ABSTRACT OF THE DISSERTATION

Let’s Build a Model: How Students Incorporate Evidence and Mechanisms in Models During Small Group Discussion

by CHI-CHUN HUNG

Dissertation Director:
Dr. Clark A. Chinn, Ph.D.

Building a model to explain a scientific phenomenon is a central activity in an inquiry environment. This study investigated three aspects of model-based inquiry and developed coding frameworks to conduct quantitative and qualitative analyses. The first set of research questions examined learners’ capability of coordinating models with evidence. The second set of research questions scrutinized learners’ competence in incorporating mechanisms into models. The third set of research questions explored the dynamics of small group discussions and types of prompts that mediated varied modeling performances. A total of 80 pre-instruction models about cellular respiration were collected from 167 seventh graders from two schools to inform the first and the second sets of research questions. Discussion logs of four small groups and the conversations between groups and two teachers were analyzed to address the third set of research questions.

The learners demonstrated proficiency in making their models consistent with an array of relevant evidence and eliminating irrelevant evidence, although a variety of prior
knowledge was also present in a large portion of models. Mechanisms were incorporated into a majority of the models, and four types of them were identified. The design of the coding scheme made it possible to discover that most learners were able to indicate sequence in the mechanism, and it was less challenging for them to present entities in models than to incorporate entities into a mechanistic system. Higher performance was found to relate to factors identified by the third set of analytical instruments. Discourse processes such as contemplating multiple evidence all at once and stipulating a mechanistic explanation were more helpful than having an interactive discussion without their application. Moreover, higher performance might benefit from specific prompts than generic prompts offered by the teachers during their conversations with small groups.

The triangulated results drawn from three aspects bring a comprehensive understanding of learners’ modeling accomplishment and how it can be influenced by certain discourse processes. The findings add to a growing body of knowledge about model-based inquiry and provide implications for small group discussions and instructional interventions in science education.
# Table of Contents

Abstract ................................................................................................................................. ii

Chapter 1 Introduction ............................................................................................................. 1

  Research questions ............................................................................................................. 5

Chapter 2 Literature Review ................................................................................................. 7

  Coordination of models and evidence ............................................................................... 9

  Incorporations of mechanisms in models ......................................................................... 14

  The group discussion processes that mediate modeling performance ...................... 19

    Group discussion ........................................................................................................... 20

    Teacher prompts .......................................................................................................... 24

Chapter 3 Methodology ......................................................................................................... 27

  The general curricular context .......................................................................................... 27

  Participants ....................................................................................................................... 29

  The cellular respiration module ....................................................................................... 30

  The materials ................................................................................................................... 31

  Data collection ................................................................................................................ 35

  Study design and procedures ......................................................................................... 36

  Data analysis .................................................................................................................... 41

    Coordination of models and evidence ......................................................................... 41

    Model and evidence consistency ................................................................................. 42

    The quality of evidence interpretation ....................................................................... 41

    Additional concepts absent from the evidence ......................................................... 49

  Incorporation of mechanisms in models ....................................................................... 49
Prompts were focused on the alignment of models and evidence .................. 126
Prompts were focused on revisiting overlooked evidence ........................... 127
Prompts were focused on examining the patterns in charts within evidence 128
Summary ........................................................................................................ 130
Summary of group discussions and teacher prompts ................................... 131
Group discussions .......................................................................................... 131
Teacher prompts ............................................................................................... 133
Chapter 6 Discussion and Conclusion ........................................................... 136
Discussion of the results .................................................................................. 136
The consistency of models and evidence ....................................................... 136
The interpretation of evidence ........................................................................ 137
The characteristics of mechanisms ................................................................. 139
Four types of models ....................................................................................... 142
Small group processes .................................................................................... 146
Teacher prompts ............................................................................................... 150
Limitations ........................................................................................................ 152
Conclusion, implication, and future directions ............................................... 154
References ....................................................................................................... 159
Appendix A: All Pieces of Evidence .............................................................. 167
Appendix B: Three Example Models in Version 2 ......................................... 173
List of Tables

Table 1. Materials in Sequence ................................................................. 32
Table 2. Activity Procedures ................................................................. 36
Table 3. The Measurement of High-Carb vs. Low-Carb ......................... 43
Table 4. The Measurement of Oxygen .................................................... 44
Table 5. The Measurement of Carbon Dioxide ....................................... 44
Table 6. The Measurement of Cellular Organelle .................................... 45
Table 7. The Measurement of Body Temperature ..................................... 46
Table 8. The Measurement of Muscle Cell Inputs .................................... 47
Table 9. The Measurement of Muscle Cell Contraction ............................ 48
Table 10. The Measurement of Physical Fitness ...................................... 48
Table 11. The Measurement of the Energy Generation Mechanisms ........... 51
Table 12. Four Small Groups ................................................................. 56
Table 13. Discourse Operation ............................................................... 58
Table 14. Dialogue Patterns ................................................................. 60
Table 15. The Percentage of Models Consistent with Different Numbers of Evidence 67
Table 16. The Percentage of Model-Evidence Consistency for Relevant Evidence ..... 68
Table 17. Performance of Evidence Interpretation Before Instruction .......... 69
Table 18. Additional Concepts Absent From the Evidence ....................... 70
Table 19. Sub-Analyses of the Energy Generation Mechanisms ............... 73
Table 20. Teacher Prompts and Subsequent Student Actions .................... 134
List of Illustrations

Figure 1. A Model as An Example to Demonstrate the Analysis .............................. 55
Figure 2. A Cellular Transformation Model .......................................................... 76
Figure 3. A Digestive Model ................................................................................. 77
Figure 4. A Non-Cellular Burning Model ............................................................. 79
Figure 5. An Exercise Model ................................................................................ 80
Figure 6. Group A1’s Model ................................................................................ 86
Figure 7. Group A2’s Model ................................................................................ 97
Figure 8. Group B1’s Model ................................................................................. 113
Figure 9. Group B2’s Model ................................................................................. 119
Chapter 1

Introduction

This study is a report of how middle school students work in small groups in a model-based inquiry environment to build models that explain a scientific phenomenon. The results were acquired from mixed-method analyses that focused on three major aspects. The first aspect examined how competent students were in coordinating models with evidence. The second aspect focused on how students incorporated mechanisms in the models. The third aspect explored the processes during group discussion that mediated students’ modeling performance.

Constructing explanations or models based on evidence to account for the observable and invisible processes of a phenomenon is the goal of model-based inquiry (Braaten & Windschitl, 2011; Windschitl et al., 2008). The activities in a model-based inquiry classroom go beyond seeking understanding of terminologies and concepts (Braaten & Windschitl, 2011), forming simple cause-effect relationships (Braaten & Windschitl, 2011), or performing defined experiments to summarize and confirm the given facts (Windschitl et al., 2008). A model-based inquiry setting emphasizes the coordination of models and various sets of data, the incorporation of visible procedures and invisible mechanisms into the explanations, and the co-construction of explanations in the presence of peer communication. These are the characteristics that make a classroom an authentic science environment resembling how scientists work in reality (Chinn & Malhotra, 2002). Scientists use models as reflective tools and they use modeling as a cognitive means to navigate through hypotheses and conjectures during inquiry (Cartier & Stewart, 2000). Students build models in a model-based inquiry
classroom and develop increasingly deeper and broader conceptual knowledge as they create and refine their models in a social community where they share ideas and evaluate each other’s explanations (Manz, 2012; Schwarz et al., 2009).

In the literature of model-based inquiry, each of these three aspects have been well investigated. There have been numerous studies reporting on students’ skills in making scientific arguments on the basis of evidence. An abundance of studies has investigated students’ mechanistic reasoning as it is revealed in conversation and in written work; many studies have also explored the reasoning processes and the structure of student talks in the group and in a classroom with their teachers. Reviews of these studies will be discussed in depth in the next chapter. There have not been studies, however, which integrate these three aspects through mixed method research to create a comprehensive understanding. This study intends to fill in this gap in research. I approached this topic by investigating seventh graders’ performance in modeling contextualized in a cellular respiration lesson.

The cellular respiration lesson is one of the units in a one-year long NSF funded project called Promoting Reasoning and Conceptual Change in Science (PRACCIS). PRACCIS provided a model-based inquiry environment to middle school students, mostly the 7th graders, to promote mastery of scientific practices and core ideas in life science. With extensive interactions with teacher participants and several professional development workshops, the PRACCIS research team collaborated with teachers to develop an array of units, including units on cell organelles, cell processes, genetics, evolution, and ecology. The units fostered scientific inquiry practices including evaluation of evidence, construction of models, coordination between models and
evidence, explanation and justification of models, argumentation, and collaborative critique techniques. In some units, students chose one model out of two model choices against evidence after they considered the degree to which the evidence was better consistent with one of the models. In other units, students engaged in experiments and generated their own data and conclusions. In still other units, students created their own models using provided pieces of evidence, as they did in the cellular respiration lesson.

Along with photosynthesis, cellular respiration is the hub when introducing the concept of energy flow among organisms, a crosscutting concept advocated in the Next Generation Science Standards [NGSS] (NGSS Lead States, 2013). Energy flows and carbon cycles are essential elements to be integrated in the notion of food webs and ecology system (Anderson et al., 1990). How an organism transforms substances through cellular respiration to produce energy and sustain life is the prerequisite knowledge for learning how energy transfers among different organisms in the ecosystem.

Researchers studying learning in the life sciences have found that certain topics are challenging for students. Cellular respiration is not easy for students to learn, like many other topics such as photosynthesis, the energy cycle, protein synthesis, enzyme structure and function, osmosis, mitosis and meiosis, and genetics (Carey, 1985; Driver et al., 1994; Lazarowitz & Penso, 1992; Wood-Robinson, 1991). These topics are challenging because, in addition to their complex concepts, they all convey mechanistic processes that are invisible, unperceivable, or often times counter-intuitive to our experiences.

Studies of cellular respiration learning have indicated that life science instruction does not yield much learning gains (Songer & Mintzes, 1994) as students across age
groups persistently entertain alternative conceptions regarding this important science topic. Some robust ideas are prevalent among elementary school children, secondary education students, undergraduates, prospective teachers, and even in-service biology teachers. Children of elementary school age tend to take breathing and respiration as synonyms (Driver et al., 1994). They either do not believe that plants “breathe,” or they believe that plants breathe like humans or animals (Cañal, 1999) in order to refill on oxygen (Driver et al., 1994). In their eyes, plants and animals use air in opposite ways as plants breathe in carbon dioxide and breathe out oxygen, whereas animals inhale oxygen, or simply air, and exhale carbon dioxide (Driver et al., 1994). Students in secondary education may still use an animalistic and anthropomorphic view to account for cell processes (Flores et al., 2003), and some of them still think breathing and respiration are equal notions (Seymour & Longden, 1991). In addition, the confusion between cellular respiration and photosynthesis is pervasive; for example, students may hold a belief that plants do not do cellular respiration or, if they do, they do it only in the dark and in the cells of the leaves only (Haslam & Treagust, 1987; Wood-Robinson, 1991). Flore and his colleagues (2003) stipulated that many alternative ideas carried by students may result from an anthropomorphic view that comes from teacher’s instruction, such as “respiration in cells needs oxygen in the same way as humans do,” and that “the process of nutrition in the cell is similar to the digestive system where food is ground and processed.” College biology instruction appears not effective in changing undergraduate students’ naïve impression about cellular respiration as they still struggle with some concepts about cellular respiration. (Anderson et al., 1990; Brown & Schwartz, 2009; Rybarczyk et al., 2007; Songer & Mintzes, 1994; Wilson et al., 2006). One of the examples is that they
may see cellular respiration as a gaseous exchange process (Anderson et al., 1990). Even the biology teachers would assume that plants only have cellular respiration at night and for animals cellular respiration occurs in the tissues and digestion is the process that provides animals with energy (Sanders, 1993).

The results gleaned from the aforementioned research demonstrated fragments of ideas that students entertained about cellular respiration. Perhaps a model-based inquiry learning experience could have enhanced their conceptual gains and helped them overcome the robust alternative conceptions about cellular respiration. Students who participated in this study navigated through the cellular respiration lesson where they were given the opportunity to mull over various sets of evidence and build a model with peers when they worked together and received prompts from the teacher during small group discussion. How well did they coordinate models with evidence? Did they integrate mechanistic explanations in their models? How did they talk together with their group members? Did the teachers’ scaffolding help? The research questions displayed below address the threefold purpose of this study.

1. How competent are students in coordinating models with evidence?

   1.1 To what extent are students’ models consistent with evidence?

   1.2 How is the quality of evidence interpretation represented in students’ models?

2. How competent are students in incorporating mechanisms in their models?

   2.1 To what extent do students’ models incorporate components of mechanisms?

   2.2 What models do students develop to explain how energy is generated?

3. What processes mediate modeling performance?
3.1 How does the discourse of groups that vary in modeling performance differ?

3.2 What prompts from teachers enhance modeling performance?
Chapter 2

Literature Review

For more than two decades in the research of science education and public understanding of science, researchers have advocated a science instruction to engage students in practices that reflect what scientists do and how the science enterprise works (Driver et al., 1996; Osborne et al., 2003). Engaging learners in inquiry activities is advocated by many contemporary education scholars as a way to develop students’ knowledge, practices, and reasoning in science (Chinn et al., 2013). The NGSS (NGSS Lead States, 2013) suggest that the essential learning goals in science classrooms should include both the understanding of scientific knowledge and the ability to appreciate and engage in scientific inquiry. Another guideline (Duschl et al., 2007) addressing learning and teaching science in grades K-8 emphasizes that by the time students enter high school, they should have developed scientific proficiency by acquiring conceptual knowledge, building and refining models based on evidence, knowing the nature and development of science, and participating in scientific practices and discourse. These perspectives of school science correspond to the most recent reform in NGSS that crosscutting concepts, such as investigating and explaining causal relationships and mechanisms, and inquiry practices, such as interpreting data and constructing models, should be integrated in science learning when students acquire the disciplinary core ideas (NGSS Lead States, 2013).

A model-based inquiry classroom is reflecting the essence of inquiry in science education. In a typical model-based project, students are encouraged to navigate through the processes of inquiry where they activate their prior knowledge and define the
problem, gather data and search patterns in data, develop models to explain data, make predictions, revise models in response of alternative evidence and critiques of others (Stewart et al., 2005). Model-based inquiry instruction stresses on model-evidence consistency, and it also urges students to build models that account for observable patterns of data and the unobservable mechanisms that might result in the performance of a certain phenomenon (Windschitl et al., 2008). The practices of coordinating models with evidence and incorporating mechanisms into a model in a model-based inquiry context are usually carried out by small groups in a classroom. The model is a thought product as well as an instrument to turn abstract concepts into a tangible representation for learners to see and reason about concepts that are not immediately visible in the topic under study (Manz, 2012). The act of making thoughts public may trigger a new round of inquiry among peers as connections between ideas become visible (Windschitl et al., 2008). It also elicits peer communication and invites mutual evaluation from different groups (Schwarz & White, 2005; White & Frederiksen, 1998). Moreover, the visualization of students’ thinking suggests instructional guidance for teachers to detect the moment where scaffolding is necessary to bridge a knowledge gap or to acknowledge students’ conceptual achievements (Larkin, 2012; Lehrer & Schauble, 2000).

In the following sections, I discuss the current understanding of learning in a model-based inquiry context. Special focuses are placed on the topics about the coordination of models and evidence, the incorporation of mechanisms in models, and small group discussion and teacher prompts during modeling.
Coordination of Models and Evidence

A guiding principle of scientific inquiry is to make a model, explanation, or claim subject to evidence (NGSS Lead States, 2013). Evidence may take in forms of texts, tables, graphs, charts, microscopic pictures, and mathematics to display data and represent relationships. Students as young as second grade were able to understand the numbers in tables representing time and its implication of the relationship between time elapse and speed when they studied motion after 10 days of inquiry activity (Hapgood, et al., 2004). Wu and Krajcik (2006) found that 7th graders were competent in creating data tables and graphs in an inquiry class where students had to sample the water from the stream behind school to determine if it was healthy. Those students were also proficient in interpreting data consistency and patterns in representations such as tables, graphs, or pictures. Translating mathematical equation into biological concepts projected no problems for college students in Svoboda and Passmore’s case study (2013) when they tried to consider if there was possible link between autism and the MMR vaccine by plotting outputs over a range of variables and parameters. Although the college students constructed models to explore and refine their ideas, they were not competent in matching their model to existing empirical data. Interpreting evidence seems less challenging than coordinating models with evidence.

School students in general do not spontaneously use evidence to generate models (Kang et al., 2014; Windschitl et al., 2008). Without a specific intervention on the notion of citing data to back up an explanation, students in an inquiry learning environment are not very skillful in including evidence in their claims even when the instruction makes students explicitly choose evidence that they consider relevant (Sandoval, 2003; study 1
in Sandoval & Reiser, 2004). For example, in the early version of BGuILE study in a computer-supported inquiry lesson about natural selection (Sandoval, 2003; Sandoval & Reiser, 2004), citation of data was rarely seen in the high school participants’ cohesively written explanations. Most student groups looked at many sets of data that they considered relevant, which indicates that they knew they needed evidence to support an explanation, but they did not necessarily cite the relevant data they examined. Perhaps students did not appreciate the value of supporting theoretical claims with figures and evidence (Takao & Kelly, 2003), or maybe groups dedicated more thinking to how to interpret data and did not discuss how to support an explanation with data (Sandoval & Reiser, 2004). Certainly telling students that they will be graded based on their use of evidence without an intervention instructing them on the techniques of doing so does not promote coordinating explanations with evidence (see Sandoval & Reiser, 2004).

Students may develop competence in using relevant evidence in their explanations after instructional interventions. In another study that also focused on learning natural selection in high school (Passmore & Stewart, 2002), students achieved higher performance in model-evidence coordination than those in the aforementioned studies. Students were taught nature-of-science concepts, after which they investigated changes in a trait over time through a case. They used a natural selection model to explain changes and supported their written arguments with appropriate pieces of evidence. They recognized patterns in data and made predictions. Moreover, they used data to develop justifications for Darwin’s model and formulate explanations about evolutionary concepts.
Intervention on coordinating models with evidence has provided the affordance for students to be more sensitive about the relationship between models and evidence. A strong awareness of the model-evidence consistency in high school students was documented by Cartier and Stewart (2000) in an inquiry-based genetics lesson. Before the lesson started, students learned the nature and attributes of models. During the lesson, students first noticed that Mendel’s model could explain some but not other sets of genetic traits. Right after they recognized the limitation of Mendel’s model, some students suggested more data be collected to test the model, and others hypothesized alternative explanations for the data they currently had when they encountered inconsistency between model and evidence. They later tested Mendel’s model against different variations to see if this model was consistent with all sets of data.

A higher frequency of citing evidence was reported in explanations by groups of students in a revised version of the BGuILE project that included an intervention about how to cite evidence in a claim (Sandoval & Millwood, 2005; study 2 in Sandoval & Reiser, 2004). In this new version of the software, a list of evidence and an explanation box were juxtaposed on the same page in order to make the link between evidence and claim more explicit. The teacher showed students an epistemic rubric of criteria for good explanations, including searching for patterns in data and citing these patterns in their explanations. Students displayed more awareness of the coordination between data and explanations. During an activity of peer critiquing, the most common comments, about a quarter of them, pointed out a lack of data in explanations (study 2 in Sandoval & Reiser, 2004). Students looked at more data sets, and they cited more data in their explanations than those who participated in the old version (Sandoval & Millwood, 2005).
The effect of intervention was also shown in Kelly and Takao’s (2002) study. They engaged college students in data exploration on a computer before they generated an argument to warrant the plant tectonic theory. More than half of the students’ arguments were supported by references to data, inferences from data, or comparisons among data.

McNeill and colleagues (2006) had middle school students receive instruction that highlighted the skills to make a claim, describe evidence, and explain how a claim was supported by evidence in their Investigating and Questioning Our World Through Science and Technology (IQWST) biology units. After instruction students’ written explanations improved on all of these three components during inquiry. The results revealed that citing evidence is more challenging than making a claim, and justifying the relationship between evidence and claims is more challenging than citing evidence. This finding was confirmed in Berland and Reiser’s (2009) study, which also adopted the IQWST framework in teaching the construction of a scientific explanation. It turned out that all of the explanations written by middle schoolers contained an accurate claim, meaning that they made a correct conclusion based on appropriate interpretation of relevant evidence. To cite evidence in their arguments was not a problem for students, but to explain why and how the evidence supported a claim was difficult for them. Students tend to “view data as something to be explained, but not necessarily a component of an argument” to support their conclusions (Sandoval, 2003, p. 41). The results of these two studies coincided with Sandoval’s argument.

In summary, learners are competent in making sense of visual representations of data from the early years in elementary school in an inquiry-based classroom. Without
intervention, students are capable of recognizing relevant evidence, but they are not adept at coordinating self-constructed models or explanations with evidence to explain a phenomenon. After learning the epistemic relationship between models and evidence students can acknowledge the limit of a model in light of evidence. They are able to make an argument in accordance with evidence after they learn how to cite evidence in an explanation. However, rarely do students explicate why a piece of evidence is important to support a claim even with scaffolding (McNeill et al., 2006).

Most inquiry studies (except the IQWST project) reviewed above explored how competent students were in coordinating models with evidence after providing them with some general information of a focal scientific theory. Students were introduced to the major ideas of the theory before generating explanations about a phenomenon based on raw data. The results could be different if students were set in a condition where they had to integrate a variety of pieces of evidence and establish a theoretical model from scratch. In addition, knowing that students can cite evidence only offers a partial answer to the question regarding how competent students are in coordinating models with evidence. What is left unknown includes: (a) students’ competence in making their self-constructed models consistent with a variety of relevant evidence yet exclude irrelevant evidence, and (b) students’ interpretation of evidence which is embedded in their models. The unknown areas lead to research question 1.1: “To what extent are students’ models consistent with evidence?” and research question 1.2: “How is the quality of evidence interpretation represented in students’ models?”
**Incorporation of Mechanisms in Models**

A scientific model has an intent to explain. In many occasions it is expected to illustrate how a cause brings about an effect in a given context and predict how it may function in a new context, often involving invisible or unobservable processes in a phenomenon. A model that explains incorporates mechanisms to address the “why” questions of a scientific phenomenon and generate the “how” account for relationships in a system of processes. A mechanism includes certain entities, their properties, and activities. It describes how entities are engaged in activities that produce regular changes from the set-up condition through intermediate stages to the termination condition (Machamer et al., 2000). A mechanism also includes the way entities interact with each other in a certain regulation in a particular condition of space, temperature, speed, pressure, spread, rotation, interval, sequence, density, humidity, etc. (Machamer et al., 2000). Changes of an entity, its properties, its interaction with another entity, or the condition in which the interaction takes place will result in a different outcome as a whole in the system (Machamer et al., 2000).

Generating a mechanistic explanation that explicates how and why something happens is different than engaging in causal reasoning that addresses the causes and effects (Berland et al., 2015; Russ et al., 2008). Nevertheless, engaging in “causal reasoning serves as a starting point for the pursuit of underlying mechanistic explanations” in a model-based inquiry classroom (Russ et al., 2009, p. 881). Students are capable of causal reasoning from an early grade year. Russ and colleagues (2008) recorded a reasoning episode during a first-grade class conversation where students demonstrated causal reasoning when they thought about the relationships between the
outcomes and the changes of the property of entities. One child argued that crumpling up a sheet of paper into a ball would only change its shape but not size, and therefore the paper was still light. This observation suggests that young grade school children might understand the causal relationship that the action of crumpling left the property of the entity unchanged and the weight remained the same. High school students in Hammer’s study (1995) had to decide whether or not a small steel ball and a cart which moved at the same speed would arrive at the same location at the same time during an initial stage of classroom discussion about Newton’s law of motion before physics instruction. Their conversation revealed their ability of isolating some factors, a crucial skill for reasoning causal relationship. In another study conducted by Sandoval (2003), high school students in a natural selection lesson were able to generate coherent causal accounts, regardless of conceptual accuracy, to explain why some finches survived and some did not in a specific condition. These examples suggest that students in a wide range of ages have the ability to engage in causal reasoning in an inquiry setting led by teachers without intervention.

The authors of the aforementioned studies argued that the causal reasoning they observed in students was an indication of mechanistic reasoning, yet the results they reported were not the best examples to imply that students were making mechanistic explanations. The evidence in the study of Russ et al. (2008) came from a girl’s argument:

If it’s balled up it’s still not heavy it’s the same size… (She picks up the crumpled paper and uncrumplles it)…It’s still the same size. (Lifts the paper up and down in front of her.) It still feels- (Crumples the paper back up) …It’s still light. (p. 519)
This girl was expressing her reasoning that changes of an entity’s shape do not result in changes of its weight, which was reasoning about a simple causal relationship instead of mechanisms. In Hammer’s (1995) study, students were thinking about several factors (the effects of air or wind, of how high the ball was thrown, and the weight of the ball) in a class discussion episode. He claimed that a student’s assertion that “there is nothing to push it back,” inferring that there must be some external force to cause the ball to slow down, was an indication of mechanism. However, mechanistic reasoning was not manifested in this student’s utterance, which instead was simple causal reasoning.

Sandoval (2003) argued that his participants actively generated mechanisms to explain data, but this argument was made based on the use of terms such as “because,” “caused,” “thus,” or “due to” in students’ writing, which rather marked causal accounts than mechanistic reasoning. Causal reasoning denotes causes and effects without explanation of the how between them, whereas mechanistic reasoning explains the process by which causes bring about effects as a system; thus one can make prediction and transfer the behavior from one case to another (Russ et al., 2009).

A model that incorporates mechanisms “makes claims about what must have happened previously to bring about the current state of things or what will happen next given that certain entities or activities are present now” (Russ et al., 2008, p. 513) to explain a chain of activities taking place in a phenomenon. The sequence is not the focal point to make a model mechanistic; instead, the sequence is embedded in the chain of changes in activities that encompass processes. Based on this definition of mechanism (Russ et al., 2008), incorporating mechanism into models appears challenging for students. Instead of explicating how entities and activities are organized to produce
regular changes from set-up to termination conditions, the superficial components and the chronological sequences are the most pronounced features in students’ models. For instance, Penner et al. (1997) found that when third graders initially tried to simulate a human elbow from a variety of materials, they only focused on the superficial aspects and attempted to make their model looked like an elbow’s shape. In their revised model after a guided discussion led by the teacher, they still predominately used perceptual clues to determine the motion function of an elbow and included at most only one aspect of the motion constraint of elbows. The directions and the range of motion in each direction were crucial factors to illuminate the mechanism of an elbow, but these factors were overlooked by them. In another study also with third graders, Forbes et al. (2015) found that the visible and invisible entities and sequential steps of the water cycle were the most salient attributes in their graphical models. Such phenomenon was also seen in Schwarz et al.’s study (2009) when 5th grader built graphical models explaining how evaporation and condensation occurred. Indeed, entities and sequence are essential parts of a mechanistic explanation to illustrate a system’s components and processes (Forbes et al., 2015). Middle school students consider sequence one of the criteria of a good model (Pluta et al., 2011; Schwarz & White, 2005). However, mechanisms comprise more than entities and sequences in order to explain how a cohesive system works.

In order to make mechanistic explanations students may have to appeal to some epistemic heuristics to help them structure the reasoning so that they can construct models that account for how elements at the unobservable level interact over space or time to bring about the observed phenomenon (Krist et al., 2019). Schwarz et al. (2014) documented two high-performing fifth-grade girls’ mechanistic accounts about their
models over a period of one and half years. During this period of time, the girls iteratively developed and refined their models about several topics. They told the researchers at multiple interview sessions that a model should explain how and why something happened, and they gave mechanistic remarks to explain how and why the given phenomena took place. For example, in one of their explanations, one girl described the onset state of air molecules and their activities. Then she described how odor molecules were evaporated from an object and traveled to our sensor. She continued her explanation by addressing how a change of temperature in the condition would bring interconnected changes to the energy, spread, and speed of molecules’ movements. In this case, the model introduces the set-up condition of unobservable entities and their activities, then gives details about the changes in the property of the entities or the condition in which activities carry out accordingly, and finishes by describing the termination condition of the processes. The results suggest that high-performing elementary students in an inquiry-based modeling learning setting enhanced by epistemic knowledge of modeling, especially on the how and why aspects of an explanation, are capable of generating mechanistic models.

Cases like what Schwarz and her colleagues reported are rare because students without instruction do not spontaneously explain a full account of mechanisms of a phenomenon. Nonetheless, students’ natural awareness of causal links is a foundation for scientific inquiry (Hammer, 1995) as “causal reasoning serves as a starting point for the pursuit of underlying mechanistic explanations, but causal reasoning alone does not define it” (Russ et al., 2009, p. 881). Students’ mechanistic accounts may not be scientifically correct, but this attention to mechanisms informs their teachers of what
peripheral causal relationships they are aware of (Russ et al., 2009) and enables the teachers to provide effective scaffolding to facilitate scientific learning.

Although not being scientifically ideal, students’ models reflect that they might have undergone a learning process in which they attend to some epistemic principles of modeling (Acher et al., 2007). In a life science classroom, students may create intermediate-level models that fall any point between the levels of cells and molecules. Students may go through epistemic reasoning that resembles what biologists practice; the expression at an intermediate level is a heuristic strategy often adopted by biology researchers (van Mil et al., 2013). Without the domain-specific knowledge of molecular mechanisms that plays a crucial role in learning cellular topics (Duncan, 2007), can students hypothesize cellular activities and mechanistic processes? Driven by this question, research question 2.1 asked, “To what extent do students’ models incorporate components of mechanisms?” and research question 2.2 asked, “What models do students develop to explain how energy is generated?”

The Group Discussion Processes That Mediate Modeling Performance

Discourse activities are core to science in which it serves as the social medium through which explanation is communicated, critique is conveyed, disagreement is identified, negotiation is undertaken, and consensus is achieved during the pursuit of conceptual and epistemic goals among the inquiry participants (Duschl, 2008). Inquiry classrooms are places where students engage in discussions with peers and teachers in attempting to construct new knowledge about natural phenomena (Chin & Osborne, 2010; Driver, et al., 1994). Discussions, especially those that take place within a small group during inquiry, offer valuable information about how a concept is developed and
how an explanation is generated jointly by group members (Woodruff & Meyer, 1997). Clear ties have been found between some factors and the quality of student exchanges, including the cognitive and social aspects of the group discussion as well as the support that students receive from their teachers (O’Donnell, 2006).

**Group Discussion**

There has been a growing body of contemporary research in forms of experimental studies and observational analyses to investigate student discourses when they are working together. Understanding has been fully developed concerning the techniques of grouping as well as the relationships between a variety of cognitive factors and learning outcomes (see an extensive discussion in O’Donnell’s 2006 article). Following that research thread, subsequent studies shed light on two major topics, the reasoning features and the patterns of the discourse movements.

**The Reasoning Features of Group Discussion.** Of those that documented students’ reasoning courses during discussions in a natural setting for an inquiry topic, most of them documented how students developed their reasoning or explanations in a small group. For instance, Richmond and Striley (1996) investigated how the 10th graders’ reasoning progressed across four experiments in small group talks. Woodruff and Meyer (1997) documented the improvement of reasoning and explanations made by the 7th graders when they tried to account for three physics scenarios. Finkel (1996) exhibited the excerpts of a high school small group’s discussions during their engagement in a model-based inquiry project. Constrained by their study designs, although these studies documented the processes of the group discussions, they did not have enough
evidence to make a conclusive argument about the necessary reasoning features in a small group discussion for successful performance.

The most inspirational studies in this cohort might be those that examined the reasoning features between high-performing and low-performing small groups. High-performing groups demonstrated a distinct set of reasoning operations that were not observed in their low-performing counterparts especially in the ability of coordinating models with evidence. Johnson and Stewart (2002) compared the reasoning and practice processes between two high school groups, one successful and one unsuccessful, when they undertook a genetics inquiry in which they had to revise Mendel’s model after exposure to anomalous data. The unsuccessful group members did not compare data patterns against two relevant original models to identify anomalies. And because the unsuccessful group members did not consider both of the original models, they also mistakenly identified normal data as anomalous. Besides, their analyses relied on expressed phenotypes instead of the invisible genotypes which accounted for variations of traits. They kept trying the same hypothesis on different populations hoping to find one that would fit the hypothesis after repeated unfruitful attempts. On the contrary, the successful group emphasized their analyses on invisible genotypes from the beginning of their inquiry, and consistently used the original two models to detect anomalous data. They revised the models and created one that could account for existing and new data patterns.

Similar to Johnson and Stewart’s comparisons between two groups, Keys (1997) highlighted the reasoning processes unfolded in two junior high school dyads during modeling. Keys had each of the dyads carry out two identical inquiry activities; a clear
distinction of reasoning operations between these two small groups was identified. In the first activity that requested general reasoning skills to resolve a black box puzzle, the higher-achieving dyad’s discussions indicated their competences in many ways, including their understanding of the tentative nature of a hypothesis, their appreciation of conducting new trials to gather more evidence for model revision, their judgment of the consistency between evidence and a model, and their integration of sets of evidence into model construction. All of these competences were absent in the discussion logs of the lower-achieving dyad. They did not question their first hypothesis’s truthfulness and did not conduct tests to confirm or refute it. The lower-achieving dyad did not compare their observations, nor did they link observations with tests. They also failed to coordinate evidence with their hypothesis. The second activity involved constructing explanations to account for a phenomenon concerning electron transfer and ion formation and bonding. The higher-achieving dyad had an incorrect preconception, but they immediately abandoned it after encountering surprising evidence. They went out to search for information from a textbook and activated relevant prior knowledge. They collected, compared, and coordinated different sets of information and evidence into a model to explain the target phenomenon. Although the lower-achieving dyad also recognized that their initial hypothesis was not consistent with the evidence, they did not seek out relevant conceptual knowledge, and only attempted to incorporate their intuitive ideas into the explanation.

The purpose of this small cohort of studies was to represent rich descriptions of the reasoning competences and the progress of practices observed in small group discussions during inquiry. Student discourse was not construed on the basis of an
existing framework. It was also not the intention of this cohort of studies to infer any patterns of group discussion.

**The Patterns of Group Discussion.** A large number of studies has probed into small group discussions to explore the patterns of reasoning or dialogue that unraveled in student conversations. Many of them followed a framework to validate or extend the theoretical assumptions. Among the frameworks, Toulmin’s layout of arguments (Toulmin, 1958) has been prevalently utilized in studies to investigate students’ reasoning competence (e.g., Erduran et al., 2004). The layout entailed several components depending on the function that each component played in an argument. Toulmin’s framework was a great tool to examine the structure of reasoning in a discourse but it fell short of capturing the depth, the epistemics, and the social aspects of group members’ interactions.

Dissimilar to Toulmin’s scheme, which stressed the structure of reasoning, other frameworks accentuated the interactional elements in group dialogues. In 1996, Mercer proposed a linguistic tool that categorized students’ speech acts into three patterns and conducted a string of studies based on this tool. The findings indicated that when peers explored a topic by challenging each other for clarification, explanation, and justification, they made more conceptual gains than when they simply accumulating each other’s ideas through confirmation, repetition, and elaboration; the latter in turn outperformed the least constructive discourse pattern, which only consisted of individual assertions and counter-assertions in a dispute (Mercer et al., 2004; Mercer & Wegerif, 1999). While Mercer and his colleagues’ framework tackled the investigation from the linguistic perspective, Hatano and Inagaki’s research (1991) pointed out that certain metacognitive processes in
group discussion were essential practices to enhance comprehension in science—for example, asking for clarification, setting up for disputation, and coordinating bits of knowledge. These metacognitive processes were later adopted by Herrenkohl and Guerra (1998) and were transformed into four categories (negotiating a shared understanding of procedures and standards, monitoring comprehension, challenging others’ perspectives, and coordinating theories with evidence) as the indicators of group members’ engagement during discussion.

Each of these analytical frameworks picked up the structure of the conversation or the metacognitive operations yet provided a less clear illustration of the reasoning operations in student talks. Some studies applied all of the three aforementioned frameworks in order to conduct comprehensive analyses (e.g., Chin & Osborne, 2010; Evagorou & Osborne, 2013); however, the analyses were mainly based on Toulmin’s model and the results in response to each framework were reported separately without noting intertwined interactions. A new analytical framework that integrates reasoning features and discourse patterns is in need for conducting more refined analyses and comprehensive understanding of group discussions contextualized in model-based inquiry. This leads to research question 3.1: “How does the discourse of groups that vary in modeling performance differ?”

Teacher Prompts

Unlike learning in a conventional direct instruction science classroom, in which factual knowledge is passed down as truth, in a model-based inquiry context students have to engage in intense reasoning to investigate the unknown and constitute explanations. The reasoning process is challenging and students are in need of teachers’
support to carry out the practice of modeling (Herrenkohl et al., 1999). As a facilitator, teachers provide scaffolding in the process to enable students to go beyond their current repertoire and build a model. Among the many effective scaffolds that teachers utilize to elevate student performance, prompting is one of them (see McNeill & Pimentel, 2010; Webb, 2009, for more).

Conflicting results have been found in different studies favoring generic or specific prompts in fostering student performance. In the study that was conducted by Davis in 2003, middle school students were assigned to either one of two prompting conditions where they carried out a writing task that involved critiquing evidence and scientific claims. Those who received generic prompts in the written packet were asked to stop and think without an instruction for what to write, and they outperformed their peers who received specific prompts which hinted at a productive avenue for reflection. Opposite results were found in McNeill and Krajcik’s study (2009) which also assessed the effectiveness of generic and specific written prompts on middle school students’ performance of writing a scientific argument. The best learning growth between pre- and post-tests was found in the condition when students received specific written prompts after their teachers gave them the general guidance of the argument framework. Kang and her colleagues (2014) further compared the effectiveness of generic and specific prompts on five types of written scaffolds embedded in scientific written activities for secondary school students. There were two scaffolding conditions, one using generic prompts and the other using specific prompts. Mixed results were obtained across different types of written scaffolds. For example, a specific checklist (such as a prompt that specified the aspects and relationships to be included in the explanation) was more effective than a
general checklist (such as a word bank), but a generic rubric (such as a sentence said, “Pick two pieces of evidence and write how they support your explanation”) worked better than a specific rubric (in which expected performances were elaborated in details in a comprehensive table). It is not easy to make a conclusive recommendation for either generic or specific prompts.

The literature reviewed above all acquired their findings from quantitative analyses on students’ writing tasks when the prompts were embedded in the writing materials. In an inquiry context, however, conversations naturally occur between peers and between students and teachers before it is time to wrap up their ideas and write an explanation. Group peers discuss during inquiry, and verbal prompts are often dropped by teachers when they lean in to provide table-side support. Research question 3.2 thus asks, “What prompts from teachers enhance modeling performance?” in light of the dialogical nature in a model-based inquiry classroom.
Chapter 3

Methodology

This chapter begins with an introduction to the general curricular environment of the project. After that I describe the participants and the cellular respiration module followed by the materials and data collection. Subsequently, I explain the study design and procedures. Finally, I discuss the analytical methods for answering my three sets of research questions.

The General Curricular Context

This cellular respiration lesson was a module of the NSF-funded project PRACCIS. PRACCIS is a research project that has developed a one-year long life-sciences curriculum engaging middle school learners to learn topical content knowledge through inquiry activities. The inquiry activities are intended to foster students’ reasoning about the relationship between models and evidence and to build models based on evidence. Students in some cases carried out their inquiry and constructed their own models to account for evidence; in other cases, they had to choose one model from several flawed ones and revise it. In still other cases, students were provided with two competing models that explained a certain topic, and they had to choose one model that fitted most evidence. In all modules, students had to give reasons to explain and justify their models based on evidence and the criteria of matching evidence with models. These evidence-model practices were situated within class and group discussions. Students also had to demonstrate how they incorporated evidence and models in individual written work.
The modules in PRACCIS during this implementation covered a wide range of topics including but not limited to cellular organelles and processes, Mendelian and molecular genetics, energy flows in photosynthesis and cellular respiration, human health, and ecology and food webs. Across those topical modules, evidence was presented in various ways. Students sometimes conducted hands-on experiments or observations, or they obtained evidence through ready-made data, computer simulations, written summaries of real studies, and so on. Each module included several pieces of evidence that varied in quality and relevance to the phenomena under study, and students used them to develop, revise, or choose among models. Different pieces of evidence frequently conflicted with each other and with different models. Students had to work out a best model that explained the phenomenon according to its fit with the most reliable evidence.

A number of instructional scaffolds were used to support students’ modeling work. In addition to the use of argumentation to orchestrate discussion and refine reasoning, another featured scaffold was the use of public “criteria for good models” introduced by the teachers on the blackboard. Examples of these criteria were “A model uses relevant evidence and fits the data” and “[A model] Captures sequence of events if applicable.” These criteria helped guide students’ attention and epistemic reasoning when they engaged in generation and evaluation of models, evidence, and arguments.

Students’ inquiry was also supported by the use of Model-Evidence Link diagrams (the MEL diagrams) before they constructed their own final models or made a decision about which model to choose. Students used different types of arrows to indicate
the relationship between each piece of evidence and each model (Rinehart et al., 2014).

The *MEL diagrams* had not been developed when this module was enacted.

Teachers who partook in PRACCIS attended a series of professional development workshops with the project team led by the principle investigators (Dr. Clark Chinn, Dr. Ravit Golan Duncan, and Dr. Richard Duschl) and their research team before the project started in the summer. Monthly workshops during the school year were held to provide participating teachers timely supports and guidance. Research team members visited one or two of each teacher’s classes weekly, discussing the classes in smaller meetings with one or two teachers and one or two research team members.

**Participants**

Four seventh-grade teachers and their classes participated in and completed this cellular respiration module. Two teachers and their students were from Middle School A, and the other two and their students were from Middle School B. School A was located in a neighborhood adjacent to a middle size city. In this school, 27% of students qualified for free or reduced-price lunch; 47% of students were Black, 27% were White, 15% were Hispanic, and 11% were Asian. About 70% of the students reached proficient or advanced proficient levels on state language arts exam, and about 65% reached these levels on state mathematics exam. School B was located in a suburban town. In this school, 1% of students qualified free or reduced-price lunch; 97% of students were White. About 90% of students reached proficient or advanced levels on state language arts and mathematics exams.

Ms. Addison from School A, Mr. Britton from School B, and their combined total of 167 students were included in this study. Those participants combined generated 80
pre-instruction models used in this study. Both teachers had more than 10 years of experience at the time of the implementation of this study. These two teachers were selected because they had the most complete data and did not leave out any crucial activities due to lack of time. They also gave their students the most autonomy in interpreting the evidence and creating their models. All of the teachers’ and students’ names were pseudonyms.

**The Cellular Respiration Module**

The cellular respiration module was developed in the spring semester of 2007 by my teammates and myself as one of the core units during that year for seventh grade teachers and students to use. The project investigators exchanged emails with participating teachers to discuss details about the enactment of this module and observed one or two of each teacher’s classes during the module. Data were collected by a crew of undergraduate students who rotated going into the classrooms depending on their available schedule. My involvement was limited to the design only, which included the instruction for students and manuals for teachers. I did not interact with teachers (apart from participating in the summer professional development), nor was I able to go to the classrooms during the time when this module was implemented in seven middle school teachers’ classrooms in the same semester of Spring 2007 and in following school year. The PRACCIS project was a near-full-year implementation that started in early October and lasted through June. The cellular respiration unit was implemented approximately two thirds of the way through the year.

Two versions of this module were enacted. Each teacher had half of their classes assigned to one version and the other half to the other. Students in each class received the
same version of the instruction. The two versions differed only in what happened in the beginning of the module. In Version 1, students were given no examples prior to model construction. In Version 2 students were provided with three non-optimal models as examples for them to evaluate before they constructed their own models. The comparison between versions was not the purpose of this study. Furthermore, no statistically significant differences between versions within each teacher’s classes were observed on models built by students ($t(60.552) = .336, p = .738$ between the means of Version 1 and 2 in Ms. Addison’s classes; $t(144) = 1.583, p = .116$ between the means of Version 1 and 2 in Mr. Britten’s classes). Therefore, I do not consider this variable further.

The Materials

Each individual student received a packet, a study worksheet, and a final individual model worksheet. Each small group had a big poster paper for them to draw and write their group models on. The classes who received Version 2 material received one additional worksheet to evaluate three given example models before they formulated their own group models on the big poster paper. The example models are described in more detail later. The materials are displayed in Table 1.

The student packet consisted of a set of driving questions, seven pieces of evidence, and a set of questions to prompt group discussions. The single-paged worksheet with the questions displayed a table that included thumbnails of each piece of evidence on the first column and three questions on the first row. The poster paper sheets were regular posters in a size that is easy for group to work together at a table. Groups ranged in size from two to four. Students were encouraged to use one color chart pen for their initial model and other colors for revisions. The final individual model worksheet was on
one single page of paper with a box on the upper half and lines below the box for further explanation or justification. Students were free to create their final individual model within the box, on the back of the paper, or on another blank paper.

Table 1

Materials in Sequence

<table>
<thead>
<tr>
<th></th>
<th>Version 1</th>
<th>Version 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Packet with evidence enclosed</td>
<td>Packet with evidence enclosed</td>
<td></td>
</tr>
<tr>
<td>Study worksheet</td>
<td>Study worksheet</td>
<td>Sheet with three example models</td>
</tr>
<tr>
<td>Blank poster for group modeling</td>
<td>Blank poster for group modeling</td>
<td></td>
</tr>
<tr>
<td>Final individual model sheet</td>
<td>Final individual model sheet</td>
<td></td>
</tr>
</tbody>
</table>

The student packet started with a set of driving questions: “Have you ever thought about why we eat and breathe? How do our bodies use food and oxygen to get energy? How does this happen within our cells?” Right after these questions, a cover story was provided. The title of the story was “Our bodies need food and oxygen to get energy. How do our cells get the energy from food and oxygen?” This title question was intended to challenge students to link the earlier driving question to what they would read subsequently in the story. In the story, two boys named Eric and Roger, the same age as most of the participants, were receiving a home assignment to figure out how the human body uses food and oxygen to get energy, which was exactly the same topic that the participants were working on in this module. The story characters engaged in some studies and presented the results of their measurements. The study results were referred to as evidence. There were seven pieces of evidence. Each piece of evidence was presented in graphs, pictures, or tables with measurements denoted in metric units. Students were
expected to utilize the information gleaned by the evidence to build their cellular respiration models.

The pieces of evidence are listed in Appendix A and the three example models used in Version 2 are shown in Appendix B. The first evidence was a table showing the similar physical conditions of the two boys in the story (see Evidence 1 in Appendix A). This evidence was not relevant to cellular respiration but was intentionally designed as a control variable for subsequent measurements.

The second evidence (see Evidence 2 in Appendix A) described muscle contraction in the presence of varying levels of sugar and oxygen. A diagram presented a microscopic view of muscle cells. The note below the microscopic view stated that muscle cells in muscles contract when we exercise, and that the harder the exercise the faster the muscles contract. In the table, the rate of muscle cell contraction was associated with the amount of oxygen inhaled and the amount of sugar (glucose) eaten.

The third evidence (see Evidence 3 in Appendix A) demonstrated the rate of exhaled carbon dioxide when the two boys were running. The graph showed sharp increase of the rate from the first to the third minute and reached a plateau at the fourth minute. No significant differences were shown between the lines of two boys.

The table in the fourth evidence (see Evidence 4 in Appendix A) showed the changes in the two boys’ body temperatures as they ran. Both boys’ body temperatures increased, but no significant differences in temperature between two boys were present.

The fifth evidence (see Evidence 5 in Appendix A) displayed the rate of oxygen inhalation by the two boys when they were running. The graph showed that the rate
increased sharply from the first to the fourth minute. No significant differences were found between the two boys.

The sixth evidence (see Evidence 6 in Appendix A) was a chart comparing how long the two boys ran after certain lapses of time based on different foods they had eaten beforehand. The boy who could run longer had high-carb food and the other one had low-carb food. Below the chart, there was a note to prompt students thinking how the difference in two boys’ physical performances could be related to how energy was generated in the cells. This was the first evidence that indicated a meaningful difference between the two boys.

The last evidence (see Evidence 7 in Appendix A) compared the densities of various organelles in the two boys’ muscle cells in a table. Next to the table, there was a microscopic picture of a generic cell and its enclosed organelles. The table showed that one boy’s mitochondrial density was slightly higher than the other boy. This was the second evidence that indicated a meaningful difference between the two boys.

On the study worksheet, there were thumbnail images of each piece of evidence on the leftmost column in the table. The thumbnail images were intended to remind students of the corresponding evidence in the evidence packet as they filled out the table. In the top row of the table, there were two questions and one prompt. At this point after students read through the evidence, they only needed to respond to the first question for each evidence. In Version 1 the first question asked, “What does this show?” The second question asked, “Is this relevant to your model?” The final prompt asked students to “Explain why it is or isn’t.”
In Version 2 the questions were a little different from those in Version 1 to accommodate the fact that Version 2 presented students three non-optimal models, whereas Version 1 did not. The first question was the same, asking what the evidence showed. The second question, “Which model is best supported by this data?” referred to the three given models on the second worksheet they received after completion of the first question. The third question, “Is this relevant to our new model?” referred to the model students co-constructed with their peers in the groups. The worksheet that contained three example models in Version 2 will be described in the section that explains study design and procedures.

**Data Collection**

Students’ pre-instruction models, worksheets, and post-instruction models were photographed or scanned from all consented participants in Ms. Addison’s six classes and Mr. Britten’s five classes. In this study, only pre-instruction models were analyzed in response to the first and second sets of research questions. Three of Ms. Addison’s classes (two using Version 1 and one using Version 2) and two of Mr. Britten’s classes (one using Version 1 and the other one using Version 2) were audio- and video-recorded throughout the yearlong project. The audio data were drawn in response to the third set of research questions. According to the teachers’ preferences, most students were organized in pairs in Ms. Addison’s classes and in groups of four in Mr. Britten’s classes. One data collector from the research team facilitated the recordings under the teachers’ supervision and filled out period reports.
Study Design and Procedures

The module spanned about four 40-minute periods of instruction on average.

Table 2 displays the activities implemented in two different versions. Before the
enactment of the module, students were not presented with any background information
on cellular respiration.

Table 2

Activity Procedures

<table>
<thead>
<tr>
<th>Activity</th>
<th>Version 1</th>
<th>Version 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Read the evidence packet</td>
<td>Read the evidence packet</td>
<td></td>
</tr>
<tr>
<td>2. Interpret evidence on the study</td>
<td>Interpret evidence on the study worksheet^a</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Evaluate three example models^b</td>
<td>Construct an initial model^c</td>
<td></td>
</tr>
<tr>
<td>4. Explain the relations of each piece</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Critique models in a classroom</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Critique models in a classroom</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Observe a burning experiment</td>
<td>Observe a burning experiment</td>
<td></td>
</tr>
<tr>
<td>8. Receive conceptual instruction</td>
<td>Receive conceptual instruction</td>
<td></td>
</tr>
<tr>
<td>9. Construct a final individual model</td>
<td>Construct a final individual model</td>
<td></td>
</tr>
<tr>
<td>10. Justify individual models in a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Justify individual models in a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Version 1 did not implement the third activity.

^Worksheet questions were not identical between versions. ^Three example models were only presented in
Version 2. ^Ms. Addison’s students in both versions constructed group models before the instruction.
Before instruction, Mr. Britten’s students in Version 1 constructed group models, whereas those in Version
2 constructed individual models.

Read the Evidence Packet

On the first day the teacher presented the class with a list of driving questions and
introduced the cover story on the evidence packet (the first activity in Table 2). The cover
story and the evidence were displayed in a cohesive story fashion to enhance students’
interest and comprehension. Contextualizing a scientific phenomenon in a situation is
evident to enhance student explanation (Kang et al., 2014). The teacher helped the
students understand the scenarios in the story and make sense of charts, tables, and pictures in the evidence.

**Interpret Evidence on the Study Worksheet**

Students received a study worksheet with thumbnail images of each piece of evidence and responded to the question “What does this show?” They were encouraged to discuss their answers with group members. Students in Ms. Addison’s classes had about 20 minutes to work on the study worksheet. Students in Mr. Britten’s classes spent 7 to 10 minutes on this task. Students who received the Version 1 packet moved on to construct a group model after this step. They did not receive three example models before constructing their initial models.

**Evaluate Three Example Models**

Students who received the Version 2 packet received an additional worksheet after they completed the first question on the study worksheet (the third activity on Table 2). This second worksheet displayed three non-optimal models in flowcharts (see *Three Example Models* in Appendix B). None of the three example models was complete or scientifically correct. In fact, each of them was inconsistent with the evidence, and none of them presented the mechanisms of cellular respiration. Students evaluated the three given models against the “criteria for good models” publicly displayed in class. Below the models, students wrote their reasons to explain why they chose one of the three models as a relatively better one and why they considered the other two less favored. After this, students responded to the second question on the study worksheet, “Which model is best supported by this data?”
**Construct an Initial Model**

Students constructed pre-instruction models (the fourth activity in Table 2). The teacher prompted students to develop their models in accordance with the “criteria for good models” that had been publicly displayed in class through the school year. Students who received the Version 1 packet constructed a model of cellular respiration with group members on a big poster, and those who received the Version 2 packet could choose to construct a model from scratch or revise one of the three example models. Ms. Addison students built their pre-lecture models on a big poster with group members. Students in Mr. Britten’s classes who received the Version 1 packet created group models on a big poster, while those who received the Version 2 packet constructed models individually on a sheet of paper.

**Explain the Relations of Each Piece of Evidence to the Model Constructed**

After students finished creating the pre-instruction model, those who received the Version 1 packet wrote their responses to two remaining questions “Is this relevant to your model?” and “Explain why it is or isn’t” in the study worksheet for each piece of evidence. Students who received the Version 2 packet answered the third question, “Is this relevant to our new model?” for each piece of evidence.

**Critique Models in a Classroom Discussion**

Students’ pre-instruction group models on posters were collected by the teacher after class and classified into three or four categories. The teacher selected a prototypical group model from each category. The following day, the teacher exhibited these prototypical group models on different tables or different corners of the blackboard and invited students to critique these models. To avoid social pressures among students or
embarrassment from having their own group models evaluated, teachers used models from other classes to enact this activity (In Mr. Britten’s Version 2 classes, he used group models from other Version 1 classes).

Each student first walked to a preferred model on a table or a corner of the blackboard, then gave reasons of their choice. Students also suggested areas where the chosen model could improve on in response to the teacher’s prompts. Students could walk to a different model if they changed their minds later. During this activity, the teacher advised students to think about the “criteria for good models” and evaluate the strengths and weaknesses of the displayed models. Prompts such as—“What is the evidence that supports this model?” “Is there any evidence this model cannot explain?” “Why don’t you choose that model?” “How would you improve this model to make it fit all evidence?” and “What do you think about somebody’s reasons?”—were recommended by the project team to stimulate students’ reasoning about the relationship between models and evidence.

**Observe a Burning Experiment**

After the class critiqued group models, the teacher burned a high-carb cracker in a beaker with a lid until the fire extinguished (the 7th activity in Table 2). This burning process was intended to serve as an analogue to how cellular respiration took place in our cells. The teacher explained that the invisible cellular transformation of oxygen and glucose into energy, carbon dioxide, heat, and water was similar to a “burning process” demonstrated. Some teachers left the explanation here and some further explained the processes at the molecular level.
Receive Conceptual Instructions

After the burning experiment, the teacher proceeded to give a benchmark lecture of cellular respiration and laid out the chemistry equation. The teachers also illustrated the connection between cellular respiration and photosynthesis and explained how energy flowed among organisms in an ecosystem. The instruction was consistent with the discipline’s core ideas and the crosscutting concepts regarding the nature of energy (NGSS Lead States, 2013).

Construct a Final Individual Model

After receiving the conceptual instruction, each student created a cellular respiration model individually on the final individual model worksheet (the 9th activity in Table 2). Students could represent their ideas in pictures, diagrams, texts, or any of these forms combined. After they were finished, students wrote reasons to justify why they thought they built a good model and indicated what pieces of evidence their model was consistent with. Ms. Addison integrated this individual modeling activity into a quiz where students responded to items that assessed their content knowledge before they proceeded to construct a final model. Students in Mr. Britten’s classes completed this activity as homework and brought their final individual models to the next class for peers to evaluate them.

Justify Individual Models in a Classroom Discussion

In the next class, students were encouraged to consider the “criteria for good models” and gave reasons to support their own final individual models in a teacher-led classroom discussion. The module ended.
Data Analysis

Three sets of analyses were conducted to address the three corresponding sets of research questions. All analyses were confined to pre-instruction data. The first set of analyses focused on how competent students were in coordinating models with evidence. The second set of analyses addressed students’ capability of incorporating mechanistic processes in their models. The third set of analyses explored the processes during group discussions that mediated the varied performances of student models.

Coordination of Models and Evidence

The first set of analyses investigated how competent students were in coordinating models with evidence. My investigation proceeded in two stages. I first asked to what extent students made their models consistent with relevant evidence and discounted irrelevant evidence. After that I evaluated students’ interpretation of each piece of evidence.

Model and Evidence Consistency. The first analysis in the first set of analyses was conducted to answer research question 1.1: “To what extent are students’ models consistent with evidence?” Students encountered seven pieces of evidence in the packet. The first evidence provided contextual information and was not directly relevant to the process of cellular respiration. I assessed the percentage of models that were consistent with different numbers of relevant evidence (from 0 to 6 pieces of relevant evidence). I also provided the percentage of model-evidence consistency for each piece of evidence.

The Quality of Evidence Interpretation. The second analysis in the first set of analyses addressed research question 1.2: “How is the quality of evidence interpretation represented in students’ models?” Students established their ideas when they were
confronted with evidence before then built a model. The contents students incorporated in a model served as a proxy to imply their interpretation of evidence. A good interpretation of evidence referred to a tight fit between a model’s contents and the key concepts embedded in evidence. A poor interpretation of evidence implied that the evidence was ignored, misunderstood, or perceived partially. Students’ interpretation of each piece of evidence was translated into a point system ranging from 0 (poor) to 5 (excellent) to reflect the degree of consistency between the model and evidence. In a case where more than one interpretation of a piece of evidence was present, the point was determined by the highest performing interpretation. Each piece of evidence had a corresponding measurement for the quality of interpretation except the muscle cell contraction evidence which contained two concepts. One concept was the relationship between muscle cell contraction and the intake of sugar and oxygen, and the other concept was the relationship between muscle cell contraction and energy generated. Therefore, there were eight measurements for seven pieces of evidence. The toly of the points from eight measurements became a model’s score of model-evidence consistency. The percentage distribution of point values within each concept was generated. The seven pieces of evidence are displayed in Appendix A. Tables 3 to 10 are the coding schemes for eight measurements.
Table 3

The Measurement of High-Carb vs. Low-Carb

<table>
<thead>
<tr>
<th>Point</th>
<th>Range of interpretation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td><strong>None.</strong> Evidence is not present.</td>
<td>“The digested food is burned in the mitochondria at the cell to produce glucose and oxygen”</td>
</tr>
<tr>
<td>1</td>
<td><strong>Glucose-output.</strong> Student states that glucose is an output.</td>
<td>“Food is taken in”</td>
</tr>
<tr>
<td>1</td>
<td><strong>Food-input.</strong> Student states that food or nutrient is an input.</td>
<td>Draw two boys each having carb and protein, respectively, before they exercise.</td>
</tr>
<tr>
<td>2</td>
<td><strong>Carbs-input.</strong> Student states that high-carb food or glucose is an input or compares inputs between high-carb versus low-carb or protein.</td>
<td>“The more carb you have the faster or longer you will run”</td>
</tr>
<tr>
<td>3</td>
<td><strong>Carbs-enable-exercise.</strong> Student states that high-carb food or glucose enables or enhances exercise.</td>
<td>“When you eat food, you get energy→ when you get energy and you breathe in O₂, your muscle cells contract”</td>
</tr>
<tr>
<td>4</td>
<td><strong>Food-energy.</strong> Student states that food is an input to a process involving energy generation.</td>
<td>“Food is transformed into glucose and transported to cells→ burning process in mitochondria→ energy”</td>
</tr>
<tr>
<td>5</td>
<td><strong>Carbs-energy.</strong> Student states that carbs, sugar, or glucose is an input to a process involving energy generation.</td>
<td>“High carb will be burned and turned into energy by the mitochondria”</td>
</tr>
<tr>
<td>5</td>
<td><strong>High-carb-energy.</strong> Student states that more carbs, or glucose transformed from high-carb food is an input to a process involving energy generation.</td>
<td></td>
</tr>
</tbody>
</table>
**Table 4**

The Measurement of Oxygen

<table>
<thead>
<tr>
<th>Point</th>
<th>Range of interpretation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td><strong>None.</strong> Evidence is not present.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td><strong>O₂-output.</strong> Student states that O₂ is an output.</td>
<td>“After you exercise, you breathe out O₂”</td>
</tr>
<tr>
<td>1</td>
<td><strong>Air-input.</strong> Student states that air is an input.</td>
<td>Food and sugar burn in stomach while “air goes to lungs”.</td>
</tr>
<tr>
<td>2</td>
<td><strong>O₂.</strong> Student mentions O₂.</td>
<td>O₂ appeared in the model.</td>
</tr>
<tr>
<td>3</td>
<td><strong>O₂-input.</strong> Student states that O₂ is an input.</td>
<td>“You breathe in O₂”</td>
</tr>
<tr>
<td>4</td>
<td><strong>O₂-exercise.</strong> Student states that O₂ is an input in a model involving exercise.</td>
<td>“They both take in more oxygen as they continue to run”</td>
</tr>
<tr>
<td>5</td>
<td><strong>O₂-energy.</strong> Student states that O₂ is an input to a process involving energy generation.</td>
<td>“O₂ from the lungs goes into cells and cells generate energy”</td>
</tr>
</tbody>
</table>

**Table 5**

The Measurement of Carbon Dioxide

<table>
<thead>
<tr>
<th>Point</th>
<th>Range of interpretation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td><strong>None.</strong> Evidence is not present.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td><strong>CO₂-input.</strong> Student states that CO₂ is an input.</td>
<td>CO₂, O₂, and food are inputs.</td>
</tr>
<tr>
<td>1</td>
<td><strong>Air-output.</strong> Student states that air is an input.</td>
<td>There is no instance for this interpretation.</td>
</tr>
<tr>
<td>2</td>
<td><strong>CO₂.</strong> Student mentions CO₂.</td>
<td>Student draws two boys and compares their CO₂ level.</td>
</tr>
<tr>
<td>3</td>
<td><strong>CO₂-output.</strong> Student states that CO₂ is an output.</td>
<td>“Breathes [in] O₂, which also makes CO₂”</td>
</tr>
<tr>
<td>4</td>
<td><strong>CO₂-exercise.</strong> Student states that CO₂ is an output in a model involving exercise.</td>
<td>“Breathe out more CO₂ as exercise increases”</td>
</tr>
<tr>
<td>5</td>
<td><strong>CO₂-energy.</strong> Student states that CO₂ is an output to a process involving energy generation.</td>
<td>CO₂ is discharged from the cells after energy is produced.</td>
</tr>
</tbody>
</table>
Table 6

*The Measurement of Cellular Organelle*

<table>
<thead>
<tr>
<th>Point</th>
<th>Range of interpretation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td><strong>None.</strong> Evidence is not present.</td>
<td>&quot;Mitochondria work with metabolism in order to break down the food&quot;</td>
</tr>
<tr>
<td>0</td>
<td><strong>Mitochondria-help-digest.</strong> Student states that mitochondria help digestion.</td>
<td>“When you burn off all the carbs mitochondria help you even out your breathing”</td>
</tr>
<tr>
<td>0</td>
<td><strong>Mitochondria-help-breathing.</strong> Student states that mitochondria help breathing.</td>
<td>“O₂ and sugar are transferred into cells and energy is generated in cells”</td>
</tr>
<tr>
<td>1</td>
<td><strong>Cell.</strong> Student states that energy is produced in cells.</td>
<td>“Roger has higher mitochondria and Eric has lower mitochondria”</td>
</tr>
<tr>
<td>2</td>
<td><strong>Mitochondria-compare.</strong> Student compares two boys’ mitochondrial densities or cellular organelles.</td>
<td>“O₂ and sugar turn into energy→ mitochondria take in the energy”</td>
</tr>
<tr>
<td>3</td>
<td><strong>Mitochondria.</strong> Student draws mitochondria next to a cell or mentions that mitochondrion is the location of any processes other than energy generation.</td>
<td>“Roger runs further because he has more mitochondria”</td>
</tr>
<tr>
<td>3</td>
<td><strong>Mitochondria-enhance-exercise.</strong> Student states that a high mitochondrial count enables or enhances muscle cell contraction or exercise.</td>
<td>“Exercise more increases the density of mitochondria”</td>
</tr>
<tr>
<td>3</td>
<td><strong>Exercise-increases-mitochondria.</strong> Student states that exercise increases mitochondrial density.</td>
<td>“Roger has higher mitochondria because he gets more energy”</td>
</tr>
<tr>
<td>3</td>
<td><strong>Energy-increases-mitochondria.</strong> Student states that mitochondrial count increases when one has more energy.</td>
<td>“Mitochondria go up (to the blood stream) and you get energy”</td>
</tr>
<tr>
<td>4</td>
<td><strong>Energy-not-in-mitochondria.</strong> Student states that mitochondria as mediators or catalysts are involved in energy generation but it does not occur in mitochondria.</td>
<td>“Mitochondria in cell is the place where energy is made”</td>
</tr>
<tr>
<td>5</td>
<td><strong>Mitochondria-energy.</strong> Student states that energy is generated in mitochondria.</td>
<td></td>
</tr>
</tbody>
</table>
Table 7

The Measurement of Body Temperature

<table>
<thead>
<tr>
<th>Point</th>
<th>Range of interpretation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>None. Evidence is not present.</td>
<td>“Good temperature allows food process to go smoothly”</td>
</tr>
<tr>
<td>1</td>
<td>Temp-influence. Student states that body temperature influences digestion, energy generation, breathing, or exercising.</td>
<td>“Eric’s temp is much warmer than Roger’s”</td>
</tr>
<tr>
<td>2</td>
<td>Compare-temperature. Student compares two boys’ body temperatures.</td>
<td>“Pasta being digested (in a picture of intestine)→ body temperature increases”</td>
</tr>
<tr>
<td>3</td>
<td>Heat-output. Student states that burning produces heat, increases body temperature, or that heat is an output.</td>
<td>“When you exercise, muscles discharge heat”</td>
</tr>
<tr>
<td>4</td>
<td>Temperature-exercise. Student states that the increase of body temperature or the release of heat is associated with muscle contraction or exercise.</td>
<td>“(An arrow pointing at a cell) gives off heat when burning off carb”</td>
</tr>
<tr>
<td>4</td>
<td>Burn-carb-heat. Student states that heat is released when carbs are burnt off.</td>
<td>“Absorb nutrients—O₂ and sugar→ digest—uses the nutrients→ burns the energy—produces heat→ Waste= CO₂—burn more CO₂ when you exercise”</td>
</tr>
<tr>
<td>5</td>
<td>Temperature-energy. Student states that temperature rises or that heat is released from a process involving energy generation.</td>
<td></td>
</tr>
</tbody>
</table>

The muscle cell contraction evidence was broken down into two measurements. The first measurement (Table 8) evaluated students’ interpretation of this evidence regarding the rate of muscle cell contraction (or exercise, interchangeably) in relation to the intake of oxygen and glucose. The second measurement (Table 9) was concerning muscle cell contraction involving energy generation.

The physical fitness evidence (Table 10) was designed to serve as a controlled variable for the two boys’ studies displayed in other evidence. It was not relevant to
cellular respiration, and therefore a point value of 5 was assigned to models that did not mention any ideas derived from this evidence.

Table 8

The Measurement of Muscle Cell Inputs

<table>
<thead>
<tr>
<th>Point</th>
<th>Range of interpretation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td><strong>None.</strong> Neither sugar nor oxygen is involving muscle contraction or exercise.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td><strong>Food-contraction.</strong> Student states that food is an input to muscle cell contraction or exercise.</td>
<td>“Exercise burns on food”</td>
</tr>
<tr>
<td>1</td>
<td><strong>Carb-contraction.</strong> Student states that carbs are an input to muscle cell contraction or exercise.</td>
<td>“High carb food makes more muscle contraction”</td>
</tr>
<tr>
<td>2</td>
<td><strong>Sugar-contraction.</strong> Student states that sugar or glucose is an input to muscle cell contraction or exercise.</td>
<td>“You eat food→ food becomes sugar in stomach→ you get energy→ you can run better”</td>
</tr>
<tr>
<td>2</td>
<td><strong>O2-contraction.</strong> Student states that oxygen is an input to muscle cell contraction or exercise.</td>
<td>“The more oxygen the more the muscles contract”</td>
</tr>
<tr>
<td>3</td>
<td><strong>Food-and-O2-contraction.</strong> Student states that food and oxygen are inputs to muscle cell contraction or exercise.</td>
<td>“Muscles increase contraction because of food and oxygen”</td>
</tr>
<tr>
<td>4</td>
<td><strong>Carb-and-O2-contraction.</strong> Student states that carbs and oxygen are inputs to muscle cell contraction or exercise.</td>
<td>“The more carb and O₂ taken in the more the muscle contract when working out”</td>
</tr>
<tr>
<td>5</td>
<td><strong>Sugar-and-O2-contraction.</strong> Student states that sugar or glucose and oxygen influence muscle cell contraction or exercise.</td>
<td>&quot;Oxygen goes to muscle with sugar creating high muscle cell contraction&quot;</td>
</tr>
</tbody>
</table>
### Table 9

The Measurement of Muscle Cell Contraction

<table>
<thead>
<tr>
<th>Point</th>
<th>Range of interpretation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>None. Muscles, muscle cell contraction, or exercise is not present.</td>
<td>“As you exercise more, you create more carbon dioxide. You breathe in oxygen”</td>
</tr>
<tr>
<td>1</td>
<td>Contraction-only. Student mentions muscles, muscle cell contraction, or exercise.</td>
<td>“When you exercise, you get energy”</td>
</tr>
<tr>
<td>2</td>
<td>Contraction-energy. Student states that muscles, muscle cell contraction, or exercise is part of an energy generation process.</td>
<td>“Sugar is converted into energy and results in rapid muscle contraction”</td>
</tr>
<tr>
<td>3</td>
<td>Energy-contraction. Student states that muscles, muscle cell contraction, or exercise follows an energy generation process.</td>
<td></td>
</tr>
</tbody>
</table>

### Table 10

The Measurement of Physical Fitness

<table>
<thead>
<tr>
<th>Point</th>
<th>Range of interpretation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Know. Student says that one should know their own body condition.</td>
<td>“You should know your body condition”</td>
</tr>
<tr>
<td>1</td>
<td>Two-figures. Student draws two figures without any explanations.</td>
<td>Student draws two boys without words comparing them.</td>
</tr>
<tr>
<td>2</td>
<td>Body-description. Student describes physical features.</td>
<td>“If you are thin and tall than short and fat”</td>
</tr>
<tr>
<td>3</td>
<td>Compare-fact. Student compares physical conditions between two figures.</td>
<td>Student draws two human figures and compares their physical conditions and exercise frequencies.</td>
</tr>
<tr>
<td>4</td>
<td>Control. Student implies that the boys’ body conditions are the same, which serves as a control variable.</td>
<td>There is no instance for this interpretation.</td>
</tr>
<tr>
<td>5</td>
<td>None. Evidence is not present.</td>
<td></td>
</tr>
</tbody>
</table>
Additional Concepts Absent From the Evidence. In addition to evaluating the quality of students’ interpretation of evidence, I also documented ideas displayed in models that were not present in the evidence. Students might obtain additional or alternative information from previous lessons or personal experiences and weave it into their models to substantiate an explanation.

Incorporation of Mechanisms in Models

The second set of analyses investigated students’ ability to incorporate mechanisms into their models. The definition of mechanisms was taken from Machamer et al.’s 2000 seminal paper which said, “Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions” (p. 3). The entities in this module included those that served as inputs from start (i.e., oxygen and glucose) and outputs in the termination condition (i.e., carbon dioxide, heat, and energy); both engaged in active processes under certain conditions at a certain location to generate energy. This module did not intend to make students transcend their grade level and create atomic-molecular models. Students were only expected to express that energy was produced in mitochondria at the cellular level, but they did not need to specify the atomic-molecular activities.

The analyses of mechanisms proceeded from two analytical perspectives. One inspected the models from a bottom-up perspective from which features of mechanisms were evaluated. The other adopted a top-down perspective and categorized the types of models based on manifested mechanisms.

The Components of Mechanisms. The first analysis in the second set of analyses addressed research question 2.1: “To what extent do students’ models incorporate
components of mechanisms?” To be commensurate with the scoring system of the evidence interpretation analysis, three levels of points were employed: 0 (no content or low quality), 3 (content was not precise), and 5 (content was good). The sum of the points from four sub-analyses became a model’s score of mechanistic explanation. The measurement is displayed in Table 11.

The first sub-analysis of mechanisms examined if all five necessary entities altogether were involved in the energy generation process regardless of correctness. A point value of 5 meant that a complete list of entities was incorporated in the process of energy generation. A point value of 0 was assigned if any one of the five entities did not participate in the energy generation process even when they were present in the model. For example, a model might explain that oxygen entered the lungs and carbon dioxide left the lungs. In this case oxygen and carbon dioxide only interacted in the lungs and did not participate in the process that generated energy.

The second sub-analysis then inspected each of the five entities’ involvement in the energy generation process, again regardless of its correctness. Point values ranged from 0, 3, to 5 for entities oxygen, glucose, carbon dioxide, and heat. A point value of 5 was released if an entity connected with other entities in the process that generated energy. If the entity interacted with other entities but was not presented in an ideal form, a point value of 3 was assigned. For example, instead of specifying oxygen, air was present, or instead of noting that heat was released, temperature was rising. No point was earned if the entity did not interact with other entities in the energy generation process. The entity energy corresponded to only two point-categories. Energy was either present in (5 points) or absent from a model (0 points).
Table 11

*The Measurement of the Energy Generation Mechanisms*

<table>
<thead>
<tr>
<th>Component of a mechanism</th>
<th>Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completion of entities</td>
<td></td>
</tr>
<tr>
<td>Input entities (glucose, sugar, carbs, or food, and oxygen) and output entities (energy, carbon dioxide, and heat or rising temperature) are complete</td>
<td>5</td>
</tr>
<tr>
<td>Input and output entities are incomplete</td>
<td>0</td>
</tr>
<tr>
<td>Entity</td>
<td></td>
</tr>
<tr>
<td>Oxygen</td>
<td></td>
</tr>
<tr>
<td>Oxygen is taken in for energy generation</td>
<td>5</td>
</tr>
<tr>
<td>Air is an input for energy generation</td>
<td>3</td>
</tr>
<tr>
<td>Neither oxygen nor air is involved in the energy generation process</td>
<td>0</td>
</tr>
<tr>
<td>Glucose</td>
<td></td>
</tr>
<tr>
<td>Carbs, sugar, or glucose is taken in for energy generation</td>
<td>5</td>
</tr>
<tr>
<td>Food is taken in for energy generation</td>
<td>3</td>
</tr>
<tr>
<td>Neither food nor carbs, sugar, or glucose is involved in the energy generation process</td>
<td>0</td>
</tr>
<tr>
<td>Carbon Dioxide</td>
<td></td>
</tr>
<tr>
<td>Carbon dioxide is released from the energy generation process</td>
<td>5</td>
</tr>
<tr>
<td>Air is an output from the energy generation process</td>
<td>3</td>
</tr>
<tr>
<td>Neither carbon dioxide nor air is released from the energy generation process</td>
<td>0</td>
</tr>
<tr>
<td>Heat</td>
<td></td>
</tr>
<tr>
<td>Heat is released from the energy generation process</td>
<td>5</td>
</tr>
<tr>
<td>Temperature rises after the energy generation process</td>
<td>3</td>
</tr>
<tr>
<td>Neither heat is an output nor rising temperature is an outcome of the energy generation process</td>
<td>0</td>
</tr>
<tr>
<td>Energy</td>
<td></td>
</tr>
<tr>
<td>Energy is an output of the energy generation process</td>
<td>5</td>
</tr>
<tr>
<td>Energy is not an output of the energy generation process</td>
<td>0</td>
</tr>
</tbody>
</table>
Component of a mechanism | Point
---|---
Mitochondria | 5
Cells | 3
Stomach or intestines | 0
Heart | 0
Blood | 0
Multiple | 0
(Not specified) | 0

Agency

A process at the cellular level turns entities into energy (processes include that entities are burned, transformed, converted into, mixed and become, or combined together) | 5
A burning process not at the cellular level turns entities into energy | 3
Entities are digested and become energy | 0
Blood circulation generates energy | 0
Exercise or burning calories generates energy | 0
Dual processes—cellular process and digestion | 0
Dual processes—cellular process and exercise | 0
Dual processes—digestion and a burning process | 0
Dual processes—digestion and exercise | 0
Dual processes—digestion and breathing | 0
Dual processes—digestion and heart pumping | 0
Triple processes—cellular process, digestion, and breathing | 0
Food, carbs, sugar, glucose, or nutrients turn(s) into energy without a process | 0
(Not specified) | 0

The third sub-analysis identified the location of energy generation or cellular respiration (see the continued section of Table 11). If a model specified mitochondria as the location of energy generation or cellular respiration, a point value of 5 was assigned. If cells were recognized as the location, a point value of 3 was assigned. A point value of 0 referred to unspecified or alternative locations. In this sub-analysis, cellular respiration was taken interchangeable with energy generation. Students who were able to hypothesize that the module’s topic cellular respiration occurred in mitochondria when
confronted with an array of evidence were acknowledged for their achievement in this sub-analysis.

The fourth sub-analysis inspected the agency that generated energy. Only models that incorporated energy were further examined for this sub-analysis. Given that molecular and atomic processes were not in the 7th grade’s curriculum, a model that explained the processes of energy production at the cellular level was the best performance expected and would receive 5 points. If a model addressed energy was derived from a burning process but it did not occur at the cellular level, a point value of 3 was assigned. For other alternative explanations (e.g., entities were digested to produce energy, see Table 11) or in the case where no explanation about the agency was provided, a point value of 0 was assigned.

**Types of Mechanistic Explanations.** The second analysis in the second set of analyses responded to research question 2.2: “What models do students develop to explain how energy is generated?” This question was resolved by the fourth sub-analysis in Table 11. The classification of different agencies of energy generation implied the types of models that students developed. Percentages of model types were acquired.

**An Example to Demonstrate Measurement Application.** In this section, I describe how I applied the coding measurements to one model to demonstrate the coding procedure (see Figure 1).

First, I display the scoring for the evidence interpretation measurement. Starting from the left side, the model addressed that carbs enhanced running, so 3 points were released for the high-carb vs. low-carb evidence interpretation (Table 3). In the second step oxygen was inhaled and in the third step it said that muscle contraction was based on
the amount of oxygen and glucose taken in (Table 4). Referring to the rule that points were assigned to the best performance for each piece of evidence, 4 points were assigned for the oxygen evidence interpretation. Carbon dioxide was exhaled, so the carbon dioxide evidence earned 3 points (Table 5). The statement in the third step awarded 5 points for the muscle cell inputs concept (Table 8) and 1 point for the muscle cell contraction concept (Table 9) because it did not express its relationship with energy. In the description below the illustration of the mitochondrion it compared two boys’ mitochondrial densities. Although it said that cellular respiration occurred in mitochondria, it did not explain that energy was formed in mitochondria; therefore, the highest score it received for the cellular organelle evidence was 3 (Table 6). The final step said that body temperature influenced energy and the breathing rate; 5 points were secured for the body temperature evidence (Table 7). Finally, the model earned 5 points for the physical fitness evidence interpretation because it did not present any ideas concerning this irrelevant evidence (Table 10).

Next, I demonstrate the procedure of mechanism analysis (see Table 11). The model did not mention the entity energy. Because at least one entity was missing for the energy generation process, the model was considered not containing complete entities. It received 0 point for the first sub-analysis. It did not explain how energy was generated; in other word, none of the entities were involved in the process of energy generation. Therefore, each of the entities received 0 point, and the sub-analysis of agency also received 0 point. However, the model's builders specified that cellular respiration occurred in the mitochondria; 5 points were assigned for the sub-analysis of location.
The Processes of Discussions From Groups Varying in Model Quality

The third set of analyses employed qualitative method to explore the discursive processes that might mediate students’ modeling performance. I evaluated small groups’ conceptual reasoning and their interaction patterns. After that, I investigated teachers’ scaffolding in the form of prompts.

The data of small group discussion were sampled from five core classes where small groups’ conversations were audio-recorded during evidence interpretation and model building. Because Mr. Britten’s core classes had only Version 1 audio data available, I confined my selection of core classes from Ms. Addison and Mr. Britten to only those who were assigned to Version 1 packet. I first separated groups based on the two teachers. Then I sorted groups into three categories based on their models’ total
scores. Within each category, I randomly chose a couple of groups among those who had most complete data of all kinds (such as the study worksheets and the final individual models of all group peers). After this I listened to the recordings of the chosen groups and picked four small groups whose recordings had better sound quality.

Drawn from Ms. Addison’s classes, Group A1 was a high-performing group and Group A2 was a medium-performing group. Drawn from Mr. Britten’s classes, Group B1 was a medium-performing group and Group B2 was a lower-performing group. Each group’s discussion was transcribed from their original audiotapes, which lasted between 16 to 35 minutes. Table 12 displays the groups’ profiles and their performances on model-evidence consistency and incorporation of mechanisms.

**Table 12**

*Four Small Groups*

<table>
<thead>
<tr>
<th>Group</th>
<th>Score</th>
<th>Student</th>
<th>Teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Evidence</td>
<td>Mechanism</td>
<td>Total</td>
</tr>
<tr>
<td>A1</td>
<td>38</td>
<td>40</td>
<td>78</td>
</tr>
<tr>
<td>A2</td>
<td>29</td>
<td>5</td>
<td>34</td>
</tr>
<tr>
<td>B1</td>
<td>28</td>
<td>13</td>
<td>41</td>
</tr>
<tr>
<td>B2</td>
<td>13</td>
<td>0</td>
<td>13</td>
</tr>
</tbody>
</table>

**Group Discussions.** The first analysis in the third set of analyses addressed research question 3.1: “How does the discourse of groups that vary in modeling performance differ?” Group discussions were first broken down into episodes if a transition led into another round of conversation when a natural pause occurred as well as when ideas were initiated, developed, and terminated in an integral segment of exchanges that contained a particular topic. Within each episode, I evaluated small groups’ discourse operations and dialogue patterns.
**Discourse Operations.** Seven categories were developed to portray discourse operations: (1) off-topic, (2) regulation and logistics, (3) model expression, (4) model criteria, (5) intra-evidence interpretation, (6) inter-evidence integration, and (7) mechanistic explanation. An episode might contain more than one discourse operations. Table 13 describes seven discourse operations’ definitions and example excerpts.

The distinction between *intra-evidence interpretation* and *inter-evidence integration* lies in whether a statement considers more than one single piece of evidence or not. “Intra” refers to within one item, and “inter” implies between different items. For example, the statement “high-carb food makes you run longer than low-carb food” is an *intra-evidence interpretation* because it only addresses the high-carb vs. low-carb evidence in one account. This statement above would become an *inter-evidence interpretation* if it instead were expanded to “high-carb food makes your run longer, and when you run you inhale more and more oxygen.” The second statement coordinates two pieces of evidence into one remark in that the high-carb vs. low-carb evidence and the oxygen evidence are considered at the same time.

The distinction between *inter-evidence integration* and *mechanistic explanation* hinges on whether there is an expression to account for a causal, sequential, or correlational relationship between evidence to suggest a preliminary mechanism. For example, the *inter-evidence integration* statement in the previous example would be coded as incorporating a *mechanistic explanation* if it were extended to become: “A high rate of oxygen intake and high-carb food make your cells able to get more energy and therefore you can run longer.”
<table>
<thead>
<tr>
<th>Operation</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Off-topic</td>
<td>Discussing irrelevant contents</td>
<td>A statement such as “What time does the period end?” was irrelevant to the task.</td>
</tr>
<tr>
<td>Regulation and logistics</td>
<td>Managing time or task roles</td>
<td>A student urged peers to manage their progress: “Hurry up, we only got one minute.”</td>
</tr>
<tr>
<td>Model expression</td>
<td>Considering drawing a particular visualization to express certain ideas</td>
<td>Students were discussing how to create their model in a way to illustrate two boys in the story, and one of them said, “Ok, so we will draw like a little stick person.”</td>
</tr>
<tr>
<td>Model criteria</td>
<td>Reflecting on criteria of a good model</td>
<td>One of the criteria for good models in this class was having a title. A student said, “We need to get a title,” referring to this class criterion, when their group model was almost completed.</td>
</tr>
</tbody>
</table>
| Intra-evidence interpretation | Conceptualizing a single piece of evidence exclusively without involving other evidence | A snippet of dialogue such as:  
  S1: All right, this shows what?  
  S2: They have basically the same amount of cells.  
  revealed a focus on the cellular organelle evidence without connecting it to other evidence. This could happen in the initial stage of modeling or throughout the entire modeling activity. |
| Inter-evidence integration    | Integrating the interpretations of different evidence without a sign of considering their mutual relationships | A statement such as “They are the same size and they had the same physical condition, so our model is showing like the muscle contractions and how the food and carbon dioxide have to do with the mitochondria” showed the coordination of multiple pieces of evidence. |
| Mechanistic explanation      | Giving a cohesive account that incorporates mechanisms                    | An explanation specifies entities and used a verb to describe how activities produced changes from the on-set to the termination conditions, regardless of its correctness. In the statement “After the food and oxygen enter…they go to the…are performed in the cells and inside that…in the mitochondria is where they are burned, and make the temperature of the body rise” entities and activities were organized to explain processes. |
Dialogue Patterns. Episodes that were associated with any discourse operations in the range from (3) *model expression* to (7) *mechanistic explanation* were further inspected on group members’ dialogue patterns. Adapting and expanding the concepts of Chi’s ICAP framework (Chi & Wylie, 2014), I identified four dialogue patterns, *interactive, constructive-constructive, active-constructive, and receptive-constructive* to classify the utterances in peer collaboration during a modeling practice. These four dialogue patterns were mutually exclusive in an episode. Table 14 defines the dialogue patterns in conditions of giving statements and raising questions.

Chi’s ICAP framework was originally designed to recognize students’ cognitive engagements through overt behaviors (Chi & Wylie, 2014). Within the acronym ICAP, *I* referred to interactive, which meant that partners co-constructed new ideas based on each other’s contributions. *C* referred to constructive, which meant that a student inferred from prior knowledge or existing materials and constructed new knowledge. *A* referred to active, which meant that a student recalled, repeated, or paraphrased information heard or presented in the instructional materials. Finally, *P* referred to passive, which meant that a student showed that they have registered the information perceived. In a later article Chi and Menekse (2015) applied the ICAP framework to an array of dialogue patterns to categorize peer collaboration. They were co-constructive, constructive-constructive, active-constructive, active-active, passive-constructive, and passive-active. I adopted these hyphenated dialogue patterns but modified them to suit the nature of my data as described below.
Table 14

Dialogue Patterns

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Giving Statements</th>
<th>Raising Questions</th>
<th>Example</th>
</tr>
</thead>
</table>
| Interactive | Interactive partners elaborate on each other’s opinions and generate new ideas. They may challenge each other or provide suggestions or reminders, which then triggers a new round of discussion for a better solution. They work together to reach a consensus before they write or draw. | Interactive partners seek explanations and comments, and respond to them. They ask questions to gather different thoughts from their partners. An initial idea may be rebutted or questioned, and the responses are taken into account to spawn subsequent ideas. | S1: Food—  
S2: We got it drawing, we got it drawing.  
S1: No, not with drawing.  
S2: Oh food—  
S1: Um, food. (Laughing)  
S2: Food—  
S1: Um, food with high carbs—  
S2: With high carb um—  
S1: High carbs. Um, gives you more energy.  
S2: And helps your performance.  
S1: Yeah. |
| Constructive-Constructive | Constructive partners each work on their own part of verbal explanation or drawing and writing without taking other partner’s ideas into consideration. In other cases, partners may check on each other’s construction to ensure the whole is achieved by the sum of the parts. | A constructive partner seeks agreement after they suggest an idea, requests another partner to generate an explanation, or monitors each other’s progress to help actualize task completion. However, the questioner does not use the response received to extend ideas. | S1: Okay, what do we…what’s the model? Food—Um, what do we do? Food?  
S2: Oxygen…Inhale…  
S3: Eric’s bad at this one.  
S4: Yeah.  
S2: Oxygen…  
S3: Eric is going faster.  
S4: Draw a person running, S2. Here, draw a person running. |
<table>
<thead>
<tr>
<th>Pattern</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Giving Statements</strong></td>
<td>An active partner repeats or summarizes the ideas heard. They may recognize knowledge gaps as well as track down the information in the evidence packet and read aloud. Few words may be expressed to repeat or summarize ideas heard during drawing or writing.</td>
<td>S1: Draw legs like this. Draw—one—leg—like this. S2: Ok. I know how to draw a leg. S1: Then raw a foot, make a sneaker… S2: Uh—huh, that’s right! S3: Like—that. S2: Now what?</td>
</tr>
<tr>
<td><strong>Raising Questions</strong></td>
<td>An active partner asks another partner to provide a fact-based short answer for clarification or confirmation. They may also request an instruction for the next step.</td>
<td>S1: Then raw a foot, make a sneaker… S2: Ok. I know how to draw a leg.</td>
</tr>
</tbody>
</table>
Due to the limitation that only audio data was collected in this study, I determined students’ dialogue patterns on the basis of their audible activities. To accommodate a model-based inquiry context with an emphasis placed upon reasoning, I applied the ICAP framework in a way that captured lines of vocal statement that expressed reasoning instead of actions of drawing or writing. This resulted in five adaptations to Chi’s framework.

First, I made a distinction between making a statement and raising a question in each dialogue pattern (see the middle columns in Table 14). I identified four levels of statements and questions, each corresponding to a specific dialogue mode.

Second, I used the term receptive to refer to the dialogue mode that engaged in the lowest level of cognitive processes. Receptive included passive (to agree or confirm received information) as well as the lower level of active processes (to request repetition or spelling) in Chi’s original framework. As I explained in the previous paragraph, my discourse analyses intended to shed light on students’ reasoning patterns, and therefore receptive dialogue mode contained statements and questions that did not indicate any sign of reasoning.

Third, I modified the definition of active dialogue mode to exclude lower level questioning that was classified as the receptive dialogue mode in Table 14. Raising a question was considered active in Chi’s ICAP framework regardless of the level of cognitive process engaged in this overt action.

Fourth, drawing some components of a model was considered constructive in the original definition of ICAP because students were creating some products. In this study when a student’s utterance indicated that they were drawing, it was considered
constructive if this action was self-directed. If in another case such action was directed by another partner’s idea, this utterance was identified as active.

Fifth, because each group had to construct a model in this task, at least one partner had to be constructive. Even though some partners were receptive or active, at least one partner had to be constructive to build a new model. Therefore, the lowest two levels of dialogue patterns in this study were receptive-constructive and active-constructive.

In the following paragraphs I explain the four dialogue patterns I utilized to code students’ group conversation. When group members had an interactive dialogue pattern, they elaborated on one another’s thoughts, took others’ thinking into consideration, sought opinions, made suggestions or reminders, responded to comments, and challenged each other’s ideas. Interactive partners corrected, incorporated, and extended their peer’s thoughts to generate additional ideas. An initial opinion might be rebutted or questioned and was followed by a counterargument or explanation, and the response triggered a new round of idea generation. In sum, interactive partners co-constructed and shared knowledge.

Being constructive meant that a partner built on their own explanation from the evidence without extending another member’s idea. Peers engaging in a constructive-constructive dialogue pattern might take turns thinking about different pieces of evidence or having different people tackling different parts of a task without checking each other’s work. When peers were taking turns interpreting evidence, each idea was built by different contributors. Different task parts were assigned to different people, and when all parts were completed so was the model. Peers in a constructive-constructive dialogue
pattern monitored each other’s progress to make sure they successfully constructed a model. They did not invite challenges from their partners after they suggested an idea; they sought agreement. They might request other people to construct an idea. For example, they might ask their partners to generate an explanation in response to their question, and the explanation was not used by the questioner to extend ideas beyond it. Members in a constructive-constructive dialogue pattern constructed knowledge independently as a team.

In an active-constructive dialogue pattern, the constructive partner was in charge of generating ideas and the active partner repeated or summarized information heard. An active partner might recognize a knowledge gap and go through the instructional material to locate relevant information, but no new ideas were generated. The purpose of asking a question by an active partner was not to expect another person to reason about explanations as a constructive partner did. They might elicit fact-based short answers from another peer in order to clarify or confirm because they were aware of conflicting or incomplete information. Another type of question raised by an active member was to request instructions for the next step in order to add specific elements in a model as advised (see the example in Table 14). In a nutshell, active students showed that they wanted to know the answer but had no intention to construct it themselves.

A receptive partner in a receptive-constructive dialogue pattern confirmed that the information generated by a constructive partner was registered. They might give comments such as “uh huh,” “okay,” or “all right” to express agreement. A receptive partner might engage in low-level cognitive processes that were originally categorized as active in the ICAP framework; for instance, requesting repetition or spelling. A request of
repetition or spelling was to help with retention in order to write it down or draw it correctly.

**Teacher Prompts.** The second analysis in the third set of analyses addressed research question 3.2: “What prompts from teachers enhance modeling performance?” I transcribed the conversations between group members and their teachers during modeling when Ms. Addison and Mr. Britten provided prompts to scaffold student performance. I first identified three major characteristics of each teacher’s prompts. After that I compared two teachers’ contrasting prompting techniques. Students’ responses to the teacher’s prompts and the follow-up performance were discussed.

**Summary**

The three sets of analyses informed my three sets of research questions by exploring whether and how students coordinate models and evidence, whether and how mechanisms are incorporated into models, and how small groups’ discourse processes mediate their model quality. The first set of analyses manifested the spread of students’ abilities to coordinate models and each piece of evidence as well as the range of interpretation of each piece of evidence. The second set of research questions was answered by analyzing the mechanistic features of models. Models were classified into different types according to their explanations of mechanisms. Based on the first two quantitative analyses that denoted students’ varied modeling performance, the third set of analyses utilized a qualitative method to examine how group discussion dynamics and teacher prompts might have mediated modeling performance. These three sets of analyses provided an integrated vision of how middle school students integrate evidence and mechanisms into models during group work in a real classroom setting.
Chapter 4

Quantitative Results: Model-Evidence Consistency and Mechanisms

This chapter presents quantitative results based on 80 models that students constructed before the benchmark instruction to explain cellular respiration based on the evidence provided in their evidence packets. The first section of this chapter responds to the first set of research questions and evaluated how competent students were in making their models consistent with evidence. The second section responds to the second set of research questions and examines how competent students were in incorporating mechanisms into their models.

A second coder and I coded 40% models together during a training session. After that we independently coded 36% of the data with an inter-rater agreement of 95.4%. Each of us coded the remaining 64% of the data after we resolved the discrepancies, and the inter-rater reliability rate rose to 97.3%. The coding was finalized after the remaining 2.7% of discrepancies were resolved.

Coordination of Models and Evidence

The analyses investigated whether the majority of the models were consistent with most or only a couple of pieces of evidence, the general quality of students’ interpretation of evidence, and alternative ideas that were absent from the evidence.

To What Extent Are Students’ Models Consistent With Evidence?

I analyzed how many models were consistent with relevant evidence and if students recognized the irrelevant evidence and eliminated it from their models. The word consistent refers to any instance in which a piece of evidence relates to a model, regardless of the quality of evidence interpretation presented in the model. As discussed
in Chapter 3, the physical fitness evidence was irrelevant to cellular respiration, and the remaining six pieces of evidence were relevant. The muscle cell contraction evidence included two distinct concepts related to cellular respiration; as long as at least one of the two concepts was consistent with a model, the model was considered relevant to this evidence. Table 15 displays the results of the 80 models collected before the instruction.

**Table 15**

The Percentage of Models That Were Consistent With Different Numbers of Evidence

<table>
<thead>
<tr>
<th>Count of relevant evidence</th>
<th>Frequency of pre-instructional models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ms. Addison (n = 36)</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The results show that 97.5% of the models were consistent with at least one piece of evidence. About half of the models (53.8%) were consistent with at least four pieces of relevant evidence. More models in Ms. Addison’s classes were consistent with at least four pieces of evidence than those in Mr. Britten’s classes.

Within these 80 models, the irrelevant evidence physical fitness was present in only 7 models (8.8%); one was from a model who was in accordance with six pieces of relevant evidence, another with four pieces of relevant evidence, and still another with two pieces of relevant evidence. Models were also inspected for any contents in relation to each of the six pieces of relevant evidence. As presented in Table 16 below, a large portion of models were consistent with the high-carb vs. low-carb evidence and the
*oxygen* evidence. Almost half of the models were consistent with the *muscle cell contraction* evidence, the *carbon dioxide* evidence, and the *cellular organelle* evidence. Only 37.5% of the models were consistent with the *body temperature* evidence.

**Table 16**

*The Percentage of Model-Evidence Consistency for Relevant Evidence*

<table>
<thead>
<tr>
<th>Evidence</th>
<th>Frequency</th>
<th>Percentage (n = 80)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-carb vs. low-carb</td>
<td>75</td>
<td>93.8</td>
</tr>
<tr>
<td>Oxygen</td>
<td>51</td>
<td>63.8</td>
</tr>
<tr>
<td>Carbon dioxide</td>
<td>39</td>
<td>48.8</td>
</tr>
<tr>
<td>Muscle cell contraction</td>
<td>37</td>
<td>46.3</td>
</tr>
<tr>
<td>Cellular organelle</td>
<td>37</td>
<td>46.3</td>
</tr>
<tr>
<td>Body temperature</td>
<td>30</td>
<td>37.5</td>
</tr>
</tbody>
</table>

**How Is the Quality of Evidence Interpretation Represented in Students’ Models?**

Students’ interpretations of evidence were translated into points ranging from 0 (*poor*) to 5 (*excellent*) for each evidence concept in order to determine the quality of evidence interpretation of each concept. Table 17 shows the means of each evidence concept and the percentage breakdown of the points.

With regard to the relevant evidence concepts, students were more competent in interpreting the *high-carb vs. low-carb* evidence and the *oxygen* evidence than in interpreting other evidence. A vast majority (81.3%) of the models expressed that food (point 4) or carbs (point 5) were related to energy generation, and 48.8% of the models stated that oxygen was involved either in exercise (point 4) or in an energy generation process (point 5). This was in stark contrast to the remaining evidence concepts. The frequencies of points 4 and 5 were much lower for each of the remaining concepts, and indeed for each of the remaining concepts more than half of the models did not mention anything concerning them (see the point-0 column for the relevant evidence concepts in
Table 17). Even though the highest frequency was found in the point-0 category for each of these remaining concepts, the second highest frequency was found in the point-5 category (or the point-3 category for muscle contraction). This indicates that although more than half of the models did not mention the concepts other than glucose and oxygen, when they did, their interpretations of the concepts were more likely to achieve the top category of performance in which the concepts were connected to the energy generated.

One notable aspect is students’ interpretation of the cellular organelle evidence. Almost one out of two models (47.2%) were credited with points ranging from 1 to 5 (see Table 17), indicating that the interpretation went down to the cellular level (see the coding scheme in Table 6). The models that were credited with 4 and 5 points for this evidence concept made up 31.3%, and this implies that about one third of the models addressed certain relationships between mitochondria and energy.

Table 17

<table>
<thead>
<tr>
<th>Evidence Concept</th>
<th>Mean</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical fitness(^a)</td>
<td>4.61</td>
<td>6.3</td>
<td>1.3</td>
<td>0.0</td>
<td>1.3</td>
<td>0.0</td>
<td>91.3</td>
</tr>
<tr>
<td>Glucose(^b)</td>
<td>3.71</td>
<td>7.5</td>
<td>7.5</td>
<td>2.5</td>
<td>1.3</td>
<td>47.5</td>
<td>33.8</td>
</tr>
<tr>
<td>Oxygen</td>
<td>2.66</td>
<td>36.3</td>
<td>3.8</td>
<td>3.8</td>
<td>7.5</td>
<td>11.3</td>
<td>37.5</td>
</tr>
<tr>
<td>Muscle inputs</td>
<td>1.21</td>
<td>60.0</td>
<td>11.3</td>
<td>6.3</td>
<td>6.3</td>
<td>2.5</td>
<td>13.8</td>
</tr>
<tr>
<td>Muscle contraction(^c)</td>
<td>0.99</td>
<td>53.8</td>
<td>15.0</td>
<td>8.8</td>
<td>22.5</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Carbon dioxide</td>
<td>2.01</td>
<td>51.3</td>
<td>1.3</td>
<td>1.3</td>
<td>12.5</td>
<td>8.8</td>
<td>25.0</td>
</tr>
<tr>
<td>Cellular organelle</td>
<td>1.79</td>
<td>53.8</td>
<td>7.5</td>
<td>2.5</td>
<td>5.0</td>
<td>1.3</td>
<td>30.0</td>
</tr>
<tr>
<td>Body temperature</td>
<td>1.35</td>
<td>62.5</td>
<td>3.8</td>
<td>1.3</td>
<td>3.8</td>
<td>13.8</td>
<td>13.8</td>
</tr>
</tbody>
</table>

*Note.* Percentages were calculated based on a total number of 80.

\(^a\)The physical fitness evidence was designed as an irrelevant evidence. Models were credited with 5 points if they left out this concept. \(^b\)The evidence concept glucose represents high-carb vs. low-carb. \(^c\)The evidence concept muscle contraction had a point range from 0 to 3.
Results drawn from Tables 16 and 17 indicate that models had higher consistency with the *high-carb vs. low-carb* evidence and the *oxygen* evidence than other relevant evidence, and they also reveal students’ superior interpretation of these two pieces of evidence to others. Perhaps students had more prior knowledge in areas regarding physical processing of food and oxygen, which resonates the findings in next section.

**Additional Concepts Absent From the Evidence**

In addition to concepts in the models that were consistent with the evidence, an extensive list of external concepts not alluded to in any evidence was found in students’ models before instruction. Table 18 shows the external concepts recognized in students’ pre-instruction models.

**Table 18**

*Additional Concepts Absent From the Evidence*

<table>
<thead>
<tr>
<th>Concept</th>
<th>Example</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digestive process breaks down food in the stomach</td>
<td>“Food is broken down in stomach into useful elements”</td>
<td>71.3</td>
</tr>
<tr>
<td>Waste is released from the cells or the body</td>
<td>“Waste leaves the cell”</td>
<td>40.0</td>
</tr>
<tr>
<td>Lungs take in $\text{O}_2$ or release $\text{CO}_2$</td>
<td>“Lungs separate oxygen”</td>
<td>18.8</td>
</tr>
<tr>
<td>Digestive process involving intestine after food is broken down</td>
<td>“Food after digestion goes to intestine and becomes waste”</td>
<td>15.0</td>
</tr>
<tr>
<td>Food or Carbs transform(s) into glucose</td>
<td>“Food is broken into glucose”</td>
<td>13.8</td>
</tr>
<tr>
<td>Stomach acid breaks down food</td>
<td>“Stomach acid breaks down the apple”</td>
<td>11.3</td>
</tr>
<tr>
<td>Nutrients from food go into the body, cells, or mitochondria</td>
<td>“The nutrients are sent to the mitochondria”</td>
<td>11.3</td>
</tr>
<tr>
<td>Blood circulates before energy is generated</td>
<td>“The oxygen and elements from food go into blood stream”</td>
<td>7.5</td>
</tr>
<tr>
<td>Sweat after exercise</td>
<td>“Sweat is an output after exercise”</td>
<td>6.3</td>
</tr>
<tr>
<td>Energy comes from food, or food is energy</td>
<td>“You eat and get energy from food”</td>
<td>5.0</td>
</tr>
<tr>
<td>Concept</td>
<td>Example</td>
<td>Frequency (%)</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Energy comes from carbs, glucose, or sugar, or one of them is energy itself</td>
<td>“Glucose is stored in mitochondria as energy”</td>
<td>3.8</td>
</tr>
<tr>
<td>Exercise creates energy</td>
<td>“You get energy after you exercise”</td>
<td>3.8</td>
</tr>
<tr>
<td>Energy is stored in cells or mitochondria for later use</td>
<td>“The energy stored in mitochondria is for the cell to do its job”</td>
<td>3.8</td>
</tr>
<tr>
<td>Energy goes to mitochondria or cells after it is generated</td>
<td>“Food and oxygen generate energy → mitochondria take in energy”</td>
<td>2.5</td>
</tr>
<tr>
<td>Water is the output</td>
<td>“The burning process produces energy, CO&lt;sub&gt;2&lt;/sub&gt;, and water”</td>
<td>2.5</td>
</tr>
<tr>
<td>Energy comes from burning the calories from food</td>
<td>“From the calories burned, it becomes energy”</td>
<td>2.5</td>
</tr>
<tr>
<td>Too many carbs will become fat</td>
<td>“If you take in too many carbs, the food will become fat”</td>
<td>1.3</td>
</tr>
<tr>
<td>Cells need O&lt;sub&gt;2&lt;/sub&gt; in order to burn the glucose</td>
<td>“After glucose enters the cell, oxygen is needed to burn the glucose”</td>
<td>1.3</td>
</tr>
<tr>
<td>Blood circulates after energy is generated</td>
<td>“Energy is delivered to cells through blood circulation”</td>
<td>1.3</td>
</tr>
<tr>
<td>ATPs are present in the model</td>
<td>“Mitochondria take in sugar and oxygen and convert them into ATPs”</td>
<td>1.3</td>
</tr>
<tr>
<td>Energy comes from nutrients</td>
<td>“Food is broken down as nutrients in the stomach=energy”</td>
<td>1.3</td>
</tr>
<tr>
<td>Energy is the fuel for cells to burn</td>
<td>“The energy in cells is burned as fuel”</td>
<td>1.3</td>
</tr>
</tbody>
</table>

Two recurring types of concept had especially high frequency. The most prevalent external concept (71.3%) was that food was processed in a digestive system after it entered the human body. For example, a statement such as “food is broken down in stomach into useful elements” might be followed by simply an arrow and “energy” or a further explanation of cellular activities that produced energy.

The second most prevalent external concept (40.0%) was the idea that waste was released from cells or a human body after certain activities (e.g., energy generation,
digestion, or exercising). For example, “oxygen is taken in to the mitochondria to help the process and CO₂ is released as waste,” and “the used-up energy is released as CO₂ and human waste.”

There were only 15 models (18.8%) that did not include additional concepts. Seven out of the eight most frequently mentioned concepts were related to the notions of food and digestion.

**Incorporation of Mechanisms in Models**

Models that were consistent with evidence might not always offer a strong mechanistic explanation to account for a scientific phenomenon. How well did students’ models explain how human cells get energy from food and oxygen? Furthermore, what were the types of students’ mechanistic explanations to account for cellular respiration? To answer these questions, I analyzed the mechanisms displayed in students’ models.

**To What Extent Do Students’ Models Incorporate Components of Mechanisms?**

The analysis of mechanisms evaluated a model’s explanation about the processes that involved energy generation. It consisted of four sub-analyses. The first sub-analysis examined if all five necessary entities were involved in energy generation. The second sub-analysis inspected each entity’s engagement in activities in the mechanistic process, regardless of the accuracy of the energy generation process. The third sub-analysis addressed the location of energy generation. The fourth sub-analysis explored the types of mechanistic processes that generated energy. Results are shown in Table 19.

It appears challenging for students to incorporate all necessary entities into models and connect them to the energy generation process. Only 8.8% of the models had all five necessary entities engaged in the mechanisms that produced energy (each of the five entities
earned at least a point value of 3). Although the rate of complete inclusion of entity is low, 68 out of 80 models (85.0%) incorporated clear sequences to explain how the processes unfolded.

Table 19

Sub-Analyses of the Energy Generation Mechanisms

<table>
<thead>
<tr>
<th>Component</th>
<th>Mean</th>
<th>0</th>
<th>3</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completion of entities</td>
<td>0.44</td>
<td>91.3</td>
<td>N/A</td>
<td>8.8</td>
</tr>
<tr>
<td>Entity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy</td>
<td>4.19</td>
<td>16.3</td>
<td>N/A</td>
<td>83.8</td>
</tr>
<tr>
<td>Glucose</td>
<td>3.11</td>
<td>18.8</td>
<td>47.5 (Food)</td>
<td>33.8 (Carbs)</td>
</tr>
<tr>
<td>Oxygen</td>
<td>1.85</td>
<td>61.3</td>
<td>2.5 (Air)</td>
<td>36.3 (O₂)</td>
</tr>
<tr>
<td>Carbon dioxide</td>
<td>1.19</td>
<td>76.3</td>
<td>0.0 (Air)</td>
<td>23.8 (CO₂)</td>
</tr>
<tr>
<td>Heat</td>
<td>0.54</td>
<td>88.8</td>
<td>1.3 (Temperature)</td>
<td>10.0 (Heat)</td>
</tr>
<tr>
<td>Location</td>
<td>1.60</td>
<td>63.8a</td>
<td>7.5 (Cells)</td>
<td>28.8 (Mitochondria)</td>
</tr>
<tr>
<td>Agency</td>
<td>1.79</td>
<td>57.5b</td>
<td>13.8 (Non-cellular)</td>
<td>28.8 (Cellular)</td>
</tr>
</tbody>
</table>

Note. Percentages were calculated based on a total number of 80.

aThe percentage was made up by 40.0% no explanation, 17.5% stomach, and 6.3% others. bThe percentage was made up by 22.5% no explanation, 17.5% digestion, 12.5% others, and 5.0% exercise.

Only models that incorporated the entity energy was further inspected for its strength of mechanisms because the phenomenon that students were requested to explain was how cells obtained energy. A model would not be considered pertinent to this phenomenon in the absence of this crucial entity. As displayed in Table 19, the results showed that among the five entities, 83.8% of the models specified energy as the outcome of the mechanistic process, and on average, energy had the highest mean score. This implies that a great number of students were aware of the intent of their models to explain how energy was produced.

Following energy, glucose had the second highest mean score among the entities that engaged in energy generation. Although this entity was present in most models, more of the models addressed food (47.5%) instead of glucose, sugar, or carbs (33.8%) as the
input of the cellular respiration process. Note that earning a point value of 5 does not suggest that glucose participated in a scientifically correct process of energy generation, but only that glucose was participating in any process of energy generation.

About a third (36.3%) of the models recognized oxygen as one of the inputs to the mechanistic process of energy generation. The frequency of the models that received 0 points (61.3%) for oxygen in the analysis of mechanisms (implying no content or low quality, see Table 19) was much higher than the frequency (36.3%) of the models that received 0 points for oxygen in the analysis of evidence interpretation analysis (implying no content, see Table 17). This indicates that in 25.0% of the models, oxygen was present but was not involved in the energy generation mechanisms. A typical example of this kind of model was that oxygen was breathed in and through the lungs carbon dioxide was breathed out; glucose was taken in and eventually became energy. In this case, oxygen and carbon dioxide were present in a model, but the preexisting knowledge about gaseous exchange through the lungs dominated the reasoning and eliminated their possible activities at the cellular level, and thus did not get involved in the mechanistic process that generated energy.

Carbon dioxide had a pattern similar to oxygen, in which some models mentioned this entity but did not explain that it participated in the activities that generated energy. Models that reflected this pattern might address oxygen and carbon dioxide in a gaseous exchange activity (for example, through the lungs or the mouth) in a separate system that did not link oxygen or carbon dioxide to the cellular respiration process. Models that received 5 points in this analysis (23.8%) addressed carbon dioxide at the end of the mechanism, however this does not necessary imply their scientific correctness. For
example, a model might explain that carbon dioxide was released from the stomach after energy was being formed there.

Only a small number of models specified heat as the byproduct of the energy generation process. It had the least frequency (10%) of the five entities.

The third sub-analysis was the location of energy generation. Table 19 shows that a total of 63.8% of the models earned 0 points for this category; 40.0% of them did not provide any explanation regarding the location of cellular respiration, and if they indicated the location, 17.5% of them incorrectly identified the stomach and 6.3% identified elsewhere. Nevertheless, for 28.8% of the models, mitochondria were labeled as the place where energy was generated and in 7.5% of models this distinction was nearly achieved by saying that this happened in cells. This suggests that, combining those who earned 3 points and 5 points in the sub-analysis of location, a total of 36.3% of the models recognized cells’ prominent role in cellular respiration before students received the benchmark instruction.

The fourth sub-analysis scrutinized the agency by which energy was produced. The results showed that 77.5% of the models incorporated mechanisms. More reports follow below.

*What Models Do Students Develop to Explain How Energy is Generated?*

Four major types of pre-instructional models were found in students’ explanations about how energy was generated. The four types were (1) the cellular transforming model, (2) the digestive model, (3) the non-cellular burning model, and (4) the exercise model. The spread of model types is displayed in the row of agency and the note in Table 19.
The Cellular Transformation Model. Cellular transformation models explained that entities went through a process in the cells and turned into energy, using terminologies such as “transformed,” “converted into,” “burned,” “processed,” “mixed and became,” “combined together,” and “turned into.” Figure 2 shows a cellular transformation model. This model addressed in the fourth step that “The oxygen and the glucose enters the mitochondria in the cell to form energy”, which was followed by the fifth step “Once the energy is formed it enters the muscle.”

Figure 2

A Cellular Transformation Model

This type of model scored a point value of 5 and was the most prevalent one (28.8%) among the major four types. In addition, almost all of the 23 models of this type
labeled mitochondria in cells as where the process took place. The participating students had learned cells and cell organelles in the same semester before this module was enacted. But they were not necessarily equipped with molecular or atomic knowledge in the 7th grade; therefore, it was considered scientifically correct to bottom out their models by describing cellular activities using terminologies quoted above. Bottoming out refers to the lowest level of mechanism from which a model stops describing further details (Machamer et al., 2000). The results show that more than a quarter of the students were competent in creating a preliminary mechanistic model that bottomed out at the cellular level prior to more specific instruction on cellular respiration.

**The Digestive Model.** Digestive models explained that energy was generated in digestive organs during a digestive process. Figure 3 instantiates a digestive model in which energy was produced in the stomach after food was digested. Oxygen and carbon dioxide did not engage in the energy generation process in spite of their presence.

**Figure 3**

*A Digestive Model*
This type made up 17.5% of the models (see Table 19 in the note). As shown in Table 17, 71.3% of the models depicted a digestive process, a very popular preexisting concept. This implies that although most of the models included this detail, only a small number of models considered digestion the agency for energy generation during cellular respiration. For those who created this type of digestive model, they might have been preoccupied by the personal experience that the body felt energized after taking in food, overlooking that the point of this model was to explain exactly how so the body got energized after food ingestion and the reason why almost all organisms needed to respire.

**The Non-Cellular Burning Model.** There were 11 models (13.8%)—all in the form of flowcharts—showing that the entities went through a burning process and produced energy, but they did not relate their explanations to the cells (see the 3-point column in Table 19).

Figure 4 presents an example of a non-cellular burning model. The text below the flowchart wrote, “After [food] goes through the digestive system it burns into energy.” In the model, food participated in the energy generation process after it was digested in the digestive system and was burned into energy, but the location was not specified.

All of the non-cellular burning models were found in the classes that received Version 2 packet, where three false example models were handed to students for correction based on evidence before they constructed their pre-instructional models. This type of model was not found in classes that received the Version 1 packet.
In the three example models, the “burning process” was framed as the agency for energy generation. Three models said that this process occurred inside the stomach. This suggests that the students who created them might have either mistaken the burning process as an alternative way to describe digestion or that they might have considered the burning process as an unfathomable mechanism that generated energy in the stomach, distinct from digestion, but did not know how to further develop the explanation. The rest of the 8 models merely designated “burning process” as the intermediate stage between input and output entities and did not specify where this was happening. It seems that students understood that their models were to explain energy generation, and they recognized the burning process as the agency of energy generation, if only they thought about cells and cellular organelles for cellular respiration.
The Exercise Model. The fourth major type consisted of 4 models (5.0%) explaining that energy was produced through exercise (see the note of Table 19). Students who created exercise models might have been preoccupied by personal experiences or alternative concepts. All of the exercise models did not specify where energy was produced.

As displayed in Figure 5, this exercise model depicted that food was transformed in a bag with liquid (referring to the stomach), then a running person was “Burning,” and finally energy and waste were the outcomes. Note that energy was produced after exercising.

Figure 5

An Exercise Model

Besides the four main types of models, there was also a variety of miscellaneous types of models, each of which only showed up once or twice. They made up 12.5% of the models in total. To name a few, one model described that energy was generated in
blood circulation, another one depicted that energy was generated during both digestion and breathing, and two models indicated that food became energy directly after it entered human bodies.

Finally, there were 19 models (22.5%) that did not give any remarks about how energy was produced, although they might present some illustration or narratives; for example, 9 of them portrayed additional information that stemmed from personal knowledge. As shown in Table 19 and its note, the frequency of models making no explanation about energy generation was only outnumbered by the frequency of the cellular transformation models. Among these 18 models that did not explain the mechanisms, 6 were from Ms. Addison’s classes, and the other 12 were from Mr. Britten’s classes. On the other hand, among the 22 cellular transformation models, 19 were from Ms. Addison’s classes and 3 were from Mr. Britten’s classes.

The full score for a model to award was 78. The overall mean score for the pre-instruction models in Ms. Addison’s classes was 47.06, and the mean score in Mr. Britten’s classes was 21.95. Perhaps such distinction of modeling performances could be explained by teachers’ scaffolding styles. In the next chapter I will analyze the qualitative differences between the two teachers’ prompts when students talked to them during pre-instruction modeling.

Summary

In response to the first set of research questions, the results showed that students were generally competent in making their models consistent with relevant evidence and eliminating irrelevant evidence from the models before they received the benchmark instruction of cellular respiration. Students also exhibited a tendency to add a wide range
of information not alluded to in the evidence to support their explanations. In terms of the quality of evidence interpretation represented in models, students generally performed better in making sense of the evidence associated with the entities that were involved in the process before energy was produced than the entities after energy was produced.

The second set of research questions addressed students’ ability to make a mechanistic model before instruction. Although it appeared challenging for students to incorporate all necessary entities into the energy generation process, about a quarter of the models presented preliminary mechanisms in which entities interacted with each other and generated energy at the cellular level. In addition, about a quarter of the models indicated mitochondria as the location where cellular respiration took place. This implies that students were more capable of explaining how and where energy was generated than incorporating all necessary entities into the mechanistic process.

Two thirds of the models provided an explanation of how energy was generated during cellular respiration. Among them, four major types of models were tracked down. In accordance with the frequency from high to low, they were (1) the cellular transformation model, which explained that entities transformed into energy in cells; (2) the digestive model, which illustrated that energy was generated through digestion in digestive organs; (3) the non-cellular burning model, which presented a simple flowchart in which entities went through a burning process and resulted in energy and byproducts; (4) the exercise model, which stated that energy was produced after exercising. All results were drawn from the models students built before they received a conceptual instruction of cellular respiration.
Chapter 5

Qualitative Results: The Processes of Group Discussions and Teacher Prompts

The third set of research questions of this study concerned the processes that mediate modeling performance. This is divided into two subordinate research questions 3.1: “How does the discourse of groups that vary in modeling performance differ?” and 3.2: “What prompts from teachers enhance modeling performance?” I address these questions through analyzing four small-group discussions and the teachers’ interactions with these groups during the modeling activity that took place before the instruction.

The first two groups (A1 and A2)—from Ms. Addison’s classes—represented high and medium modeling performances among all participants. The latter two groups (B1 and B2)—from Mr. Britten’s classes—represented medium and low modeling performances. Group A1 was one of Ms. Addison’s top-performing groups, and Group A2 was among those who performed at the bottom level of her classes, even though Group A2 outperformed most groups in Mr. Britten’s classes. Group B1 was Mr. Britten’s top-performing group, and Group B2 was his lowest-performing group.

Modeling performance was determined by the consistency of models and evidence as well as by the incorporation of mechanisms in students’ group models as described in the methodology chapter. I first overview the coding schemes employed for group discussion analyses. After that I present conversation excerpts from Groups A1 and A2, and I report on how Ms. Addison provided prompts. Then I present conversation excerpts from Groups B1 and B2, followed by the display of prompts offered by Mr. Britten.
Review of Coding Schemes

Discourse Operations

The contents of each episode of group discussion were analyzed in terms of seven discourse operations (see Table 13 in Chapter 3 for the full explication). Off-topic referred to conversations irrelevant to the inquiry task or the topic of cellular respiration. Regulation and logistics concerned discourse about time and task management. Model expression referred to conversations that considered a particular visualization to represent certain ideas. Model criteria referred to the reflection on the use of certain criteria to create a good model. Intra-evidence interpretation referred to contemplating a single piece of evidence exclusively in one statement. Inter-evidence integration referred to considering multiple pieces of evidence and integrating the interpretations of evidence into a cohesive statement. Mechanistic explanation was an account that specified entities and used a verb to describe how activities produced changes from the on-set to the termination conditions. An episode might comprise more than one discourse operation.

Dialogue Patterns

Group members’ dialectical interactions when engaged in conceptual reasoning were classified into four dialogue patterns (see Table 14 in Chapter 3 for full details). In an interactive dialogue pattern, group members elaborated upon and challenged each other’s ideas and co-constructed knowledge. Peers in a constructive-destructive dialogue pattern worked together as a team, but each individual person undertook an independent area in the task without corroborating and extending each other’s ideas. Their contributions might take the form of requesting others to construct an explanation or reminding the team of their progress. In an active-destructive dialogue pattern, the
constructive partner was in charge of generating ideas, and the active partner sometimes repeated or summarized the information heard and on other occasions asked questions to clarify or confirm the information. An active partner might recognize a knowledge gap and locate the sought information in the instructional materials. Next step might be asked. Finally, in a receptive-constructive dialogue pattern, a receptive partner confirmed that the information generated by a constructive partner was retained in the moment. A receptive partner might request repetition or the spelling of certain words, but next step was not sought.

The High-Performing Group: A1

Group A1 was a representative of those who created high quality models in Ms. Addison’s classes before the benchmark lecture. A majority of Ms. Addison’s pre-lecture groups achieved high performance; they labeled the entities in the models, indicated a sequence of processes, and explained how cellular respiration happened at the cellular level. These achievements might develop during group discussions. I first present Group A1’s model, and then I analyze this pair’s communication during their group modeling.

Group A1’s Model

Group A1’s models provided a mechanistic explanation that involved entities and activities in a cohesive system. Oxygen and food (specified as “high carb”) entered the body from the mouth. From this set-up condition, the entities proceeded with intermediate activities. They arrived at a cell and were burned into energy in a mitochondrion. Besides energy, this process also produced carbon dioxide and heat, which then departed the cell. The heat from the cell made the body temperature rise. After the “finished product,” energy, went into the arm muscles, it enabled the muscles to
contract. Muscle cell contraction was the termination condition, and the contraction would intensify if there was a higher intake of both oxygen and glucose in the set-up condition. The model included special notes that oxygen intake, carbon dioxide “outtake,” and body temperature increased during exercise.

**Figure 6**

*Group A1’s Model (A High-Performing Group)*

This model was consistent with the evidence concerning oxygen, glucose, carbon dioxide, muscle cell contraction, cellular organelle, and body temperature. It described the activities that linked relevant entities together from the set-up condition through the intermediate activities taking place at certain locations, and eventually led to the final
products in the final stage. The blue ink indicated the revised parts. All construction and revision were completed before the benchmark lecture. This model had a full total score of 78, with 38 for evidence interpretation and 40 for mechanistic explanation.

**Group A1’s Discussions**

Group A1 was composed of two partners, S1 and S2. In the analyses that follow, I highlight episodes of their conversation on Day 1 and Day 2 of modeling, and then I discuss how the model was different before and after they received scaffolding from Ms. Addison.

**Day 1 of Modeling.** This group wasted no time. Right after they briefly introduced themselves in the recording, S1 immediately delved into *intra-evidence interpretation* for each piece of evidence.

1  S1: So now, well, we have to go through these first.
2  Um the more sugar and oxygen, the more, the more the muscles contract when you work out.
3  The more, the more, uh, the more you exercise, the more CO$_2$ you exhale.
4  The more you exercise, the um, your body temperature goes up.
5  And the more you exercise, the more oxygen you inhale ‘cause it goes in.
6  And also, um, if you have a high carb meal or diet it helps your performance.
7  And organelles can affect your final…
8  S2: That’s wrong.
9  S1: I know…So…
10 S2: What about looking at it in this way…Let’s draw a cell.
11 S1: Well, well we’re showing how um the um how taking in stuff and output.
12  I think first, well, can you draw first of all?
13  S2: I’ll try.
14  S1: Okay, we can go through the text again.
15  Well, I think we should first draw like one body, draw a body. (Drawing)
16  S2: Question—a body?
17  S1: A body. A big one. (Drawing)
18  S1, S2: (Discussing drawing a mouth)
19  S1:  Now write like...feels like...feels CO$_2$ coming out and oxygen going in [the mouth].
20  S1, S2: (Borrowing a pencil)
21  S2:  ¡Andale! ¡Andale! [“Hurry up” in Spanish]
22  S1:  I am drawing. (Long pause)
23  S2:  It doesn’t need to be perfect.
24  S2:  I know, but that wasn’t, that wasn’t even a body.
25  S1:  Now draw the line and then write “CO$_2$ being exhaled” and then write like— (Interrupted by Ms. Addison)

Group A1 was off to a good start, as their intra-evidence interpretations of each evidence recognized the oxygen and carbon dioxide as well as their activities, and they did not dawdle on aesthetic details. From lines 1 to 6, S1 expressed his intra-interpretation, in which each evidence was interpreted independently without thinking about other evidence. In lines 7, 8, and 9, the students recognized a knowledge gap—they could not make sense of the cellular organelle evidence. A suggestion was made by S2 in line 10 to move on and draw a cell, which is the appropriate scale to model how cellular respiration happens. However, in the constructive-constructive dialogue pattern of this episode, opinions were only generated but not extended by the other person. S1 dominated the conversation in this group, and he ignored S2’s suggestion and asked her to draw a body (the organism level) to start with. They spent only about one minute on drawing (model expression discourse operation) as they knew “It doesn’t have to be perfect.” They pressed more on the expression of the activities of the entities through the mouth as shown in lines 11, 19, and 25. These three lines indicated a primary operation of inter-evidence integration when S1 thought about oxygen going in to the mouth and through there carbon dioxide being exhaled. S1 took multiple evidence into consideration and did not think about the causal relationship between these two entities. For example, he did
not think about changing one entity’s rate or intensity might bring out certain changes in its interaction with another entity in a systematic manner.

In the next episode, the pair continued their conversation after Ms. Addison stopped by and told them to think about the next activities that involved both oxygen and food. S1 apparently read aloud what he was writing, and S2 challenged his ideas (“We need to draw the muscles? Why?”), monitored the performance (referring to a typo “Xerisr?”), and gave reminders (“You gotta get a title for this thing. Are you almost done?”). However, S1 did not respond, nor did he appear to take his partner’s ideas into consideration. This pattern of constructive-constructive discussion went on for about six minutes until a brief brainstorming episode, as presented below. S1 hesitated about what to put down on their model for the food, and S2 joined in the reasoning.

1  S2:  Good! We are done?
2  S1:  No. (Long pause)
3       Um…with that? No, the food in…that’s a good point. (Long pause)
4       Food—
5  S2:  We got it drawing, we got it drawing.
6  S1:  No, not with drawing.
7  S2:  Um food—
8  S1:  Um food. (Laughing)
9  S2:  Food—
10  S1: Um, food with high carbs—
11  S2: With high carbs um—
12  S1: High carbs…um…gives you more energy!
13  S2: And helps your performance!
14  S1: Yeah

This episode started with S2’s question about if they were finished and ended with an interactive dialogue pattern in which they worked together and extended upon each other’s ideas to generate an additional part of the model. S2 started the intra-evidence
interpretation in line 7, and S1 repeated her word, also starting his sentence with “food.” The same situation reoccurred in the subsequent turns. S1 stretched the initial idea in line 4 a little bit along the course and eventually completed it in line 12. S2 built an additional idea in line 13 based on S1’s sentence in line 12 and arrived at a consensual conclusion, which made this episode interactive. In their model, the final product read, “Food—with high carbs gives more energy and helps your performance.” This was a good addition to their earlier writing, “Foods enter the mouth.”

At the end of Day 1, Group A1 had a conversation with Ms. Addison. This conversation revealed that they had a mental model of external respiration to account for cellular respiration in which gaseous exchanges occurred in the mouth and lungs. Although S1 recognized mitochondria as the key point in the cellular organelle evidence as shown in this expression, “Okay, now I don’t remember how cells and mitochondria…,” the external respiration framework made it hard to fit this cellular organelle evidence into their explanation. When encouraged by the teacher to “tie that in,” S1 first said that “all of this happened because of mitochondria,” and later changed it to “the mitochondria performed all of that,” but eventually “[all these things] they just go back into the pod [mouth].” This final response indicates that even though their model denoted the entities and activities, it was not organized at the cellular level at all. This made Ms. Addison raise a question: “If it just stays in your mouth, would anything else, would you get cellular?” Students did not answer, and the recording was turned off at this point.

**Day 2 of Modeling.** The next day students were to revise their models. S1 decided to “go over it” and started from the first evidence, asking S2 to “Make a note that
Roger and Eric are the same size because…if they were like different sizes and shapes and stuff, the tests might not be so accurate.” This suggests that S1 understood the concept of controlled variables. An active-constructive dialogue pattern ran through most of their conversation on this day as S1 dictated his opinions and S2 transcribed S1’s ideas on their model with some clarification questions asked. Right after S2 wrote the note about physical condition, S1 double checked to see if their model included all evidence; they had all but the cellular organelle evidence.

1 S1: So, now about the cell organelles. Um…so, after the food—(writing) after the—(writing) food and oxygen enter—(writing) they go in the cells and inside that, that in the mitochondria…
2 (S2 interrupts him and the pair discuss some logistics issue.)
3 S1: Anyway, um, yeah. So, after food and oxygen enter—(writing)
4 S2: They go to and perform what?
5 S1: They go to and are performed in the cells.
6 And inside that and the mitochondria is where they’re burned, where they are burned—(writing)
7 S2: Where—where they are—(writing)
8 S1 & S2: Where they are burned—(writing)
9 S1: And they make the temperature of the body rise. (Pause)
10 So, yeah, that’s important”

This was the first time Group A1 showed an understanding of the cellular organelle evidence. After S1 figured out that mitochondria were the place where food and oxygen were burned, he demonstrated an initial sign of mechanistic explanation, in which entities and activities were organized by processes to describe how changes occurred from the on-set to the termination conditions.

In most episodes, S2 did not join in the reasoning with S1; instead she played a supportive role assisting task tracking, which made some episodes of their discussion a constructive-constructive pattern. The conversation below is one of these instances.
S1: So, we have their physical fitness condition.
Um, their muscle cell contraction.
Carbon dioxide that's taken in.
Their body temperature that comes from the food and, that comes from
the oxygen and food that entered into the mitochondria that makes it
burn off to create a rising temperature.
The oxygen that's taken in...I'm sorry, the carbon dioxide is given off.
S2: (Interrupting) Right.
S1: (Continued from line 5) And the oxygen comes in.
S2: (Interrupting) Did we do the oxygen already?
S1: (Continued from line 7) That a high carb meal and diet can help you
perform better.
S2: Did we do oxygen already?
S1: Yeah, we did.
S2: Where is it? (Pause) Ok, distance is that.
S1: And then the distance is, remember, with the high carb meal.
(S1 and S2 are arguing what exactly they have written down in their
model.)
S1: With high carbs gives more energy—
S2: And helps your performance.

In this conversation, S1 described that oxygen and food entered the mitochondria, where
they were burned off to make the body temperature rise, and carbon dioxide was let off.
In addition, food rich in carbs gave more energy. Although this statement mentioned that
high-carb food gave more energy, it did not address how energy was generated in the
cells to make muscle cell contraction possible. Therefore, S1’s statement failed to fully
account for cellular respiration. Nevertheless, S1’s words in line 4 (“the oxygen and food
that entered into the mitochondria that makes it burn off to create a rising temperature.
…the carbon dioxide is given off”) tapped into mechanistic explanation again. This time
he incorporated more entities and activities into his account than his statement in the
previous episode about the processes. Later in the group’s conversation with the teacher,
it became clear that S1’s mechanistic explanation was to describe a model of metabolism
where the term “burning off” referred to burning off calories.
**Revision of the Model.** The group’s model ended up demonstrating an explanation that incorporated mechanisms of cellular respiration, which was slightly different from the conversation contents between S1 and S2 as displayed in previous sections. For example, the physical fitness evidence was removed after it was challenged by Ms. Addison asking, “Were we actually using the study to compare them, or were we using the study to see how they behave?” S1 quickly said, “To see how they behave,” and S2 erased it.

Oxygen and carbon dioxide were originally labeled by Group A1 as an input and an output through the mouth; after being questioned by Ms. Addison, S1 still thought that the activities enacted by these two entities took place in the lungs. This indicated that when it came to gaseous exchange, it was hard for the students to resist making the association with the previously learned framework of respiration occurring in the lungs.

Group A1 was able to identify mitochondrion as the key organelle in cellular respiration because they pointed out that mitochondrion was the only organelle in the chart that had a different density rate between the two boys. But it was challenging for them to find a logical way to blend it in the model. After being prompted by Ms. Addison to think about the word “cellular,” S1 considered mitochondrion a device to accelerate metabolism and burn off food and oxygen quicker. During another follow-up conversation with Ms. Addison’s prompts, he finally proposed that the whole process of cellular respiration happened in the mitochondria after he eliminated the locations of mouth and lungs.

The final product energy had not been present in the model before Group A1 added the blue-ink portions. The finished product in the previous draft was “The more
sugar and oxygen taken in, the more the muscle contracts when working out,” read aloud by S1 when he was writing it down. Ms. Addison asked, “What does the sugar and oxygen give you?” Without hesitation, “Energy,” said S1. Afterwards the group took Ms. Addison’s suggestion and drew a mitochondrion in a cell to support their original written part. Except for the false response to the question regarding the location of gaseous exchange during cellular respiration, the students were able to provide correct answers to Ms. Addison’s questions about the entities and activities in the on-set and termination conditions, respectively. Their final model achieved high quality that incorporated mechanisms and was consistent with evidence. This suggests that Group A1 was able to construct a mechanistic model based on an array of evidence in a model-based inquiry setting only if some necessary scaffolding was provided to stretch their competence in formulating a cohesive explanation at the cellular level.

At the very end of their modeling S1 proposed the title “Cell Respiration” for their model. But it was rejected by S2, and she wrote, “Is food Needed for our Body?” After this, a researcher asked S2 what their model explained. S2 responded:

How Eric and Roger basically get more energy and oxygen. They’re the same size, and they had the same physical condition, so our model is showing like the muscle contractions and how the food and CO₂ have to do with the mitochondria. The title S2 made in addition to her description of what their model explained indicate that S2 did not fully grasp the mechanisms of cellular respiration, despite making an initial suggestion to start their model from a cell.
Summary

Group A1 displayed their high competence in grasping the key point of each piece of evidence and made their model consistent with relevant evidence. The challenge they encountered was to fit mitochondria in their model. A possible reason for this challenge was because they held a conceptual framework that was closer to external respiration at the human-body level. This concept that gaseous exchange took place at the mouth and lungs at first hindered their potential competence in creating a mechanistic model to account for cellular respiration. They followed the advice given by the teacher to fine tune their reasoning—at first about the activities that food and oxygen engaged in, later about processes that took place at the cellular level, and finally about how and where energy was generated in the cells—and eventually revised their model drastically to show cellular processes and incorporate the mechanisms of cellular respiration.

A big portion of Group A1’s discussion fell into the constructive-constructive dialogue pattern in which S1 was leading their group through model expression, intra-evidence interpretation, inter-evidence integration, and mechanistic explanation discourse operations; S2 often engaged in metacognitive thinking by monitoring their progress to help actualize task completion. S2 was never in a role that initiated new reasoning, and when she challenged S1’s ideas, she was mostly ignored. Although S2 often reminded S1 of the knowledge gaps they encountered and the procedure of their tasks, she did not acknowledge any criteria of modeling that were displayed on the blackboard. The second most common dialogue pattern in this group was active-constructive, which was followed by a couple of interactive episodes. In the active-constructive episodes, S2 repeated what she heard from S1 and asked for confirmation at
times during her transcribing. They had two *interactive* episodes. In one they cogenerated their interpretation of the *high-carb vs. low-carb* evidence, and in the other one they confirmed that they had earlier reached a consensus about this evidence.

As a high-performing group, Group A1 was not characterized consistently by *interactive* conversation; rather their good performance was ascribed to the teacher’s scaffolding (discussed later) as well as outstanding discourse operations of *intra-evidence interpretation, inter-evidence integration, and mechanistic explanation*, which were mostly constructed by only one dominating student through constructive-constructive episodes of discourse.

**The Medium-Performing Group: A2**

Group A2 exemplified the minority of groups who produced only intermediate-level models in Ms. Addison’s classes before the benchmark lecture. A description of this group’s model is followed by analyses of this group’s discussion.

**Group A2’s Model**

Group A2 had two members, S3 and S4. Group A2’s model was moderately consistent with all relevant evidence and had a very clear sequence of steps, but the steps were mutually independent from each other. The model started from the statement that high-carb food made one run faster. Next the group wrote that oxygen inhalation and carbon dioxide exhalation increased during exercise. The third step described that the rate of muscle cell contraction was based on the intake of oxygen and glucose, but the description did not specify the relationship between the entities and the rate of muscle contraction. After this, the model presented a mitochondrion and denoted that cellular respiration occurred in mitochondria, citing study results from the table in the *cellular*
organelle evidence. Finally, the last step showed a thermometer and wrote, “Body temperature affected your energy and breathing rate.”

This model labeled steps and clearly stated that cellular respiration took place in mitochondria. However, the activities or statements in one step did not derive from its previous step, nor did it affect its following one. Moreover, the steps were not integrated into a cohesive process. Overall, the model did not explain how cellular respiration took place in mitochondria and the related mechanisms. This model earned a score of 29 for evidence interpretation and 5 for mechanistic explanation, which summed to 34 as the total score.

**Figure 7**

*Group A2’s Model (A Medium-Performing Group)*
Group A2’s Discussions

Group A2 had a different dynamic from their peers in the previous group. The pair exhibited a shift in how they interacted across the two days. On the first day of modeling, they spent a great deal of time discussing the details of drawing in an active-constructive dialogue pattern, in which S3 played the constructive role and S4 did what he was directed to do and asked drawing related questions. On the following day, after S4 was asked by S3 to participate in discussion, their conversation became interactive. They determined the relevance of all evidence to their model and double checked their model against the “criteria for good models” on the blackboard.

Day 1 of Modeling. Group A2’s discussion started as a dispute about logistics right after the modeling activity began. S3 wanted to persuade S4 to cooperate with her, and S4 said he was not a good drawer and would be “just agreeing with everything you say.” S3 moved on nevertheless and completed the first step of their model by herself. She wrote, “The more carbs you eat the faster you run.” S4 did not participate in reasoning or drawing at all and complained, “This is so boring,” when S3 was working alone.

This situation went on until another minor dispute arose. S3 expressed her vexation, “It’s not just me doing the work. We’re both doing the work!” “I can’t draw,” said S4. “So? Try! I see you drawing all the time on the back of your hand,” S3 rejoined. After S3 gave him a little bit of push, S4 agreed to get involved in drawing. They started a lengthy conversation in which S3 had to demonstrate each drawing action for S4 to follow accordingly.
Okay, so the more you um, maybe draw running, so the more you exercise, the more…

Drawing what running?

No—people. Like leg running.

Okay—running.

Yeah, so like we can put the more you exercise, the more you exercise the more CO\textsubscript{2} and oxygen you inhale the more CO\textsubscript{2} you exhale.

Okay, there is one problem: I can’t draw a body running.

Try! Try!

What am I supposed to do?

Draw legs.

Is it going to be sticks?

Draw legs like this. Draw—one—leg—like this. (Drawing)

Okay, I know how to draw a leg.

Then draw a foot, make a sneaker. (Drawing)

Uh—huh, that’s right!

(Drawing) Like—that.

Now what?

Now draw another foot for two legs.

But next to it...(Drawing) Like—this.

They’re supposed to be running.

I know.

In this active-constructive episode S3 was in charge of intra-evidence interpretation (lines 1 and 5), as well as the visual design of the model (model expression). She did not explain how she came up with her interpretation. S4 asked for clarification and instruction and did exactly what he was told. He was barely involved; he did not ask questions, challenged S3’s ideas, or participated in reasoning.

After they finished the second step of the model, Ms. Addison stopped by. Upon receiving Ms. Addison’s question, “Why do we need those legs?” S3 explained her rationales, “To, um, make it seem like they’re running so, so like someone’s running so we can put ‘The more you exercise the more oxygen you inhale, and the more CO\textsubscript{2} you exhale.’” Ms. Addison prompted them, “Where are you going to put ‘the more oxygen?’ And where is that oxygen coming from? The legs?” Ms. Addison intended to make them
think about entities and activities, but S3 did not reflect on the implications of the teacher’s questions and missed the opportunity to leverage the direction pointed out by the teacher into an alternative way to structure their model. S4 was again only concerned about the superficial details of the model right after the teacher left: “Yo yo yo, I’m about to fix something. Wouldn’t the front leg be bent? And the back leg be straight? That would make it look more like he’s running.” Although S3 responded, “Nah—it’s okay,” S4 fixed the legs as shown in their model. S3, on the other hand, wrote down exactly what she told Ms. Addison she was about to write and drew two mouths inhaling O₂ and exhaling CO₂ with arrows. This episode instantiates a constructive-constructive dialogue pattern. S3 engaged in intra-evidence interpretation, considering each evidence one at a time, while S4 only offered ideas and comments on the extraneous details of model expression—how the model looked on the paper. They did not build new ideas from or extend each other’s thoughts.

As S3 was about to head onto the third step of the model, S4 finished his drawing of the sneakers and started shuffling his playing cards and counting down the time left. Amid the chatting between S4 and other students, S3 was completely absorbed in intra-evidence interpretation and model construction until the end of the recording.

1  S3: Okay, so now we can put ‘The more oxygen and sugar the muscle cell takes in the more the muscle cell will contract.’
2  S4: Okay.
3  S3: Contract is more like…Oh Ms. Addison?
4  S4: (Laughing) One minute!
5  S3: Please don’t worry about no bells for lunch and worry about what we’re doing in the project!
6  S4: Two, one, okay, one more minute. (Under breath) Yes!
7  S3: Ms. Addison! I have a question.
8  S4: Is it lunch related?
9  (A couple of off-topic exchanges)
S3 interpreted the muscle cell contraction evidence and tried to think about how to construct the third step of the model; meanwhile S4 displayed minimal engagement in line 13 asking for repetition. Group A2’s discussion ended in the receptive-constructive dialogue pattern, which was the epitome of the first day of modeling.

**Day 2 of Modeling.** The second day began as S3 checked the first three steps in the model. She then reasoned about the fourth step, “Okay, so now we put, um, mitochondria, we have read…um…because it shows Roger has more mitochondria than Eric…the mitochondria is the place where cell respiration happens.” She drew a mitochondrion and put down her intra-evidence interpretation on the model. S4’s reluctance to help upon S3’s request finally created a stir. A portion of a long yet mild-toned dispute is displayed below.

| 1  | S3:  | It’s not my job to give you something to do. |
| 2  | S4:  | Well, let’s see, you’re the one drawing, you’re the one giving all the directions, and you’re the one writing. |
| 3  | S3:  | No, I’m not! I’m trying to do something here. I’m talking. I’m waiting for you to give me some of your ideas but you’re just sitting there drawing [referring to drawing something else than the model]. |
| 4  | S4:  | I don’t have any ideas. |
| 5  | S3:  | Try to think though, at least try to help. I mean to think of ideas in which you may think that, um, how we should discuss it. |
| 6  | S4:  | All right. That works for me. |
This discussion was a turning point, after which their discussion took an upturn and became mostly interactive.

The shift in their interaction was abrupt. They started to give feedback in response to one another’s ideas when they filled out the last two questions in the study worksheet. They evaluated if evidence was relevant to their model and explained the reason why. The following excerpt shows how they relied on the evidence to resolve a disagreement.

1  S3:  I think it’s in the mitochondria…that’s where, um, cellular respiration happens.
2   What do you think?
3  S4:  I think that cell respiration happens in the nucleus because that’s where it happens.
4  Cell respiration is an important thing for the cells and why wouldn’t it happen in the nucleus?
5  S3:  But, however, when we look at the, um the data that we have, it looks like it’s happening inside the nucleus, but when you look at the chart, it says that um, shows that Eric, I mean Eric has less mitochondria than Roger, which tells Roger um has more mitochondria, if you read the data in the packet.
6  S4:  Uh huh.
7  S3:  So, okay, it is mitochondria, now?
8  S4:  Uh huh.

In this interactive discussion S3 offered her opinion, and S4 approached the same topic from an alternative viewpoint that important cellular processes were supposed to take place in the nucleus. S3 rebutted by giving reference to the cellular organelle evidence. She used her intra-evidence interpretation about this evidence to justify her opinion. On the study worksheet they wrote “Yes” (that this evidence was relevant), and the reason they gave was “We think that the mitochondria is where cell respiration takes place.”

Later when they tackled the carbon dioxide evidence, S3 decided that carbon dioxide was not relevant to the cellular respiration, “Because that’s not what it is taking
in during the process; that is what’s produced after the process is done...that’s what’s produced.” This rationale revealed an interesting belief of hers that only the entities in the set-up condition had relevancy, and the end products in the termination condition were not considered part of the mechanistic process. S3 repeated her statement after she was requested by S4 to clarify, and this made this short conversation active-constructive.

In another instantiated interactive discussion shown below, the pair had different opinions about the relevance of the body temperature evidence. This time it was S4 who offered justification for his stance.

1 S3: Okay, body temperature.
2 We said “No” because the body temperature really doesn’t have any impact, um...
3 S4: Sure it does!
4 You cold, you aren’t going to run too fast.
5 If you’re hot, you aren’t going to run too fast.
6 But if you’re, like, warm, then you’re going to run pretty fast.
7 And if you’re hot and you drink water, it’ll make you run faster.
8 So, yeah, that got some effects.
9 S3: Well...
10 S4: Body temperature affects your energy rates.
11 Like, if you’re really cold, you’re losing a lot of energy, same with if you’re too hot.
12 S3: So, body temperature, so, changes to “Yes.” (Writing)
13 S4: (Interrupting) Uh huh.
14 S3: (Continued from line 12) And we’ll say body temperature why?
15 S4: Because body temperature affects your energy cause an example...and then put an example.
16 It also affects your breathing rate too!
17 S3: (Writing) It—affects—your—energy—and—breathing—rates.
18 S4: And then put examples, and then put examples.
19 S3: We don’t need to put an example, just to explain why.
20 Um, and then since we think it’s “Yes,” we have to put that on the poster.

S3’s reason for thinking that the body temperature evidence was irrelevant was based on a similar belief that she held toward the carbon dioxide evidence; she believed that the
end product in the termination condition was not relevant. Her decision was challenged by S4’s opinion which was derived from personal experiences. S4 was convinced. In line 20 she said because it was relevant, their model had to include this evidence. This statement tapped into the *modeling criteria* discourse operation and demonstrated her epistemic understanding that model and evidence had to be consistent.

When Ms. Addison stopped by for the final time, she questioned this group about the presence of the thermometer in their model. Afterwards the pair went through the list of the “criteria for good models” on the blackboard and checked if their model needed any final revision. They affirmed that their model met all of the criteria. No revision was performed, and the feet and the thermometer stayed in the model. At the end, another student stopped over, “How’s your model doing?” S4 acclaimed, “Ooo, we’re done! I drew the thermometer! I drew the thermometer!” which was immediately followed by S3, “I drew this one. He drew this one.” Their recording ended after S4 suggested they play Blackjack when he was flashing a deck of cards.

**Summary**

Group A2’s discussion unfolded in opposite ways on the first and the second days of modeling. The shift of their dialogue patterns appeared to result from S4’s drastic change in his engagement upon being rebuked by S3 seven minutes into the second day of modeling.

Having to collaborate with an uncooperative partner on the first day, S3 interpreted evidence, designed the construction of the model, and directed the task procedures mostly in monologue. Her conversation with S4 was predominately confined to instructions about *model expression* in an *receptive-constructive* dialogue pattern. She
did not get any support from S4 about her reasoning, and therefore the construction of their model was not corroborated.

On the second day S4 became engaged in the conversation with S3 especially when they had conflicting opinions about evidence interpretation or the relevance of evidence to their model. In their interactive dialogues, they rebutted each other’s opinions with reasons before arriving at an agreement. S3 always referred to evidence to support her ideas, and this performance might have been advantaged by her being the person who generated *intra-evidence interpretations* for model construction. On the contrary, S4, who did not demonstrate any *intra-evidence interpretation* during the entire modeling activity, heavily relied on prior knowledge or personal experiences to justify his opinion but nonetheless contributed substantively to the model’s drawing.

Their model had a good title, indicated sequence, incorporated entities and activities, and recognized mitochondria the place for cellular respiration to occur. Having accomplished so much in their model, it is a pity that their model missed a crucial account to explain how exactly energy was generated in mitochondria. Should they have reflected on Ms. Addison’s prompts, thought about the entities and activities rather than the aesthetic details, and considered how to integrate information from different evidence to address an integral process, Group A2 could have had a chance to incorporate mechanisms in the model and elevate their modeling performance to a higher level.

**Prompts Offered by Ms. Addison**

All of the groups in Ms. Addison’s classes achieved a high or medium level in their models. Ms. Addison characteristically tailored guidance and connected her prompts
to (a) entities and cellular activities, (b) the key concepts of evidence, and (c) the modeling criteria.

**Prompts Were Tailored to Connect to Entities and Cellular Activities**

Ms. Addison regularly directed students’ attention away from expressing extraneous details in their models and further encouraged them to present their explanations at the cellular level, incorporating necessary entities, activities, and the location where the processes happened.

The excerpt below took place when Ms. Addison checked in with Group A1 for the first time during their model construction. In the beginning students were drawing a big body and a mouth to denote that “CO₂ coming out and oxygen going in” and “CO₂ being exhaled.” Ms. Addison challenged them to think about a better visual expression to explain cellular respiration. In the excerpts that follow, Ms. Addison is represented by A and students are represented by S.

1. A: Do you need a whole body to show cell respiration?
2. S1: No.
3. A: What’s the most important thing you’ll need to know?
4. S1: Mouth.
5. A: Ok, so what comes into the mouth?
6. S1, S2: Um, oxygen.
7. A: And food.
8. A: Ok, so maybe you can put a box.
9. Alright. So where do the oxygen and food go? That’s the next step.
10. What happens, according to these things?

After being questioned by Ms. Addison, the students acknowledged that it was not necessary to show a body, but they apparently thought that the mouth they had drawn should be kept. Instead of correcting the student’s response, Ms. Addison raised a question in line 5 to guide students’ attention to entities and their activities. She urged
students to reconsider their expression of cellular respiration. Ms. Addison prompted that where the entities went and what activities happened after the intake of oxygen and food should be crucial compartments in the explanation of cellular respiration.

After Ms. Addison left, students picked up her advice and initiated a long conversation in which they mulled over the roles that oxygen and food played in generating energy and their contribution to exercise: “The muscle. Oxygen going in.” “High carb gives you more energy…and helps your performance.” The prompt was effective. In the end of students’ discussion, an awareness of a knowledge gap surfaced: “We gotta do mitochondria. You’re talking about mitochondria.” “I know because when they assessed it Roger had the slightly high mitochondria.” “So…” “We don’t understand this part.” Ms. Addison came to their group at this moment.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A: Ok. These are all your facts you’re writing here but how does this show me a model of what’s taking place in people? All of it?</td>
</tr>
<tr>
<td>2</td>
<td>S1: Um.</td>
</tr>
<tr>
<td>3</td>
<td>A: Ok, so you’re saying food goes in, oxygen goes in, CO(_2) comes out. So you do have that.</td>
</tr>
<tr>
<td>4</td>
<td>Heat, temperature goes out—put an arrow there.</td>
</tr>
<tr>
<td>5</td>
<td>Ok. Now where does this take place?</td>
</tr>
<tr>
<td>6</td>
<td>In the mouth? So all of this process takes place in the mouth?</td>
</tr>
<tr>
<td>7</td>
<td>S1: In the lungs.</td>
</tr>
<tr>
<td>8</td>
<td>A: Ok, the oxygen comes in from your lungs, but does it say anything about oxygen coming in from your lungs here?</td>
</tr>
<tr>
<td>9</td>
<td>We need to show that on the diagram.</td>
</tr>
<tr>
<td>10</td>
<td>Ok. You have the heat coming out, the CO(_2) coming out, you have the food going in, you have the oxygen going in.</td>
</tr>
<tr>
<td>11</td>
<td>But more importantly, what did you recognize as a place where this might be going on inside the cell?</td>
</tr>
<tr>
<td>12</td>
<td>S1: Because like the mitochondria—</td>
</tr>
<tr>
<td>13</td>
<td>A: (Interrupting and yelling at the class) Umn…excuse me! Everything to remember is…shh…hello?</td>
</tr>
<tr>
<td>14</td>
<td>Remember the word “ce-llu-lar”, cellular respiration? Keep that in mind!</td>
</tr>
</tbody>
</table>
Ms. Addison’s utterances from lines 1 to 4 suggest that this model displayed the activities entities engaged in a way that reflected on each piece of evidence. But this was a model explaining external respiration without incorporating the cellular organelle evidence. This corresponds to their earlier confusion about mitochondria. Ms. Addison told them to think about a place inside the cells which hold the activities. She also prompted the entire class that they should think about the processes at the cellular level.

**Prompts Were Tailored to Connect to the Key Concepts of Evidence**

Ms. Addison constantly reminded students to revisit the evidence and make sure their model was consistent with the key concepts in the evidence. For example, in a class announcement in the middle of the modeling activity, she said, “May I suggest you look at each study? There are some questions there they ask you. You gotta get the model. I don’t see you do that. You’re just pulling things out of your hat without making any sense of the data.” Ms. Addison asked the students to interpret the evidence first and follow the prompting questions on their study worksheet (“What does this show?” and “Is this relevant to your model?”) to help them make the evidence and model consistent. In another example, when she questioned students why an extraneous thermometer was shown in their model, she tailored her prompt to the evidence to challenge students to think about what the evidence was really saying.

1  A:  Do you have a thermometer? Do you have a thermometer in your body?
2  S:  No.
3  A:  How do you know that?
4  A:  What evidence is it in that study that shows your body temperature?
5  A:  What happened to your body temperature?
When students experienced difficulty figuring out the role a certain piece of evidence played in the system and did not refer to this evidence in their model, instead of asking them to simply include the missing part, Ms. Addison engaged students in a conversation in which she provided hints until they generated the answer themselves. In the excerpt below, Ms. Addison facilitated the students who were not able to make sense of the cellular organelle evidence.

1  A: In the study did we use a comparison except for one thing?  
2    Yes, we did, one place.  
3  S: Yeah, in the beginning, and then when they ran.  
4  A: And that’s because it was based on the different diets. But what about this one?  
5  S: Oh, yeah because Roger had like slightly high like a slightly high mitochondria level.  
6  A: Ok, so what do you think the mitochondria might have to do?  
7  S: Well, it might burn more food and oxygen quicker.  
8  S: So, like have a faster metabolism.

Ms. Addison wanted to direct their attention to the cellular organelle evidence, but the students did not get it. Instead of telling them what it was, she continued on the conversation and pointed at the cellular organelle evidence in line 4. Although students saw that one boy had slightly higher count of mitochondria than the other, which was a direct interpretation of the table, they did not further relate this evidence to cellular respiration. Ms. Addison encouraged them to think about the role mitochondria might have played. Finally, in lines 7 and 8 students incorporated the potential function of mitochondria into the process of cellular respiration. Their statement suggested the influence of prior knowledge about the metabolism concept on their interpretation of this evidence.
**Prompts Were Tailored to Modeling Criteria**

The publicly displayed criteria in Ms. Addison’s classes included standards such as a model should have clearly labeled process or sequence, show only relevant details, explain all the evidence from data, keep a certain degree of abstraction, and have consistency between drawing and words. When students’ models approached completion, Ms. Addison adapted her prompts from supporting model construction (as displayed in previous sections) to reinforcing model revision by referring to the “criteria for good models.”

1. S2: Ms. Addison, we are done.
2. A: You don’t feel like this needs correcting?
3. S2: Uh—
4. A: Is this clear to me? I just see a big jumbo.
5. Is this a process?
7. A: Ok, where do we start? Because I don’t know where to begin.

Ms. Addison critiqued this group’s model and reminded them of the direction to refine it. The prompts that she provided matched one of the “criteria for good models”—a clearly labeled process should be displayed in a model.

The modeling activity did not end until after Ms. Addison pressed for a final round of revision. “Before we wrap,” Ms. Addison announced, “look at your criteria and speak to each other about those things shown on the model, because tomorrow we are going to do a class critique, and we want to make sure that when we are discussing you have a sense of what the criteria are, so take a minute to do that with each other.” She picked a few criteria on the list, “Do I have a title? Is the process clear? Is it understandable to somebody? Does it show sequence of events?” Her prompts
successfully engaged students in cross-checking their models along the “criteria for good models” and evaluated if each piece of evidence was relevant to their model.

1  S4: Yep, label. Label those processes.
2  S3: Label those processes clear—we have that.
3  S4: Ching-ching!
4  S3: Show any relevant ideas—relevant that we said “Yes” to.

Students also explained the reasons whether certain evidence was or was not relevant to their models in response to the third question on the study worksheet.

**Summary**

Ms. Addison customized her prompts to accommodate specific student needs for better performance in the visual presentation of entities and activities at the cellular level, the incorporation of key concepts of evidence into models, and the revision of models to accord with model criteria. She guided students’ focus away from superficial and aesthetic details to entities and activities when they were considering what to be presented in their pre-lecture models. All of the groups in her classes depicted a certain degree of cellular processes in their models, and many incorporated mechanisms. She reminded students to construct their models in accordance with the crucial evidential concepts rather than the superficial aspects of evidence. When students needed help with evidence interpretation, she highlighted the key points in evidence to facilitate students’ comprehension of evidence. She clearly pointed out the areas of a model that needed improvement and referred the students to certain modeling criteria for possible revision. Overall, Ms. Addison’s prompts were specific in pointing out the way for reasoning and improvement and her scaffolding effectively elevated students’ performance in modeling.
The Medium-Performing Group: B1

The medium-performing Group B1 represented the best group from Mr. Britten’s classes in School B. School B was a much more proficient school than School A in general terms; however most of Mr. Britten’s students exhibited inferior performance than their peers in School A. Groups in Mr. Britten’s classes brainstormed and distributed thinking and drawing tasks between four members in one period, whereas pairs in Ms. Addison’s classes had two periods across two days to work on modeling, but they spent the second period mostly on filling up the study worksheets. Each of the four sampled small groups was distinctive in how they reasoned and collaborated during the modeling.

Group B1’s Model

Group B1’s model presented a conceptual formula which wrote, “food + oxygen inhaled = energy.” A boy was running in the center of their poster. His breathing and temperature were increasing, and carbon dioxide was let out. In the top right corner, they drew another boy at a much small scale and wrote, “Run = energy,” implying that energy enabled one to run. Overall, the model illustrated an exercise scenario in which a boy took in food and oxygen to get energy to run. During his running, he burned food he ate and increased his breathing rate and body temperature. No explanation was given about how and where food and oxygen might interact to generate energy, and the concept of the cellular organelle evidence was left out of this model.

In contrast to Group A2’s medium-quality model, which labeled the sequence and location of cellular respiration but did not explain by which entities and activities energy was generated, Group B1’s model specified that food and oxygen generated energy, but the location and process of energy generation were absent. The notation of entities and
energy in the formula demonstrated that this group was aware of the first half of the process of cellular respiration. Group B1’s medium-quality model was credited with a total score of 41, within which 28 was for evidence interpretation and 13 was for mechanistic explanation.

**Figure 8**

*Group B1’s Model (A Medium-Performing Group)*

*Group B1’s Discussions*

Two minutes into the audio recording the *intra-evidence interpretation* of the *muscle cell contraction* evidence commenced Group B1’s modeling. They tried to make sense of a chart that displayed how the rate of muscle cell contraction related to the amounts of glucose and oxygen taken in.

1. S5: Um, I don’t get the second one [referring to the evidence].
2. S7: I don’t get the third one, let me go…
3 S6: S8 said that you average the amount of glucose and the amount of oxygen.
4 S7: Wait, get it off of me! Oxygen what?
5 S6: Uh, yeah.
6 S8: I can see the second one.
7 S5: That’s what we write? You average them out?
8 S6: Yeah.
9 S5: You average what?
10 S6: Um, the amount of glucose and the amount of oxygen inhaled.

In line 3, S6 gave a terminal answer that she saw on S8’s worksheet, and the rest of the group accepted the answer after a couple of questions for repetition and confirmation. This interaction constituted an active-constructive dialogue pattern, in which S8 was the constructor, and the rest were active participants. With regard to S8’s interpretation, S8 emphasized how the quantitative measurement was calculated rather than the evidence’s key point that the rate of muscles contraction was correlated to the amount of oxygen and sugar. None of the other group members challenged this interpretation and the group wrote down the words they heard on their study worksheets.

Immediately after the previous episode, students engaged in a brief discussion. It concerned the first piece of evidence, which provided the physical conditions of the two boys in the story displayed in a simple table. The evidence clearly showed that the two boys were very similar to each other—as a controlled variable to validate their subsequent measurements—yet Group B1 cared more about “Which one is the fat one?” They misread the information and could not reach a consensus for this simple but irrelevant issue. Their distraction by this trivial element in the table did not further their efforts to interpret the evidence properly, and the question “Who is the fat one?” became the most reoccurring topic in the rest of their discussions.
The third episode of on-task discussion took place after a 5-minute off-topic chatting. Here they exchanged ideas about their inter-evidence interpretation of the physical fitness and high-carb vs. low-carb evidence.

1 S7: Roger and…
2 (A group member starts singing.)
3 S8: CO₂ not C₂O!
4 S7: Sorry!
5 S6: That's the fat one, right?
6 S8: Yeah
7 S5: No—it’s Roger! You see?
8 S7: No.
9 S5: He couldn’t come up with Eric’s running pace, so…
10 S6: No Roger’s the fat one.
11 S7: He had a big steak beforehand!
12 S6: What an idiot!
13 S7: You are right!
14 S6: Roger had pasta, that’s why he’s better.
15 S5: And he’s taller!

This conversation concerned the high-carb vs. low-carb evidence, but their argument still focused on who the “fat” one was. The discussion between lines 7 to 11 indicates that they flipped the data about two boys’ diet and physical performances displayed in the evidence (it was Eric who was left behind in the evidence). In line 14, S6 finally touched on the main point of the high-carb vs. low-carb evidence. Unfortunately, rather than extending S6’s statement and thinking about the relationship between carb intake and physical performance, the rest of the group made no more comments but circled back to the boys’ overall physical condition from the first evidence. In a nutshell, they had an intense interactive episode, but they only elaborated at the superficial level of the evidence and misread the information in the evidence, so that the interactive discussion did very little to enhance their reasoning.
The previous episode was followed by another minute of off-task chatting until Mr. Britten told them that they had to “create a model to show how our body and our cells use the food and oxygen to get energy.” He also reminded them to “give reasons to explain why your model is good and make sure your model explains the data.” At this point, Group B1 had five minutes left so they quickly made some decisions about logistics and moved into a constructive-constructive dialogue pattern for model construction. Each group member silently contributed to a certain portion of the model and did not discuss how to put everything together into a model. They regrouped from time to time to comment on the aesthetic aspect of the model.

2. S7: (Laughing) This is great. (Laughing)
3. S5: Ew, what is that?
4. S7: (Laughing)
5. S6: Okie Dokie, I’m done!
6. S5: Ew, what is that?
7. S8: Ew, S7!
8. S7: (Laughing hard) That’s food!
9. S5: Oh—that’s food.
11. (Group members commenting on S8’s neat handwriting)
12. S5: What is that?
13. S7: It’s a guy running [referring to the small-scaled human figure].
14. S5: No, ew, S7!
16. S8: Guys, actually I see that’s where the head is.
18. S5: No, no, no, no, he’s completely drawn, S7, complete that one [referring to the small-scaled human figure] or else it’s going to look stupid!

This excerpt above gives a synopsis of how this group created their model. Their dialogue showed that monitoring and suggesting were made exclusively for model
expression and no new ideas were developed between group members. From their model’s drawing it shows that S7 took S6’s suggestion and drew a big person running. S6 generated the idea and S7 generated the drawing; no explanation or discussion was exchanged.

After S7 announced, “I’m finished. I drew most of that,” Mr. Britten came to this group and listed all the entities that were not present in their model, “What about oxygen? What about time? What about carbon dioxide? Temperature? What about the other studies?” Group B1 quickly assembled their team and revised the model accordingly.

1  S7:  Body temperature? Like how can we do that?
2  S8:  Ugh!
3  (Loud noise from working and movements)
4  S7:  Breathing is increasing?
5  S6:  Yeah, amount inhaled that goes up.
6  (Discussing who should write their names down on the poster)
7  S6:  We good!
8  S5:  Uh—I’m almost finished.

In this final episode of constructive-constructive dialogue pattern, each team member added a bit into the poster with minimal verbal exchanges except for when they sought for conformation in lines 4 and 5. At the end Group B1’s model displayed additional explanations: “Temp is getting higher,” “CO₂ let out,” and “Breathing is increase because you are working hard.” These newly added entities and activities improved the original model, but still did not explain cellular respiration.

**Summary**

A majority of Group B1’s conversation was off-topic. The analyses were established on the episodes when the group members were talking about evidence and modeling. Their discussion in general revealed an active-constructive dialogue pattern, in
which the person who first figured out the intra-evidence interpretation of a certain piece of evidence was seen by the rest of the group as the provider of the answer; no challenges or inquiry for explanations were observed.

In the only one interactive episode they had, they considered two pieces of evidence simultaneously (physical fitness and high-carb vs. low-carb) and they reached a conclusion after an initial disagreement. Unfortunately, their inter-evidence integration in this interactive discussion focused on the most trivial aspects of the evidence. Regardless of the dialogue pattern utilized, they tended to either miss the key points or misinterpret the evidence. For instance, the two boys’ body weight was never a key point despite their recurring concern with which boy was the fat one. The chart in the high-carb vs. low-carb evidence clearly showed that Roger out-paced Eric and that he ate high-carb food, but misinterpretations such as “[Roger] couldn’t come up with Eric’s running pace” and “He had a big steak beforehand” still arose.

During their relatively brief model construction time, their collaborative effort was displayed in a constructive-constructive dialogue pattern featuring a phenomenon where each person took up a portion of the model without reaching out to one another for feedback or critique. Following Mr. Britten’s prompts about overlooked evidence, they added a few entities and activities but never considered them at the cellular level. Their quick revision in response to the teacher’s advice and the inclusion of energy as the end product helped their model to achieve a medium-level quality.
The Lower-Performing Group: B2

Group B2’s Model

Group B2’s model depicted two stick figures and made several comparisons between them regarding their physical conditions and diets. The model set these two stick figures in a soccer game context and indicated that their body temperatures were “practically” the same during the game. The left stick figure ate a big pasta dinner and a fruit cup. He scored five goals and his oxygen inhalation level was high from all the running. The other stick figure had barbeque ribs for dinner. He got tired easily, breathed heavily, and was benched. His oxygen inhalation was low from sitting on the bench.

Figure 9

Group B2’s Model (A Lower-Performing Group)
This model was a description of two persons’ physical performances in a sports game rather than an explanation of cellular respiration. It did not mention energy, mitochondria, or a full range of entities and activities in the energy generation process. This model’s total score was 13, with 13 for evidence interpretation and 0 for mechanistic explanation.

**Group B2’s Discussions**

Group B2’s four members from Mr. Britten’s class were very engaged throughout the day of modeling. Yet they were the group who received the lowest score in Mr. Britten’s classes.

Group B2 moved quickly through the initial four pieces of evidence with minimal communication heard on the recording (I assume that they were silently checking out each other’s responses to the first question on the study worksheet) until they engaged in *intra-evidence interpretation* about the oxygen evidence. S11’s idea that one of the boys could have asthma triggered an intense *interactive* discussion for two minutes that tapped into group members’ personal knowledge and experiences about asthma and breathing. But the dialogue pattern suddenly became *active-constructive* when they resumed their focus on the evidence. One student announced her correct interpretation that “the more time exercising, the rate of oxygen inhaled increases,” and the rest of the group repeated some words and asked for clarification.

In most episodes, every member of this group committed to reasoning about evidence in an *interactive* dialogue pattern; for example, when they discussed the *high-carb vs. low-carb* evidence.

1 S12: (Reading the evidence) “However, for some reason, Eric could
not keep up with Roger’s running pace after a while. They’re in similar physical condition.” So…

S11: It could be because of asthma. That would totally make sense.

S10: But it says they’re in the same physical condition.

S12: No, no but look, “Eric had a big steak beforehand and Roger had pasta and a fruit cup instead.”

S9: (Screaming for an emerging breakthrough) Oh carbs make you more energetic. So, write that!

S10: Carbs make you more energetic and enable you to run further…and enable you to to more…um…run more…more for longer.

S11’s suggestion was refuted by S10 and S12 in lines 3 and 4 whose ideas were based on the evidence. Hearing S12 reading aloud the evidence that identified the boys’ different diets, S9 in line 5 abruptly concluded that “carbs make you more energetic.” This statement was further extended by S10’s remark in line 6, which then was included on their study worksheets.

A similar dialogue pattern reoccurred when they discussed the cellular organelle evidence. But this time their intra-evidence interpretation was not as strong as it was on the high-carb vs. low-carb evidence.

S11: I think Roger might be stronger, in just the slightest way.

S9: We should say they have the same mus—they have really similar muscle cells.

S12: No, they have similar muscle cells but Roger’s mi-to-chon-dria is high...slightly higher.

S10: They have the same…like they have basically the same amount of cells—

S12: Except Roger has higher mitochondria.

S11: Quickly put that down because we have five minutes.

S12: Alright, this shows what?

S10: Um, they have basically the same um amount of cells.

S12: Well, density…

S10: But cell den—oh, S11. (S11 is talking to someone else from another group)
This *interactive* episode started from S11’s hunch. S9 had an alternative idea, and S12 disagreed with S9’s idea by pinpointing the most vital information in this evidence that the boys’ mitochondrial densities were different. However, this awareness displayed in lines 3 and 5 was not acknowledged by other members right away. In lines 4 and 8, S10 thought instead that the boys had a similar number of muscle cells, and S12 seemed unsatisfied with S10’s idea and brought up “density” again in line 9. At this point S10 wanted to say “cell density” but did not finish her sentence. No further discussion about this evidence was found in the recording. In their final interpretation of this evidence, their study worksheet read, “Similar amount of cell density.” In this episode, everybody expressed their ideas in response to other peers’ thoughts. Yet they decided to take S10’s opinion as their conclusion which missed the key point of the evidence. This conclusion was made before they had a chance to take everybody’s ideas into consideration.

When Group B2 proceeded to generate a model, they maintained their *interactive* dialogue pattern and exchanged ideas about *model expression* based on their excellent conceptualization of certain evidence. They ensured that the better-performing stick figure had high-carb food. “Alright, big pasta dinner, and a banana!” “And a salad?” “No, no, fruit cup, fruit cup.” They made the other stick figure have only proteins. “Barbeque ribs?” “French fries. Oh no, French fries is carbs.” “Barbeque ribs.” New ideas were developed by different peers from an initial thought and were consolidated into the final construction of the model. But all of the ideas they exchanged were confined to *model expression* in a way to portrait two “little stick person[s]” in a soccer game. No reference was made to the evidence, nor did any discussion arise to integrate the key points found from different evidence.
After Group B2’s members decided what to draw, they split themselves into two subgroups. One worked on the model, and the other evaluated if every piece of evidence was relevant to their model in response to the third question in the study worksheet. Their model was not consistent with most evidence, but they thought all evidence was relevant to their model. For example, they put “Yes” on the study worksheet to indicate that the *cell organelle* evidence was relevant, and the reason was “Energy was brought in by the organelle mitochondria. [The better-performing boy] had more energy.” Regardless of its accuracy, this statement clearly implied that energy was a function of mitochondria.

The only disagreement they experienced was whether the *body temperature* evidence was relevant. “No, it’s definitely not,” said S10. “Body temperature, ‘No’?” S9 asked. S12 jumped in from the drawing subgroup, “Yeah, ’cause it’s not the further they run the hotter they get.” After this brief interactive discussion, Mr. Britten told them, “I need to see oxygen and carbon dioxide…I don’t see anything about temperature.” Students replied that because they already determined that the *body temperature* evidence was not relevant, they did not need to address it in their model. They also stood firm on their opinion that the boys’ body temperature did not change, because “it can’t do that.” Mr. Britten insisted, “No, you should include that.” He made a few attempts to guide them to revisit the evidence, “Look at the temperature chart…Those numbers aren’t all the same numbers…They are close, but they are not the same, isn’t it?” However, the teacher’s guidance was rejected by the group.

1. S10: (Replying to the teacher) Barely…No, body temperature is barely affected.
2. S12: It’s not the longer he ran the higher it was.
3. S9: I just wrote that your body temperature really doesn’t change...like it can’t do that.
It appeared that Group B2 believed that their model was consistent with all evidence despite that their model merely put two stick figures side by side and listed the food they ate before a soccer game and the oxygen inhaled. They called Mr. Britten over, “You have body temperature the same? But doesn’t it say that body temperature changed?” This time the group admitted that it raised just a little, and eventually put “Both body temp practically same” on their model. Although Group B2 engaged in interactive discourse patterns when they built a model, they never revisited the evidence during model construction until Mr. Britten indicated that evidence was missing from their model; yet they still did not weave overlooked evidence into the model. Moreover, when Mr. Britten told them his interpretation of the evidence, this group refused to adopt his ideas. If they had used their time to improve their model to accord with the evidence instead of generating reasons to defend their biased prior knowledge, they might have been able to enhance the quality of their model.

Summary

Group B2’s members were fully engaged and committed to all aspects of their task. When they were trying to make sense of the evidence, sometimes their discussion was in an interactive dialogue pattern where everybody contributed ideas and ideas were responded to and expanded. At other times their interactions were the epitome of an active-constructive conversation in that there could be one member who determined the conclusive interpretation of a piece of evidence and the rest of the group repeated and copied it. Regardless of the dialogue patterns, Group B2 often jumped into a conclusion
in the middle of someone’s reasoning course, even at the moment when a peer was just coming to the realization of a key point in the evidence. Apparently, their conclusion was not drawn after taking everyone’s opinions into account despite that ideas were elaborated on between peers.

During model construction, Group B2 consistently exchanged their thoughts until they reached a mutual consensus about how to draw a model. For the evidence that they conceptualized well, they were able to incorporate it into the model in a way that resonated with their understanding. However, their model was not consistent with many pieces of evidence. Also, their prior knowledge tended to override and distort their interpretation when it conflicted with a piece of evidence. Therefore, they misinterpreted the evidence and even refused to change their views after receiving Mr. Britten’s guidance. The rejection of taking advice made their model inconsistent with evidence. Group B2’s model turned out to be an elaborated illustration involving two boys in a sports context. Their model did not present entities or activities at the cellular level, nor did it incorporate mechanisms. This result implies that even though Group B2 regularly engaged in interactive discussions, this dialogue pattern was not necessarily productive for producing a good model. Compared to the high-performing Group A1 whose talk was mostly in the active-constructive pattern, Group B2’s interactive discussion was limited to the intra-evidence interpretation and model expression discourse operations, while Group A1 touched upon inter-evidence integration and mechanistic explanation discourse operations. Perhaps the reasoning about multiple evidence all at once and the pondering of the mechanisms based on multiple evidence are the stipulation for an interactive group discussion to be productive.
Prompts Offered by Mr. Britten

Students in Mr. Britten’s classes demonstrated medium and low competence in their modeling. Contrary to Ms. Addison’s specific prompts which covered a wider range of focal modeling practices such as expressing the entities and activities at the cellular level, Mr. Britten did not deviate in his assistance from reminding students to (a) make models fit all evidence, (b) revisit overlooked or misinterpreted evidence, and (c) examine the patterns in charts within evidence. Mr. Britten’s scaffolding was expressed in generic terms, and did not guide students to think about the cellular processes and the activities that entities involved. Students did not receive prompts to build a model at the cellular level, nor were they encouraged to integrate concepts across evidence or to incorporate entities and activities in their models to explain the phenomenon of cellular respiration.

Prompts Were Focused on the Alignment of Models and Evidence

The period began as Mr. Britten rotated between groups and repeated the task goal that they had to create “a model that clearly shows how you think cells are consuming, turning food and oxygen into energy that is supported by each of these things [evidence].” Groups did not initiate modeling until he stopped by to check on their performance.

1    S8:   So can we use this page to finish?
2    B:    It’s no problem. This model which is based on all of this [evidence] will be done before you leave. (Mr. Britten leaves)
3    S7:   What do we do with all of that?
4    S8:   (Gasp) We only have five minutes?!
5    S6:   (Calling out) Mr. Britten, what do we draw?
6    B:    (Returning to this group) You’re going to, on the last page…so you’re going to create a model to show how our body and our cells use the food and oxygen to get energy.
Give reasons to explain why your model is good and make sure your model explains the data.

So you guys are getting through Step 3 today. (Mr. Britten leaves)

S5: I say S7 draws this. (Group members start to discuss the logistics about who draws what with which color of highlighters)

Mr. Britten’s reminder emphasized building a model based on the evidence provided. To another group, a prompt was given after Mr. Britten’s scaffolding was sought, “Ok, so you gotta finish making your model and show it to me. Show me how it ties it to each of the data in the studies before we go.” Students initiated drawing on the poster right after this instruction was released and talked aloud what they were drawing: “Alright, big pasta dinner.” “And a banana!”

Prompts Were Focused on Revisiting Overlooked Evidence

Inconsistency between evidence and models could have resulted from not taking evidence into account in a model. Mr. Britten especially urged students to consider the evidence of oxygen, carbon dioxide, and temperature, but he did not remind them to think about how the entities might interact with each other in processes. During the entire modeling activity, Mr. Britten did not help students pay attention to the muscle cell contraction and the cellular organelle evidence, the two pieces of evidence that addressed the cells.

S7: I’m finished, I drew most of that.
B: Okay, um…great, the problem is that this doesn't take in a lot of the other pieces of data in the studies.
You don’t see oxygen, you don’t see carbon dioxide.
Carbon dioxide…
What about oxygen, what about time, what about carbon dioxide, temperature?
What about the other studies? (Pause)
There’s a whole bunch of things that you wrote up in the charts
After Mr. Britten left, this group picked up his prompts and included all of the evidence he mentioned. Despite Mr. Britten’s pressing students to think about evidence, he only focused on some pieces of evidence and ignored others. Mr. Britten did not bring up the evidence that required reflection on cellular activities—the evidence that displayed the relationship between the rate of muscle cell contraction and the amount of oxygen and sugar taken in as well as the evidence that depicted cellular organelles including mitochondria. The former evidence guided students’ attention to a microscopic picture of cells and suggested that there was hidden intermediate activities behind the intake of oxygen and glucose and the performance of muscle cell contraction. The latter evidence demonstrated a microscopic picture showing cellular organelles and hinted at mitochondria by making it the only discrepant organelle density between two boys. The students had learned cellular organelles and their functions before this present module, which means it is reasonable to assume that students were prepared to reason at the cellular level. However, it did not come to students’ awareness that cellular respiration was a cellular energy exchange mechanism that took place in certain cellular organelles. Without receiving scaffolding on this particular point, no groups in Mr. Britten’s classes built a model that explained the topic at the cellular level.

_Prompts Were Focused on Examining the Patterns in Charts Within Evidence_

Besides indicating missing evidence in a model, Mr. Britten also directed students to look at the patterns in the chart within the overlooked evidence.
your model.

I don’t see anything about temperature, or that sort of stuff.

S12: That’s why we wrote “No.”

B: No, you should include that.

S12: But none of you guys are helping me.

S9: Okay, write his body temperature was the same the whole time and write his was—

S10, 11, 12: No. No—

S9: No [refuting the “No” above], body temperature is the same!

B: Look at the temperature though.

S9: They are the same!

B: Look at the temperature chart…wait, look at the temperature chart.

Those numbers aren’t all the same numbers, are they?

S9: I feel Frank has …oh yeah.

S12: I’ll say that they both have the same…

B: They are close, but they are not the same, isn’t it? (Mr. Britten leaves)

S10: Barely…No, body temperature is barely affected.

S12: It’s not the longer he ran the higher it was.

S9: I just wrote that your body temperature really doesn’t change…like it can’t do that.

S10: Alright, guys we have to have everything in our model.

S9: We did. Just write our names on it.

Although students added body temperature in their model after Mr. Britten’s direction in line 4, they had an alternative interpretation of this evidence. S9’s group partners disagreed with her idea that they should write that body temperature was the same the whole time, but they also disagreed with Mr. Britten’s reading of the chart that the temperatures were different along the time. S10’s response in line 17 shows that she did perceive the differences in the chart. The assertion “it can’t do that” made by S9 in line 19 reveals her belief that humans maintain a steady body temperature unconditionally, and this belief was imposed on this group’s interpretation of the data.

After students finished their model, they asked Mr. Britten for comment. Mr. Britten asked them to reconsider the body temperature evidence.
1 S12: That’s a pretty darn good model! That’s a pretty darn good model!
2 S9: Wait, Mr. Britten, do you have a minute?
3 B: You are gonna keep…whether it gives you…Okay, let’s take a look.
4 S10: We have body temperature the same [on the model].
5 But doesn’t it say that body temperature changed [on the evidence]?
6 B: But that [the group model] says the body temperature is the same.
7 S12: By like 1.1!
8 S11: The same.
9 S9: Practically, she said practically. (Students put “Both body temp practically same” on the model)

Although Mr. Britten once again notified students of the misinterpreted evidence and students demonstrated that they had the ability to understand the chart as seen in line 8, students’ belief in invariable body temperature was so robust that they resisted from taking his advice. This group was capable of observing tiny differences in charts, and this was not a trivial issue for them. They spent a great deal of time discussing the possible causes of the tiny differences observed in two boys’ oxygen and carbon dioxide breathing rates, but they refused to acknowledge the differences shown in the body temperature evidence due to the belief that “your body temperature really doesn’t change…like it can’t do that.” Mr. Britten’s prompting style did not successfully help students put aside their prior knowledge and recognize their perceiving of the evidence.

Summary

Mr. Britten primarily focused his prompts on the generic modeling criterion that models should fit with evidence. He reminded students of overlooked entities in their models and addressed them the generic modeling criterion that models should be consistent with evidence. When his advice was accepted by students, he did not push
them further for explicating the cellular processes. When he was confronted with students’ resistance of taking his advice, he was patient and directed students to the absent evidence to further inspect the patterns. Overall, he did not scaffold them to think about the cellular entities and activities, nor did he hint on how to translate evidence into a cohesive virtualization that incorporates mechanisms. Without receiving prompts about considering cellular entities and activities, Mr. Britten’s small groups all failed to generate a mechanistic model.

**Summary of Group Discussions and Teacher Prompts**

The findings from the discourse data addressed my third set of research questions. The processes that mediated modeling performance were discussed from two aspects. The first aspect was what and how students talked to each other when they worked together. The second aspect was how teachers carried out their scaffolding during small group work.

**Group Discussions**

The results suggest that *interactive* discussions need certain conditions in place before they are productive. To be more specific, the discussion that contains *inter-evidence integration* and *mechanistic explanation* discourse operations by a less interactive group can vastly outperform another intensively interactive group who only exchanges ideas about *intra-evidence interpretation* and *model expression*.

This conclusion can be drawn from the contrast between the four groups. The high-performing group A1 was the only group that integrated multiple evidence in a single reasoning course (*inter-evidence integration*) and expounded mechanisms in their statements (*mechanistic explanation*). Most of Group A1’s discussion was in
constructive-constructive and active-constructive patterns with a couple of interactive episodes, but their model outscored the interactive Group B2, a lower-performing group, who only talked about the interpretation of each evidence exclusively (intra-evidence interpretation) and the drawing of their model (model expression) in their conversation. Moreover, neither of the two medium-performing groups, Groups A2 and B1, had dialogue patterns that were as interactive as Group B2’s. Group A2 conversation was receptive-constructive at first and later turned interactive. Group B1 was characterized with the active-constructive and constructive-constructive patterns. Groups A2 and B1 tapped into similar discourse operations, intra-evidence interpretation, and model expression, as did Group B2, but these two groups developed their models after listening to peers’ complete reasoning, whereas Group B2 had prior knowledge impose on data interpretation and often jumped to a conclusion in the middle of the interactive discussion before all of the opinions were weighed.

The medium performers, Groups A2 and B1, outperformed Group B2 even though they did not have a mostly interactive dialogue pattern in their discussions as it was in Group B2. Although Groups A2, B1, and B2 all engaged in intra-evidence interpretation and model expression exclusively, they differed from each other according to several contrasting practices. Group B1 revised their model following Mr. Britten’s prompts. Group A2 did not revise their model but they referred to evidence when they built their model. Group B2 did not revise their model, nor did they check the evidence when they created a model, which might be one of the reasons to account for their model’s inferior quality.
Group A1 and Group A2 were similar in that they each had only one leader who determined the interpretation of evidence and the construction of the model. Both groups always went back to evidence when they thought about how to build their models. Group A1 integrated concepts from multiple evidence (inter-evidence integration) and thought about the mechanisms (mechanistic explanation). They reflected on Ms. Addison’s prompts and improved their model. Group A2, on the contrary, discussed each piece of evidence alone (intra-evidence interpretation) and emphasized on the drawing of their model (model expression). They affirmed that their model did not need any revision. In the end, Group A1’s model received the highest score, and Group A2’s performance fell on the medium level.

**Teacher Prompts**

Ms. Addison and Mr. Britten were different in how they offered prompts to facilitate students’ practice of modeling. The teacher prompts and the corresponding subsequent student conversations are summarized in Table 20.

Ms. Addison tailored her prompts to connect to specific modeling skills. She began her scaffolding with critiquing students’ visual representation by advising them to present the entities and activities at the cellular level in their models. She challenged students to think about the key concepts of evidence. If a piece of evidence was overlooked or if students had difficulty making sense of it, instead of pointing out the missing parts, Ms. Addison engaged students in a conversation with her until they perceived her hints and figured out the key concepts of the evidence. Finally, when students thought their models were complete, she reminded them of the modeling criteria by which they were expected to examine and refine their models. Perhaps it was because
of the specific prompting techniques that Ms. Addison leveraged that most of her
students’ group models performed better than the medium level.

**Table 20**

**Teacher Prompts and Subsequent Student Actions**

<table>
<thead>
<tr>
<th>Teacher prompt</th>
<th>Student action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ms. Addison’s specific prompts connected to Entities and cellular activities</td>
<td>Group A1’s talk shifted from breathing to the cellular activities.</td>
</tr>
<tr>
<td></td>
<td>Group A2 ignored a challenge from Ms. Addison concerning the relationship between legs and oxygen, and did not discuss it.</td>
</tr>
<tr>
<td>The key concepts of evidence</td>
<td>Group A1 hypothesized about the possible functions of mitochondria.</td>
</tr>
<tr>
<td></td>
<td>Group A2 discussed what the body temperature evidence meant.</td>
</tr>
<tr>
<td>Modeling criteria</td>
<td>Students examined their models against the “criteria for good models.”</td>
</tr>
<tr>
<td>Mr. Britten’s generic prompts focused on Fitness of models and evidence</td>
<td>Students started building models immediately.</td>
</tr>
<tr>
<td>Overlooked evidence but did not mention cellular evidence</td>
<td>Students added the information from the overlooked evidence but left out the evidence concerning cellular organelles and activities because they were not prompted.</td>
</tr>
<tr>
<td>Patterns in evidence charts</td>
<td>Students refused to recognize the pattern they perceived.</td>
</tr>
</tbody>
</table>

Mr. Britten pressed on a generic principle that models had to be consistent with evidence. Based on this principle, he listed the evidence that students overlooked, indicated the pattern embedded in evidence, and urged students to include the evidence in their model. However, he did not direct students to look at the evidence concerning cells, nor did he encourage them to think about how entities might interact at the cellular level.
Students did not receive hints to attend to the key concepts of evidence or to hypothesize the processes of cellular respiration. A majority of group models in Mr. Britten’s classes performed lower than the medium level.
Chapter 6

Discussion and Conclusion

This dissertation study investigated seventh-grade students’ competence in incorporating evidence and mechanisms into models as well as the processes of small group discussions and teacher prompts in a model-based inquiry class. In this chapter I discuss the key findings addressing three sets of research questions, which is followed by the limitations. In conclusion I discuss the significance of this study, its future instructional implications, and future research directions.

Discussion of the Results

The Consistency of Models and Evidence

One of the major findings of this study is that middle school students were capable of building models consistent with relevant evidence while at the same time eliminating the irrelevant evidence when working with a wide variety of complex evidence. Earlier studies found that high school and college students were not competent in citing evidence presented as charts in their explanations in an inquiry setting (Sandoval, 2003; Sandoval & Reiser, 2004; Takao & Kelly, 2003); the practice of incorporating evidence into an explanation was not spontaneous. The superior modeling achievement of this study’s seventh graders in comparison to older students could be ascribed to the introduction of the epistemic metacognitive knowledge (Barzilai & Chinn, 2018) of modeling. The introduction about the nature and attributes of models (Cartier & Stewart, 2000; Passmore & Stewart, 2002) or the rubrics of modeling (Sandoval & Millwood, 2005) can promote model-evidence consistency, and the provision of the rationale of making a scientific explanation by teachers has been shown to enhance
students’ inclusion of a claim, evidence, and justification in an explanation (McNeill & Krajcik, 2008). During the school year when students participated in the PRACCIS project, they engaged in activities that were engineered to build their content knowledge immersed in an inquiry setting that stressed developing epistemic knowledge about modeling. The activities ranged from the nature and features of models and evidence, to different relationships between models and evidence, to the criteria for good models (including the alignment of evidence and models, the explanation of processes supported by evidence, the indication of sequence, etc.). It is very likely that the epistemic knowledge of modeling might play a critical role in promoting students’ capability of constructing models consistent with several pieces of relevant evidence in the absence of irrelevant evidence.

The Interpretation of Evidence

In this study, the drawing and words presented in models were serving as a proxy for understanding students’ interpretation of the evidence concepts. The findings pointed out that middle school students’ interpretation of data could be under the influence of their familiarity of the information. On the one hand, they were skillful in interpreting evidence that expressed concepts with which they were familiar. In this study, students performed very well in interpreting the high-carb vs. low-carb evidence and the oxygen evidence. Students’ knowledge about how human bodies eat and breathe at the organic level not only enabled them to infer the information presented in the evidence but also fill in the gaps between evidence in their explanation of the cellular respiration process. Their additional illustration was primarily aimed at the digested process and the external respiration process through the lungs, although these details were not displayed in the
evidence. They also intertwined sophisticated concepts into their models in which they elaborated that food as big molecules was transformed into glucose as smaller molecules before participating in the energy generation process. They could have transferred this concept from an earlier model-based inquiry lesson about cellular organelles and osmosis designed by the PRACCIS team. No model was found to further address facilitated diffusion for glucose movement through the cell membrane as this concept was to be introduced in a later grade. This phenomenon is consistent with my claim that prior knowledge takes a place in students’ evidence interpretation.

On the other hand, when encountering with novel ideas, students might not know how to integrate these ideas into their models although they might be able to make sense of the data. Research has suggested that primary school kids and middle school learners are able to understand the interplay between two variables and recognize the patterns of data revealed in different forms of graphical representations (Hapgood et al., 2004; Wu & Krajcik, 2006). The students in this study should have acquired the skills to interpret evidence by the point when this cellular respiration module was enacted in the second semester of the one-year-long project. Their less competent interpretation of certain pieces of evidence might be because that they did not know how to weave some unfamiliar ideas into their models. For instance, more than half of the models omitted the explanation of the concept in the muscle input evidence. It might be the first time for the students to think about muscles at the cellular level. Moreover, the muscle contraction process was independent from the cellular respiration mechanism. Therefore, students were not adept at incorporating this concept into their models. Even for models that scored high for interpreting the muscle input concept, the idea was often represented as a
separated side note to describe the relationship between sugar, oxygen, and muscle contraction. This shows that students were not sure how to intertwine this concept into their models.

**The Characteristics of Mechanisms**

The majority of the participating students were able to formulate a mechanistic explanation about an invisible and unperceivable natural phenomenon upon reading the evidence before instruction. They were proficient in speculating sequential processes that organized entities and their interactions in certain locations. This result resonates with earlier studies which pointed out that components and sequences were the most prominent features included in students’ models (Penner et al., 1997; Zangori & Forbes, 2016; Zagori et al., 2017). Students in these earlier studies were all third graders and were not competent in explaining mechanistic processes. More than three quarters of the middle school students in this study were able to express preliminary mechanistic explanations after they examined an array of evidence. This contrast implies that some growth increment about reasoning skills may occur during late elementary or early middle school stage.

Mechanisms refer to a type of systems schema in which entities with specified properties are organized in particular activities to generate continuous changes from the beginning to the end of the process in a certain condition (Craver, 2006; Machamer et al., 2000). This definition guided my analytical approach, and only entities and activities they were involved in the energy generation process were considered included in the mechanistic explanation. This analytical approach made it possible to find that a portion of the models displayed components and activities but did not organize them into a
mechanistic explanation about how certain input entities brought about end products. It was less challenging for students to display entities in models than to incorporate entities into a mechanistic system. The results also indicated that input entities were more likely to be incorporated into the mechanisms than output entities. The predominante incorporation of input entities in the mechanisms could result from students’ considering output entities not relevant to the mechanistic process. The belief that an element is irrelevant to a theory often leads students to exclude its possibility and ignore it accordingly (Kuhn et al., 1992). This phenomenon was observed in the discussion episode demonstrated in Chapter 5 when student S3 from Group A2 gave her rationale for why she thought that it was not necessary to include carbon dioxide in their group model, but it was necessary to include oxygen.

In addition, varied frequencies of incorporation were detected between different input entities in students’ models. The frequency of oxygen participating in the mechanisms was not as high as food and carbs combined. Students’ daily knowledge might be the reason for this performance. For example, it is a common experience that a hungry person feels energized after taking in some food. Also, students may often be reminded by their school coaches to eat high-carb food some time prior to a sport competition. The role that food or carbs plays in the energy producing process could be more compatible with students’ personal experiences, but it is less likely for one to physically feel that energy is recharged on the function of oxygen.

In previous studies, secondary school students expressed the alternative conception that cellular respiration is an exchange of gases in plants’ leaves (Flores, et al., 2003) or is synonymous with breathing in the cells of animals’ respiratory organs
(Alparslan et al., 2003), even after taking high school biology courses. In contrast, none of the pre-instruction models in this study mistook cellular respiration as breathing or gaseous exchange. Nearly half of the models in this study mentioned some aspects of the *cellular organelle* evidence, and more than a third of the models incorporated the core of this evidence and went down to the cellular level to explain the mechanisms. Besides, energy was present in a great portion of the models. These performances might be a confluence of multiple features with respect to the module’s contextualizing instruction.

Contextualizing instruction refers to situating science phenomena within a real world context that takes place outside of the classroom in order to motivate students and activate relevant knowledge (Rivet & Krajcik, 2008). Such designs have been found to promote students’ explanation competence (Kang et al., 2014). Because cellular respiration in human cells cannot be directly experienced in classrooms, phenomena-based evidence that consisted of second-hand empirical data was utilized to provide direct links to the natural phenomenon (McNeill & Berland, 2017). The cellular respiration module was contextualized in a cover story in which two boys who were the same age as the participating students were collecting data before, during, and after they exercised. The boys were doing homework in a running context that attempted to address the same driving question as the participating students were faced with in class. Such a contextualizing design might have helped students project themselves into the boys’ situation and sustain their attention and interest (Rivet & Krajcik, 2008). The relevancy between the exercise scenario and students’ personal lives might have made the inquiry meaningful for students to figure out why one boy was able to outperform the other, and
from there students might find that the answer could be found in the assortment of evidence.

Besides the utilization of contextualizing instruction, the evidence was also designed to facilitate reasoning during model building. The evidence covered a range of familiar concepts as well as novel information. The familiar concepts could have enabled students to draw on some initial ideas to make inference from their preexisting knowledge about oxygen, carbon dioxide, glucose, and temperature. The novel information oriented students’ attention to the unperceivable cellular level and guided them to reason about the possible role that mitochondria played in the cellular respiration process. Following this approach to design, the muscle cell contraction evidence delivered a new message that the rate of cell contraction was related to oxygen and glucose, and the cellular organelle evidence made mitochondrion the only variable dissimilar between two boys at the cellular level. These design features might have set the seventh graders in the right direction as they worked on this topic and facilitated their reasoning as they created models.

**Four Types of Models**

Four major types of coherent models were identified amid those that incorporated mechanisms to explain cellular respiration. Among the four types, the most common one was the cellular transformation model. Students who constructed this type of model merged their prior knowledge and their interpretations of evidence. In addition to the process of cellular respiration, they included model elements on how glucose and oxygen went through the body system and eventually interacted in the cells to produce energy. The elaboration of the digestion process was supplemental information and did not
compromise the integrity of the mechanism of cellular respiration. Without the knowledge of molecular and protein interactions, students’ models bottomed out at the cellular organelles and were not able to generate more complicated mechanisms. Yet the experience of making models at the cellular level may form stepping stones for future learning (Mohan et al., 2009). A prospective learning benefit may emerge when students later acquire advance knowledge in molecular biology ahead in the learning progression of science education.

The second most popular type of model was the digestive model, which had a simple and intuitive claim that food gets digested and becomes energy. Students who crafted such models appealed to personal knowledge about digestion and ignored the cellular organelle evidence. When students encounter anomalous data that contradict their original belief, one of the reactions can be to ignore them (Chinn & Brewer, 1993). In response to how the body gets energy, these students kept intact their digestive theory and showed no trace of incorporating the cellular organelle evidence in their models.

Students had different reactions to the oxygen, carbon dioxide, and body temperature evidence, though. These entities also conflicted with their knowledge of digestion, but in order to make their models consistent with the evidence, they found ways to incorporate them into the digestion process, a reaction categorized by Chinn and Brewer (1993) as making peripheral theory changes. The intuitive ideas that animals get energy from food through digestion and that the purpose of respiration is to provide oxygen and to release carbon dioxide are so robust that they have been found prevalent even among in-service biology teachers (Sanders, 1993). It is not wrong to think that energy is transferred from food to humans; however, cellular respiration comprises
chemical reactions that are different from those in a digestive process. To conflate cellular respiration with digestion suggests a conceptual gap at the cellular level.

Non-cellular burning models stated that energy was released after a “burning process” in the stomach, or no location was specified. This type of model was found only in classes that received the three example models. The example models were intended to prepare students with some initial ideas by offering the “burning process” sketches that stood for the to-be-solved cellular respiration mechanism. But the example models did not help students expand their reasoning beyond the term “burning process,” nor did they help them think about possible activities that might occur during this “burning process.” Students who created this type of model seemed to model the vocabulary rather than the underlying concepts. They grabbed the keyword “burning process” in the material and either presented it alone in the model or integrated it with their prior knowledge about digestion. The possible mechanism that the keyword encompassed was overlooked. Furthermore, instead of thinking about how to correct the initial model, they may have built their explanation upon an incorrect example model. An embedded scaffold aiming to assist students with modeling on top of a given foundation may have been ineffective in advancing their reasoning. Zangori et al. (2015) found that in general, third graders in a scaffolded condition who had to draw their explanations about the water cycle based on a given picture expressed the sequences but did not articulate the mechanisms. In contrast, the students who had to draw models on a blank paper combined their scientific knowledge with alternative conceptions to explain the mechanistic processes of the water cycle. Scaffolds that are embedded in the modeling task before students start off may not be effective in supporting modeling; it may even adversely constrain students’ reasoning.
Another possible reason for students to model the vocabulary rather than the concepts is that the word burning by itself might already be regarded as a mechanistic process. When students saw the word “burning” in the example models, they might regard it as the target process itself rather than only a sketch to stand for the target process. In the example models, “burning process” was what Craver (2006) called a “filler term,” which suggests that a potential activity is taking place in the process but does not explain how that activity is carried out. When students perceive the filler term as the mechanism, teachers may need to inform them of the metaphorical function of the word and encourage students to speculate about the target mechanism behind the veil of the word.

A small fraction of models presented an exercise explanation in which energy was produced during or after muscle contraction and did not point out where this process took place. This explanation could have derived from the personal experience that one could feel more energized after exercising. Similar to those who constructed the digestive models, students who constructed the exercise models may also have been led by alternative theories to ignore the cellular organelle evidence, which impeded their further reasoning at the cellular level. The oxygen, carbon dioxide, and heat were well woven into this type of model though.

A common belief among primary, secondary, and even college students is that cellular respiration is equivalent to external respiration (Anderson et al., 1990; Driver et al., 1994; Seymour & Longden, 1991). This alternative belief was not observed in students’ pre-instruction models in this study. Instead, this study provided a new result: a great percentage of the models integrated details about digestion and gaseous exchange into the processes of energy generation in light of the evidence. Regardless of the types of
mechanisms incorporated into the models and whether or not there was a mechanism, students’ models displayed the conspicuous phenomenon that the prior knowledge was not neglected, nor was the information embedded in evidence ignored. Instead, these elements were fused together in students’ models.

**Small Group Processes**

Small group discussion in a collaborative setting has been a topic of research for decades. Recent studies conducted by Chi and her colleagues suggested that more interactive interactions between peers result in better cognitive growth (Chi, 2009; Chi & Menekse, 2015; Chi & Wylie, 2014), but the findings of this study pointed out a caveat: the extent of interactive collaboration between peers is not the sole factor determining the cognitive gains or lack thereof. In a model-based inquiry environment, engaging in discourse operations such as reasoning about the mechanisms and integrating different evidence in a less interactive conversation are more productive than only talking about the way to present a model and cogitating one piece of evidence at a time in an interactive conversation. In this study, Group B2 was a thoroughly interactive group yet their model was of lesser quality than those of three other groups. Their group members’ interactive exchanges did not enable productive reasoning as a team. On the contrary, Group A1 did not engage in a more interactive dialogue pattern than their peers, but their reasoning about multiple evidence all at once and their mechanistic explanation put Group A1 ahead of the rest.

Besides discourse operations and dialogue patterns, there could be other potential explanations to account for why four groups’ performances varied. The unequal group size was one of them. Groups in School A consisted of two members, while groups in
School B consisted of four members. Group size in small group discussions has been suggested to yield variations in learning accomplishments. Alexopoulou and Driver (1996) found that students who discussed a physics topic in pairs were more likely to dispute over an absolutely right or wrong answer, whereas those in quartets respected different opinions and tended to negotiate the meaning of concepts. They concluded that groups of four could develop better reasoning skills than those in groups of two. In this study, groups’ conversational dynamics only partially coincided with Alexopoulou and Driver’s observation. The pairs in Groups A1 and A2 occasionally challenged each other’s opinions, yet the disagreements did not deter them from making productive exchanges. The quartets in Groups B1 and B2 had harmonious discussions, yet this did not help them advance their modeling performance and the interpretation of the evidence. Although School A students’ pre-lesson models in average outperformed those of School B, there is not enough evidence from only four groups to claim that working in a group of two is more effective than in a group of four or to suggest whether group size is in relation to their discrepant models. Group size may have played a role on some aspects, but it is beyond this study’s scope to determine its influence.

Other possible explanations for the groups’ contrasting performances include whether group members entertained different ideas, how they resolved the disagreements, and whether they were aware of knowledge gaps. Carrying conflicting opinions into a group discussion has been found to promote learning (Howe et al., 1992), but the conflicting opinions do not lead to productive outcome if they are not evaluated and are simply ignored or rejected (Chan, 2001). Small groups who are aware of their knowledge gaps and treat contradictory scientific information as a problem to be explained have a
higher chance to succeed in conceptual growth (Chan, 2001). Such arguments are not consistent with the four cases in this study. The students in high-performing Group A1 did not express conflicting opinions, but they were conscious of their challenge that they could not make sense of the *cellular organelle* evidence. Medium-performing Group A2 occasionally encountered different opinions when they interpreted evidence, and they relied on personal experiences to resolve them. But no sign of awareness of a knowledge gap was found in their discussion. Another medium-performing Group B1 did not experience any conflicting ideas between peers. A few members explicitly expressed their knowledge gaps when they tried to interpret evidence, and they immediately received instructions from others what to write down without exchanges of ideas. The lower-performing Group B2 encountered different opinions from different members when they discussed what evidence meant. They appeared to be confident about their evidence interpretation and modeling and did not perceive any potential challenges in the inquiry activity. Based on the behaviors observed from the four groups, it is hard to judge whether having conflicting ideas, treating conflicting ideas as a problem to solve, and being explicit of knowledge deficiency contribute to students’ performances on modeling.

It is likely that how students perceived the model-based inquiry task influenced their modeling performances. High-performing Group A1’s searching for the meaning of the anomalous evidence, seeking connections between evidence, and the zigzag modeling process mirrors the inquiry practices scientists do in the research realm. They were engaging in what Jiménez-Aleixandre et al. (2000) called “doing science.” When students are “doing science,” they build knowledge and assess the knowledge they have produced
(Hutchison & Hammer, 2010). They frame what they are doing as making sense of a scientific phenomenon (Hutchison & Hammer, 2010). Lower-performing Group B2 might have thought that they were doing science, too. They assessed the evidence from multiple viewpoints, and they coordinated their model with selected evidence. Unfortunately, their model was rather describing a soccer game than explaining a scientific phenomenon, and they declined to improve its quality. In contrast, medium-performing Group B1’s discussions showed that they entertained an epistemic frame that could be called “doing the lesson” (Jiménez-Aleixandre et al., 2000). When students are “doing the lesson,” they perform for the teacher and treat the inquiry activity as an assignment (Berland & Hammer, 2012). They follow the rules of the task, fulfill the teacher’s expectations, and just get the work done. Indifference was the epitome of Group B1’s attitude. If someone came up with an interpretation of evidence, the rest of the team hastily copied it down to their sheets without further evaluation of the idea heard. They completed the model together by distributing different parts of the task amongst themselves, but they did not discuss how to build a model or what the model was for. When Mr. Britten asked them to add overlooked information, they did accordingly. In the end, they met the objective—they filled out the table of the study worksheet and wrote down a model that met the teacher’s standard. Group A2 was made up of two members who had contrasting attitudes. S4 wanted to get involved to a minimal extent; he only wanted to pass time. S3 was in charge of evidence interpretation and model construction, but she did not express her reasoning so it was difficult to tell how she framed the inquiry activity. The analyses of the group conversations imply that how students frame their
scientific inquiry may not be the sole predictor of the quality of the artifacts they
collect during an inquiry activity.

With respect to the discussion above, the most probable factor that accounts for
how the four groups’ performances vary might be whether they took multiple evidence
into consideration and engaged in mechanistic explanation. Students’ modeling
performance might be predicted by the discourse operations used by a small group in
their discussions. Science teachers who implement inquiry lessons may consider holding
reflective discussions with students and addressing the importance of thinking about
different evidence all at once and reasoning about the invisible mechanisms behind
phenomena. These two discourse operations may play a significant role to enhance their
learning outcomes.

Teacher Prompts

Students in School A significantly outperformed their peers in School B on all of
the assessments addressing the research questions, despite coming from a school that was
lower-performing on state tests scores. This phenomenon led my scrutiny to the
instructional strategies implemented by Ms. Addison and Mr. Britten. The major
difference between the two teachers’ scaffolding was manifested in their prompts when
students were interpreting evidence and building models. Building a mechanistic model
requires three arenas of reasoning, including considering the invisible level, identifying
the relevant elements, and coordinating the elements over space or time at the invisible
level (Kris et al., 2019). Ms. Addison guided students to think below the scale of organs
and contemplate the cellular entities, the activities, and the place in a cohesive sequence.
Her prompts were specific and covered the reasoning essentials for giving rise to
mechanistic explanation. In contrast, Mr. Britten’s prompts were generic, centering on one general principle, which was to make models fit the evidence, and no further hints were given to attend to the evidence concerning cells and cellular activities. The contrastive modeling outcomes between the two teachers’ classes suggest that Ms. Addison’s specific prompts could be a more effective scaffold than Mr. Britten’s generic prompts.

Prior research has yielded mixed results on the effects of specific versus generic prompts. Specific prompts were found to be more effective than generic prompts in helping middle-school students write scientific explanations that included claims, evidence, and justifications of evidence (McNeill & Krajcik, 2009). Such findings conflicted with another study in which middle school learners who received generic prompts developed more coherent understandings of the science project they worked on (Davis, 2003). In both of these previous studies, the prompts were presented in written form before students generated responses. In this study, the prompts occurred in conversations between the teachers and small groups, and students’ ability to create mechanistic models was better when they received specific prompts rather than generic prompts from the teachers. My conjecture is that the nature and context of the task may affect the efficacy of either specific or generic prompts. In the aforementioned studies, students’ achievements were assessed through students’ writing of scientific explanations, whereas in this study students created models, which were evaluated by both the coordination of models and evidence as well as the incorporation of mechanisms into models. Constructing a mechanistic explanation while reflecting on novel evidence to account for abstract processes is challenging, so that even the high-performing group
A1 during the second semester of an inquiry-based research project still experienced difficulty in making sense of the *cellular organelle* evidence and in abandoning the preexisting conceptual models of breathing, digestion, and metabolism. The reasoning skills needed to develop a mechanistic explanation (see Krist et al., 2019, for details) are undoubtedly learned, and most challenging strategies probably require extensive practice and coaching to reach proficiency. Specific prompts may thus be necessary in this sense to provide metacognitive assistance and direct learners’ reasoning to the components of a mechanism and to hypothesize the seemingly unfathomable cellular-level processes. The fact that Ms. Addison’s students outperformed Mr. Britten’s students not only suggests that specific prompts might be more effective for student learning when such prompts were delivered by the teachers through conversation during the practice of inquiry, but also indicates that students from schools that perform lower on standardized tests can do well in a model-based inquiry environment with the appropriate scaffolds.

**Limitations**

A general limitation of this study is that it is difficult to be sure whether the models that students constructed were completely based on the evidence even when a model was consistent with all relevant evidence. A pretest questionnaire to collect what students think about cellular respiration before they receive the evidence packet would have helped identify if some contents in a model that derive from prior knowledge are reinforced by the evidence. In addition, a pretest may explain if some model types are reflecting students’ pre-existing conceptual frameworks about the mechanisms of the topic.
Due to inability to link study worksheets to the students who completed them, students’ competence with evidence interpretation had to be indirectly observed from their performance with model-evidence coordination as shown in their models.

Time was another limitation in this study. The total time each teacher spent on the module was different and hence the length of the time each teacher spent on the pre-lecture modeling activity varied. In the teacher’s manual the estimated duration of this module was three periods each lasting about 40 minutes. Mr. Britten stuck with the schedule, but Ms. Addison decided to expand the duration of the module to five periods. Students in Ms. Addison’s classes spent one period discussing the evidence and two other periods building models, while students in Mr. Britten’s classes discussed the evidence and built models altogether within one period. Surely a longer length of time allocated for students to work in groups means there is higher probability for students to come up with good reasoning and build a better model. Although in this study the characteristics of reasoning and the dynamics of interaction of each small group emerged in the early phase of their discussions and made each group distinct, it is possible that Mr. Britten’s students could have read the evidence more carefully, made sense of the evidence better, and built better models if they had more time to work on the project.

Finally, there was no video-recorded data for small group discussions, and therefore not all interactions were captured. When students were writing aloud the action of writing was identified and denoted in the transcriptions, but there was no way to know what they were writing if they erased the initial remarks. Only whatever remained on the poster was analyzed, not the process of their modeling. Sometimes there were pauses during a discussion. When there was a lengthy period of silence, it was usually not
possible to infer what they were doing. Also, it was impossible to know if all group
members were on task when they were not talking.

Conclusion, Implication, and Future Directions

In the research canopy that studies inquiry-based learning in science, the
coordination of evidence and explanations, the incorporation of mechanisms in visual
representations or verbal expressions, and the reasoning courses manifested in peer
discussions are three of the major branches. Each of the three branches has been well
scrutinized in earlier research, but never have they been investigated all together in one
study to see how pre-instruction group dynamics relate to varied modeling performances.
Derived from the quantitative and qualitative analyses on the three aspects, this study
provides a comprehensive understanding of student competences during model-based
inquiry in a natural setting. It also indicates that with appropriate scaffolds and group
reasoning processes, students in all schools can perform well.

Making explanations consistent with evidence has been found challenging for
students in various age groups. Among the studies that have investigated the coordination
of evidence and explanations, one strand has investigated the relationship between
students’ justification of the evidence mentioned in their explanation and the quality of
the explanation (for instance, McNeill et al., 2006), and another strand has inspected the
citations of evidence in students’ written explanation (for instance, Sandoval &
Millwood, 2005). This study extended the second strand of research by evaluating how
middle school students perform on model-evidence consistency in a condition when the
epistemic knowledge of scientific inquiry had been co-constructed in class during
teacher-led activities and discussions before the cellular respiration module was enacted.
The results suggest that after students have been immersed in the epistemic notions and the practice of scientific inquiry through several topical lessons for a prolonged period of time, they are able to construct models highly consistent with relevant evidence and eliminate irrelevant evidence before they obtain any conceptual instruction of the scientific topic.

The results from this study also suggest that creating preliminary mechanistic explanations before instruction is within most middle school students’ capacity. In this study, the concepts abstracted from the evidence as well as prior knowledge held by the students were fused into an integrated account that expressed how entities interacted at some location and unfolded processes step by step. This finding emerged from a fine-grained coding technique that distinguished entities and activities that participated in the mechanistic processes from those that did not. Moreover, this analytic technique examined the quality spread of each mechanistic component and the level at which the mechanistic explanation bottomed out. Mechanistic reasoning has been an increasingly well-probed topic in the field of science education. In some studies, the expression of causal reasoning was counted as an indicator that students had composed a mechanistic explanation (for example, Hammer, 1995). In others, the quality of mechanism incorporated in a model was determined by how many components and sequences were shown and how many links of the cause and the effect was displayed (for example, Zangori et al., 2017). Still others evaluated the quality of a student’s model based on the mechanism incorporated, but did not explain how the evaluation proceeded (for example, Penner et al., 1997). And finally, there were also studies that created an exhaustive coding scheme spanning a wide range of mechanistic reasoning yet provided evidence
from students’ causal thinking to instantiate mechanistic reasoning (for example, Russ et al., 2008). This study delivers a clear definition of mechanism and displays a pragmatic coding scheme to precisely assess each model’s mechanistic features, and accordingly enables the further analysis of the mechanistic types.

Conversations between peers and between student and teacher are two major types of dialogue during model-based inquiry. The topic of peer discussion has been scrutinized by a substantial number of research for over three decades. Among studies that investigated this topic in an inquiry setting, some documented group conversations and reported valuable descriptions of students’ reasoning features but did not explain any frameworks or rubrics by which they analyzed the dialogues (for example, Johnson & Stewart, 2002). Another cohort of studies examined the relationship between the structure of students’ conversation and learning performance. Toulmin’s (1958) argumentation framework has been prevalently adopted to study the structure of reasoning (for instance, Erduran et al., 2004). But Toulmin’s framework was not able to illuminate the social and epistemic aspects of a discussion. On the other hand, Mercer and his colleagues developed categories to explore the socio-linguistic patterns displayed in group discussions (for example, Mercer et al., 2004), and Chi and her colleagues proposed a framework to assess the relationship between cognitively engaged social interaction and conceptual growth (for example, Chi & Wylie, 2014). However, both Mercer’s and Chi’s methods did not aim to analyze the reasoning processes in discourse. This study synthesized the strengths of preexisting frameworks, refined and modified Chi’s framework, and developed two analytic methods to evaluate both the reasoning operations revealed in conversation logs and the groups’ dialogue patterns as shown in
their social interactions. The results suggest that better modeling achievements are associated with how small group members reason about the evidence and mechanisms, more than how interactively they converse during group discussions. Students may build better models when they jointly think about multiple pieces of evidence all at once and reason about the mechanisms. These critical reasoning processes may have to take place before an interactive dialogue pattern can work its magic. It is appropriate for teachers and education practitioners to provide students with pre-discussion training sessions that target the features of a mechanistic explanation as well as the strategies to weigh and consider different evidence altogether, followed by a team building process that reinforces an interactive dialogue pattern between peers.

Another implication of this study is that, with appropriate scaffolds, the students in lower-performing schools can do well and even outperform their peers in higher-performing schools. Specific prompts that direct students’ attention to the invisible mechanisms of a scientific phenomenon are more productive than generic prompts that stress general principles such as the alignment of model and evidence. In this study, the two-way conversational exchanges between the small groups and their teachers provided a rich context to advance the understanding of the influences different prompts had on students’ reasoning processes and on the varied modeling outcomes.

The findings and the frameworks developed in this study suggest many possibilities for future research. First, students’ capability of making models consistent with evidence needs to be further examined in other science topics to ensure that students’ strong performance in this study was not mainly due to their prior knowledge of this topic. In addition, helping students develop epistemic knowledge about inquiry and
modeling can be helpful for model-evidence coordination in a model-based inquiry learning context. This is an area that can be scrutinized further. Another appropriate direction for future research is to apply the mechanism coding scheme to inspect students’ mechanistic reasoning in written passages. By continuing to try this scheme with different forms of data, it will help identify places for modification and increase the reliability of this analytical instrument. In order to fully explore how the dynamics of small group discussion may mediate modeling achievement and to refine the analytical frameworks, it is necessary to try the frameworks in more group discussions both within each teacher’s classes and between different teachers’ classes. Researchers should study groups that have similar discourse operations and different dialogue patterns as well as groups that have different discourse operations and similar dialogue patterns to strengthen the validity of the finding that the effect of discourse operation overrides the effect of dialogue pattern. Finally, an empirical study that randomly assigns half of the classes to specific prompts and the other half to generic prompts from the same teacher and between different teachers may help determine the efficacy of specific and generic prompts during teacher-student conversation in an inquiry-based science classroom.
References


Appendix A

All Pieces of Evidence

Evidence 1

The following table is a comparison of Eric and Roger’s physical condition.

Table (1) Physical Fitness

<table>
<thead>
<tr>
<th></th>
<th>Eric</th>
<th>Roger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height (cm.)</td>
<td>158</td>
<td>161</td>
</tr>
<tr>
<td>Weight (kg.)</td>
<td>52</td>
<td>50</td>
</tr>
<tr>
<td>Exercise</td>
<td>Regularly</td>
<td>Regularly</td>
</tr>
</tbody>
</table>
Evidence 2

This study describes muscle contraction in the presence of varying levels of sugar and oxygen.

Figure (1) Muscle Cell Contraction During Exercise

Microscopic View of Muscle Cells

**Note: When we use muscles, like doing exercise, the muscle cells in the muscle contract (tighten). The harder the exercise, the faster the muscles contract.

Rate of Muscle Cell Contractions

<table>
<thead>
<tr>
<th>Amount of Oxygen inhaled</th>
<th>Amount of Sugar (Glucose) eaten</th>
<th>Rate of Muscle Cell Contraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>+</td>
<td>Low</td>
</tr>
<tr>
<td>High</td>
<td>+</td>
<td>Medium</td>
</tr>
<tr>
<td>Low</td>
<td>-</td>
<td>Medium</td>
</tr>
<tr>
<td>High</td>
<td>+</td>
<td>High</td>
</tr>
</tbody>
</table>
This figure compares the amount of carbon dioxide Eric and Roger exhaled while exercising.

Figure (2) Exercise and CO$_2$ Exhaled (breathed out)
Evidence 4

This table is the comparison of body temperature for Eric and Roger as they ran.

Table (2) Body Temperature During Run

<table>
<thead>
<tr>
<th></th>
<th>10 min.</th>
<th>20 min.</th>
<th>30 min.</th>
<th>40 min.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eric</td>
<td>36.7</td>
<td>36.9</td>
<td>37.4</td>
<td>38.0</td>
</tr>
<tr>
<td>Roger</td>
<td>36.6</td>
<td>36.9</td>
<td>37.2</td>
<td>37.8</td>
</tr>
</tbody>
</table>

**Note: Body temperature in degree Celsius**

Evidence 5

This figure compares the amount of oxygen Eric and Roger inhaled while exercising.

Figure (3) Exercise and O₂ Inhaled (breathed in)
However, for some reason Eric could not keep up with Roger’s running pace after a while. They are in similar physical condition, so what could be the reason? Eric thought that it might have to do with what they ate before running. Eric had a big steak beforehand, and Roger had pasta and a fruit cup instead. Below is the graph describing how long they ran based on different foods they ate.

Figure (4) Low-Carb Runner vs. High-Carb Runner

**Note- Remember Eric and Roger’s assignment was to link their measurements to how our bodies use food and oxygen to get energy and how this happens within our cells.**
Evidence 7

This figure compares the density (amount/cell) of various organelles in Eric and Roger’s muscle cells.

**Figure (5) Cellular Organelles**

<table>
<thead>
<tr>
<th>Organelles</th>
<th>Eric</th>
<th>Roger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleus</td>
<td>normal</td>
<td>normal</td>
</tr>
<tr>
<td>Mitochondria</td>
<td>normal</td>
<td>slightly high</td>
</tr>
<tr>
<td>Ribosomes</td>
<td>normal</td>
<td>normal</td>
</tr>
<tr>
<td>Endoplasmic reticulum</td>
<td>normal</td>
<td>normal</td>
</tr>
</tbody>
</table>
Appendix B

Three Example Models in Version 2

*Three Example Models*