complex systems principles. *The Journal of the Learning Sciences*, *17*(4), 465–516.

Gregory, J. (2018). *Game engine architecture*. crc Press.

Guarino, N., & Welty, C. (2002). Evaluating ontological decisions with OntoClean. *Communications of the ACM*, *45*(2), 61–65.

Guo, Y., & Wilensky, U. (2014). Beesmart: A microworld for swarming behavior and for learning complex systems concepts. *Proceedings of the Constructionism 2014 Conference. Vienna, Austria*.

Guo, Y., & Wilensky, U. (2018). Mind the Gap: Teaching High School Students about Wealth Inequality through Agent-Based Participatory Simulations. *Proceedings of the Constructionism 2014 Conference. Vienna, Austria*.

Gutiérrez, K. D., & Jurow, A. S. (2016). Social design experiments: Toward equity by design. *Journal of the Learning Sciences*, *25*(4), 565–598.

Hammer, D. (1996). Misconceptions or p-prims: How may alternative perspectives of cognitive structure influence instructional perceptions and intentions. *The Journal of the Learning Sciences*, *5*(2), 97–127.

Healey, J. (2014). Physiological sensing of emotion. *The Oxford Handbook of Affective Computing*, 204.

Healey, J. (2011). Gsr sock: A new e-textile sensor prototype. *2011 15th Annual International Symposium on Wearable Computers*, 113–114.

Heisenberg, W. (1942). *Philosophie: Le Manuscript de 1942* (Edition Seuil: Heisenberg, Germany, 1998.).

Hmelo-Silver, C. E., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive Science*, *28*(1), 127–138.

Holbert, N., & Wilensky, U. (2019). Designing educational video games to be objects-to-think-with. *Journal of the Learning Sciences*, *28*(1), 32–72.

Holland, D., & Lachicotte, W. (1998). *W., Skinner, D., & Cain, C.(1998). Identity and agency in cultural worlds*. Cambridge, MA: Harvard University Press.

Hölldobler, B., & Wilson, E. O. (2009). *The Superorganism: The Beauty, Elegance, and Strangeness of Insect Societies*. W. W. Norton & Company.

Horn, M. S., Brady, C., Hjorth, A., Wagh, A., & Wilensky, U. (2014). Frog pond: A codefirst learning environment on evolution and natural selection. *Proceedings of the 2014 Conference on Interaction Design and Children*, 357–360.

Horn, M. S., Phillips, B. C., Evans, E. M., Block, F., Diamond, J., & Shen, C. (2016). Visualizing biological data in museums: Visitor learning with an interactive tree of life exhibit. *Journal of Research in Science Teaching*, *53*(6), 895–918.

Hornecker, E., Marshall, P., Dalton, N. S., & Rogers, Y. (2008). Collaboration and interference: Awareness with mice or touch input. *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, 167–176.

Humphrey, T., & Gutwill, J. (2005). *Fostering Active Prolonged Engagement. The Art of*

Creating APE Exhibits. Exploratorium. Left Coast Press, Walnut Creek, CA.

Hölldobler, B., & Wilson, E. O. (1990). *The Ants* (1st edition). Belknap Press: An Imprint of Harvard University Press.

Isen, A. M., Daubman, K. A., & Nowicki, G. P. (1987). Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology*, *52*(6), 1122.

Jackson, K. M., & Trochim, W. M. (2002). Concept mapping as an alternative approach for the analysis of open-ended survey responses. *Organizational Research Methods*, *5*(4), 307–336.

Jacobson, M. J. (2001). Problem solving, cognition, and complex systems: Differences between experts and novices. *Complexity*, *6*(3), 41–49.

Jacobson, M. J., Kapur, M., So, H.-J., & Lee, J. (2011). The ontologies of complexity and learning about complex systems. *Instructional Science*, *39*(5), 763–783.

Jacobson, M. J., & Wilensky, U. (2006). Complex systems in education: Scientific and educational importance and implications for the learning sciences. *The Journal of the Learning Sciences*, *15*(1), 11–34.

Jenkins, A., Jupiter, S. D., Capon, A., Horwitz, P., & Negin, J. (2020). Nested ecology and emergence in pandemics. *The Lancet. Planetary Health*, *4*(8), e302–e303. https://doi.org/10.1016/S2542-5196(20)30165-0

Jonassen, D. H. (1991). Evaluating constructivistic learning. *Educational Technology*, *31*(9), 28–33.

Jung, C. G., Montague, D. E., & Association, S. in W. (1969). Routledge and K. *Paul,* *London*.

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension.* Allyn & Bacon.

Kafai, Y. B. (2006). Playing and making games for learning: Instructionist and constructionist perspectives for game studies. *Games and Culture*, *1*(1), 36–40.

Kafai, Y. B., & Resnick, M. (1996). In *Constructionism in Practice: Designing, Thinking, and Learning in A Digital World* (1 edition). Routledge.

Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A Flaw in Human Judgment*.

Kanade, T. (1974). *Picture processing system by computer complex and recognition of human faces*.

Kanade, T. (1977). *Computer recognition of human faces* (Vol. 47). Birkhäuser Basel.

Kelly, A. (2004). Design research in education: Yes, but is it methodological? *The Journal of the Learning Sciences*, *13*(1), 115–128.

Kinchin, I. M. (2003). Effective teacher↔ student dialogue: A model from biological education. *Journal of Biological Education*, *37*(3), 110–113.

Kinchin, I. M. (2014). Concept mapping as a learning tool in higher education: A critical analysis of recent reviews. *The Journal of Continuing Higher Education*, *62*(1), 39–49.

Klopfer, E., Osterweil, S., Groff, J., & Haas, J. (2009). Using the technology of today in the classroom today: The instructional power of digital games, social networking, simulations and how teachers can leverage them. *The Education Arcade*, *1*, 20.

Koehrsen, W. (2018, August 31). *An Implementation and Explanation of the Random Forest in Python*. Medium. https://towardsdatascience.com/an-implementation-and-explanation-of-the-random-forest-in-python-77bf308a9b76

Kuhn, T. (1962). The structure of scientific revolutions (Chicago and London, 1962). *While in the Physical Sciences There Are Few Paradigms at Any One Time, and Competing Paradigms Cannot Co-Exist for Long, in Art Criticism There Continues to Be a Wide Number of Competing Paradigms*.

Lambert-Mogiliansky, A., & Busemeyer, J. (2012). Quantum type indeterminacy in dynamic decision-making: Self-control through identity management. *Games*, *3*(2), 97–118.

Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist*, *50*(5), 372.

Langton, C. G. (1997). *Artificial life: An overview*. Mit Press.

Larson, K., & Picard, R. (2005). The aesthetics of reading. *Appears in Human-Computer Interaction Consortium Conference, Snow Mountain Ranch, Fraser, Colorado*.

Laughlin, R. B. (2005). *A Different Universe: Reinventing Physics from the Bottom Down* (1st edition). Basic Books.

League, E., Kabat, P., Egerton, P., Baddour, O., Paterson, L., Nullis, C., Castonguay, S., & Walsh, M. (2019). *United in Science: High-level Synthesis Report of Latest Climate Science Information convened by the Science Advisory Group of the UN Climate Action Summit 2019*.

Lee, V. R., Krakowski, M., Sherin, B., Bang, M., & Dam, G. (2006). *Methodological challenges for identifying and coding diverse knowledge elements in interview data*.

Leinhardt, G., & Crowley, K. (1998). Museum learning as conversational elaboration: A proposal to capture, code, and analyze talk in museums. *Report Available at Http://Mlc. Lrdc. Pitt. Edu/Mlc*.

Leinhardt, G., Crowley, K., & Knutson, K. (2003). *Learning conversations in museums*. Taylor & Francis.

Lenhart, A., Kahne, J., Middaugh, E., Macgill, A. R., Evans, C., & Vitak, J. (2008). Teens, Video Games, and Civics: Teens' Gaming Experiences Are Diverse and Include Significant Social Interaction and Civic Engagement. *Pew Internet & American Life Project*.

Levenson, R. W. (1992). Autonomic Nervous System Differences among Emotions. *Psychological Science*, *3*(1), 23–27. https://doi.org/10.1111/j.1467-9280.1992.tb00251.x

Levy, S. T., & Wilensky, U. (2009). Students' learning with the connected chemistry (CC1) curriculum: Navigating the complexities of the particulate world. *Journal of Science Education and Technology*, *18*(3), 243–254.

Linton, M., Norman, D. A., & Rumelhart, D. E. (1975). *Explorations in cognition*.

Logo, F. (2015). *Logo Foundation*. https://el.media.mit.edu/logo-foundation/what_is_logo/history.html

Louw, M., & Crowley, K. (2013). New ways of looking and learning in natural history museums: The use of gigapixel imaging to bring science and publics together. *Curator: The Museum Journal*, *56*(1), 87–104.

Lovato, S. B., & Piper, A. M. (2019). Young Children and Voice Search: What We Know From Human-Computer Interaction Research. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00008

Margolis, E., & Laurence, S. (1999). *Concepts and cognitive science. Concepts: Core readings*. Cambridge: The MIT Press.

Marston, W. M. (1938). *The lie detector test.*

Martin, K., Horn, M., & Wilensky, U. (2019). Prevalence of Direct and Emergent Schema and Change after Play. *Informatics in Education*, *18*(1).

Martin, K., Horn, M., & Wilensky, U. (2020). Constructivist Dialogue Mapping Analysis of Ant Adaptation. *Informatics in Education*, *19*(1), 77–112. https://doi.org/10.15388/infedu.2020.05

Martin, K., Pierson, A., & Sengupta, P. (2018). Ant Food Grab: Integrating Block Based Programming and Biology in 8th Grade Science. *Proceedings of the Narst 2018 Conference. Atlanta, United States. Http://Www. Vimapk12. Net*.

Martin, K., Wang, E. Q., Bain, C., & Worsley, M. (2019). Computationally Augmented Ethnography: Emotion Tracking and Learning in Museum Games. *International Conference on Quantitative Ethnography*, 141–153.

Martin, K., & Wilensky, U. (2019). *NetLogo Models Library: Ant Adaptation*.

https://ccl.northwestern.edu/netlogo/models/AntAdaptation

Martin, P., Bateson, P. P. G., & Bateson, P. (2007). *How good are your measures. Measuring Behaviour: An Introductory Guide, 3rd Edition*. Cambridge University Press.

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, *21*(9), 1281–1289.

Mavadati, S. M., Mahoor, M. H., Bartlett, K., Trinh, P., & Cohn, J. F. (2013). Disfa: A spontaneous facial action intensity database. *IEEE Transactions on Affective Computing*, *4*(2), 151–160.

McCormack, & Wright. (1991). *Sim Ant*.

McDuff, D., El Kaliouby, R., Kodra, E., & Picard, R. (2013). Measuring voter's candidate preference based on affective responses to election debates. *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 369–374.

McGaugh, J. L. (2003). *Memory and emotion: The making of lasting memories*. Columbia University Press.

McGaugh, J. L. (2004). The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annu. Rev. Neurosci.*, *27*, 1–28.

McIntosh, D. (1969). *The Foundations of Human Society*. Chicago: The University of Chicago Press.

McQuiggan, S. W., Robison, J. L., & Lester, J. C. (2010). Affective transitions in narrative-centered learning environments. *Journal of Educational Technology & Society*, *13*(1), 40–53.

Mead, G. (1932a). *The Philosophy of the Present*. Chicago: University of Chicago Press.

Mead, G. (1932b). The Physical Thing. In *The Philosphy of ther Present* (pp. 119–139). University of Chicago Press.

Mead, G. (1934). *Mind, Self and Society*. University of Chicago Press.

Medin, D. L., & Rips, L. J. (2005). *Concepts and Categories: Memory, Meaning, and Metaphysics.* Cambridge University Press.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook*.

Minsky, M. (1988). *Society of mind*. Simon and Schuster.

Minsky, M. (2007). *The emotion machine: Commonsense thinking, artificial intelligence, and the future of the human mind*. Simon and Schuster.

Montuori, A. (2004). Edgar Morin: A partial introduction. *World Futures*, *60*(5–6), 349–355.

Morin, E. (1992). *Method: Towards a study of humankind*. Lang.

Morin, E. (2001). *Seven complex lessons in education for the future*. UNESCO.

Morin, E. (2008). *On Complexity*. Hampton Press.

Murphy, G. (2004). *The big book of concepts*. MIT press.

Nemirovsky, R., & Tierney, C. (2001). Children creating ways to represent changing situations: On the development of homogeneous spaces. *Educational Studies in Mathematics*, *45*(1–3), 67–102.

Novak, J. D. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching*, *27*(10), 937–949.

Novak, J. D., Bob Gowin, D., & Johansen, G. T. (1983). The use of concept mapping and knowledge vee mapping with junior high school science students. *Science Education*, *67*(5), 625–645.

Ochoa, X., & Worsley, M. (2016). Editorial: Augmenting Learning Analytics with Multimodal Sensory Data. *Journal of Learning Analytics*, *3*(2), 213–219. https://doi.org/10.18608/jla.2016.32.10

Ocumpaugh, J., Andres, J. M., Baker, R., DeFalco, J., Paquette, L., Rowe, J., Mott, B., Lester, J., Georgoulas, V., & Brawner, K. (2017). Affect dynamics in military trainees using vmedic: From engaged concentration to boredom to confusion. *International Conference on Artificial Intelligence in Education*, 238–249.

Okada, A. (2008). Scaffolding school pupils' scientific argumentation with evidence-based dialogue maps. In *Knowledge Cartography* (pp. 131–162). Springer.

Okada, A., & Buckingham Shum, S. (2008). Evidence-based dialogue maps as a research tool to investigate the quality of school pupils' scientific argumentation. *International Journal of Research & Method in Education*, *31*(3), 291–315.

Olson, I. C., & Horn, M. S. (2011). Modeling on the table: Agent-based modeling in elementary school with NetTango. *Proceedings of the 10th International*

Conference on Interaction Design and Children, 189–192.

Oviatt, S., Cohen, A., & Weibel, N. (2013). Multimodal learning analytics: Description of math data corpus for ICMI grand challenge workshop. *Proceedings of the 15th ACM on International Conference on Multimodal Interaction*, 563–568.

Pagels, H. R. (1989). *The Dreams of Reason: The Computer and the Rise of the Sciences of Complexity*. Bantam.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books, Inc.

Papert, S. (1986). *Constructionism: A new opportunity for elementary science education*. Massachusetts Institute of Technology, Media Laboratory, Epistemology and ….

Papert, S. (1993). *The children's machine: Rethinking school in the age of the computer.* ERIC.

Papert, S. (1996). A Word for Learning. In Y. B. Kafai & M. Resnick (Eds.), *Constructionism in practice: Designing, thinking, and learning in a digital world* (p. 12). Mahwah, N.J. : Lawrence Erlbaum Associates. https://trove.nla.gov.au/version/46590127

Papert, S., & Harel, I. (1991a). Situating constructionism. *Constructionism*, *36*(2), 1–11.

Papert, S., & Harel, I. (1991b). Situating constructionism. *Constructionism*, *36*(2), 1–11.

Parke, F., & Waters, K. (1996). Computer Facial Animation. *AK Peters*.

Perkins, D. N., & Grotzer, T. A. (2000a). *Models and Moves: Focusing on Dimensions of Causal Complexity To Achieve Deeper Scientific Understanding.*

Perkins, D. N., & Grotzer, T. A. (2000b). *Models and Moves: Focusing on Dimensions of Causal Complexity To Achieve Deeper Scientific Understanding.*

Phillips, D. C., & Dolle, J. R. (2006). From Plato to Brown and beyond: Theory, practice, and the promise of design experiments. *Instructional Psychology: Past, Present and Future Trends. Sixteen Essays in Honour of Erik De Corte*, 277–292.

Piaget, J. (1926). *The language and thought of the child* ((Preface by E. Claparéde; translated by Marjorie and Ruth Gabain)). Routledge and Kegan Paul.

Piaget, J. (1929). *The child's conception of the world* ((translated by Joan and Andrew Tomlinson).). Kegan Paul, Trench, Taubner, & Company.

Piaget, J. (1933). *The moral judgement of the child*. Simon and Schuster.

Piaget, J. (1952a). *Conservation of Continuous Quantities*.

Piaget, J. (1952b). *The origins of intelligence in children*. International Universities Press, Inc.

Piaget, J. (1983). *Piaget's Theory*.

Picard, R. W. (2000). *Affective computing*. MIT press.

Picard, R. W. (2014). The promise of affective computing. *The Oxford Handbook of Affective Computing*, 11–20.

Picard, R. W., & Healey, J. (1997). Affective wearables. *Personal Technologies*, *1*(4), 231–240.

Pollner, M., & Stein, J. (1996). Narrative mapping of social worlds: The voice of experience in Alcoholics Anonymous. *Symbolic Interaction*, *19*(3), 203–223.

Pontual Falcão, T., & Price, S. (2011). Interfering and resolving: How tabletop interaction facilitates co-construction of argumentative knowledge. *International Journal of Computer-Supported Collaborative Learning*, *6*(4), 539–559. https://doi.org/10.1007/s11412-010-9101-9

Posner, G. J., & Gertzog, W. A. (1982). The clinical interview and the measurement of conceptual change. *Science Education*, *66*(2), 195–209.

Pratt, S. C. (2005). Quorum sensing by encounter rates in the ant Temnothorax albipennis. *Behavioral Ecology*, *16*(2), 488–496.

Pratt, S. C., Sumpter, D. J. T., Mallon, E. B., & Franks, N. R. (2005). An agent-based model of collective nest choice by the ant Temnothorax albipennis. *Animal Behaviour*, *70*(5), 1023–1036. https://doi.org/10.1016/j.anbehav.2005.01.022

Renkl, A. (2009). Why constructivists should not talk about constructivist learning environments: A commentary on Loyens and Gijbels (2008). *Instructional Science*, *37*(5), 495–498. https://doi.org/10.1007/s11251-009-9098-5

Resnick, M. (1994). *Turtles, Termites, and Traffic Jams, Explorations in Massively Parallel Worlds*. The MIT Press, Cambridge (MA).

Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., & Silverman, B. (2009). Scratch: Programming for all. *Communications of the ACM*, *52*(11), 60–67.

Resnick, M., & Wilensky, U. (1992). *Starlogo.*

Resnick, M., & Wilensky, U. (1993). Beyond the deterministic, centralized mindsets:

New thinking for new sciences. *Annual Meeting of the American Educational Research Association, Atlanta, GA*.

Resnick, M., & Wilensky, U. (1998). Diving into complexity: Developing probabilistic decentralized thinking through role-playing activities. *The Journal of the Learning Sciences*, *7*(2), 153–172.

Rideout, V. (2017). The Common Sense census: Media use by kids age zero to eight. *San Francisco, CA: Common Sense Media*, 263–283.

Rikers, R. M. J. P., van Gog, T., & Paas, F. (2008). The effects of constructivist learning environments: A commentary. *Instructional Science*, *36*(5), 463–467. https://doi.org/10.1007/s11251-008-9067-4

Robinson, E. J., Ratnieks, F. L., & Holcombe, M. (2008). An agent-based model to investigate the roles of attractive and repellent pheromones in ant decision making during foraging. *Journal of Theoretical Biology*, *255*(2), 250–258.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context.* Oxford university press.

Rogoff, B. (2003). *The cultural nature of human development*. Oxford university press.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*(4), 296–320.

Sandoval, W. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences*, *23*(1), 18–36.

Sandoval, W. A. (2004). Developing learning theory by refining conjectures embodied in educational designs. *Educational Psychologist*, *39*(4), 213–223.

Schachter, S. (1964). The interaction of cognitive and physiological determinants of emotional state. In *Advances in experimental social psychology* (Vol. 1, pp. 49–80). Elsevier.

Schlosberg, H. (1952). The description of facial expressions in terms of two dimensions. *Journal of Experimental Psychology*, *44*(4), 229.

Schlosberg, H. (1954). Three dimensions of emotion. *Psychological Review*, *61*(2), 81.

Schneider, B., & Blikstein, P. (2015). *Unraveling Students' Interaction Around a Tangible Interface Using Multimodal Learning Analytics*. *7*(3), 29.

Segers, M. S. (1996). Assessment in a problem-based economics curriculum. In *Alternatives in assessment of achievements, learning processes and prior knowledge* (pp. 201–224). Springer.

Sengupta, P., Dickes, A., Farris, A. V., Karan, A., Martin, D., & Wright, M. (2015). Programming in K-12 science classrooms. *Communications of the ACM*, *58*(11), 33–35. https://doi.org/10.1145/2822517

Sengupta, P., & Wilensky, U. (2009). Learning electricity with NIELS: Thinking with electrons and thinking in levels. *International Journal of Computers for Mathematical Learning*, *14*(1), 21–50.

Sherin, B. L., Krakowski, M., & Lee, V. R. (2012a). Some assembly required: How scientific explanations are constructed during clinical interviews. *Journal of Research in Science Teaching*, *49*(2), 166–198. https://doi.org/10.1002/tea.20455

Sherin, B. L., Krakowski, M., & Lee, V. R. (2012b). Some assembly required: How scientific explanations are constructed during clinical interviews. *Journal of Research in Science Teaching*, *49*(2), 166–198.

Shettel, H. H. (1976). *An evaluation of visitor response to" Man in his Environment."*. American Institutes for Research.

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: Effects of motivation, interest, and academic engagement. *The Journal of Educational Research*, *95*(6), 323–332.

Slotta, J. D., & Chi, M. T. H. (2006). The impact of ontology training on conceptual change: Helping students understand the challenging topics in science. *Cognition and Instruction*, *24*(2), 261–289.

Snibbe, S. S., & Raffle, H. S. (2009). Social immersive media: Pursuing best practices for multi-user interactive camera/projector exhibits. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1447–1456.

Strohecker, C. (1995a). Embedded microworlds for a multiuser environment. *Archives & Museum Informatics*, 57–66.

Strohecker, C. (1995b). A Model for Museum Outreach Based on Shared Interactive Spaces. *ICHIM, Multimedia Computing and Museums*, 57–66.

Sutherland, I. E. (1965). The ultimate display. *Multimedia: From Wagner to Virtual Reality*, *1*.

Taylor, M. C. (2003). *The moment of complexity: Emerging network culture*. University of Chicago Press.

Taylor, S., Jaques, N., Chen, W., Fedor, S., Sano, A., & Picard, R. (2015). Automatic identification of artifacts in electrodermal activity data. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1934–1937.

Thaler, R. H., & Ganser, L. J. (2015). *Misbehaving: The making of behavioral economics*. WW Norton New York.

Tierney, J., & Baumeister, R. F. (2019). *The Power of Bad: How the Negativity Effect Rules Us and How We Can Rule It*. Penguin Press.

Toulmin, S. E. (1972). *The collective use and evolution of concepts*. Princeton University Press.

Trochim, W. M. (1989). An introduction to concept mapping for planning and evaluation. *Evaluation and Program Planning*, *12*(1), 1–16.

Tronick, E. Z. (1989). Emotions and emotional communication in infants. *American Psychologist*, *44*(2), 112.

Turkle, S., & Papert, S. (1992). *Epistemological Pluralism: Styles and Voices Within the Computer Culture*. 22.

Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International Journal of Computer Vision*, *57*(2), 137–154.

Vlahos, J. (2015, September 16). Barbie Wants to Get to Know Your Child. *The New York Times*. https://www.nytimes.com/2015/09/20/magazine/barbie-wants-to-getto-know-your-child.html

Vom Lehn, D., Heath, C., & Hindmarsh, J. (2001). Exhibiting interaction: Conduct and collaboration in museums and galleries. *Symbolic Interaction*, *24*(2), 189–216.

Von Neumann, J. (1951). The general and logical theory of automata. *1951*, 1–41.

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, *24*(4), 535–585.

Vossen, P., Agerri, R., Aldabe, I., Cybulska, A., van Erp, M., Fokkens, A., Laparra, E., Minard, A.-L., Aprosio, A. P., & Rigau, G. (2016). Newsreader: Using knowledge resources in a cross-lingual reading machine to generate more knowledge from massive streams of news. *Knowledge-Based Systems*, *110*, 60–85.

Vygotsky, L. S. (1966). Play and its role in the mental development of children. *Soviet Psychology*, *1*(2), 6.

Wagh, A., Cook-Whitt, K., & Wilensky, U. (2017). Bridging inquiry-based science and constructionism: Exploring the alignment between students tinkering with code of computational models and goals of inquiry. *Journal of Research in Science Teaching*, *54*(5), 615–641.

Wagh, A., & Wilensky, U. (2018). EvoBuild: A Quickstart Toolkit for Programming Agent-Based Models of Evolutionary Processes. *Journal of Science Education and Technology*, *27*(2), 131–146.

Weintrop, D., Holbert, N., Wilensky, U., & Horn, M. (2012). Redefining constructionist video games: Marrying constructionism and video game design. *Proceedings of* *the Constructionism 2012 Conference. Athens, Greece*.

Wertsch, J. (1991). *Voices of the mind: A sociocultural approach of mediated action Harvester*. Cambridge, MA: Harvard University Press.

Wertsch, J. V. (1997). *Narrative tools of history and identity*. Sage Publications Sage CA: Thousand Oaks, CA.

White, H. C. (1995). Where do languages come from?-Switching talk. *Unpublished Manuscript, Columbia University, New York*.

Whitehill, J., Serpell, Z., Foster, A., Lin, Y.-C., Pearson, B., Bartlett, M., & Movellan, J. (2011). Towards an optimal affect-sensitive instructional system of cognitive skills. *CVPR 2011 WORKSHOPS*, 20–25.

Wiener, N. (1948). *Cybernetics or control and communication in the animal and the machine*. Technology Press.

Wilensky, U. (1997a). NetLogo Ant Lines model. *Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. Http://Ccl. Northwestern. Edu/Netlogo/Models/AntLines*.

Wilensky, U. (1997b). Netlogo ants model. *Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL*.

Wilensky, U. (1997c). What is normal anyway? Therapy for epistemological anxiety. *Educational Studies in Mathematics*, *33*(2), 171–202.

Wilensky, U. (1999a). GasLab—An extensible modeling toolkit for connecting micro- and macro-properties of gases. In *Modeling and simulation in science and* *mathematics education* (pp. 151–178). Springer.

Wilensky, U. (1999b). *Netlogo*. Center for Connected Learning and Computer-Based Modeling. http://ccl.northwestern.edu/netlogo/.

Wilensky, U. (2020, April). New Developments in Restructuration Theory and Understanding Complex Systems Through Agent-Based Restructurations. *Restructuring Concepts and Tools Through a Complexity Perspective*. AERA, San Francisco. http://tinyurl.com/rvb86xr

Wilensky, U., & Jacobson, M. J. (2014). Complex Systems and the Learning Sciences. In *The Cambridge Handbook of the Learning Sciences* (pp. 319–338). Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.020

Wilensky, U., & Papert, S. (2010). Restructurations: Reformulations of knowledge disciplines through new representational forms. *Constructionism*.

Wilensky, U., & Rand, W. (2015). *Introduction to agent-based modeling: Modeling natural, social and engineered complex systems*. The MIT Press, Cambridge MA.

Wilensky, U., & Reisman, K. (2006). Thinking like a wolf, a sheep, or a firefly: Learning biology through constructing and testing computational theories—an embodied modeling approach. *Cognition and Instruction*, *24*(2), 171–209.

Wilensky, U., & Resnick, M. (1995). New thinking for new sciences: Constructionist approaches for exploring complexity. *Annual Meeting of the American Educational Research Association, San Francisco, CA*.

Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and Technology*, *8*(1), 3–19.

William James. (1893). *Principles of Pyschology*. Harvard University Press.

Wilson, E. O. (1984). The superorganism. In *Biophilia* (pp. 23–38). Harvard University Press.

Wolfram, S. (2002). *A new kind of science* (Vol. 5). Wolfram media Champaign, IL.

Yoon, S. A. (2018). Complex systems and the Learning Sciences: Implications for learning, theory, and methodologies. *International Handbook of the Learning Sciences*, 157–166.

Yoon, S. A., Elinich, K., Wang, J., Steinmeier, C., & Tucker, S. (2012). Using augmented reality and knowledge-building scaffolds to improve learning in a science museum. *International Journal of Computer-Supported Collaborative Learning*, *7*(4), 519–541.

Zajonc, R. B. (1994). Evidence for nonconscious emotions. *The Nature of Emotion: Fundamental Questions*, 293–297.

Zhang, X., Yin, L., Horowitz, A., Reale, M., Canavan, S., Liu, P., & Cohn, J. F. (2013). A 3d spontaneous dynamic facial expression database. *IEEE International Conference on Automatic Face and Gesture Recognition.*

6 Vita

Web Portfolio: www.Kitcmartin.com

Kit Martin

Education

2015-2021 **Northwestern University**, Ph.D. in Learning Sciences.

committee: Michael Horn, Marcelo Worsley, Uri Wilensky (Chair)

dissertation *Ant Adaptation: Measuring model based learning with multiple data streams*

2010–2012 **University of Minnesota**, M.D.P., Poverty Reduction.
Quantitative and Qualitative Methods in International Development

thesis *Sustainable Development in Sierra Leone: OneVillage Partners' Strategic Plan*

2004–2008 **Bard College**, *B.A., History*.

thesis *Silk Spirals and the Law, on Women in Islamic Governance in 16th Century Ottoman Turkey*

Research Experience

2021 **Assistant Professor**, *Penn State*, State College, Pennsylvania.

Design and implement education interventions in entomology, computer science, physics, public health, and agent-based models.

Developed block based programming environments that lower the difficulty threshold for students to develop scientific theory.

Publish and disseminate the Ant Game that teaches about integrated ecology for understanding complex systems `http://www.gettheantgame.com`.

Develop the Bee Game to teach about integrated ecology for pollinator health.

Developed AI-powered affect tracking technology for improved team coordination in the science classroom.

2015–2021 **PhD research assistant**, *Northwestern University*, Evanston, Illinois.

Design and implement education interventions in biology, multilingual education, historical literacy, computer science, and computational thinking.

Developed block based programming environments that lower the difficulty threshold for students to develop scientific theory.

Conducted intensive educational experiment at the Field Museum to test if short game play could lead to new insight into complex systems thinking. Found that it did for a third of players.

Published 24 peer reviewed papers and pieces of education software.

2020 **AR Transition Lead**, *Northwestern University*, Evanston, Illinois.

Design and implement augmented and virtual reality interventions in Northwestern classrooms.

Develop literature review of AR in education.

Deploy and evaluate VR/AR prototypes.